

Elaborazione numerica dei segnali

Fabio Rocca

*Dipartimento di Ingegneria Elettronica
e Informazione
Politecnico di Milano*

Milano, 2010

Indice

Prefazione	xiii
1 Il Campionamento	1
1.1 Introduzione	1
1.2 Ricostruzione di un segnale	6
1.3 Campionamento passa banda	7
1.4 Il filtro antialias	8
2 La trasformata di Fourier discreta	11
2.1 Campionare in tempo e in frequenza	11
2.2 Proprietà della DFT	13
2.3 Filtraggi con la DFT	14
2.3.1 Convoluzioni aperiodiche: sequenze limitate	15
2.3.2 Sequenze illimitate: overlap-save; overlap-add	15
2.4 Antitrasformazione di Fourier discreta	18
2.4.1 Il derivatore e la costante di tempo	20
2.5 Interpolazione con la DFT	22
2.5.1 Inserzione di zeri e repliche spettrali	22
2.6 Il calcolo della DFT: la FFT	25
2.6.1 Le farfalle	27
2.6.2 Occupazione di memoria e <i>bit reversal</i>	28
2.7 Filtraggio di sequenze illimitate	29
2.7.1 Costo computazionale della FFT	29
2.7.2 Segmentazione di sequenze illimitate	30
2.8 Trasformata di Fourier in base 4	31
3 Ricostruzione di segnali continui	33
3.1 Introduzione: sequenze limitate	33

INDICE

3.2	Ricostruzione di polinomi	37
3.3	Interpolatori derivati dal mantentore	38
3.3.1	Interpolatore cubico troncato	40
3.3.2	Interpolatori con filtro ideale troncato	40
3.4	Rigidità e sovraelongazioni	40
3.5	La linea elastica o "spline"	42
3.5.1	Analisi dello "spline"	43
3.5.2	L'equazione dello spline	44
3.5.3	La risposta all'impulso dello spline	46
3.6	Il campionamento irregolare	48
3.6.1	Conversione di frequenze di campionamento	48
3.6.2	Interpolazione irregolare con la funzione <i>sinc</i>	50
4	La trasformata z	51
4.1	Introduzione: sequenze e polinomi	51
4.2	Sistemi lineari discreti	54
4.3	Sinusoidi campionate; zeri	54
4.4	Il piano z	56
4.4.1	Le sequenze nel piano z	57
4.4.2	Zeri a fase minima e fase massima	59
4.4.3	Sequenze ribaltate: $\tilde{A}(z)$	62
4.4.4	Sequenze a fase minima, massima o mista	63
4.4.5	Trasformata di Fourier e trasformata z	64
4.4.6	Proprietà della trasformata zeta	65
4.4.7	Durata di sequenze con varia caratteristica di fase	65
4.5	Poli	66
4.6	Equazioni alle differenze finite	70
4.7	Filtri passa banda	72
4.7.1	Sfasatori puri	74
4.7.2	Oscillatori, divani, splines	77
4.8	Serie geometriche troncate	81
4.9	Appendice: Il calcolo dei residui	83
5	Filtri numerici	87
5.1	Introduzione	87
5.2	Finestre rettangolari e triangolari	88
5.3	Finestre cosinusoidali	90
5.4	Altri tipi di finestre	92

INDICE

5.5	Progetto con la tecnica MINIMAX	93
5.6	Lunghezza dei filtri FIR passa - basso	95
5.7	Filtri IIR	97
5.7.1	Derivazione numerica	97
5.7.2	La trasformazione bilineare	103
5.8	L'invarianza della risposta impulsiva	106
5.9	Sintesi di filtri IIR dall'autocorrelazione	107
5.9.1	Le equazioni di Yule Walker	107
5.9.2	L'autocorrelazione dai coefficienti del filtro	111
5.9.3	Equazioni di Yule Walker in forma matriciale	112
5.9.4	La ricorsione di Levinson	114
5.9.5	Calcolo dei c_n dagli $a_{h,N}$; la verifica di Schur	116
5.9.6	La sintesi del filtro	117
5.10	Filtri da combinazione di sfasatori puri	118
5.11	Filtri di Butterworth, Chebyshev, ellittici	120
6	Campionamento in frequenza	123
6.1	Introduzione	123
6.2	Progetto	128
6.3	Campionamento in frequenza e finestrazione	131
6.4	Effetti del ritardo	131
6.5	Quantizzazione dei coefficienti	135
6.6	Generalizzazioni	137
7	Strutture canoniche e riverberanti	141
7.1	Strutture canoniche di filtri numerici	141
7.2	Quantizzazione dei coefficienti	142
7.2.1	Sfasatori puri	145
7.3	Strutture riverberanti: introduzione	147
7.4	Relazione con i sistemi numerici	150
7.5	Cascata di tronchi di linea	152
7.6	Il filtro a traliccio ricorsivo	153
7.7	La ricorsione di Levinson	156
7.7.1	Aggiunta di tronchi di linea	160
7.8	Calcolo dei coefficienti di riflessione	161

INDICE

8	Filtri polifase	165
8.1	Filtri per decimazione o interpolazione	165
8.2	Decimazione e interpolazione	167
8.2.1	I filtri polifase	177
8.3	Cascata di interpolazione e decimazione	178
8.4	Banchi di filtri	179
8.4.1	La matrice di modulazione	179
8.5	La ricostruzione perfetta	182
8.5.1	Banchi di due filtri	184
8.5.2	I filtri in quadratura a specchio (QMF)	186
8.5.3	Equivocazione e ricostruzione perfetta	187
8.6	Banchi di M filtri	190
8.7	Le trasformate lapped: la MDCT	191
9	Quantizzazione dei segnali	195
9.1	La quantizzazione dei segnali	195
9.1.1	Rumore di quantizzazione dei segnali	196
9.1.2	Retroazione dell'errore di quantizzazione	200
9.1.3	Effetti di trabocco (overflow)	203
9.2	Quantizzazione e DFT	204
9.3	I cicli limite	205
10	Le trasformate di Hilbert e chirp z	209
10.1	Trasformate di sequenze causali	209
10.2	Sfasatori puri di $\pi/2$	212
10.3	Sequenze periodiche e segnali continui	214
10.3.1	Singolarità nella trasformazione di Hilbert	215
10.4	Funzioni a fase minima	216
10.5	Il segnale analitico	218
10.6	La trasformata chirp z	221
10.7	Il segnale chirp	222
10.8	Interpolazione con fattore qualsiasi	224
11	Processi casuali: generalità	227
11.1	Processi casuali discreti	227
11.2	Matrici di covarianza di sequenze filtrate	229
11.2.1	Matrici di covarianza di trasformate	232
11.3	Matrici di Toeplitz: autovalori	237

INDICE

11.3.1 Autovettori ed autovalori di matrici di covarianza circolanti	239
11.3.2 Ancora la densità di probabilità di un vettore casuale	240
11.3.3 Un esempio numerico	241
11.4 Stime di autovalori di matrici di covarianza	244
11.5 Fluttuazione dei corsi dei titoli di borsa	247
11.6 Stima a massima verosimiglianza (Wikipedia)	248
11.6.1 Il metodo	248
11.6.2 Difficoltà del metodo; stima di una media	251
11.6.3 La stima lineare MMSE	252
11.6.4 Stima della potenza di segnali	253
11.7 Le direzioni di arrivo	257
11.7.1 Autocorrelazione di una sequenza spaziale	258
11.8 Appendici	259
11.8.1 Autocorrelazione e spettri per variabili continue	259
11.8.2 Dimensione di variabili continue e discrete	260
11.8.3 Teorema del limite centrale	261
12 Analisi spettrale non parametrica	263
12.1 Introduzione	263
12.2 Analisi spettrale non parametrica	264
12.3 Risoluzione spettrale e dispersione	266
12.3.1 La dispersione della stima spettrale	266
12.3.2 Polarizzazione della stima spettrale dovuta al tron- camento dei dati	267
12.4 Il periodogramma	268
12.5 La stima dell'autocorrelazione	270
12.6 Errori da evitare	274
12.7 Modifiche del periodogramma	275
12.7.1 Smussamento dello spettro	275
12.7.2 Media di periodogrammi di sottosequenze	278
12.7.3 La tecnica WOSA	279
12.8 Sommario	280
12.9 Un esempio	281
13 Analisi Spettrale Parametrica	285
13.1 Introduzione	285
13.2 Analisi tutti zeri (MA)	286

INDICE

13.2.1	Troncamento della funzione di autocorrelazione	289
13.3	Analisi tutti poli (AR)	291
13.4	La stima spettrale AR	294
13.5	Un esempio	295
13.5.1	La stima della funzione di autocorrelazione	295
13.5.2	Stima MA	295
13.5.3	Stima AR	296
14	La predizione lineare	301
14.1	L'errore di predizione	303
14.2	Predizione lineare e analisi AR	304
14.2.1	Allungamento del predittore	305
14.3	La ricorsione di Levinson	307
14.3.1	Vantaggi computazionali	310
14.4	Il filtro a traliccio non ricorsivo	310
14.5	Descrizioni alternative dell'autocorrelazione	311
14.6	Ordine del predittore	314
14.6.1	Risoluzione spettrale	314
14.6.2	Lunghezza dei dati e criterio di Akaike	316
14.7	Condizionamento di \mathbf{R}	317
14.8	Stima con rumore	318
14.8.1	Il caso $\sigma_z^2 = 0$	322
14.8.2	Il caso $\sigma_z^2 \neq 0$	323
14.8.3	Predizione, stima e sbiancamento: un riassunto	323
14.9	La deconvoluzione nel dominio delle frequenze	327
14.9.1	Deconvoluzione predittiva nelle frequenze	330
14.10	Le equazioni Y - W nel dominio degli autovettori	333
15	Stima lineare ottima	337
15.1	Fattore di scala di un vettore noto	337
15.1.1	Effetti del rumore colorato	339
15.1.2	Stima di una costante in rumore bianco	341
15.1.3	Stima di una costante in rumore colorato	341
15.1.4	Stima di una pendenza in rumore bianco	344
15.2	Stime di una frequenza	345
15.2.1	Stima lineare di una rampa di fase	345
15.2.2	Discussione	346
15.2.3	Stima non lineare di una frequenza in rumore colorato	347

INDICE

15.2.4 Effetti della lunghezza limitata dei dati	349
15.3 Stima di processi multidimensionali	350
15.3.1 Stima simultanea di una costante e di una pendenza	352
15.4 La separazione di sorgenti	352
16 Filtraggio adattativo	357
16.1 La tecnica di Widrow e Hoff	359
16.2 Dispersione degli autovalori	362
16.3 Rapidità della convergenza	364
16.3.1 Dispersione del filtro errore \underline{g}_i	364
16.4 Ottimizzazione del passo di aggiornamento	371
16.4.1 <i>Genesi e collasso delle società idrauliche</i>	372
16.5 Un esempio di filtro adattativo	373
17 Le direzioni di arrivo	377
17.1 DOA con il periodogramma	379
17.2 Stima AR: la risoluzione angolare	380
17.2.1 2 sensori, una sorgente	381
17.2.2 N+1 sensori, una sorgente	381
17.2.3 Numero dei sensori della schiera	385
17.3 Reiezione direzionale del disturbo	387
17.4 Stima spettrale di Capon (minimamente polarizzata) . .	389
18 Il segnale vocale	393
18.1 Tecniche di codifica del segnale vocale	396
18.2 Codifica delle forme d'onda	396
18.2.1 La Modulazione Delta	397
18.2.2 L'ADPCM	400
18.2.3 La codifica lineare predittiva	401
18.3 La codifica del segnale musicale (MP3)	403
18.4 Cenni al riconoscimento del parlato	405
19 Segnali bidimensionali	407
19.1 Le singolarità a 2D	407
19.2 Sistemi lineari 2D	410
19.2.1 Le funzioni cilindriche	412
19.3 Le sinusoidi 2D	413
19.4 L'antitrasformazione di Fourier 2D	414

INDICE

19.4.1	Funzioni a simmetria circolare	415
19.5	Proprietà della trasformazione	418
19.5.1	Funzioni cilindriche e separabili	418
19.6	Filtri ideali 2D	420
19.7	Filtri a ventaglio	421
20	Filtri 2D di natura fisica	423
20.1	L'equazione dei potenziali	423
20.1.1	Continuazione in basso e stima lineare	425
20.2	La propagazione di onde	427
20.2.1	Il problema del telerilevamento	429
20.2.2	Le onde piane	430
20.3	Il principio della fase stazionaria	434
20.4	Il <i>chirp</i> , la lente sottile e l'ottica di Fourier	435
20.5	L'equazione della diffusione	440
21	Sequenze 2D	443
21.1	L'equivocazione a 2D	444
21.1.1	Effetti di moiré	445
21.1.2	Il campionamento esagonale	447
21.2	Il campionamento del segnale televisivo	450
21.2.1	Equivocazione spazio - temporale	451
21.2.2	Lo spettro del segnale televisivo campionato	453
21.3	Ricostruzione 2D	454
21.4	Trasformata di Fourier discreta a 2D	458
21.5	Filtraggi con prodotto di trasformate	459
21.6	Filtri 2D	459
21.6.1	La trasformazione di Mc Clellan	460
21.6.2	Filtri per la continuazione di campo	462
21.7	Interpolazione mediante stima lineare	463
21.7.1	Processi casuali a 2D	464
21.7.2	Spettri di potenza a 2D	464
21.7.3	Stima lineare a 2D	465
21.7.4	Stima lineare ed interpolazione	467
21.8	Poligoni di Voronoi e triangoli di Delaunay	468
21.9	Il campionamento a 3D	470

22 La Tomografia	473
22.1 Introduzione	473
22.2 Trattamento nei numeri d'onda	474
22.3 Il sinogramma	479
22.4 La retroproiezione	481
22.5 La retroproiezione filtrata	483
22.5.1 Equivocazione angolare	486
22.5.2 Retroproiezioni e numeri d'onda	487
22.5.3 Proiezioni incomplete o mosse	491
22.6 Tomografia 3D	492
22.7 Tecniche iterative	494
22.7.1 Retroproiezione e gradiente	494
22.7.2 Retroproiezione e trasposizione	495
22.7.3 Interpretazione stocastica della pseudo inversione	497
22.7.4 Interpretazione tomografica del gradiente	498
23 Tecniche di ottimizzazione	501
23.1 Il gradiente	501
23.2 Il gradiente coniugato	503
23.2.1 Ortogonalità coniugata e incorrelazione	503
23.2.2 Connessione con la ricorsione di Levinson	506
23.3 Gradiente e gradiente coniugato	509

Siamo lieti che l'autore abbia gentilmente concesso
di offrire in download questo eccellente materiale presso
teoriadeisignali.it
dove è possibile trovare altre risorse su argomenti correlati

Prefazione

Il corso di "Elaborazione numerica dei segnali" viene tenuto per gli allievi del corso di laurea in Ingegneria delle Telecomunicazioni al Politecnico di Milano dal 1985. Gli allievi hanno seguito in precedenza un corso sui segnali, in cui vengono studiati i sistemi lineari, la trasformazione di Fourier continua e discreta, i principi di teoria delle probabilità ed i processi stocastici. Inoltre, nel successivo biennio della laurea specialistica in Telecomunicazioni sono disponibili molti altri corsi dove vengono sviluppate con maggiore respiro le possibili applicazioni, siano esse dedicate alle immagini ed al segnale audio, al telerilevamento radar, ai segnali geofisici. La verifica viene fatta con un esame scritto; il compito a metà semestre copre il programma fino alla stima spettrale esclusa.

Nel testo, ho cercato per quanto possibile di partire da considerazioni elementari dando immagini intuitive (*spesso scritte in corsivo*) di molti tra gli argomenti trattati. Certamente la trattazione analitica è una guida più sicura e la via intuitiva non è sempre la più semplice, ma poi è più facile da memorizzare. Mi auguro poi che lo studente possa avere strumenti di tipo ingegneristico più che matematico per comprendere le relazioni intercorrenti tra le ipotesi presunte ed i risultati ottenuti e possa valutare cosa può accadere nei frequenti casi in cui queste ipotesi non sono più attendibili.

Il corso è di prevalente natura metodologica, anche se vengono sempre proposte delle applicazioni.

Questo testo è stato fatto per essere un ausilio alla didattica in aula, piuttosto che uno strumento per l'autodidattica. In questo caso, consiglio di cominciare con i libri di McClellan o di Mitra, inseriti tra i rifer-

PREFAZIONE

imenti bibliografici, ricchi di esempi *MATLABTM*. Il testo di Proakis e Manolakis è piú completo, ma può risultare piú difficile ad una prima lettura.

In questo testo non si troveranno esercizi perchè questi, insieme a complementi e a numerosi testi d'esame sono disponibili sul sito

http://risorse.dei.polimi.it/dsp/courses/ens_l1/index.htm

Qui, gli studenti troveranno indirizzi di posta elettronica per semplificare l'interazione docenti/studente, esercizi vari con i relativi codici *MATLABTM*, data e luogo degli esami, esiti, soluzioni dei compiti, e materiale utile in generale.

Aggiungo in fine una breve bibliografia commentata perchè lo studente possa trovare riferimenti completi per ogni argomento del corso.

Bibliografia.

- Mitra, S., 1998, *Digital Signal Processing, a computer - based approach*, McGraw-Hill, pp 864. Libro elementare, ricco di esempi basati su *MATLABTM*.
- Mc Clellan, J., Schafer, R., Yoder, M., 1998, *DSP first, A Multimedia Approach*, Prentice Hall, pp.514. Libro molto elementare, con esempi basati su *MATLABTM*.
- Proakis J., Manolakis D., 1995: *Digital Signal Processing*, Prentice Hall, 947 pp. e quattro appendici. Libro introduttivo didattico, semplice ma molto diffuso.
- Mitra S., Kaiser J., 1993: *Handbook for Digital Signal Processing*, Wiley Interscience, 1168 pp. Completo manuale che copre ampiamente il testo.
- Fliege, N., 1994, *Multirate Digital Signal Processing*, J. Wiley & Sons, pp. 340. Un buon libro sui filtri polifase e sui problemi di decimazione e interpolazione.
- Ingle, V., Proakis, J., *Digital Signal Processing Using MATLABTM*, Thomson, 2007: Un libro sulla prima parte del corso, con un gran numero di esercizi.

Capitolo 1

Il Campionamento

1.1 Introduzione

I segnali naturali sono definiti per valori continui della variabile tempo o spazio: è possibile misurare il loro valore in ogni istante o posizione, in un intervallo assegnato. Volendo fare un'elaborazione è necessario effettuare un **campionamento** e cioè passare da una funzione definita su un asse (ad esempio spaziale) continuo, ad una sequenza di numeri corrispondenti alle misure del segnale in posizioni assegnate (il segnale campionato). L'operazione di campionamento è corretta se è possibile ricostruire il segnale continuo a partire dai campioni misurati. E' ovvio che in generale ciò è impossibile.

Per fissare le idee, si consideri un sistema di allarme antincendio (le dimensioni possono essere qualsiasi: una foresta, un edificio, un sistema elettronico); è necessario campionare e controllare le temperature ovunque per lanciare l'allarme quando è necessario. Se avessimo a disposizione un solo ambiente, illimitato, sarebbe ragionevole spaziare i sensori di temperatura ad intervallo regolare, dipendente dalle dimensioni dei focolai di incendio che si vogliono poter rivelare (l'ordine di grandezza dell'intervallo di campionamento sarà il chilometro per le foreste, il decimetro per i sistemi elettronici).

Usualmente, si suppone di dover campionare un segnale definito su un dominio spaziale (o temporale) illimitato e di essere in condizioni di stazionarietà, e cioè in una situazione in cui le caratteristiche statistiche del segnale studiato non cambino con lo spazio (o con il tempo). Sia

CAPITOLO 1 IL CAMPIONAMENTO

$x(t)$ il segnale da campionare (nel caso precedente, la temperatura x , funzione della posizione t del sensore); il segnale sia reale e supponiamo di volerne fare misure ad intervallo T ; cioè si supponga nota la sequenza x_n ; $0 \leq n \leq N - 1$:

$$x_n = x(nT + T_o) \quad (1.1)$$

Il segnale continuo $\hat{x}(t)$ ricostruito a partire dalla sequenza dipenderà in generale da T_o e cioè dalla fase del reticolo di campionamento (nel caso del sistema antincendio, i sensori di temperatura siano disposti a distanza regolare T , e sia T_o l'ascissa del primo sensore). Il campionamento è corretto quando la ricostruzione non dipende da T_o (vogliamo valutare correttamente la consistenza di un incendio anche se il focolaio capita tra due sensori).

Gli errori della ricostruzione sono chiamati equivocazione (aliasing) e intuitivamente si possono dividere in due categorie: errori dovuti alla mancata registrazione di un evento, che è caduto "tra due misure" (*l'incendio non osservato*) o viceversa misure di eventi che, pur essendo di durata molto limitata, vengono interpretati come eventi lunghi circa T (vedi fig. 1.1) (si pensi al detto popolare "prendere lucciole per lanterne" e cioè credere grande (una lanterna) una cosa piccola (la lucciola)). Un modo ovvio per controllare il corretto campionamento di un segnale è nel verificare che il segnale continuo ricostruito non dipende dalla fase (l'istante iniziale) della trama del campionamento. Se si vogliono evitare fenomeni di equivocazione sarà opportuno che anche gli eventi di durata minima siano campionati almeno due volte perchè siano delimitati (nel tempo o nello spazio) e ne venga evitata sia la perdita che l'esagerazione. Diventa quindi essenziale garantire che nel segnale misurato ad intervallo T siano assenti eventi brevi che creerebbero equivocazione. Questo si ottiene mediando (smussando) il segnale prima del campionamento e cioè **prefiltrando il segnale da campionare** (filtraggio antiequivocazione o antialias); in altre parole, lo strumento la cui misura è destinata ad essere campionata ad intervallo T deve essere "sensibile" non al valore locale del segnale, bensì al valore "medio" del segnale in quell'intervallo.

In questo modo l'errore di equivocazione viene eliminato sistematicamente. Non sempre ciò è facile o possibile. Ad esempio, sarebbe facile far equivocare il sistema d'allarme antiincendio considerato precedentemente, anche se questo fosse complesso; basterebbe accostare un fiammifero ad un sensore. Infatti in questo caso il campione di temperatura

1.1 INTRODUZIONE

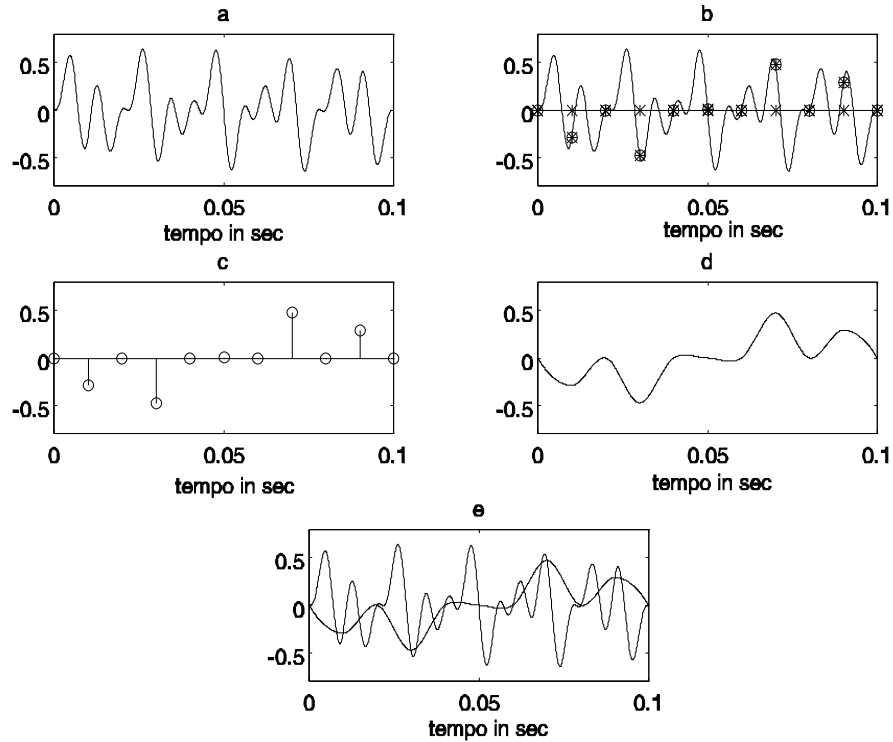


Figura 1.1: Errori di ricostruzione di un segnale dovuti ad un campionamento insufficiente

misurato dal sensore avrebbe valore elevato e molto differente da quello medio dell'ambiente; tuttavia, se vi è un solo sensore per intervallo, il sistema ricostruisce da quella misura la temperatura di tutto l'intervallo, equivoca e lancia l'allarme.

Nel caso in cui si operi su segnali funzione del tempo, il filtraggio anti equivocazione (antialias) è facile. Infatti è implicita la disponibilità di tutti i valori del segnale, per qualsivoglia istante di misura. Non è affatto così se si considerano variabili funzione dello spazio: in questo caso ogni campione del segnale comporta il costo di una postazione di misura. Il filtraggio anti equivocazione deve essere fatto usando strumenti opportuni che fanno a basso costo la media del segnale nell'intervallo; non deve essere fatto campionando fittamente (il che sarebbe costosissimo) facendo seguire eventualmente l'operazione di media e il successivo cam-

CAPITOLO 1 IL CAMPIONAMENTO

pionamento a frequenza inferiore (sottocampionamento). Una misura è in genere ben più costosa che non la sua memorizzazione.

Passando ad un'analisi quantitativa, se si campiona un segnale ad intervallo T a partire da un istante qualsiasi T_0 , sarà necessario, per evitarne l'equivocazione, rimuovere tutte le componenti a frequenza superiore alla frequenza di Nyquist f_N , metà della frequenza di campionamento f_c ,

$$f_N = \frac{f_c}{2} \quad (1.2)$$

con un filtro opportuno. Infatti il segnale campionato è equivalente ad una sequenza di impulsi ideali con ampiezze proporzionali a quelle dei campioni $\{x_n\}$, equispaziati nel tempo (fig.1.2):

$$x_c(t) = x(t) \sum_{n=-\infty}^{+\infty} \delta(t - nT - T_0) = \sum_{n=-\infty}^{+\infty} x(nT) \delta(t - nT - T_0) \quad (1.3)$$

Moltiplicare nel dominio del tempo il segnale originario per il pettine di impulsi equivale a fare la convoluzione delle trasformate nel dominio delle frequenze; la trasformata del pettine di impulsi spaziati di T è un pettine di impulsi nelle frequenze, spaziati di $f_c = 1/T$. In formule si ha, indicando con la freccia il passaggio nel dominio coniugato delle frequenze:

$$\sum_{n=-\infty}^{+\infty} \delta(t - nT - T_0) \leftrightarrow \frac{1}{T} \sum_{m=-\infty}^{+\infty} e^{-j2\pi m f_c T_0} \delta(\omega - 2\pi m f_c); \quad (1.4)$$

La trasformata di Fourier di $x_c(t)$ è quindi:

$$X_c(\omega) = \int_{-\infty}^{+\infty} x_c(t) e^{-j\omega t} dt = X(\omega) * \frac{1}{T} \sum_{m=-\infty}^{+\infty} e^{-j2\pi m f_c T_0} \delta(\omega - 2\pi m f_c) = \quad (1.5)$$

$$= \frac{1}{T} \sum_{m=-\infty}^{+\infty} e^{-j2\pi m f_c T_0} X(\omega - 2\pi m f_c) \quad (1.6)$$

Con $X(\omega)$, $X_c(\omega)$ sono rispettivamente indicate le trasformate di Fourier di $x(t)$, $x_c(t)$ e con $*$ il prodotto di convoluzione. Le repliche si sommano con sfasamenti proporzionali a T_0 .

1.1 INTRODUZIONE

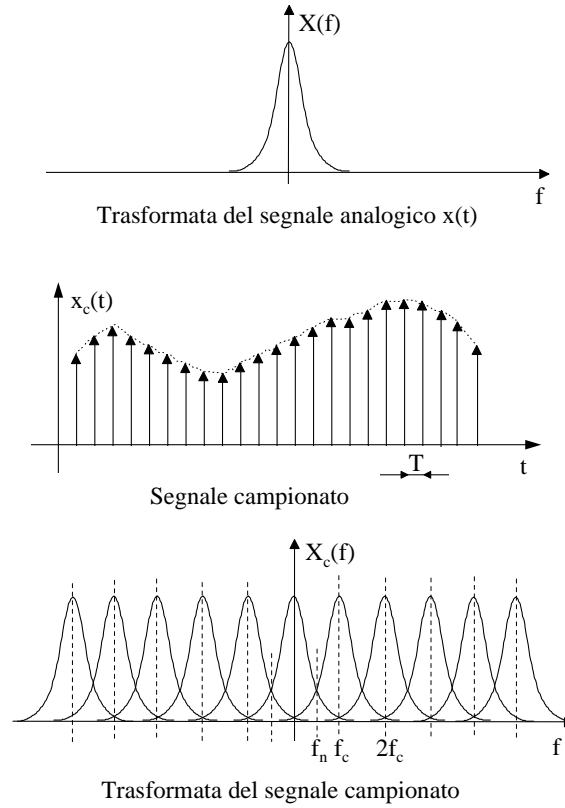


Figura 1.2: Relazione tra trasformata del segnale continuo e trasformata del segnale campionato

Rimuovendo da $x(t)$ tutte le componenti a frequenza superiore a quella di Nyquist, si impedisce la sovrapposizione, nel dominio delle frequenze dello spettro originale con le sue repliche centrate a $m f_c$; $m \neq 0$ e quindi si permette la ricostruzione.

Un altro tipico effetto dell'equivocazione è il fenomeno per cui una sinusoide di frequenza f qualsiasi, se campionata ad intervallo T qualsiasi, può essere ricostruita (equivocando) come una sinusoide della stessa ampiezza ma a frequenza f_{eq} , inferiore alla frequenza di Nyquist.

$$f_{eq} = f - \frac{m}{T}; \quad 0 \leq f_{eq} \leq f_N \quad (1.7)$$

Per rendersi conto del fenomeno, basta pensare di illuminare con una luce

CAPITOLO 1 IL CAMPIONAMENTO

regolarmente intermittente un disco metà bianco e metà nero che ruoti a velocità angolare costante. Se due lampi di luce successivi mostrano il disco nella stessa posizione, l'angolo di rotazione può essere stato nullo o un qualsiasi numero intero di giri m ; quindi si ha incertezza sulla velocità angolare effettiva del disco, a meno che non si abbiano delle informazioni in più, ad esempio sull'intervallo in cui si prevede sia la velocità di rotazione; questo punto sarà discusso tra poco nel paragrafo sul campionamento passa banda.

1.2 Ricostruzione di un segnale

L'operazione di ricostruzione di un segnale continuo dalla sequenza dei suoi campioni (interpolazione) verrà analizzata nel prossimo capitolo; per ora è sufficiente osservare che dallo spettro $X_c(\omega)$ sembra semplice ottenere lo spettro del segnale ricostruito $\hat{X}(\omega)$ rimuovendo con un filtro ideale passa basso tutte le repliche di $X_c(\omega)$ centrate alle frequenze mf_c ; $m \neq 0$ (fig. 1.3). Che la cosa non sia poi così facile lo si vedrà in

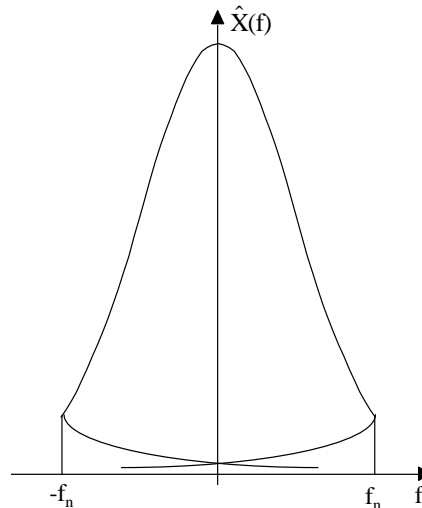


Figura 1.3: Spettro del segnale ricostruito, ottenuto rimuovendo le repliche dello spettro $X(f)$ con un filtro ideale

seguito; qui basti osservare che il filtro ideale ha una risposta all'impulso

1.3 CAMPIONAMENTO PASSA BANDA

di durata illimitata; se la sequenza a disposizione é di lunghezza limitata, come sempre accade, dovremo risolvere il problema delle "condizioni ai limiti" e cioé cosa ipotizzare al di fuori dell'intervallo in cui la funzione é nota, per effettuare una ricostruzione "corretta".

1.3 Campionamento passa banda

Nel paragrafo precedente non é stata fatta alcuna ipotesi sullo spettro di $x(t)$ tranne che esso occupi frequenze inferiori a f_N , condizione necessaria per ottenere una corretta ricostruzione; inoltre, si é supposto $x(t)$ reale. Generalizzando invece il caso precedentemente introdotto (il disco rotante) del campionamento di una sinusoida a frequenza maggiore della

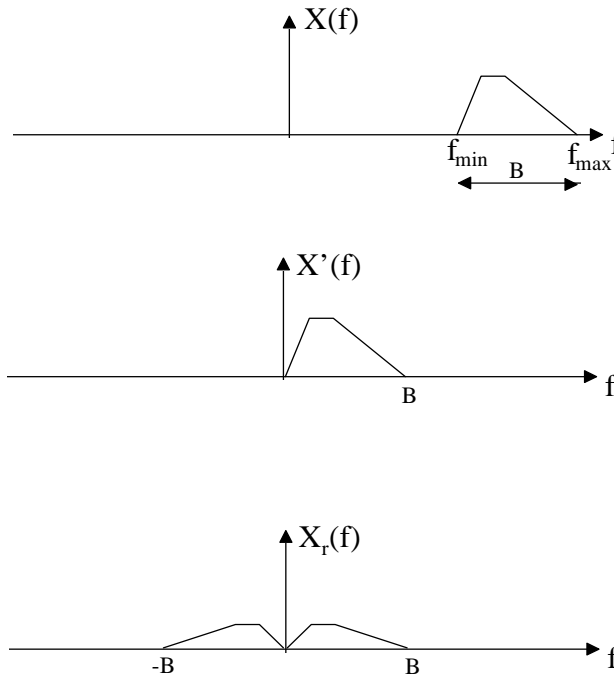


Figura 1.4: Spettri relativi al campionamento passa-banda

frequenza di Nyquist che ruota di un angolo $\varphi > \pi$ tra un campione ed

CAPITOLO 1 IL CAMPIONAMENTO

il successivo, supponiamo che $x(t)$ sia un segnale complesso che occupi una banda B compresa tra le frequenze $B \ll f_{min} \div f_{max} = f_{min} + B$; (lo spettro non occupi le frequenze negative).

$$\begin{aligned} X(f) \neq 0; & \quad f_{min} \leq f \leq f_{max}; \\ f_{max} - f_{min} = B; & \quad f_0 = \frac{f_{max} + f_{min}}{2} \end{aligned} \quad (1.8)$$

E' ovvio che usare come frequenza di campionamento $2f_{max}$ sarebbe uno spreco; possiamo operare in vario modo. Se conosciamo f_{min} e moltiplichiamo il segnale $x(t)$ per la sinusoidale complessa $e^{-j2\pi f_{min}t}$ otteniamo un segnale complesso $x'(t)$ il cui spettro é contenuto tra 0 e B e non ha componenti a frequenze negative, come già osservato.

$$X'(f) \neq 0; \quad 0 \leq f \leq B \quad (1.9)$$

Lo spettro della parte reale di questo segnale occupa le frequenze da $-B \div B$; lo spettro della parte immaginaria occupa lo stesso intervallo di frequenza; sommato a quello della parte reale, risultano eliminate le componenti spettrali a frequenze negative.

Campionando ad intervallo $1/2B$ la parte reale e cioè usando $2B$ campioni per secondo, otterremo una sequenza da cui potremmo fare la ricostruzione del segnale originale (fig. 1.4).

Similmente, potremmo centrare lo spettro del segnale convertito sull'origine e cioè moltiplicare il segnale $x(t)$ per la sinusoidale complessa $e^{-j2\pi f_0 t}$ e ottenere un segnale complesso $x''(t)$ di banda $B/2$

$$X''(f) \neq 0; \quad -\frac{B}{2} \leq f \leq \frac{B}{2} \quad (1.10)$$

Questo segnale è da campionare con B campioni complessi al secondo; in un modo o nell'altro, si vede che un segnale di banda B si traduce in $2B$ numeri reali al secondo. E' tuttavia necessario, per potere fare correttamente il campionamento passabanda, conoscere a priori l'intervallo $f_{min} \div f_{max}$.

1.4 Il filtro antialias

In questa sezione, discuteremo brevemente del dimensionamento di un filtro antiequivocazione. Supponiamo di avere un segnale sinusoidale a

1.4 IL FILTRO ANTIALIAS

pulsazione ω , da campionare ad intervallo T .

$$x(t) = e^{j\omega t}; \quad x_n = e^{j\omega nT}$$

Ovviamente, il filtro antiequivocazione ideale è:

$$h(t) = \frac{\sin(\pi t/T)}{\pi t/T}$$

$$H(\omega) = T; \quad |\omega| < \pi/T$$

di lunghezza illimitata ma di difficile adozione nella pratica; infatti, per determinare un campione del segnale, si dovrebbe conoscerne l'andamento per tutti i valori del tempo. Il più semplice filtro antialias è il mantenitore, che ha bisogno dei dati solo nell'intervallo di durata T , :

$$h_0(t) = 1; \quad |t| < T/2$$

$$H_0(\omega) = T \frac{\sin(\pi f T)}{\pi f T}$$

questo filtro tuttavia attenua bene solo le componenti del segnale a frequenza $1/T$ e multiple e non, come sarebbe necessario, tutte le componenti a frequenze al di sopra della frequenza di Nyquist, $1/2T$. Un filtro più accettabile è un filtro gaussiano

$$h_g(t) = e^{-t^2/2\sigma^2}$$

$$H_g(\omega) = \sqrt{2\pi}\sigma \times e^{-\omega^2\sigma^2/2}$$

di lunghezza limitata e di banda limitata. Per avere un dimensionamento, si può pensare di imporre un'attenuazione di 40dB alla frequenza di campionamento:

$$e^{-\omega^2\sigma^2/2} = e^{-4,6}; \quad \omega = \frac{2\pi}{T} \longrightarrow \sigma = .48T$$

Nei tempi si vede che:

$$h_g(T) = .11$$

e cioè che il filtro si estende fino ai campioni adiacenti. Volendo invece un'attenuazione di 40 dB alla frequenza di Nyquist, la banda si dimezza, la durata raddoppia e si ha:

$$h_g(T) = .58; \quad h_g(2T) = .11;$$

CAPITOLO 1 IL CAMPIONAMENTO

In altre parole, un filtro antialias che sopprima abbastanza bene le componenti alla frequenza di Nyquist, ha approssimativamente una forma triangolare e una durata di circa 4 campioni. Peraltro, un filtro triangolare di base 4 campioni ha funzione di trasferimento:

$$H(\omega) = \left[\frac{\sin(4\pi fT)}{4\pi fT} \right]^2 16T^2$$

e alla frequenza di Nyquist vale:

$$H\left(\frac{\pi}{T}\right) = .025H(0)$$

e cioè circa -32dB rispetto al valore a frequenza 0.

Capitolo 2

La trasformata di Fourier discreta

2.1 Campionare in tempo e in frequenza

Consideriamo una sequenza di N impulsi di ampiezze x_n , $0 \leq n \leq N-1$, e la loro trasformata di Fourier:

$$X(f) = \sum_{n=0}^{N-1} x_n e^{-j2\pi f n T} \quad (2.1)$$

Poichè la $X(f)$ è trasformata di una sequenza di impulsi e quindi di un segnale campionato, è funzione periodica della frequenza, con periodo $1/T$. Se campioniamo anche $X(f)$ **nelle frequenze** ad intervallo Δf la sequenza di impulsi si replicherà nei tempi ad intervallo $1/\Delta f$. Scegliendo $\Delta f = 1/NT$ le repliche della sequenza di impulsi non si sovrappongono, ed il campionamento anche nel dominio della frequenza non crea equivocazione. Tuttavia, l'infinita ampiezza degli impulsi, sia nel dominio delle frequenze che nel dominio dei tempi, correlativa dell'illimitata periodicità delle funzioni nell'altro dominio e pertanto di un'energia illimitata, complica inutilmente il problema. E' peraltro intuitivo che debba esistere una relazione tra le N "ampiezze" dei campioni nel dominio dei tempi che quelle dei campioni nel dominio trasformato delle frequenze.

Definiremo ora una trasformazione lineare, **la trasformazione discreta di Fourier** (in inglese: Discrete Fourier Transform DFT); essa lega una sequenza di N **campioni** complessi (non più impulsi ma numeri)

CAPITOLO 2 LA TRASFORMATA DI FOURIER DISCRETA

ad un'altra sequenza di N numeri complessi. Questa seconda sequenza corrisponde peraltro alle ampiezze degli impulsi di cui è costituita la trasformata di Fourier della sequenza di impulsi originaria; gli uni (nel dominio dei tempi) sono distanti tra loro dell'intervallo T e gli altri (nel dominio delle frequenze) dell'intervallo $1/NT$. Dimosteremo che la trasformazione così definita è invertibile; peraltro, la cosa è intuibile, visto che le due operazioni di campionamento non provocano equivocazione. Quindi, la trasformata di Fourier discreta di una sequenza di N numeri complessi è costituita da altrettanti numeri complessi che danno una rappresentazione alternativa della sequenza data nel dominio delle frequenze, come somma di altrettante sinusoidi.

La coppia di relazioni che legano una sequenza alla sua trasformata di Fourier discreta possono essere derivate sia direttamente, sia dalla trasformata di Fourier consueta sostituendo alle variabili continue t, f le variabili discretizzate $nT, k\Delta f$. La trasformata di Fourier consueta è:

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt \quad (2.2)$$

L'integrale diventa una sommatoria, se la $x(t)$ è una somma di impulsi ai tempi nT : il fattore T servirebbe poi anche a ricordare la dimensionalità dell' integrale di Fourier.

$$X(f) = T \sum_{n=-\infty}^{+\infty} x_n e^{-j2\pi fnT} \quad (2.3)$$

Tuttavia, nel testo useremo sempre la definizione di formula (2.1). Inoltre in questo caso la $X(f)$ è periodica di periodo $1/T$; infatti:

$$X(f + m/T) = X(f) \quad (2.4)$$

Se poi la funzione $X(f)$ è campionata a intervallo $\Delta f = 1/NT$ la sequenza di impulsi di ampiezze $\{x_n\}$ è replicata con periodo NT . Definiamo allora come Trasformata di Fourier Discreta di una sequenza di N numeri complessi, la sequenza di N numeri complessi:

$$X(k\Delta f) = X_k = \sum_{n=0}^{N-1} x_n e^{-j2\pi k\Delta fnT} = \sum_{n=0}^{N-1} x_n W_N^{-nk}; \quad (2.5)$$

2.2 PROPRIETÀ DELLA DFT

$$W_N = \exp\left(j\frac{2\pi}{N}\right)$$

Si verifica che la Trasformazione discreta di Fourier è invertibile e la antitrasformata è:

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X(k) W_N^{nk} \quad (2.6)$$

Per convincersene, basta applicare il principio della sovrapposizione degli effetti e quindi limitarsi a supporre che la sequenza da trasformare $\{\tilde{x}_n\}$ sia costituita da un solo impulso ad un istante qualsiasi:

$$\tilde{x}_n = 1; n = n_0; \tilde{x}_n = 0; n \neq n_0$$

Allora:

$$\begin{aligned} \tilde{X}_k &= \sum_{n=0}^{N-1} \tilde{x}_n W_N^{-nk} = W_N^{-n_0k} \\ \frac{1}{N} \sum_{k=0}^{N-1} \tilde{X}_k W_N^{nk} &= \frac{1}{N} \sum_{k=0}^{N-1} W_N^{-n_0k} W_N^{nk} = \frac{1}{N} \sum_{k=0}^{N-1} W_N^{(n-n_0)k} = \tilde{x}_n \end{aligned}$$

Infatti, il versore W_N corrisponde ad una rotazione di un angolo pari a $2\pi/N$; se $n \neq n_0$ questo versore fa un numero intero di giri e la somma genera un poligono chiuso e quindi a somma zero. Se invece $n = n_0$ il versore non si sposta dal valore unitario ed il risultato è ovvio.

2.2 Proprietà della DFT

La discretizzazione dell'integrale di Fourier comporta un fattore di dimensione $[T]$ che però si omette nella definizione consueta della trasformata di Fourier discreta (DFT). Nel seguito cercheremo però di ricordare sempre l'origine fisica della trasformazione e quindi ci riferiremo alle dimensionalità originali degli assi di frequenze e tempi, per permettere un più efficace uso dell'intuizione. Nel dominio trasformato (che continuiamo pertanto ad indicare come dominio delle delle frequenze), vista la periodicità della trasformata, si osserva che i valori dell'indice $0 \leq k \leq N/2 - 1$ corrispondono alle frequenze positive; i valori $N/2 \leq k \leq N - 1$ corrispondono alle frequenze negative; in particolare, $k = N - 1$ corrisponde alla frequenza

CAPITOLO 2 LA TRASFORMATA DI FOURIER DISCRETA

$f = -1/NT$; la frequenza di Nyquist corrisponde al valore dell'indice $N/2$ se N è pari.

Similmente a quanto accade per la trasformata di Fourier continua si hanno diverse relazioni utili, tra cui la relazione di Parseval:

$$\frac{1}{N} \sum_{k=0}^{N-1} |X(k)|^2 = \sum_{n=0}^{N-1} |x_n|^2 \quad (2.7)$$

il ritardo (circularizzato: $x(n + N) = x_n$):

$$x_{n+m} \leftrightarrow X(k)W_N^{-mk} \quad (2.8)$$

la convoluzione circolare:

$$\frac{1}{N} \sum_{n=0}^{N-1} x_n y_{n-m} \leftrightarrow X(k)Y(k) \quad (2.9)$$

il coniugio e l'inversione dell'asse temporale:

$$x_{-n}^* \leftrightarrow X^*(k) \quad (2.10)$$

2.3 Filtraggi con la DFT

Abbiamo visto che la trasformazione di Fourier discreta corrisponde alla decomposizione di una sequenza di N campioni complessi nella somma di N sinusoidi complesse di frequenze k/NT ; $0 \leq k \leq N - 1$; è possibile convolvere due sequenze **entrambe di N campioni** moltiplicando le loro trasformate di Fourier.

Bisogna però ricordare (o verificare, se non si vuole sempre risalire alle origine fisiche della trasformazione) che così si realizza una **convoluzione periodica** delle due sequenze; infatti, essendo gli spettri campionati nelle frequenze, le sequenze di impulsi sono replicate nel tempo. Inoltre, il prodotto degli spettri campionati ha significato solo se i campioni nel dominio delle frequenze sono rilevati alle stesse frequenze, e pertanto le due sequenze (nei tempi e nelle frequenze) devono avere eguale numero di campioni, perchè l'operazione di prodotto abbia senso fisico.

Poi, il prodotto dei due spettri corrisponde alla convoluzione delle due sequenze originali, periodicizzate e cioè replicate dopo N campioni;

la sequenza risultante sarà ancora una sequenza periodica su N campioni, diversa dalla sequenza di $2N - 1$ campioni corrispondente alla convoluzione **aperiodica** delle due sequenze. Tutto passa quindi come se l'asse temporale fosse circolare e il risultato della convoluzione periodica è inquinato dal mescolamento dei primi e degli ultimi campioni della sequenza, come se fossero vicini tra di loro.

Anche se in alcuni casi (sistemi in coordinate polari) l'asse circolare può essere utile, in generale interessano solo convoluzioni aperiodiche; anch'esse sono possibili con la trasformata di Fourier discreta a patto di usare accortezze opportune o di accettare delle approssimazioni.

2.3.1 Convoluzioni aperiodiche: sequenze limitate

Si voglia effettuare la convoluzione di una sequenza $\{x_n\}$ lunga kN campioni, (si pensi a kN grande, dell'ordine di migliaia di campioni) con un filtro $\{h_m\}$ di lunghezza M campioni, con M dell'ordine delle centinaia. Come prima cosa, abbiamo osservato che **prima della trasformata di Fourier** si devono portare le due sequenze ad avere eguale lunghezza, perché le due trasformate siano campionate agli stessi valori della frequenza. Per fare ciò aggiungiamo $M - 1$ campioni nulli (che indicheremo con zeri, sperando che non vi sia confusione con gli zeri della trasformata z che non hanno nulla a che vedere) alla sequenza lunga N ed $N - 1$ zeri alla sequenza lunga M . Dopo questa operazione di allungamento con zeri (zero padding, in inglese) le due sequenze hanno eguale lunghezza $N + M - 1$ e si procede alle loro trasformate di Fourier (fig. 2.1).

Poi si moltiplicano le trasformate (che noi indicheremo spesso come gli spettri) e si antitrasforma; il risultato è la convoluzione periodica delle due sequenze allungate con zeri e quindi è esattamente la convoluzione aperiodica delle due sequenze originali. Infatti aggiungendo zeri abbiamo creato lo spazio, per fare sì che la convoluzione non si riavvolgesse sulla parte iniziale dei dati.

2.3.2 Sequenze illimitate: overlap-save; overlap-add

Qualora, come spesso accade, la sequenza originaria è di lunghezza illimitata, per procedere alla convoluzione con un filtro di lunghezza M

CAPITOLO 2 LA TRASFORMATA DI FOURIER DISCRETA

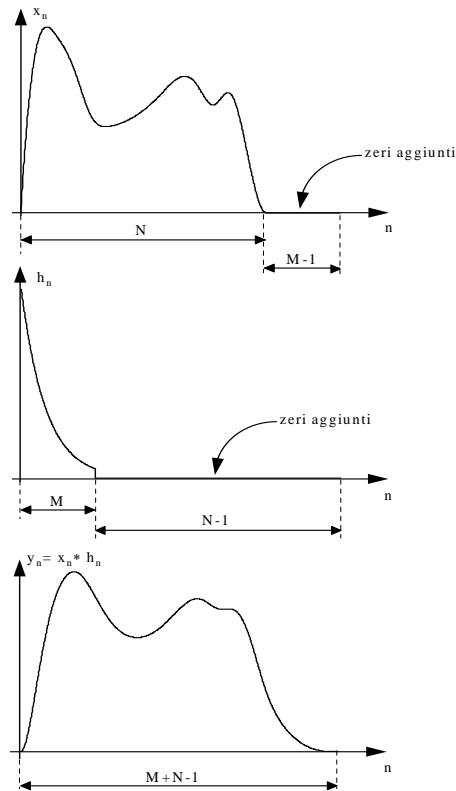


Figura 2.1: Convoluzione aperiodica con la DFT

dobbiamo scegliere tra due metodi simili ma alternativi per effettuare la convoluzione aperiodica dell'intero segnale. I metodi sono chiamati **Overlap and add** e **Overlap and save**. In sostanza in entrambi i casi si deve:

- a) segmentare il segnale di lunghezza illimitata in sottosequenze lunghe N campioni,
- b) tenere conto degli effetti della convoluzione periodica, nel moltiplicare le trasformate di segnale e filtro.

Supponendo che il filtro occupi i primi M campioni della sequenza, si può:

Overlap and add: azzerare gli ultimi $M - 1$ campioni della sequenza lunga N , moltiplicare gli spettri ed antitrasformare. L'antitrasforma-

2.3 FILTRAGGI CON LA DFT

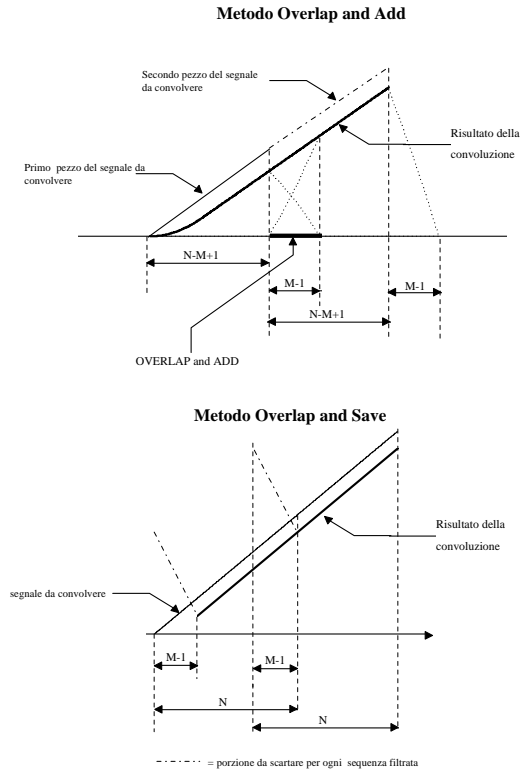


Figura 2.2: Filtraggio di sequenze illimitate con la DFT

ta contiene la convoluzione corretta del segnale parzializzato; nei primi $M - 1$ campioni però manca la coda dei campioni della sottosequenza precedente dove il segnale di ingresso era stato azzerato. Per arrivare al risultato corretto, si deve quindi riprendere il segnale originale dal campione di posizione $N - M + 1$, prenderne altri $N - M + 1$ campioni, allungarli con $M - 1$ zeri e così via. Il risultato finale si ottiene dai risultati parziali sommando le sequenze sovrapposte (overlap and add).

Overlap and save: in questo caso non si azzerano gli ultimi $M - 1$ campioni della sequenza lunga N campioni e ne risulta la creazione di $M - 1$ campioni errati **all'inizio** della sequenza convoluta che risulta dalla antitrasformazione di Fourier del prodotto degli spettri. Questi campioni del risultato finale, errati perchè inquinati dagli ultimi campioni

per effetto della circolarità della convoluzione, vengono annullati. Ogni sottosequenza deve quindi iniziare con $M - 1$ campioni di anticipo come nel caso overlap - add discusso prima. Le sottosequenze successive, che contribuiscono ognuna $N - M + 1$ campioni validi (gli ultimi), vanno giustapposte e non sommate.

2.4 Antitrasformazione di Fourier discreta

Nella sezione precedente si è discusso come convolvere tra loro sequenze di N campioni, create nel dominio dei tempi. Spesso accade però di avere dei segnali da antitrasformare, definiti nel dominio delle frequenze; nel decidere la lunghezza N bisogna stare attenti che sia rispettata, almeno approssimativamente, la condizione di non sovrapposizione delle repliche nei tempi (assenza di equivocazione temporale). Se si ha a che fare con risposte all'impulso illimitate (ad esempio come capita con filtri ideali), questo non è possibile e pertanto è necessario convivere con l'equivocazione ed accettare approssimazioni. Per fare un esempio semplicissimo, si voglia ritardare di τ la sequenza (fig. 2.3); in altre parole, si vuole la sequenza corrispondente al segnale continuo originario ricampionato con un pettine ritardato di τ rispetto al pettine di campionamento originale: per fare ciò basta moltiplicare la trasformata della sequenza per $e^{-j\omega\tau}$. Anzitutto, i campioni della trasformata vanno collocati nella posizione corretta (N sia un numero pari):

$$R(2\pi k\Delta f) = R_k = \exp(-j\frac{2\pi k}{NT}\tau); \quad (2.11)$$

$$R_{N-k} = R_k^*; \quad 0 \leq k < N/2; \quad R_{N/2} = R_{N-N/2}^* = \cos \frac{\pi\tau}{NT}$$

Inoltre, non si deve dimenticare l'effetto del campionamento dell'asse delle frequenze; l'antitrasformata della sequenza nel dominio delle frequenze è la sequenza nel dominio dei tempi che si ottiene come segue;

a) si considera l'impulso ideale, lo si ritarda di τ e lo si filtra passa basso con un filtro ideale di banda $B = 1/2T$; questo filtraggio è realizzato implicitamente dal limitare i valori della trasformata alla pulsazione di Nyquist $|\omega| \leq \pi/T$ (fig. 2.4).

b) si campiona questo segnale con intervallo T (la periodicità nel dominio delle frequenze è implicita nella trasformata di Fourier discreta).

2.4 ANTITRASFORMAZIONE DI FOURIER DISCRETA

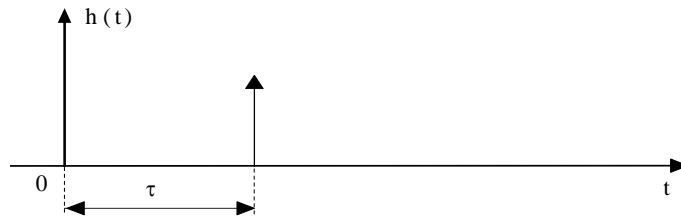


Figura 2.3: Risposta all'impulso ideale di un ritardatore

c) si periodicizza ad intervallo NT (si ritarda di mNT la sequenza di lunghezza infinita e si sommano tutte le repliche). Quest'effetto é dovuto al campionamento nelle frequenze.

Si osservi ora che la risposta nei tempi del ritardatore, non periodizzata, ha durata infinita, e converge a zero con l'andamento del sinc e cioè come $1/n$; gli errori potrebbero non essere trascurabili.

Cosa si deve fare per evitare l'errore di equivocazione, temporale questa volta? Si deve smussare la trasformata nel dominio delle frequenze, prima di procedere al campionamento.

Ad esempio, nel caso del ritardatore indicato, prima di campionare, si può convolvere la trasformata, **prima del campionamento** con un triangolo di durata significativa, in modo che la risposta all'impulso, prima della replica, decada come $1/n^3$. La convoluzione va fatta in via analitica oppure numerica, partendo da un campionamento della funzione $\exp(-j2\pi f\tau)$ più fitto che $\Delta f = 1/NT$. Ciò corrisponde a considerare la sequenza corrispondente al ritardatore su un intervallo molto lungo, moltiplicarla per una funzione decrescente con n , e poi replicarla, per effetto del campionamento.

Sarebbe utile poter fare gli smussamenti nel dominio delle frequenze convolvendo con dei sinc, per fare veri e propri troncamenti nel dominio dei tempi, ma ovviamente la cosa é molto difficile numericamente; quindi il tutto equivale a trasferire il problema del ripiegamento dal dominio dei tempi dove avviene, a quello delle frequenze dove dovrebbe essere evitato con un filtraggio ideale, altrettanto difficile da effettuare, se non di più.

Vedremo ora questi effetti con due esempi semplici.

2.4.1 Il derivatore e la costante di tempo

Nella sezione precedente abbiamo visto come sia semplice realizzare un ritardatore di τ ma anche che è inevitabile accettare delle approssimazioni. Vedremo in questa sezione come realizzare un derivatore; basterebbe moltiplicare la trasformata per $j\omega$. Così facendo, si devono ripetere le osservazioni fatte nella sezione precedente. Pertanto: derivare operando nel dominio delle frequenze con la DFT equivale (operando nel dominio dei tempi) a convolvere il segnale con la versione **campionata e replicata** del derivatore ideale filtrato passa basso $d_f(t)$. Quest'ultimo è ottenuto campionando la derivata della risposta all'impulso del filtro ideale passa basso (fig. 2.5).

$$d_f(t) = \frac{d \sin \pi t/T}{dt \pi t/T} = \frac{1}{\pi t^2} \left(\pi t \cos \frac{\pi}{T} t - T \sin \frac{\pi}{T} t \right);$$

$$d_f(nT) = \frac{1}{nT} (\cos n\pi) = \frac{(-1)^n}{nT}; \quad n \neq 0; \quad d_f(0) = 0$$

Per prefiltrare prima del campionamento in frequenza dobbiamo smusare nelle frequenze l'operatore di derivazione, che peraltro è molto liscio tranne che nell'origine e alla frequenza di Nyquist. In particolare il punto angoloso nell'origine è fortemente influenzato dagli effetti di troncamento o di ripiegamento nei tempi.

Consideriamo ad esempio l'effetto del troncamento sulla componente continua e cioè il valore medio dell'operatore nei tempi; essendo i campioni successivi dell'operatore a segni alterni nei tempi, qualsiasi troncamento comporta un valor medio non nullo e quindi la perdita dello zero alla frequenza zero. Per riportare a zero la funzione di trasferimento in continua si deve fare qualcosa che comunque deforma l'operatore di derivazione; il ripiegamento invece non perturba il valor medio dell'operatore, ma ha altri effetti dannosi. In conclusione, l'unica cosa fattibile semplicemente è scegliere N sufficientemente grande così che l'inevitabile effetto di ripiegamento nei tempi comporti conseguenze accettabili.

Come ultimo esempio consideriamo l'antitrasformata di Fourier dell'esponenziale, che ha funzione di trasferimento:

$$A(\omega) = \frac{1}{1 + j\omega\tau} \quad (2.12)$$

2.4 ANTITRASFORMAZIONE DI FOURIER DISCRETA

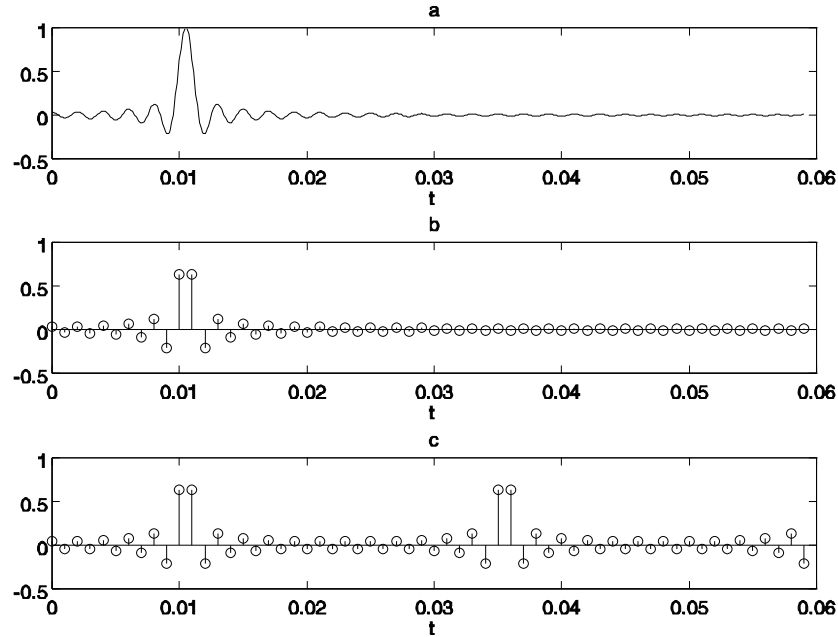


Figura 2.4: a) Risposta impulsiva di un ritardatore filtrata passabasso b) Campionamento nei tempi c) Campionamento nelle frequenze (Periodicizzazione nei tempi)

campionando nelle frequenze si replica l'esponenziale nei tempi; quindi la durata totale NT deve essere ad esempio almeno $NT > 3\tau$ (fig. 2.6). Ancora dobbiamo fissare il valore di T e questo verrà fatto in base alla necessità di evitare equivocazione nelle frequenze e cioè per campionare correttamente nei tempi: alla frequenza di Nyquist, dove sarà sovrapposta la replica proveniente dal campionamento, è necessario che lo spettro sia già molto ridotto; p. e. ad un 1/100 del valore assunto a frequenza 0; si ricava allora

$$\frac{\pi\tau}{T} \simeq 100 \rightarrow T = .0314\tau; \quad \frac{N\pi}{100} > 3 \rightarrow N > 95 \quad (2.13)$$

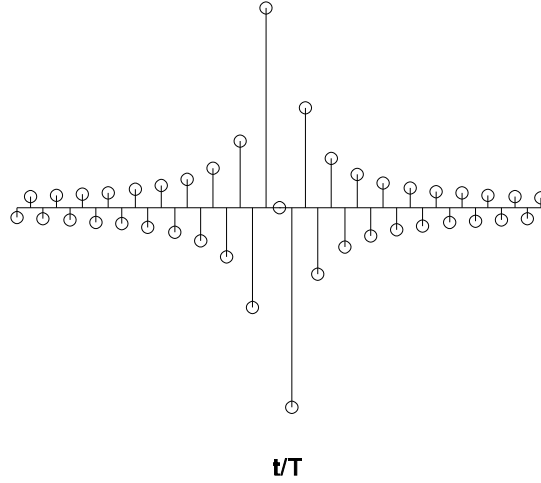


Figura 2.5: Derivatore ideale a banda limitata, non causale, campionato

2.5 Interpolazione con la DFT

2.5.1 Inserzione di zeri e repliche spettrali

L'interpolazione di sequenze per ottenere un segnale ad intervallo di campionamento più piccolo può essere fatta in un modo molto semplice mediante la DFT. Sia assegnata una sequenza $\{x_n; n = 0, \dots, N-1\}$ campionata ad intervallo T ; per fare un esempio di duplicazione della frequenza di campionamento e quindi di interpolazione su un intervallo di durata metà, consideriamo la sequenza di lunghezza doppia ma di eguale durata $\{x'_m; m = 0, \dots, 2N-1\}$ che si ottiene dalla sequenza precedente inserendo campioni di ampiezza nulla ad intervallo $T/2$. Si ha allora:

$$x'_{2n} = x_n; \quad x'_{2n+1} = 0;$$

La DFT della nuova sequenza $\{x'_m\}$ è:

$$X'_k = \sum_{m=0}^{2N-1} x'_m W_{2N}^{-mk} = \sum_{n=0}^{N-1} x_n e^{-j\frac{2\pi nk}{N}} = X_k; \quad k = 0, \dots, 2N-1 \quad (2.14)$$

Ma si ha anche:

$$X'_{k+N} = X_k; \quad k = 0, \dots, N-1$$

2.5 INTERPOLAZIONE CON LA DFT

Quindi, $\{X'_k\}$, trasformata della sequenza di $2N$ campioni $\{x'_m\}$ coincide con la sequenza X_k DFT della $\{x_n\}$, però replicata 2 volte nel dominio delle frequenze. In generale, se in una sequenza vengono inseriti $M - 1$ campioni nulli tra un campione ed il successivo (e quindi la lunghezza totale della sequenza viene moltiplicata per M), la DFT risulta replicata M volte.

Soppressione delle repliche spettrali con filtraggio ideale

Sopprimiamo ora tutte le repliche spettrali moltiplicando la sequenza X'_k per degli zeri disposti in modo opportuno ed anche per un opportuno fattore di scala M (supponiamo N pari: per conservare la simmetria della componente alla frequenza di Nyquist originale $\pm 1/2T$ si adotta il fattore $M/2$):

$$X''_k = MX'_k; \quad k = 0, \dots, N/2 - 1; \quad k = MN - N/2 + 1, \dots, MN - 1$$

$$X''_k = MX'_k/2; \quad k = N/2; \quad k = MN - N/2$$

$$X''_k = 0; \quad k \neq 0, \dots, N/2; \quad k \neq MN - N/2, \dots, MN - 1$$

allora la sequenza antitrasformata $\{x''_m\}$ è una sequenza della stessa durata ma di MN campioni, e sarà l'interpolata della sequenza originale su un intervallo di campionamento pari a T/M . L'interpolazione avviene mediante la convoluzione della sequenza inframmezzata con zeri per il filtro ideale campionato e periodicizzato la cui DFT è appunto:

$$F_k = M; \quad k = 0, \dots, N/2 - 1; \quad k = MN - N/2 + 1, \dots, MN - 1$$

$$F_k = M/2; \quad k = N/2; \quad k = MN - N/2$$

$$F_k = 0; \quad k \neq 0, \dots, N/2; \quad k \neq MN - N/2, \dots, MN - 1$$

I campioni della sequenza che costituisce la risposta all'impulso del filtro interpolatore, antitrasformata della $\{F_k\}$, sono facilmente calcolabili dalla solita espressione della serie geometrica troncata (N pari) con l'aggiunta dell'effetto del termine di ampiezza $M/2$:

$$f_n = \frac{M}{MN} \frac{\sin(\pi n N / MN)}{\tan(\pi n / MN)} = \frac{1}{N} \frac{\sin(\pi n / M)}{\tan(\pi n / MN)} \quad (2.15)$$

CAPITOLO 2 LA TRASFORMATA DI FOURIER DISCRETA

Il fattore di scala M nasce dal fatto che si desidera un'interpolazione e pertanto si vuole che:

$$f_0 = 1$$

Quindi, la componente continua della DFT non si conserva con l'interpolazione ma viene accresciuta di un fattore M . Si osservi che se la sequenza di partenza avesse avuto una durata illimitata, il filtro da utilizzare come interpolatore sarebbe stato il filtro ideale passa basso con frequenza di taglio $1/2T$ campionato con intervallo di campionamento $1/MT$ e cioè:

$$f'_n = \frac{\sin(\pi n/M)}{\pi n/M}$$

L'effetto della periodizzazione è modesto se MN è grande; peraltro, se $M = N = 2$ si ha $\{f_i\} = \{1, .5, 0, .5\}$ e si trova un interpolatore a triangolo.

Nel caso in cui N è dispari si ha la consueta espressione del *sinc* discreto, antitrasformata di uno spettro "rettangolare" centrato intorno a frequenza 0:

$$F_k = M; \quad k = 0, \dots, (N-1)/2; \quad k = MN - (N-1)/2, \dots, MN - 1$$

$$F_k = 0; \quad k \neq 0, \dots, (N-1)/2; \quad k \neq MN - (N-1)/2, \dots, MN - 1$$

e risulta infine (N dispari):

$$f_n = \frac{1}{N} \frac{\sin(\pi n/M)}{\sin(\pi n/MN)}$$

L'interpolazione con aggiunta di zeri (zero padding)

In conclusione, la tecnica per interpolare con la DFT una sequenza di lunghezza N (pari) per farne una sequenza di lunghezza MN è la seguente:

- si calcola la DFT della sequenza di lunghezza N ;
- si allunga lo spettro aggiungendo $(M-1)N - 1$ campioni nulli, inseriti tra i campioni di posto $N/2$ e $N/2 + 1$;
- si moltiplica per M e, per mantenere le simmetrie, si suddivide il valore del campione di posto $N/2$ tra i campioni di posto $N/2$ e $MN - N/2$, e infine si antitrasforma.

Il risultato sarà la sequenza di partenza, interpolata con la funzione *sinc* periodizzata descritta nella formula (2.15).

2.6 Il calcolo della DFT: la FFT

Per calcolare la sequenza $\{X_k\}$ a partire dalla sequenza $\{x_n, 0 \leq n \leq N - 1\}$ dobbiamo applicare la matrice $\underline{\mathbf{W}}_N^*$ di dimensioni $[N \times N]$:

$$X_k = \sum_{n=0}^{N-1} x_n W_N^{-nk}; \quad W_N = e^{j2\pi/N} \quad (2.16)$$

$$\underline{\mathbf{X}} = \underline{\mathbf{W}}^* \underline{\mathbf{x}} \quad (2.17)$$

dove con $\underline{\mathbf{x}}$, $\underline{\mathbf{X}}$ abbiamo indicato due vettori colonna contenenti rispettivamente la sequenza e la sua trasformata discreta di Fourier.

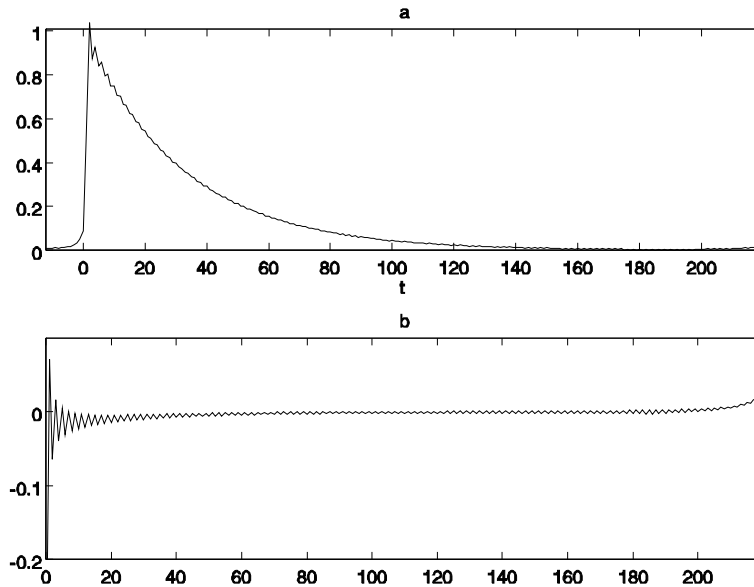


Figura 2.6: a) Antitrasformata discreta di una F.d.T con un polo reale; b) differenza tra antitrasformata discreta e antitrasformata continua

La matrice $\underline{\mathbf{W}}$ appare trasposta (o coniugata essendo Hermitiana),

CAPITOLO 2 LA TRASFORMATA DI FOURIER DISCRETA

per i segni negativi negli esponenti della trasformazione diretta.

$$\begin{bmatrix} X(0) \\ X(1) \\ X(2) \\ \dots \\ X(N-2) \\ X(N-1) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & W_N^{-1} & W_N^{-2} & \dots & W_N^2 & W_N \\ 1 & W_N^{-2} & W_N^{-4} & \dots & W_N^4 & W_N^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & W_N^{-(N-2)} & W_N^4 & \dots & W_N^{-4} & W_N^{-2} \\ 1 & W_N & W_N^2 & \dots & W_N^{-2} & W_N^{-1} \end{bmatrix} \times \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \dots \\ x_{N-2} \\ x_{N-1} \end{bmatrix} \quad (2.18)$$

I campioni della riga n -esima corrispondono alle successive posizioni di un vettore (una sinusoidale complessa) che fa $n - 1$ giri in N campioni.

La seconda riga è una sinusoidale complessa che fa un ciclo in senso orario in N campioni (frequenza $f = 1/NT$) e così via fino all'ultima riga che corrisponde ad una sinusoidale complessa che fa $N - 1$ giri in N campioni; in altre parole, ruota in senso antiorario di un giro in N campioni (frequenza $f = -1/NT$). Se N è pari, la frequenza di Nyquist è alla riga $N/2 + 1$; è necessario ricordare come la frequenza 0 sia nella prima riga. La matrice $\underline{\mathbf{W}}$ non è una matrice ortogonale vera e propria perchè in quel caso dovrebbe essere:

$$\underline{\mathbf{A}}^{-1} = \underline{\mathbf{A}}^* \quad (2.19)$$

Invece si ha, combinando trasformazione ed antitrasformazione di Fourier:

$$\frac{1}{N} \underline{\mathbf{W}} \underline{\mathbf{W}}^* = \underline{\mathbf{I}}; \quad \rightarrow \quad \frac{1}{N} \underline{\mathbf{W}}^* = \underline{\mathbf{W}}^{-1} \quad (2.20)$$

Per calcolare $\underline{\mathbf{X}}$ è necessario applicare questa matrice e la cosa sarebbe numericamente assai costosa (circa N^2 moltiplicazioni) se non fosse stata scoperta la trasformata di Fourier veloce (FFT: Fast Fourier Transform) che permette di fare questo calcolo con un numero assai limitato di operazioni, crescente solo logicamente con N . Alla base della FFT c'è l'idea che una qualsiasi rotazione di k/N giri si può ottenere come combinazione di $m = \log_2 N$ rotazioni di $1/2$ giro, $1/4$ di giro ecc. fino ad $1/N$ giro, sempre che N sia una potenza di 2 e cioè $N = 2^m$.

2.6.1 Le farfalle

Il meccanismo che porta alla FFT può essere analizzato in breve: si consideri la trasformata di Fourier discreta (DFT) di una sequenza di N campioni e per semplicità si supponga N potenza di 2. Si decomponga la sommatoria in due sommatorie che si riferiscono la prima ai termini di posto pari della sequenza, e la seconda a quelli di posto dispari. Si osserva che le due sommatorie che compaiono sono ancora delle DFT, che si riferiscono questa volta a sequenze di $N/2$ campioni l'una. Indichiamo con $G(k)$, $H(k)$ queste due trasformate; risulta:

$$X(k) = \sum_{n=0}^{N/2-1} x_{2n} W_N^{-2nk} + \sum_{n=0}^{N/2-1} x_{2n+1} W_N^{-(2n+1)k} = \quad (2.21)$$

$$\sum_{n=0}^{N/2-1} x_{2n} W_{N/2}^{-nk} + W_N^{-k} \sum_{n=0}^{N/2-1} x_{2n+1} W_{N/2}^{-nk} = G(k) + W_N^{-k} \times H(k)$$

$$X(k + N/2) = G(k) - W_N^{-k} \times H(k)$$

In scrittura matriciale:

$$\begin{bmatrix} X(k) \\ X(k + N/2) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \times \begin{bmatrix} G(k) \\ W_N^{-k} H(k) \end{bmatrix} \quad (2.22)$$

dove si è fatto uso della proprietà delle trasformate di sequenze di $N/2$ campioni di essere periodiche di periodo $N/2$ e quindi:

$$H(k + N/2) = H(k); \quad G(k + N/2) = G(k)$$

inoltre:

$$W_N^{k+N/2} = -W_N^k$$

Nasce così il grafo di flusso che dá gli elementi di una trasformata di Fourier di N campioni, combinando quelli di due trasformate di Fourier di $N/2$ campioni l'una; in particolare viene indicato con il nome di *farfalla*, (*butterfly*) (fig. 2.7) il grafo che connette $X(k)$, $X(k+N/2)$ a $G(k)$, $H(k)$. La decomposizione si può ripetere fino ad avere sequenze di 2 campioni.

Allo stesso risultato si può arrivare ripetendo le considerazioni fatte in precedenza per l'interpolazione di una sequenza 2:1 eseguita inserendo

CAPITOLO 2 LA TRASFORMATA DI FOURIER DISCRETA

una sequenza di zeri in posizioni alterne e poi sopprimendo la replica spettrale. In questo caso, invece, la sequenza originale è la somma di **due sequenze** che hanno campioni nulli in posizioni alterne, e cioè:

- la sequenza dei campioni pari (con campioni nulli in posizioni dispari) che ha trasformata discreta di Fourier G_k , periodica di periodo $N/2$;

- la sequenza dei campioni dispari, ritardata di T rispetto alla sequenza dei campioni pari, con campioni nulli in corrispondenza delle posizioni pari, che ha trasformata discreta di Fourier $H_k W_N^{-k}$, periodica di periodo $N/2$. Infatti, nel dominio trasformato, il **ritardo di un campione** tra le due sequenze è ottenuto moltiplicando la H_k per il termine $\exp(-j2\pi k \Delta f T)$ e poichè $\Delta f = 1/NT$ questo ritardo corrisponde appunto al termine W_N^{-k} .

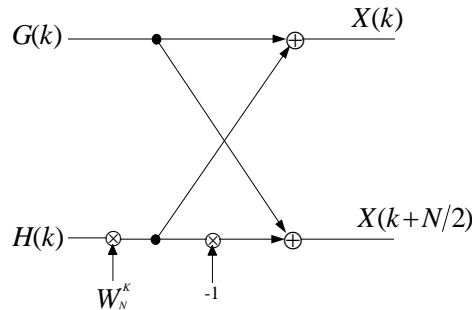


Figura 2.7: Grafo a "farfalla"

2.6.2 Occupazione di memoria e *bit reversal*

La trasformazione di Fourier veloce ha il vantaggio di potere essere fatta sul posto: in altre parole, non c'è bisogno di ulteriore memoria per effettuare la trasformazione. Una volta memorizzata la sequenza di N campioni, si possono realizzare i primi grafi a farfalla 2×2 e successivamente gli altri scambiando i campioni memorizzati con quelli via via calcolati. Alla fine, ci si troverà con la trasformata memorizzata là dove era prima memorizzata la sequenza. Supponiamo ora che $N = 2^m$, e che quindi l'indice del campione sia un numero binario. Si verifica che, per avere i dati della trasformata (antitrasformata) memorizzati in ordine

2.7 FILTRAGGIO DI SEQUENZE ILLIMITATE

progressivo, i dati originali vanno ordinati disponendo prima i campioni di posto pari (ultimo bit 0), e tra quelli, prima quelli di posto pari (penultimo bit 0) e poi ancora tra quelli, quelli di posto pari (terzultimo bit 0) ecc. e così per tutti gli stadi del calcolo. Quindi, per avere la trasformata nell'ordine naturale, i dati vanno scritti nell'ordine inverso dei bit del loro indice e cioè *bit reversed*. Se i dati sono trascritti nell'ordine *bit reversed*, partendo da dati in ordine naturale, l'ordinamento della trasformata risulterà progressivo. Similmente, se i dati sono scritti in ordine naturale, sarà la trasformata ad essere disposta in ordine *bit reversed*. Ad esempio, con $m = 3$, l'ordine dei campioni del segnale *bit reversed* è:

$$000, 100, 010, 110, 001, 101, 011, 111$$

Come osservato, si può ottenere quest'ordine da quello naturale senza ulteriore occupazione di memoria, a parte due registri per ogni coppia di numeri complessi da scambiare di posizione.

2.7 Filtraggio di sequenze illimitate

2.7.1 Costo computazionale della FFT

Determiniamo anzitutto il costo computazionale di una trasformata di Fourier veloce (trascurando il costo computazionale delle somme). Ad ogni stadio della trasformazione devono essere calcolati $N/2$ grafi a farfalla che costano una moltiplicazione complessa l'uno. (fig.2.8) Sono da calcolare $\log_2 N$ stadi e quindi un totale di $(N \log_2 N)/2$ moltiplicazioni complesse ognuna corrispondente a 4 moltiplicazioni reali l'una. Se poi si considera che con una trasformata di Fourier veloce si possono fare 2 trasformate di Fourier di due sequenze reali di N campioni l'una, disponendo una delle due sequenze sull'asse reale e la seconda sull'asse immaginario, e poi combinando in modo opportuno le parti reale ed immaginaria della trasformata, si conclude che il costo computazionale di una trasformata di Fourier di una sequenza reale è di $\log_2 N$ moltiplicazioni per campione. Infatti, assegnate le funzioni reali $s_i(t)$; $i = 0, 1$, consideriamo il segnale complesso:

$$s(t) = s_0(t) + js_1(t) \leftrightarrow S(\omega) = S_0(\omega) + jS_1(\omega) \quad (2.23)$$

CAPITOLO 2 LA TRASFORMATA DI FOURIER DISCRETA

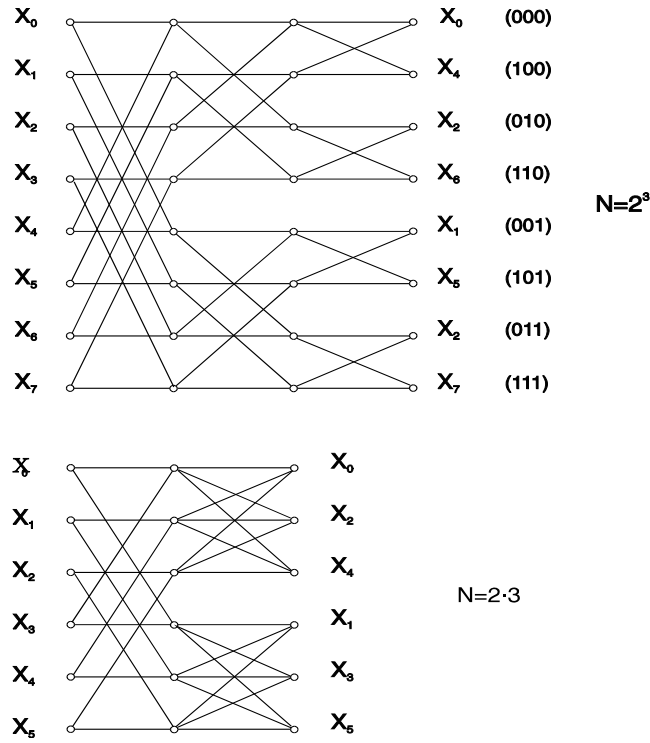


Figura 2.8: Grafi corrispondenti a trasformazioni di Fourier veloci

per le simmetrie delle trasformate delle funzioni reali:

$$S_i(-\omega) = S_i^*(\omega); \quad i = 0, 1; \quad (2.24)$$

Pertanto:

$$2j^i S_i(\omega) = S(\omega) + (-1)^i S^*(-\omega); \quad i = 0, 1 \quad (2.25)$$

2.7.2 Segmentazione di sequenze illimitate

Supponiamo di volere filtrare una sequenza di lunghezza illimitata con un filtro di M campioni; come scegliere N , lunghezza delle trasformate di Fourier? Se N è troppo piccolo, sarà eccessivo il numero dei campioni da eliminare o ricalcolare a causa del meccanismo overlap save o overlap add. Peraltro, se N è troppo grande, sprecheremo calcoli per fare trasformate di Fourier troppo lunghe.

2.8 TRASFORMATATA DI FOURIER IN BASE 4

L'ottimo si calcola subito: il numero di campioni utili per trasformata è pari a: $N' = N - M + 1$; il costo di ogni trasformata è $N \log_2 N$ moltiplicazioni reali; il costo effettivo per campione sarà quindi

$$c = \frac{N \log_2 N}{N - M + 1} \quad (2.26)$$

che ha un valore minimo per $N \simeq 10M$. Tenendo conto del costo computazionale di trasformata, antitrasformata e prodotto di spettri, l'uso della trasformata di Fourier risulta quindi preferibile non appena il filtro è più lungo di una ventina di campioni; tuttavia ciò comporta una latenza (ritardo della disponibilità) di N campioni (invece che M), ritardo che è inaccettabile in alcuni casi.

2.8 Trasformata di Fourier in base 4

Nella sezione precedente abbiamo visto come si effettuano trasformate di Fourier veloci con N potenza di 2. Lo stesso ragionamento si estende al caso in cui N è scomponibile in fattori primi piccoli (2, 3, 5); allora non avremo più farfalle ma strutture di calcolo un po' più complesse: un caso particolare si ha per la base 4, per cui ora faremo i conti per esteso. Suddividendo la sequenza in quattro sottosequenze, il calcolo della FFT procede così:

$$\begin{aligned} X(k) &= \sum_{n=0}^{N/4-1} x_{4n} W_N^{-4nk} + \sum_{n=0}^{N/4-1} x_{4n+1} W_N^{-(4n+1)k} + \\ &+ \sum_{n=0}^{N/4-1} x_{4n+2} W_N^{-(4n+2)k} + \sum_{n=0}^{N/4-1} x_{4n+3} W_N^{-(4n+3)k} = \\ &= \sum_{n=0}^{N/4-1} x_{4n} W_{N/4}^{-nk} + W_N^{-k} \sum_{n=0}^{N/4-1} x_{4n+1} W_{N/4}^{-nk} + \\ &+ W_N^{-2k} \sum_{n=0}^{N/4-1} x_{4n+2} W_{N/4}^{-nk} + W_N^{-3k} \sum_{n=0}^{N/4-1} x_{4n+3} W_{N/4}^{-nk} \end{aligned} \quad (2.27)$$

CAPITOLO 2 LA TRASFORMATA DI FOURIER DISCRETA

Inoltre, poiché $W_N^{N/4} = j$

$$\begin{aligned}
 X(k + mN/4) &= \sum_{n=0}^{N/4-1} x_{4n} W_{N/4}^{-nk} + j^m W_N^{-k} \sum_{n=0}^{N/4-1} x_{4n+1} W_{N/4}^{-nk} + \\
 &+ (-1)^m W_N^{-2k} \sum_{n=0}^{N/4-1} x_{4n+2} W_{N/4}^{-nk} + (-j)^m W_N^{-3k} \sum_{n=0}^{N/4-1} x_{4n+3} W_{N/4}^{-nk}; \quad m = 0, 1, 2, 3
 \end{aligned} \tag{2.28}$$

In forma matriciale:

$$\begin{bmatrix} X(k) \\ X(k + N/4) \\ X(k + N/2) \\ X(k + 3N/4) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -j & -1 & j \\ 1 & -1 & 1 & -1 \\ 1 & j & -1 & -j \end{bmatrix} \times \begin{bmatrix} G(k) \\ W_N^{-k} H(k) \\ W_N^{-2k} L(k) \\ W_N^{-3k} M(k) \end{bmatrix} \tag{2.29}$$

dove con G, H, L, M sono state indicate le trasformate delle quattro sottosequenze, ciascuna di lunghezza $N/4$. Moltiplicare per j non costa numericamente, e quindi si osserva che la FFT in base 4 ha 3 moltiplicazioni complesse per farfalla, $N/4$ farfalle per stadio, $\log_4 N$ stadi: in totale il numero di moltiplicazioni complesse é

$$n'' = 3 \times \frac{N \log_2 N}{4 \cdot 2} = 0.75 \times \frac{N}{2} \log_2 N \tag{2.30}$$

con un risparmio, rispetto alla trasformata di Fourier veloce in base 2, del 25%, che é significativo .

Capitolo 3

Ricostruzione di segnali continui

3.1 Introduzione: sequenze limitate

Nel trattamento del segnale numerico, vi sono numerosissime occasioni per effettuare delle interpolazioni: il caso ovvio è quello in cui abbiamo un segnale che è noto in certe posizioni e che deve essere valutato in altre, dove non è disponibile una misura diretta. È anche utile mescolare sequenze numeriche campionate con diverse frequenze di campionamento, oppure sequenze campionate con orologi imperfetti: si pensi ad esempio ad una registrazione numerica di un segnale musicale, a cui si voglia sommare un segnale audio di commento, raccolto con un altro registratore, magari dotato di una mediocre temporizzazione. Operando su segnali analogici, basta un nodo sommatore ed il problema è banale: non così nel caso di segnali campionati, in cui si deve trasformare un segnale campionato su un pettine di campionamento in un segnale campionato su un altro pettine. In più i due pettini di campionamento possono essere entrambi irregolari a causa di imperfezioni negli orologi. È chiaro che le regole del campionamento vanno rispettate, e pertanto i segnali vanno filtrati ognuno nella propria banda di Nyquist; uno dei due sarà quindi sovracampionato e cioè non occuperà tutta la banda disponibile. Si vedrà che la precisione con cui va fatta l'operazione di interpolazione deve essere molto elevata, se si desidera conservare un'elevata qualità del segnale campionato. In questo capitolo, analizzeremo l'operazione di interpo-

CAPITOLO 3 RICOSTRUZIONE DI SEGNALI CONTINUI

lazione intesa come ricostruzione di un segnale continuo, eventualmente ricampionabile con le regole esposte nel capitolo 1. Invece, nel capitolo 7, discuteremo di tecniche che trasformano direttamente una sequenza in un'altra campionata ad altra frequenza senza passare, nemmeno concettualmente, per il momento in cui il segnale è continuo.

La ricostruzione di un segnale continuo a partire da una sequenza di campioni posizionati regolarmente lungo un asse è possibile inserendo un filtro ideale passa basso che opera sulla sequenza di impulsi ideali corrispondenti ai campioni, che opera come interpolatore. Tuttavia, nascono problemi con il ritardo infinito apportato da questo filtro; in altre parole, per fare la ricostruzione ad un dato tempo servono infiniti campioni del segnale sia prima che dopo quell'istante. Il segnale continuo $x(t)$ è ottenuto dalla sequenza dei campioni come:

$$x(t) = \sum_{n=-\infty}^{+\infty} x_n \frac{\sin \frac{\pi(t-nT)}{T}}{\frac{\pi(t-nT)}{T}} = \sum_{n=-\infty}^{+\infty} x_n \operatorname{sinc} \frac{t-nT}{T} \quad (3.1)$$

Abbiamo adottato la definizione

$$\operatorname{sinc}(x) = \frac{\sin(\pi x)}{\pi x} \quad (3.2)$$

E' ora opportuno determinare tecniche di ricostruzione di un segnale continuo che siano piú semplici ed in particolare diano un risultato anche con sequenze corte; quali sono i requisiti da imporre, visto che non sarà comunque possibile garantire di ottenere di nuovo la $x(t)$ iniziale? Infatti questa corrisponderebbe al risultato dell'interpolazione fatta con un filtro ideale, disponendo di un illimitato numero di campioni prima e dopo l'istante di interesse. Se i campioni disponibili sono limitati, e quindi "pochi", come accadrá in genere, cosa é opportuno fare?

Un interpolatore è quindi un filtro definito su un asse temporale o spaziale continuo che, alimentato da una sequenza limitata di impulsi ideali di ampiezza proporzionale a quella dei campioni della sequenza numerica, permette di ricostruire un segnale $\hat{x}(t)$ quanto possibile simile al segnale originario $x(t)$. In sostanza sarà un buon filtro passabasso, in grado di rimuovere al meglio le repliche dello spettro del segnale generate dal campionamento (fig. 3.1). Diremo che un interpolatore "rispetta" (to honor, in inglese) il dato, se il segnale continuo ricampionato restituisce

3.1 INTRODUZIONE: SEQUENZE LIMITATE

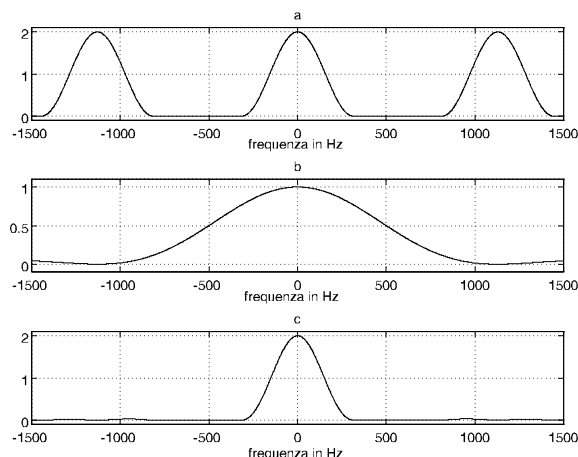


Figura 3.1: Rimozione delle repliche spettrali di un segnale campionato con un filtro ricostruttore triangolare

il campione assegnato. Non sempre gli interpolatori rispettano i dati; a volte, specialmente quando si può supporre che alla sequenza di ingresso sia sovrapposto del disturbo, può essere preferibile che il segnale continuo in uscita, ricampionato, restituisca una sequenza non identica a quella originale, ma smussata, proprio per ridurre l'effetto del disturbo. Si suppone cioè che sui campioni da interpolare sia sommata una sequenza indipendente che va rimossa.

Ancora, per descrivere un interpolatore, si può affermare che rispetta la costante, la retta, la parabola ecc, quando si ha che, se la sequenza di ingresso è ottenuta campionando una costante, una retta, una parabola ecc, il segnale continuo di uscita è appunto lo stesso segnale di partenza; si osservi che un interpolatore può rispettare le rette ma non i dati o viceversa (fig. 3.2).

Rispettare la costante significa che la funzione interpolata è costante, se tutti i campioni della sequenza sono eguali tra loro; più in generale, se la funzione campionata fosse un polinomio di grado k , la funzione ricostruita dall'interpolatore, di complessità crescente con k , dovrà essere identica al polinomio di partenza.

Indichiamo con ordine dell'interpolatore il grado dei polinomi utilizzati per effettuare l'interpolazione. L'interpolatore di ordine $k = 0$ è detto mantenitore: è quello che conserva le costanti (e solo quelle). Possi-

CAPITOLO 3 RICOSTRUZIONE DI SEGNALI CONTINUI

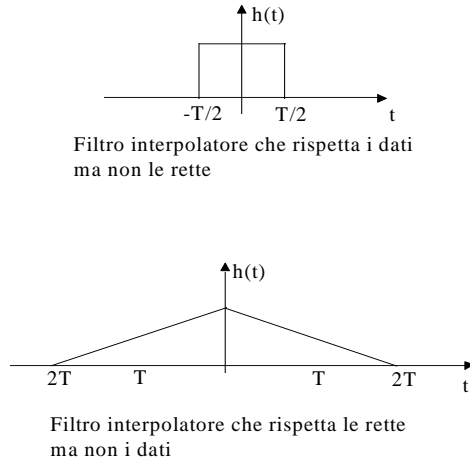


Figura 3.2:

amo poi considerare l'interpolatore lineare (ordine $k = 1$) che conserva le rette e tra queste quella con coefficiente angolare nullo e cioè la costante. L'interpolatore quadratico (ordine $k = 2$), conserverà le parabole, e cioè, quando i dati di ingresso sono disposti su una parabola campionata, il risultato dell'interpolazione sarà proprio la stessa parabola continua nel tempo. Si considera anche l'interpolatore cubico (ordine $k = 3$); gli interpolatori utilizzati in pratica poi sono il mantenitore, l'interpolatore lineare e un particolare interpolatore cubico detto anche bacchetta elastica o "spline".

Per caratterizzare un interpolatore, utilizzeremo la risposta all'impulso, e cioè la funzione tempo continua $h_k(t)$, uscita dell'interpolatore, quando all'ingresso vi è la sequenza corrispondente all'impulso unitario $x_n = \delta_n$:

$$\delta_n = 0; \forall n \neq 0; \delta_0 = 1 \tag{3.3}$$

La risposta all'impulso $h_k(t)$ di una classe di interpolatori di ordine k , che ora considereremo, ha la durata $(k + 1)T$ ed è costituita da $k + 1$ archi di polinomi di grado k -esimo. Indicheremo con $H_k(\omega)$ la trasformata di Fourier di $h_k(t)$. Indicata con $x_k(t)$ la funzione interpolata con

3.2 RICOSTRUZIONE DI POLINOMI

l'interpolatore di ordine k si ha, se la funzione ricostruita rispetta il dato:

$$x_k(t) = \sum_{n=-\infty}^{+\infty} x_n h_k(t - nT) \quad (3.4)$$

$$x_k(nT) = x_n; \quad \rightarrow h_k(0) = 1; \quad h_k(nT) = 0; \quad n \neq 0$$

Si noti che in genere $x_k(t) \neq x(t)$ e cioè la funzione interpolata non è eguale alla funzione originale, a meno che questa non fosse polinomiale del grado k .

3.2 Ricostruzione di polinomi

È facile trovare le condizioni che devono essere soddisfatte dalla risposta all'impulso di un interpolatore perché un polinomio di grado k sia ricostruito correttamente. Partiamo da $k = 0$; basta applicare la sovrapposizione degli effetti: se tutti i campioni della sequenza sono eguali tra loro ($x_n = x_0$), la funzione interpolata sarà la sovrapposizione di tante repliche eguali della risposta all'impulso $h_k(t)$, sfalsate di T l'una rispetto all'altra. Ma replicare con passo T nel dominio del tempo implica campionare a passo $f_c = 1/T$ nel dominio delle frequenze e quindi la trasformata della funzione interpolata sarà eguale alla trasformata della risposta all'impulso campionata alle frequenze m/T ; perché poi questa funzione campionata si riduca ad un impulso a $f_c = 0$ e quindi sia la trasformata di una costante, bisogna che:

$$H_k(m\omega_c) = 0; \quad \omega_c = \frac{2\pi}{T}; \quad m \neq 0 \quad (3.5)$$

Quindi per rispettare le costanti basta che la trasformata della risposta all'impulso dell'interpolatore si annulli in corrispondenza della frequenza di campionamento e delle sue armoniche. Tra tutti gli interpolatori che rispettano la costante, ce ne è uno, più semplice di tutti, chiamato mantentore.

Se il segnale $x(t)$ ha uno spettro che non occupa solo la frequenza zero come una costante, ma è concentrato alle basse frequenze come una rampa, non basterà che la trasformata della risposta all'impulso dell'interpolatore sia nulla solo alle frequenze $m f_c$ per ottenerne una interpolazione corretta. Essa deve essere nulla in tutto un intorno di queste frequenze,

in modo da *azzerare* le repliche dello spettro. Si verifica che non solo $H_k(\omega)$ deve essere nulla alle frequenze $m f_c$ ma devono anche essere nulle almeno le prime k derivate rispetto alla pulsazione ω . Parimenti, in un intorno dell'origine, k derivate rispetto a ω della $H_k(\omega)$ devono essere nulle, e cioè la $H_k(\omega)$ deve essere quanto piú costante possibile, intorno a frequenza zero. Se k tende ad infinito, si ottiene il filtro ideale.

Nel caso dell'interpolatore a bacchetta elastica, accade appunto che le derivate della funzione di trasferimento (fino alla derivata terza compresa) sono nulle alle armoniche della frequenza di campionamento; a frequenza zero, la funzione è costante fino a termini di quarto ordine. Quindi la bacchetta elastica onora i polinomi di terzo grado.

3.3 Interpolatori derivati dal mantenitore

L'interpolatore piú semplice che onora la costante é il mantenitore cui corrisponde la risposta all'impulso (fig. 3.3,a):

$$h_0(t) = 1; \quad |t| \leq \frac{T}{2} \quad \leftrightarrow \quad H_0(\omega) = T \frac{\sin \pi f T}{\pi f T} = T \operatorname{sinc}(fT) \quad (3.6)$$

Infatti,

$$H_0(m\omega_c) = T \frac{\sin \pi m}{\pi m} = 0; \quad \omega_c = \frac{2\pi}{T}; \quad m \neq 0$$

Immediatamente piú complesso, ma ancora molto semplice é l'interpolatore lineare che ha risposta all'impulso e funzione di trasferimento (fig. 3.3,b):

$$h_1(t) = 1 - \frac{|t|}{T}; \quad |t| \leq T; \quad h_1(t) = \frac{1}{T} (h_0(t) * h_0(t)) \quad (3.7)$$

$$H_1(\omega) = T \operatorname{sinc}(fT)^2 \quad (3.8)$$

Si vede subito che

$$H_1(m\omega_c) = T \left(\frac{\sin \pi m}{\pi m} \right)^2 = 0; \quad \frac{dH_1}{d\omega}(m\omega_c) = 0 \quad \omega_c = \frac{2\pi}{T}; \quad m \neq 0$$

In entrambi i casi la risposta all'impulso é molto corta (T o $2T$); poi possono essere interessanti gli interpolatori

$$H_2(\omega) = T \operatorname{sinc}(fT)^3; \quad \leftrightarrow \quad h_2(t) = \frac{1}{T} (h_1(t) * h_0(t)) \quad (3.9)$$

3.3 INTERPOLATORI DERIVATI DAL MANTENITORE

$$H_3(\omega) = T \operatorname{sinc}(fT)^4; \quad \leftrightarrow \quad h_3(t) = \frac{1}{T} (h_2(t) * h_0(t)) \quad (3.10)$$

(il valore della costante si trova imponendo pari a T il guadagno a frequenza zero). Il problema di questi ultimi due interpolatori (fig. 3.3 c,d) é che rispettano polinomi di grado k ma non sequenze qualsiasi di dati; infatti le risposte all'impulso $h_2(t)$ e $h_3(t)$ non si annullano per $t = T$ (vedi fig 3.3); infatti la funzione di trasferimento decade per valori bassi della frequenza e quindi la sequenza risultato del ricampionamento del segnale interpolato é una versione **smussata** della sequenza assegnata.

Interpolatori di ordine k derivati dal mantenitore

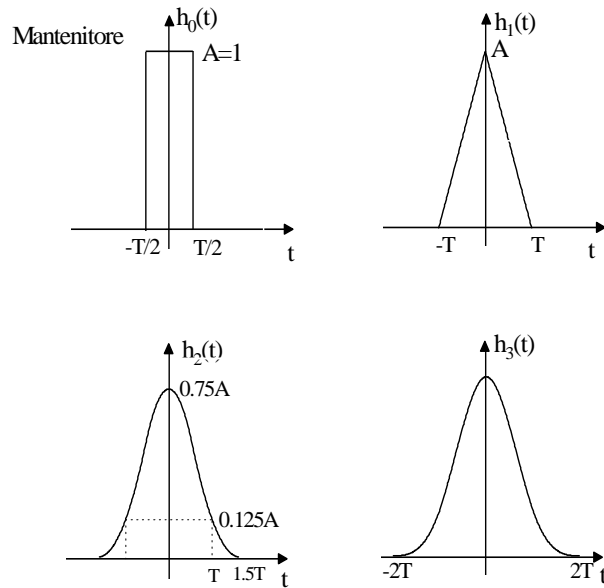


Figura 3.3: Interpolatori di ordine k derivati dalla convoluzione con se stesso del mantenitore: a) $k = 0$ mantenitore; b) $k = 1$ lineare; c) $k = 2$ parabolico; d) $k = 3$ cubico

3.3.1 Interpolatore cubico troncato

Un interpolatore interessante per la sua semplicità ed accuratezza è il seguente interpolatore cubico, che ha durata $4T$. Le equazioni sono:

$$y(-x) = y(x) = 1 - (1 + a)x^2 + ax^3; \quad 0 \leq x = \frac{t}{T} \leq 1 \quad (3.11)$$

$$y(-x) = y(x) = (a - 2)(x^3 - 5x^2 + 8x - 4); \quad 1 \leq x \leq 2 \quad (3.12)$$

Il passaggio per lo 0 a $\pm T$ garantisce che i dati sono onorati. inoltre, per tutti i valori del parametro a la derivata prima è continua; la derivata seconda è discontinua agli estremi e pertanto è inutile imporre la continuità all'interno. Per $a = 3/2$ la funzione di trasferimento è massimamente piatta (fino al 4° ordine) nell'origine e gli zeri alle frequenze multiple di $1/T$ sono di secondo ordine; questo quindi è il valore consigliato per a .

3.3.2 Interpolatori con filtro ideale troncato

Altri interpolatori si possono trovare ricorrendo al filtro ideale, e limitandone la durata troncandolo con una finestratura p.e. a coseno rialzato come si vedrà nel capitolo 5.

3.4 Rigidità e sovraelongazioni

Consideriamo ora vari interpolatori, dal mantenitore o dall'interpolatore lineare, che non hanno memoria da un campione al successivo, al filtro ideale, dove l'influenza di un campione si estende all'infinito. Nel caso più semplice (il mantenitore) non è garantita nemmeno la continuità della funzione; poi, via via, si accresce la *rigidità* dell'interpolatore e cioè si impone la continuità della funzione (interpolatore lineare) e poi di un crescente numero di derivate. Nel caso estremo del filtro ideale, vi è continuità di tutte le derivate in corrispondenza dei campioni; quindi non vi è più possibilità di ritrovare la posizione del reticolo di campionamento dal segnale ricostruito, e di fatto il segnale ricostruito è quello di partenza, indipendente dalla posizione del reticolo di campionamento.

Le condizioni agli estremi sono tanto più influenti quanto più l'interpolatore è rigido. Nel caso del mantenitore e dell'interpolatore lineare,

3.4 RIGIDITÀ E SOVRAELONGAZIONI

non sono rilevanti; nel caso del filtro ideale invece, se la sequenza da interpolare è limitata, si possono arbitrariamente assegnare tutte le derivate del segnale da interpolare in corrispondenza del primo e dell'ultimo campione. Ciò influenza l'andamento del segnale all'interno dell'intervallo, pur senza alterare i valori dei campioni, all'interno dell'intervallo. **Pertanto, è arbitrario usare interpolatori rigidi (come il filtro ideale) con sequenze corte.** È ragionevole utilizzare interpolatori con lunghezza della risposta all'impulso molto più breve della lunghezza della sequenza da interpolare. In questo modo, gli effetti di bordo sono limitati.

Nel caso dell'interpolatore lineare, i valori minimi e massimi del segnale interpolato sono quelli dei campioni; interpolando invece i dati come con una bacchetta flessibile l'escursione dei valori è maggiore. Utilizzando il filtro ideale, imponendo la continuità di tutte le derivate e quindi una risposta all'impulso dell'interpolatore di lunghezza infinita, può addirittura accadere che l'uscita sia illimitata per alcune sequenze di dati.

Per fare un caso evidente, consideriamo una sequenza alla frequenza di Nyquist, che moltiplichiamo per la funzione $sign(n) = n/|n|$; $sign(0) = 0$ (fig. 3.4). Il segnale interpolato con il filtro ideale tende all'infinito in corrispondenza del cambio di segno. Poiché la singolarità nasce dalla somma dei *sinc* (una serie armonica) che si combinano con segni eguali, l'ampiezza cresce con il logaritmo del numero di campioni della sequenza interpolata. In proposito, si ricordi che

$$\sum_{n=1}^N |sinc(\frac{t}{T} - n)| = |sinc(\frac{t}{T})| \sum_{n=1}^N \frac{1}{1 + n\frac{T}{t}}$$

e poi che la serie logaritmica è:

$$\sum_{n=1}^{+\infty} \frac{x^n}{n} = \log \frac{1}{1-x}; \quad |x| < 1 \quad (3.13)$$

Al limite, per $x = 1$, si trova la serie armonica. Si vedrà poi in un capitolo successivo quando si discuterà della trasformazione di Hilbert, che l'involuppo di un segnale interpolato dipende dalla trasformata di Hilbert del segnale stesso; la trasformata di Hilbert della funzione *sign* ed in generale di ogni gradino ha una singolarità di tipo logaritmico che si ritrova sia nell'involuppo che nel segnale interpolato.

CAPITOLO 3 RICOSTRUZIONE DI SEGNALI CONTINUI

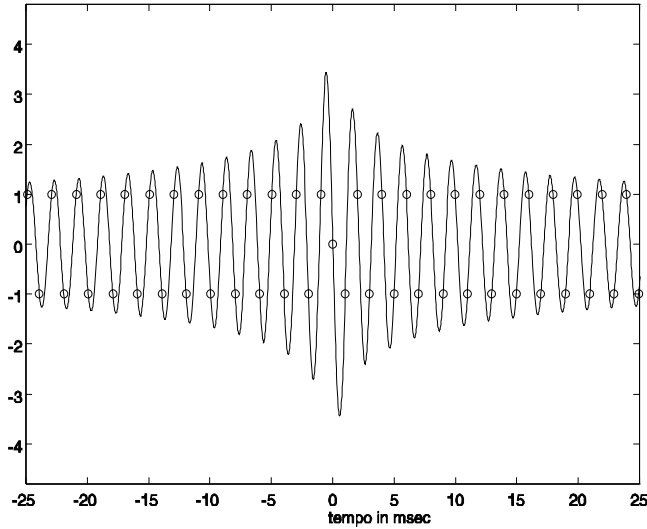


Figura 3.4: Esempio di sovraelongazioni dovute all'eccessiva rigidità del filtro ideale usato come interpolatore

3.5 La linea elastica o "spline"

Esiste un altro interpolatore cubico, diverso da $H_3(\omega)$ e dall'interpolatore cubico troncato visto in precedenza che non solo rispetta rette, parabole e cubiche ma rispetta anche i dati, cosa che non accade per $H_3(\omega)$. Questo interpolatore, noto come linea o bacchetta elastica (inglese: *spline* = bacchetta elastica), è molto utilizzato sia in sistemi monodimensionali che multidimensionali.

Intuitivamente, si può pensare di tracciare su un foglio di carta le posizioni dei campioni del segnale e puntarvi degli spilli. Per costruire una curva continua che li congiunga, si può usare un filo teso tra uno spillo e l'altro (ed allora risulta l'interpolazione lineare) oppure usare come curvilineo una bacchetta flessibile (un filo di acciaio, p.e.) la cui posizione di equilibrio rappresenta il segnale interpolato (fig. 3.5), o magari ancora una via di mezzo come una bacchetta elastica sottoposta a tensione.

L'interpolazione effettuata con la bacchetta elastica (la linea elastica) onora costanti e rette (ma anche parabole e cubiche) ed anche il dato. Peraltro, il risultato dell'interpolazione dipende anche da come la bac-

3.5 LA LINEA ELASTICA O "SPLINE"

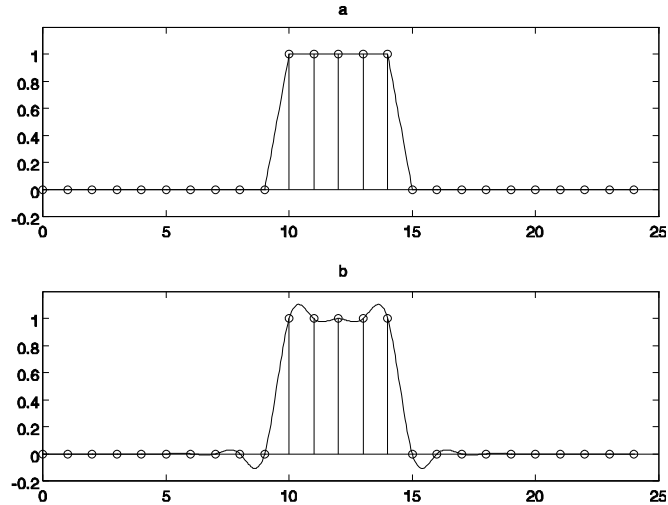


Figura 3.5: Interpolazione di un impulso rettangolare: lineare (a); con bacchetta elastica (b).

chetta è fissata agli estremi dell'intervallo. Lasciando liberi gli estremi, la posizione di equilibrio della bacchetta corrisponde al minimo dell'energia elastica immagazzinata e pertanto la curvatura media della linea è minima; il segnale interpolato è allora, pur sempre onorando i dati, il più *liscio* possibile. La bacchetta proseguirà fuori dell'intervallo in cui è nota la sequenza fornendo una particolare estrapolazione dei dati. Altrimenti, si può incastrare agli estremi la bacchetta fissando così la derivata prima del segnale interpolato agli estremi dell'intervallo (non necessariamente a zero). Ancora, si può pensare di sottoporre a tensione gli estremi della bacchetta. Si ha quindi un parametro libero in più: se la tensione è fortissima, siamo di nuovo nel caso dell'interpolazione lineare. Se la tensione è nulla, siamo nel caso esposto di seguito. In sostanza, si dispone di una grande libertà nello scegliere tra un'interpolazione a tratti di rette (interpolazione lineare) fino ad un'interpolazione con archi di cubica a derivata seconda continua nei nodi (splines).

3.5.1 Analisi dello "spline"

Passando ad un'analisi quantitativa della posizione di equilibrio della bacchetta, osserviamo che la sua curvatura è proporzionale al momento

CAPITOLO 3 RICOSTRUZIONE DI SEGNALI CONTINUI

flettente in quel punto; le reazioni vincolari dei nodi sono forze che la inflettono: il momento flettente dovuto all'insieme delle reazioni vincolari, ha un andamento lineare a tratti, cambiando pendenza in corrispondenza di ogni nodo e di ogni nuova reazione vincolare. In conclusione, la posizione di equilibrio della bacchetta e pertanto la curva interpolante ha una curvatura che varia linearmente tra un nodo e l'altro, ma in modo continuo. Se si fa l'ipotesi addizionale che gli scostamenti $y(x)$ della bacchetta dall'asse delle ascisse siano molto piccoli, si può dire che la sua curvatura è proporzionale alla derivata seconda della curva, che è quindi lineare a tratti. Infatti si ha che la curvatura di una linea, inverso del raggio di curvatura $r(x)$ è:

$$\frac{1}{r(x)} = \frac{y''(x)}{[1 + y'(x)^2]^{3/2}} \simeq y''(x); \quad |y'(x)| < 1 \quad (3.14)$$

In ogni tratto tra due campioni, la curva ricostruita sarà pertanto una cubica, definita cioè con quattro parametri, ad esempio i valori della funzione e delle derivate in corrispondenza dei campioni. Le derivate sono incognite e vanno determinate utilizzando le condizioni di continuità delle derivate prime e seconde da un tratto di curva al successivo.

3.5.2 L'equazione dello spline

Indichiamo con y_i, y'_i i valori della funzione da interpolare e della sua derivata in corrispondenza del campione i -esimo. La funzione interpolante cercata è un polinomio di terzo grado (e cioè con una derivata seconda che varia linearmente con la variabile x). Nell'intervallo $i, i + 1$ la funzione interpolante $f(x)$ è descritta dalla seguente equazione:

$$f(x) = y_i(1 - 3x^2 + 2x^3) + y_{i+1}(3x^2 - 2x^3) + y'_i(x - 2x^2 + x^3) + y'_{i+1}(-x^2 + x^3) \quad (3.15)$$

Per verificare l'equazione precedente, si supponga diverso da zero solo uno dei quattro parametri $y_i, y_{i+1}, y'_i, y'_{i+1}$, e si suppongano nulli gli altri tre. I polinomi che risultano, hanno appunto valore nullo in loro corrispondenza: ad esempio, nel caso di y_i , il polinomio $1 - 3x^2 + 2x^3$ ha valore 1 in $x = 0$ è nullo per $x = 1$ e le sue derivate sono nulle in

3.5 LA LINEA ELASTICA O "SPLINE"

$x = 0, 1$. Nel caso di y'_{i+1} , il polinomio $-x^2 + x^3$ ha valore 0 in $x = 0$; 1 ha derivate prime nulle in 0 ed unitaria in 1. Pertanto la $f(x)$ descritta dall'equazione precedente è il più generale polinomio di terzo grado che abbia assegnati i valori della funzione e delle derivate agli estremi in 0, 1.

E' possibile calcolare la derivata seconda del polinomio ed imporre che al nodo i -esimo le derivate seconde calcolate per l'intervallo $i, i + 1$ e $i - 1, i$ coincidano. Ne risulta la seguente equazione alle differenze finite, che ci permette di calcolare i valori delle derivate della funzione da interpolare in corrispondenza dei nodi:

$$y'_{i-1} + 4y'_i + y'_{i+1} = 3(y_{i+1} - y_{i-1}) \quad (3.16)$$

Questa equazione alle differenze finite sarà studiata con le tecniche della trasformata z nel capitolo 4. Qui osserviamo anche che imponendo che la derivata seconda sia nulla ai nodi terminali per imporre la condizione di immagazzinamento della minima energia elastica, si ottiene:

$$y'_0 = \frac{-y'_1 + 3(y_1 - y_0)}{2} \quad (3.17)$$

$$y'_N = \frac{-y'_{N-1} + 3(y_{N-1} - y_N)}{2} \quad (3.18)$$

Abbiamo quindi un sistema di N equazioni nelle N incognite (le y'_i) che è risolvibile facilmente essendo tridiagonale (e cioè avendo nella matrice dei coefficienti solo la diagonale e le due subdiagonali adiacenti diverse da zero). La soluzione viene fatta per sostituzione, a partire dalla prima equazione.

$$\begin{bmatrix} 2 & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & 4 & 1 & \dots & 0 & 0 & 0 \\ 0 & 1 & 4 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 4 & 1 & 0 \\ 0 & 0 & 0 & \dots & 1 & 4 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} y'_0 \\ y'_1 \\ y'_2 \\ \dots \\ y'_{N-2} \\ y'_{N-1} \\ y'_N \end{bmatrix} = \quad (3.19)$$

CAPITOLO 3 RICOSTRUZIONE DI SEGNALI CONTINUI

$$= \begin{bmatrix} -3 & 3 & 0 & \dots & 0 & 0 & 0 \\ -3 & 0 & 3 & \dots & 0 & 0 & 0 \\ 0 & -3 & 0 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 3 & 0 \\ 0 & 0 & 0 & \dots & -3 & 0 & 3 \\ 0 & 0 & 0 & \dots & 0 & -3 & 3 \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \dots \\ y_{N-2} \\ y_{N-1} \\ y_N \end{bmatrix}$$

Una volta trovate le N incognite y'_i esse vanno sostituite nell'equazione (13.8) e quindi si determinano i vari archi di polinomi di terzo grado.

Si osservi anche che il concetto di bacchetta elastica si estende con facilitá al caso di campioni non equidistanti; in questo caso, i coefficienti lungo la diagonale e le subdiagonali non sono costanti.

3.5.3 La risposta all'impulso dello spline

Vogliamo ora determinare la risposta all'impulso dello spline $s_0(t)$ cioè la funzione che si ottiene quando solo $y_0 = 1$ e tutti gli altri $y_i = 0$; $-\infty < i < \infty$.

Lo spettro di $s_0(t)$ si può calcolare con il seguente metodo. Anzitutto, dall'equazione alle differenze finite (13.8) si trova la trasformata di Fourier della sequenza delle derivate della $s_0(t)$ in corrispondenza dei campioni; infatti indicando rispettivamente con $Y'_c(\omega)$, $Y_c(\omega)$ le trasformate di Fourier delle sequenze delle derivate e dei campioni.

$$y'_{i-1} + 4y'_i + y'_{i+1} = 3(y_{i+1} - y_{i-1}) \rightarrow \quad (3.20)$$

$$Y'_c(\omega)[e^{j\omega T} + 4 + e^{-j\omega T}] = 3Y_c(\omega)[e^{j\omega T} - e^{-j\omega T}]$$

$$Y'_c(\omega) = Y_c(\omega) \frac{3j \sin \varphi}{2 + \cos \varphi}; \quad \varphi = \omega T \quad (3.21)$$

Nel caso particolare in cui $Y_c(\omega) = 1$, e cioè per la risposta all'impulso si ha:

$$Y'_c(\omega) = \frac{3j \sin \varphi}{2 + \cos \varphi} \simeq j\varphi; \quad (3.22)$$

3.5 LA LINEA ELASTICA O "SPLINE"

Si osservi che la sequenza di campioni y'_n con trasformata $Y'_c(\omega)$ è corrispondente ad un derivatore numerico. Questa funzione di trasferimento somiglia molto alla funzione di trasferimento desiderata per un derivatore senza avere valori infiniti, come accadrà per il derivatore ricorsivo. Infatti

$$Y'_c(\omega) \simeq j\omega Y_c(\omega) = j\omega T Y_c(\omega) \left(1 - \frac{1}{180} \omega^4 T^4 \dots \right); \quad |\omega T| < 1 \quad (3.23)$$

Vedremo poi nel seguito altre tecniche per trovare le derivate di una

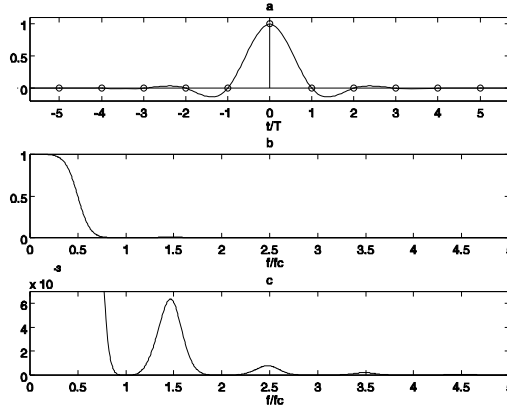


Figura 3.6: Risposta all'impulso (a) e trasformata (b) dello spline; in (c) è ingrandita la trasformata per vederne i valori in corrispondenza delle frequenze $n f_c$.

sequenza numericamente e cioè trovare, agli istanti di campionamento, la sequenza delle derivate del segnale continuo ottenuto filtrando la sequenza iniziale con un filtro ideale.

Utilizzando poi le equazioni che danno la funzione a partire dai valori ai nodi dei campioni e delle derivate, si trova la trasformata della risposta all'impulso cercata. Risulta che la risposta all'impulso della bacchetta elastica ha trasformata $S_0(\omega)$ (fig. 4.40) con equazione:

$$S_0(\omega) = \frac{3 \operatorname{sinc}^4(fT)}{1 + 2 \cos^2 \pi fT} \quad (3.24)$$

In particolare $S_0(\omega)$ ha zeri di quarto ordine alle frequenze multiple di $1/T$ ed è costante fino al quarto ordine nell'origine.

3.6 Il campionamento irregolare

3.6.1 Conversione di frequenze di campionamento

In molte occasioni può essere utile convertire un segnale da una frequenza di campionamento ad un'altra: p.e. inserire un segnale audio da un CD in una linea telefonica numerica e quindi passare da circa 44KHz a 8KHz di frequenza di campionamento; a parte il necessario filtraggio per sopprimere l'alias, sarà anche necessario interpolare il segnale su un altro pettine di campionamento, totalmente indipendente; ovviamente, sarebbe possibile ritornare ad un segnale analogico e ricampionare, ma ciò comporta delle perdite di rapporto segnale rumore. Ancora, è utile mescolare sorgenti numeriche di origini diverse, anche se nominalmente alla stessa frequenza di campionamento; per esempio, si vogliono mescolare segnali audio registrati da varie posizioni in un teatro, durante lo stesso concerto, da diversi registratori numerici. I vari clock usati per le registrazioni possono essere imperfetti e le loro frequenze possono pendolare rispetto alla frequenza nominale: i segnali audio devono essere allineati temporalmente e poi i segnali numerici corrispondenti devono essere ricampionati in modo simultaneo. I vari campioni devono essere interpolati. Un ulteriore caso è la tecnica analogica varispeed in cui l'asse temporale viene accelerato o ritardato per ottenere degli effetti audio; è opportuno realizzare la stessa funzione con tecniche numeriche.

Inoltre, vi sono variazioni casuale della frequenza di orologio dovute a imperfezioni del sistema di trasmissione numerico indicate con il nome di **jitter**. È approssimativamente possibile rimuovere il jitter di trasmissione da una sequenza; non quello di registrazione, tuttavia, a meno di non disporre di altre fonti di sincronizzazione come ad esempio i segnali stessi. La sequenza ricevuta viene memorizzata su una memoria FIFO (First In First Out), e quindi questa memoria viene riletta al ritmo opportuno. Questo ritmo deve essere un pò più veloce se la memoria va svuotata, più lento se la memoria tende a riempirsi. I tempi di restituzione sono confrontati con i tempi dell'orologio di campionamento, registrati, ed i campioni ai tempi necessari sono interpolati.

In ogni modo, la conversione della frequenza di campionamento va fatta in due passi: per prima cosa si ottiene una sequenza, approssimativamente campionata al ritmo desiderato, ma che conserva le eventuali imprecisioni del ritmo di campionamento iniziale. Ad esempio la si sovra-

3.6 IL CAMPIONAMENTO IRREGOLARE

campiona di un fattore M per poi sottocampionarla di un fattore N . Per fare ciò, si adottano le tecniche di interpolazione e decimazione discusse nel capitolo 7. Quindi, si fa un'operazione di precisione, calcolando i campioni di una sequenza esattamente temporizzata.

In sostanza, si dispone di una sequenza registrata a tempi "errati" ed irregolari ma noti; vogliamo ricostruire la sequenza in modo corretto. È sorprendente la precisione dell'interpolazione necessaria per garantire un buon rapporto segnale - rumore p.e. 100 dB, corrispondente ai 16 bit con cui i campioni sono registrati su un CD. Il segnale desiderato è:

$$s(nT); \quad -\infty < n < \infty$$

Il segnale disponibile sia invece :

$$s(t_n); \quad t_n - nT = \tau_n \ll T$$

$$s(nT) = s(t_n - \tau_n) \approx s(t_n) - \tau_n \left. \frac{ds(t)}{dt} \right|_{t=t_n}$$

L'errore quadratico medio è pertanto:

$$\overline{\varepsilon^2} = \overline{\tau^2} \times \overline{\left| \frac{ds}{dt} \right|^2}$$

Nel caso di sequenze a spettro uniforme, utilizzando il teorema di Parseval si ottiene, indicando con σ^2 il valore quadratico medio del segnale:

$$\left| \frac{ds}{dt} \right|^2 = \sigma^2 \frac{\pi^2}{3T^2} \longrightarrow SNR^{-1} = 10^{-10} = \frac{\overline{\varepsilon^2}}{\sigma^2} = \frac{\pi^2}{3} \times \frac{\overline{\tau^2}}{T^2} \longrightarrow \sqrt{\frac{\overline{\tau^2}}{T^2}} = .55 \times 10^{-5}$$

L'errore di temporizzazione nell'interpolazione deve essere dell'ordine del **milionesimo di campione** se i successivi errori temporali sono perfettamente casuali e quindi creano un disturbo casuale. Se invece supponiamo che i successivi valori τ_n siano tutti eguali tra loro (o lentamente variabili) allora l'errore diventa praticamente inosservabile perchè equivale ad avere un errore nella funzione di trasferimento pari a:

$$\Delta(\omega) = e^{j\omega\tau}$$

Tuttavia questo ritardo (o anticipo) indesiderato va ridotto a valori accettabili (ben maggiori del milionesimo di campione) se si vogliono temporizzare più segnali assieme.

3.6.2 Interpolazione irregolare con la funzione *sinc*

Rimane da discutere il problema dell'interpolazione di funzioni campionate irregolarmente e cioè ad intervalli ineguali di tempo o spazio. In linea teorica, per N tanto grande da poter applicare in pratica un filtro ideale, se si hanno a disposizione N campioni nell'intervallo NT , anche se non equispaziati, è tuttora possibile determinare gli N campioni equispaziati originali e di qui l'interpolazione con un filtro ideale; basta scrivere un sistema di N equazioni in N incognite dove i termini noti sono gli N valori assunti dalla funzione x_n nei tempi assegnati (t_n irregolari) e le incognite i valori dei campioni a_n ai tempi nT . Il sistema:

$$x(t) = \sum_{n=-N/2}^{+N/2-1} a_n \operatorname{sinc}\left(\frac{t-nT}{T}\right); \quad x(t_n) = x_n; \quad n = -N/2, \dots, N/2-1 \quad (3.25)$$

ha tante equazioni quante incognite. In termini matriciali si ha:

$$\underline{\mathbf{G}} \underline{\mathbf{a}} = \underline{\mathbf{x}}$$

dove $\underline{\mathbf{x}}$, $\underline{\mathbf{a}}$ sono i vettori di componenti $\{x_n\}$, $\{a_n\}$ rispettivamente e la matrice $\underline{\mathbf{G}}$ ha componenti

$$g_{i,j} = \operatorname{sinc}\left(\frac{t_i - t_j}{T}\right)$$

Il sistema diventa mal condizionato non appena il campionamento diventa fortemente irregolare; se vi sono due campioni molto vicini, due righe della matrice $\underline{\mathbf{G}}$ saranno molto simili tra loro ed un autovalore almeno della matrice $\underline{\mathbf{G}}$ sarà molto basso, ecc. Quindi, questa tecnica di interpolazione, apparentemente infallibile, è destinata a dare risultati poco utili.

Capitolo 4

La trasformata z

4.1 Introduzione: sequenze e polinomi

Nell'elaborazione numerica dei segnali viene fatto continuo uso della Trasformata z , e cioè di una corrispondenza che viene stabilita tra sequenze di campioni (complessi o reali) e polinomi in una variabile complessa che verrà indicata con z . In questo capitolo cercheremo di riassumere le caratteristiche della trasformata z usando solo concetti elementari. Anzi tutto ricordiamo che la trasformata z di una sequenza di lunghezza finita $\{a_{1,n}\}$ contenuta tra le posizioni indicizzate con $-M, \dots, N$ è il "polinomio" (con esponenti sia positivi che negativi) nella variabile complessa z :

$$A_1(z) = \sum_{n=-M}^N a_{1,n} z^{-n} \quad (4.1)$$

Ad esempio, la sequenza costituita da un campione di ampiezza 2 al posto caratterizzato dall'indice $n = 18$ e di un altro campione di ampiezza -3 al posto caratterizzato dall'indice $n = -1$ è il "polinomio":

$$A_1(z) = -3 \times z + 2 \times z^{-18}$$

Quindi la trasformata z di una sequenza $\{a_{1,n}\}$ di lunghezza finita a meno del fattore z^M è il "polinomio" nella variabile complessa z^{-1} :

$$A_1(z) = z^M P(z^{-1})$$

dove $P(z)$ è un polinomio vero e proprio (contenente solo potenze positive cioè) di ordine $M+N$. Ritardare di K campioni una sequenza corrisponde

CAPITOLO 4 LA TRASFORMATA Z

a moltiplicarne la trasformata z per z^{-K} e viceversa anticiparla di H campioni corrisponde a moltiplicarne la trasformata per z^H .

Vediamo ora perchè la trasformata z serve: consideriamo la sequenza $\{c_n\}$ convoluzione di due sequenze $\{a_n\}$, $\{b_n\}$. Si abbia cioè:

$$\{c_n\} = \{a_n\} * \{b_n\}$$

$$c_n = \sum_k a_{n-k} b_k$$

La trasformata z della sequenza $\{c_n\}$ è il prodotto delle trasformate z delle sequenze $\{a_n\}$, $\{b_n\}$. Ciò è immediatamente verificabile mediante la sovrapposizione degli effetti, p.e. rappresentando la sequenza $\{b_n\}$ come somma di impulsi ritardati di k campioni, con k variabile. La sequenza $\{c_n\}$ è la somma delle repliche della sequenza $\{a_n\}$ progressivamente sfalsate di k campioni e moltiplicate per b_k . Dette infatti:

$$A(z) = \sum_l a_l z^{-l}$$

$$B(z) = \sum_k b_k z^{-k}$$

$$C(z) = \sum_n c_n z^{-n}$$

si ha:

$$C(z) = \left(\sum_k b_k z^{-k} \right) \left(\sum_l a_l z^{-l} \right) = B(z) \times A(z)$$

Poi, è anche vero che moltiplicando due "polinomi" se ne ottiene un terzo che ha per sequenza dei coefficienti la convoluzione delle sequenze dei coefficienti dei primi due. A prodotti di convoluzione di sequenze (indicati con $*$), estremamente utili da calcolare, corrispondono quindi i prodotti delle rispettive trasformate z .

$$C(z) = \left(\sum_k b_k z^{-k} \right) \left(\sum_l a_l z^{-l} \right) = \tag{4.2}$$

$$= B(z) \times A(z) = \sum_k \sum_l a_l b_k z^{-(k+l)} \tag{4.3}$$

4.1 INTRODUZIONE: SEQUENZE E POLINOMI

e posto:

$$n = k + l$$

si verifica che:

$$C(z) = \sum_n \left(\sum_k a_{n-k} b_k \right) z^{-n} = \sum_n c_n z^{-n}$$

Per semplicitá supporremo che le sequenze studiate abbiano sempre il primo campione nell'origine ($M = 0$) e pertanto le loro trasformate z siano polinomi che contengono il termine di grado zero.

Per il teorema fondamentale dell'algebra ogni polinomio di ordine N a coefficienti complessi ha N radici nel campo complesso z_1, z_2, \dots, z_N ; quindi **ogni sequenza di $N + 1$ campioni complessi che inizia a partire dall'origine dei tempi (il primo campione è posto nella posizione di indice 0) si può decomporre nella convoluzione di N sequenze elementari costituite da 2 campioni complessi di cui il primo è di ampiezza unitaria ed è posizionato nell'origine ed il secondo, posizionato al posto di indice 1, ha valore $-z_n, n = 1, \dots, N$.** Infatti una qualsiasi sequenza di N campioni che inizia nell'origine ha trasformata z pari a:

$$A(z) = \sum_{n=0}^N a_n z^{-n}$$

Ma, per il teorema fondamentale dell'algebra possiamo scrivere:

$$A(z) = \sum_{n=0}^N a_n z^{-n} = a_0 \prod_{n=1}^N (1 - z_n z^{-1}); \quad (4.4)$$

Abbiamo mostrato che la sequenza di $N+1$ campioni $\{a_n; n = 0, 1, 2, \dots, N\}$ è perciò decomponibile nel prodotto di convoluzione che segue:

$$\{a_n\} = a_0 * \{1, -z_1\} * \{1, -z_2\} * \{1, \dots\} * \{1, -z_N\} \quad (4.5)$$

Questo fatto, che **non** si estende a due o piú dimensioni, ci porta a grandi semplificazioni; infatti, qualsiasi sequenza di lunghezza finita potrà essere vista come la convoluzione di sequenze elementari di due campioni, che chiameremo zeri per il motivo indicato nella sezione 3.

4.2 Sistemi lineari discreti

Un sistema lineare, a cui si applica la sovrapposizione degli effetti, è caratterizzato da un ingresso ed un'uscita; la linearità permette di calcolare l'uscita supponendo che l'ingresso sia costituito da una successione di segnali impulsivi, che possono essere considerati uno per volta; nel caso discreto, che è quello di interesse nel seguito, il sistema ha per ingresso la sequenza illimitata di campioni $\{x_n\}$ e per uscita la sequenza $\{y_n\}$; **l'intervallo di campionamento in ingresso ed uscita è identico.**

La risposta del sistema ad un impulso nella posizione i -esima sarà una sequenza $h_{i,k}$ con l'indice k compreso tra i valori $-M_i$ e N_i ; la sequenza $h_{i,k}$ può cambiare al variare dell'indice i . Sempre per la sovrapposizione degli effetti, l'uscita sarà

$$y_n = \sum_{k=-M_{n-k}}^{N_{n-k}} x_{n-k} h_{n-k,k} = \sum_{l=n+M_l}^{n-N_l} x_l h_{l,n-l}$$

Ora, supponiamo che la risposta all'impulso del sistema non dipenda dall'indice ma sia tempo (spazio) invariante. L'uscita è

$$y_n = \sum_{k=-M}^N x_{n-k} h_k = \sum_{l=n+M}^{n-N} x_l h_{n-l}; \quad \{y_n\} = \{x_n\} * \{h_n\}$$

e cioè la convoluzione tra le due sequenze $\{x_n\}$ e $\{h_n\}$.

Nel caso che la sequenza di ingresso ad un sistema tempo (spazio) invariante sia una sinusoide campionata è noto (ma lo dimostreremo di nuovo nella prossima sezione) che la risposta è ancora una sinusoide campionata della stessa frequenza di quella di ingresso ma di diversa ampiezza e fase, dovute al transito per il sistema; le caratteristiche di ampiezza e di fase del sistema lineare (al variare della frequenza della sinusoide di ingresso) ne corrispondono alla **funzione di trasferimento**.

4.3 Sinusoidi campionate; zeri

Abbiamo visto come ogni sequenza si può decomporre in una convoluzione di tante sequenze elementari di due campioni; per trovare la funzione di trasferimento di una sequenza di $N + 1$ campioni e quindi cosa accade

4.3 SINUSOIDI CAMPIONATE; ZERI

quando convolviamo una sinusoidi (campionata) con una sequenza di $N + 1$ campioni, basta convolvere questa sinusoidi successivamente con tutte le N sequenze elementari componenti (la funzione di trasferimento della sequenza è quindi il prodotto delle funzioni di trasferimento delle sequenze elementari). Ci riferiremo a sinusoidi generalizzate, e cioè esponenziali complessi il cui modulo può crescere o decrescere geometricamente con l'indice n . Queste sinusoidi generalizzate sono delle sequenze $\{x_n\}$ illimitate descritte dalla formula:

$$x_n = B\rho^n \times \exp(j2\pi fnT) ; \quad -\infty \leq n \leq \infty \quad (4.6)$$

dove B è un numero complesso e la frequenza f può avere un valore qualsiasi; peraltro è inutile fare uscire f dall'intervallo $-f_N < f < f_N$ dove $f_N = 1/2T$ è la frequenza di Nyquist, e ciò per evitare equivocazione. Nel caso $\rho = 1$ si ritrovano le sinusoidi complesse; le consuete sinusoidi e cosinusoidi ne costituiscono le parti reali ed immaginaria: si ha cioè:

$$B = B_R + jB_I = |B|e^{j\varphi} ; \quad |B| = \sqrt{B_R^2 + B_I^2}; \quad \tan \varphi = \frac{B_I}{B_R}$$

$$Be^{j2\pi fnT} = |B| [\cos(2\pi fnT + \varphi) + j \sin (2\pi fnT + \varphi)];$$

Indichiamo ora con z_1 la variabile complessa

$$z_1 = \rho e^{j2\pi fT}$$

Una sinusoidi generalizzata é quindi la sequenza:

$$\{x_n\} = \{Bz_1^n\}; \quad -\infty \leq n \leq \infty \quad (4.7)$$

Se si convolve una sinusoidi generalizzata con la sequenza di due campioni $1, -z_0$ posizionati nell'origine e nel campione successivo, si osserva che il segnale in uscita è:

$$\{Bz_1^n\} * \{1, -z_0\} = \{Bz_1^n - Bz_0z_1^{n-1}\} = \{B(1 - z_0z_1^{-1})z_1^n\} \quad (4.8)$$

e cioè ancora la stessa sinusoidi generalizzata, moltiplicata per il fattore complesso $1 - z_0z_1^{-1}$, nullo quando

$$z_1 = z_0$$

CAPITOLO 4 LA TRASFORMATA Z

Piú in generale, utilizzando la formula (4.4) si verifica che la convoluzione di una sinusoide generalizzata con una sequenza qualsiasi può essere ottenuta come la cascata delle convoluzioni con le sequenze elementari di due campioni in cui la sequenza è stata decomposta.

Pertanto, il risultato della convoluzione di una sinusoide generalizzata campionata ed una sequenza qualsiasi $\{a_n\}$ è ancora la stessa sinusoide generalizzata moltiplicata per un fattore complesso che è definito come la **funzione di trasferimento della sequenza**; questo fattore è $A(z_1)$ e cioè il valore che assume la trasformata z della sequenza, assegnando all'argomento il valore z_1 della frequenza generalizzata della sinusoide $z_1 = \rho e^{j2\pi fT}$.

Abbiamo quindi trovato un significato fisico alla variabile z . Peraltro, il significato ovvio della variabile z è l'anticipo di un campione. Con ciò si intende che se una trasformata di una sequenza è moltiplicata per z , i campioni corrispondenti alla nuova trasformata sono quelli di prima, ma anticipati di un intervallo T . Un qualsiasi valore di z può rappresentare l'argomento di una sinusoide generalizzata; il valore $A(z_1)$ della trasformata z di una sequenza $\{a_n\}$ calcolata per il valore z_1 è il fattore per cui deve essere moltiplicata la sinusoide generalizzata di argomento z_1 quando viene convoluta con la sequenza $\{a_n\}$.

Si comprende perché una sequenza di due campioni viene indicata come uno **zero**: infatti la radice (lo zero) del polinomio di primo grado corrispondente alla sua trasformata z è l'argomento della sinusoide generalizzata che, convoluta con la sequenza elementare, viene annullata.

4.4 Il piano z

L'argomento delle sinusoidi generalizzate è rappresentabile come un punto di un piano complesso (figura 4.2) detto piano z ; la distanza dall'origine corrisponde a ρ ; l'angolo formato con l'asse delle ascisse è $\omega T = 2\pi fT$. In questo piano, viene tracciato un cerchio di raggio unitario, che è il luogo degli argomenti corrispondenti a sinusoidi campionate vere e proprie. La sinusoide a frequenza 0, (il segnale costante) corrisponde al punto $z_1 = 1 + j \times 0$; la sinusoide alla frequenza di Nyquist è invece corrispondente al punto $z_2 = -1 + j \times 0$.

4.4 IL PIANO Z

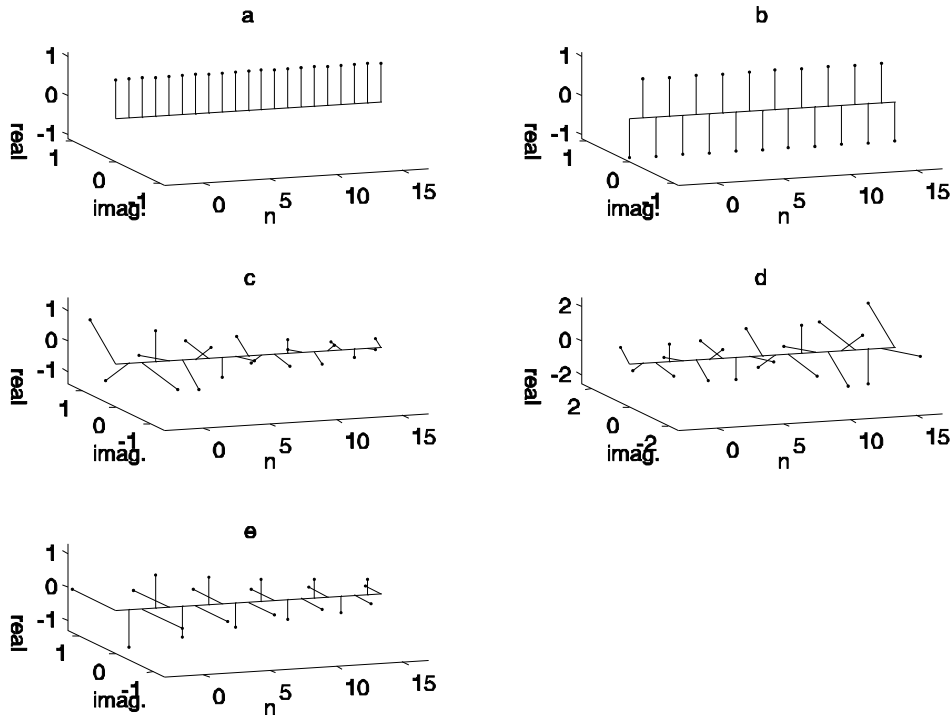


Figura 4.1: Esempi di sinusoidi generalizzate campionate $x_n = A\rho^n e^{jn\varphi}$; $\varphi = 2\pi f/f_c$: a) $\rho = 1$, $f/f_c = 0$; b) $\rho = 1$, $f/f_c = 0.5$; c) $\rho = 0.95$, $f/f_c = 2/5$; d) $\rho = 1.05$, $f/f_c = 2/5$; e) $\rho = 0.95$, $f/f_c = 1/4$.

4.4.1 Le sequenze nel piano z

Una qualsiasi sequenza di $N + 1$ campioni è rappresentabile nel piano z con la disposizione degli N zeri del polinomio corrispondente alla sua trasformata z ; compare anche una singolarità (un polo di ordine N nell'origine del piano z) che però non ha un vero e proprio significato fisico e nasce dalla scelta dell'esponente negativo per rappresentare i ritardi nella trasformata z .

Come già osservato, è immediato trovare la funzione di trasferimento corrispondente ad una sequenza qualsiasi; infatti, basta calcolare il valore della trasformata z della sequenza sul punto del cerchio unitario corrispondente alla frequenza desiderata. Procedendo per via geometrica, si vede che nel caso di una sequenza di due campioni, la funzione di

CAPITOLO 4 LA TRASFORMATA Z

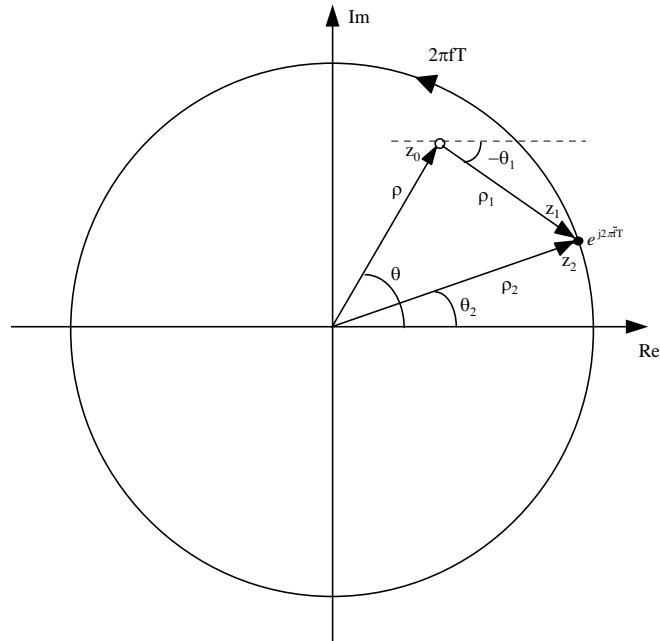


Figura 4.2: Calcolo della F.d.T. di una sequenza relativa ad uno zero per via geometrica

trasferimento é (fig. 4.2):

$$A(z) = 1 - z_0 z^{-1} = \frac{(z - z_0)}{z}$$

e quindi corrisponde al rapporto tra i due vettori complessi che congiungono il punto rappresentativo della sinusoidale sul cerchio unitario e lo zero e l'origine del piano z rispettivamente. Le caratteristiche di ampiezza e di fase sono:

$$|A(z)|_{z=\exp(j\omega T)} = \frac{|z - z_0|}{|z|} = |z - z_0|$$

$$\angle A(z)_{z=\exp(j\omega T)} = \angle (z - z_0) - \angle z$$

E' quindi possibile tracciare semplicemente la caratteristica di ampiezza corrispondente a zeri variamente posizionati nel piano z ed a sequenze piú complesse; ovviamente, la funzione di trasferimento di una sequenza é il prodotto delle funzioni di trasferimento delle sequenze elemen-

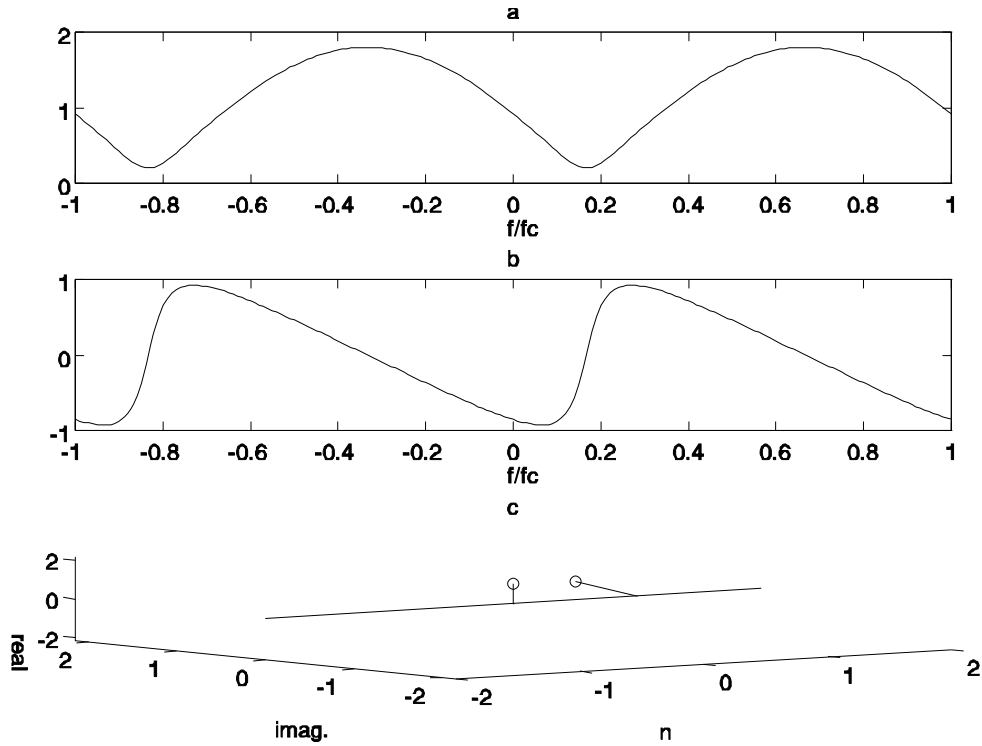


Figura 4.3: Caratteristiche di ampiezza e fase dello zero complesso della figura 4.2

tari così come la sequenza è convoluzione delle sequenze elementari. Sequenze reali hanno zeri reali o complessi coniugati, come ovvio visto che corrispondono a radici di polinomi a coefficienti reali.

4.4.2 Zeri a fase minima e fase massima

Si possono avere zeri sia sul cerchio unitario (ed allora annullano sinusoidi vere e proprie) che fuori (annullano sinusoidi generalizzate crescenti con l'indice) che infine dentro (annullano sinusoidi generalizzate decrescenti con l'indice).

Gli zeri posizionati dentro il cerchio unitario sono chiamati zeri a fase minima e zeri a fase massima quelli all'esterno (fig. 4.4,4.5). Il motivo si evince dalle caratteristiche di fase delle rispettive funzioni di

CAPITOLO 4 LA TRASFORMATA Z

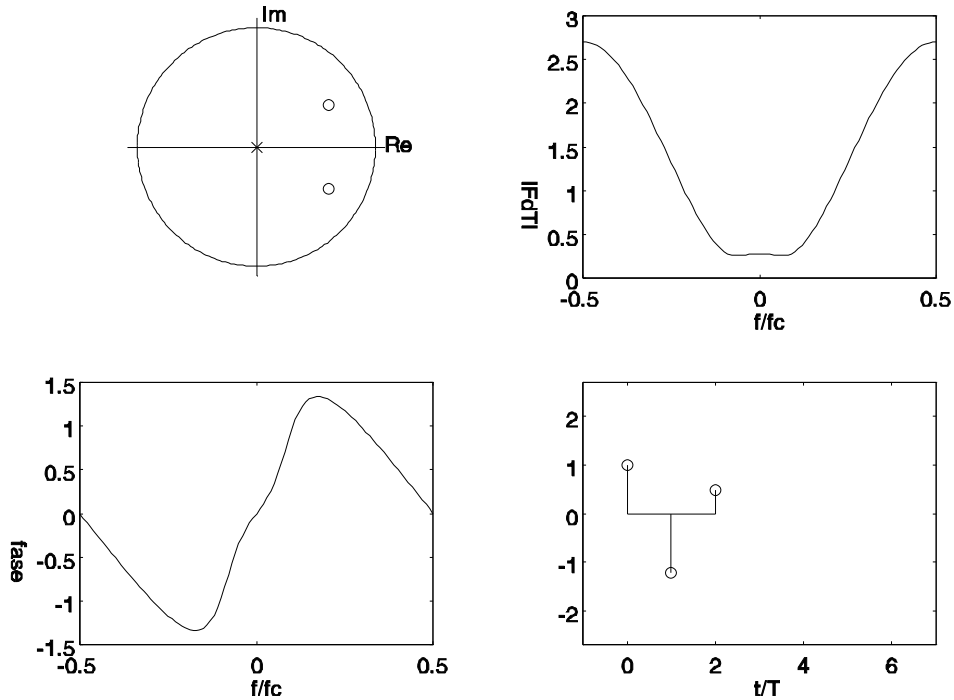


Figura 4.4: Caratteristiche di fase ed ampiezza di una F.d.T a fase minima

trasferimento. Calcoliamo anzitutto la caratteristica di fase di uno zero a frequenza zero, posto sul cerchio unitario. Si abbia cioè:

$$A(z) = 1 - z^{-1}$$

La caratteristica di fase e cioè lo sfasamento apportato a sinusoidi di pulsazione ω è:

$$\begin{aligned} \angle(A(z)|_{z=\exp(j\omega T)}) &= \angle(1 - e^{-j\omega T}) = \\ &= \angle\left(2e^{-j\omega T/2}j \sin \frac{\omega T}{2}\right) = \left(\frac{\pi}{2} - \frac{\omega T}{2}\right) \text{sign}(\omega) \end{aligned}$$

discontinua nell'origine. Si consideri ora uno zero sempre a frequenza zero, ma non sul cerchio unitario. Sia quindi:

$$A(z) = 1 - \rho z^{-1}$$

4.4 IL PIANO Z

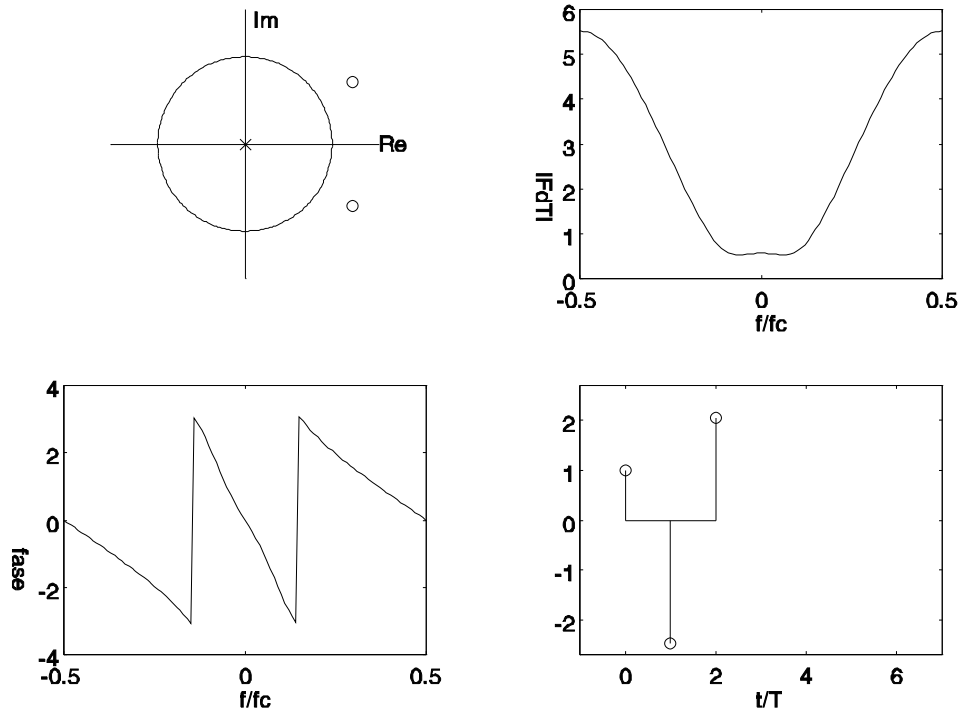


Figura 4.5: Caratteristiche di fase ed ampiezza di una F.d.T a fase massima

$$\angle (A(z) |_{z=\exp(j\omega T)}) = \angle (1 - \rho e^{-j\omega T}) ;$$

Se $\rho < 1$ e quindi lo zero è interno al cerchio unitario, la fase a frequenza 0 è 0 (vedi figura 4.6) e la caratteristica di fase a frequenza zero è continua. Se lo zero è fuori del cerchio unitario e cioè se $\rho > 1$, la fase a frequenza zero è π . Infatti, la somma dei campioni della sequenza è negativa, e infatti si osserva in figura una discontinuità di fase nell'origine pari a 2π .

Finora abbiamo implicitamente supposto ρ reale. Se invece si ha

$$\rho = |\rho| e^{j\omega T}$$

la frequenza dello zero è ω , e l'intero grafico di figura 4.6, periodico di periodo 2π , rimane immutato a parte una traslazione di un angolo ωT .

Il motivo della discontinuità della caratteristica di fase di 2π , nel caso di zero a fase massima, può essere anche compreso con la seguente considerazione. Quando costruiamo una caratteristica di fase, il punto

Figura 4.6: Caratteristiche di fase di uno zero a frequenza zero sia a fase minima che a fase massima ($\rho = 0,9; 1,1.1$). Nel caso dello zero a fase massima, la fase a frequenza zero è π . Quando lo zero è assai prossimo al cerchio unitario, variazioni minime dei valori dei campioni della sequenza possono portare a variazioni molto forti della caratteristica di fase.

rappresentativo della frequenza della sinusoide fa un intero giro lungo il cerchio unitario, dalla frequenza di Nyquist negativa a quella positiva ($-\pi \Leftrightarrow \pi$). Nel caso di uno zero interno al cerchio unitario e quindi a fase minima, la fase della funzione di trasferimento parte e ritorna al valore iniziale: infatti, sia il vettore z che il vettore $z - z_0$ fanno entrambi un intero giro intorno all'origine. Nel caso di uno zero a fase massima, la fase ritorna al valore iniziale ma deve avere fatto un salto di 2π . Infatti in questo secondo caso, il vettore z fa un intero giro intorno all'origine mentre il vettore $z - z_0$ no. Nel caso di uno zero a fase massima pertanto vi sarà sempre una discontinuità di 2π della fase della funzione di trasferimento.

4.4.3 Sequenze ribaltate: $\tilde{A}(z)$

E' interessante osservare che due zeri reciproci e coniugati (ossia due sequenze di 2 campioni le cui trasformate z sono caratterizzate da zeri $z_0, 1/z_0^*$ in posizione reciproca e coniugata, e cioè in posizione opposta rispetto al cerchio unitario) hanno eguale andamento della caratteristica di ampiezza (**ma solo sul cerchio unitario**) e diversa caratteristica di

fase. Ponendo $z = e^{j\varphi}$

$$|A(z)|^2 = |1 - z_0 e^{-j\varphi}|^2 = 1 + |z_0|^2 - 2\operatorname{Re}(z_0 e^{-j\varphi});$$

la sequenza con lo zero ribaltato $A_r(z)$, ha la seguente caratteristica di ampiezza (immutata, a meno di un fattore):

$$|A_r(z)|^2 = \frac{|z_0^* - e^{-j\varphi}|^2}{|z_0|^2} = \frac{1 + |z_0|^2 - 2\operatorname{Re}(z_0 e^{-j\varphi})}{|z_0|^2} = \frac{|A(z)|^2}{|z_0|^2}$$

Consideriamo una sequenza causale $\{a_n\}$; $0 \leq n \leq N$, e quindi $a_n = 0$; $\forall n < 0$ la cui trasformata z è un polinomio $A(z)$ a coefficienti complessi di grado N . Indicheremo con $\tilde{A}(z)$ lo stesso polinomio $A(z)$ ma con i coefficienti coniugati e nell'ordine inverso; esso è trasformata z della sequenza $\{\tilde{a}_n\}$ che si ottiene dalla sequenza $\{a_n\}$, coniugandola e ribaltandola rispetto all'istante 0 e ritardandola in modo da riavere una sequenza causale. Risulta che, se la sequenza corrisponde ad un polinomio $A(z)$ di grado N ed è quindi lunga $N + 1$ campioni, si ha:

$$\tilde{A}(z) = z^{-N} A^*(1/z); \quad |A(z)|_{z=e^{j\omega T}} = |\tilde{A}(z)|_{z=e^{j\omega T}} \quad (4.9)$$

Bisogna rammentare che il campione di posto 0 della sequenza $\{\tilde{a}_n\}$ è $\tilde{a}_0 = a_N \neq 1$.

4.4.4 Sequenze a fase minima, massima o mista

Una sequenza $A(z)$ la cui trasformata z ha tutti gli zeri dentro il cerchio unitario è detta una sequenza a fase minima. Se tutti gli zeri sono fuori del cerchio unitario, la sequenza è detta a fase massima; ad esempio, $\tilde{A}(z)$ è a fase massima, se $A(z)$ è a fase minima. La sequenza è detta a fase mista nei casi in cui parte degli zeri sono dentro e parte sono fuori del cerchio.

Convolgendo una sequenza $\{a_n\}$ di $N + 1$ campioni con la sequenza $\{\tilde{a}_n\}$ (quella che si ottiene ribaltando e ritardando la sequenza $\{a_n\}$ come visto al paragrafo precedente) si ottiene una sequenza lunga $2N + 1$ campioni, simmetrica rispetto al campione di posto N ; ciò si può verificare con una sequenza corrispondente ad uno zero e poi osservando che convoluzioni di sequenze simmetriche sono sequenze simmetriche. Nel dominio delle trasformate, si passa da $A(z)$ ad $\tilde{A}(z)$ ribaltando tutti gli N

CAPITOLO 4 LA TRASFORMATA Z

zeri della trasformata nella posizione reciproca rispetto al cerchio unitario (e coniugata se si tratta di zero complesso). Pertanto la caratteristica di fase di $A(z)\tilde{A}(z)$ sarà quella corrispondente ad un ritardo pari a NT e cioè:

$$\angle[A(z)\tilde{A}(z)]|_{z=e^{j\omega T}} = -N\omega T$$

Quindi la caratteristica di fase della trasformata z di una sequenza a fase massima $\tilde{A}(z)$ si deduce facilmente da quella della sequenza a fase minima $A(z)$ corrispondente.

$$\angle\tilde{A}(z)|_{z=e^{j\omega T}} = -\angle A(z)|_{z=e^{j\omega T}} - N\omega T$$

Ribaltando gli zeri rispetto al cerchio unitario, non cambia la caratteristica di ampiezza, come si è visto, ma si ribalta (e si ritarda) la sequenza nel tempo.

4.4.5 Trasformata di Fourier e trasformata z

E' evidente che la trasformata z di una sequenza di N campioni e cioè:

$$X(z) = \sum_{n=0}^{N-1} x_n z^{-n}$$

e la trasformata di Fourier della stessa sequenza:

$$X_F(f) = \sum_{n=0}^{N-1} x_n e^{-j2\pi f n T}$$

sono legate tra loro ed in particolare:

$$X_F(f) = X(z = e^{j2\pi f n T})$$

e cioè i valori assunti dalla trasformata z lungo il cerchio unitario sono appunto i valori della trasformata di Fourier della sequenza stessa. I valori della trasformata di Fourier discreta si otterranno poi campionando lungo il cerchio unitario, ad angoli crescenti multipli di $2\pi/N$, i valori della trasformata z .

$$X_k = \sum_{n=0}^{N-1} x_n e^{-j2\pi k \Delta f n T} = X(z = e^{j2\pi n k \Delta f T}) = X(z = e^{j\frac{2\pi n k}{N}})$$

4.4.6 Proprietà della trasformata zeta

Si verifica che la trasformata z, tra le altre, ha le seguenti proprietà:

Linearità

$$Z(a\{x_n\} + b\{y_n\}) = aX(z) + bY(z) \quad (4.10)$$

Spostamento temporale

$$Z(\{x_{n-m}\}) = z^{-m}X(z); \quad \text{infatti} \quad (4.11)$$

$$\sum_{n=-\infty}^{\infty} x_{n-m}z^{-n} = z^{-m} \sum_{k=-\infty}^{\infty} x_kz^{-k} = z^{-m}X(z) \quad (4.12)$$

Scalamento in z

$$X\left(\frac{z}{a}\right) \rightarrow \{a^n x_n\} \quad (4.13)$$

ed in particolare se $a = e^{j\omega_0 T}$, la moltiplicazione per una sinusoide complessa nei tempi, comporta la rotazione della trasformata z di un angolo corrispondente alla rotazione di fase della sinusoide in un campione.

Differenziazione in z

$$\begin{aligned} -z \frac{d}{dz} X(z) &\rightarrow \{nx_n\} && \text{infatti} \\ -z \frac{d}{dz} X(z) &= -z \frac{d}{dz} \sum_{n=-\infty}^{\infty} x_n z^{-n} = \sum_{n=-\infty}^{\infty} nx_n z^{-n} \end{aligned} \quad (4.14)$$

4.4.7 Durata di sequenze con varia caratteristica di fase

Una sequenza la cui trasformata z sia a fase zero sul cerchio unitario (abbiamo appena visto che la trasformata z sul cerchio unitario coincide con la trasformata di Fourier) è simmetrica rispetto all'origine dei tempi e quindi **non è causale**; tra tutte le sequenze con un'assegnata caratteristica di ampiezza è quella che ha la *durata minima* nel senso che vedremo tra breve. Inoltre, sempre a pari caratteristica di ampiezza, tra le sequenze causali la sequenza a fase minima è quella che ha *ritardo minimo* e cioè ha l'energia massimamente concentrata presso l'origine dei tempi, ma solo per $n \geq 0$. La dimostrazione è elementare e verrà fatta per

CAPITOLO 4 LA TRASFORMATTA Z

segnali continui, ma è valida anche per segnali campionati. Utilizziamo come misura della durata τ di una funzione del tempo $a(t)$ l'espressione:

$$\begin{aligned}\tau &= \sqrt{\frac{\int_{-\infty}^{+\infty} t^2 a^2(t) dt}{\int_{-\infty}^{+\infty} a^2(t) dt}} = \\ &= \sqrt{\frac{\int_{-\infty}^{+\infty} \left| \frac{dA(\omega)}{d\omega} \right|^2 d\omega}{\int_{-\infty}^{+\infty} |A(\omega)|^2 d\omega}}\end{aligned}\quad (4.15)$$

L'equazione precedente risulta dal teorema di Parseval applicato alle trasformate di $ta(t)$ e $a(t)$. Ponendo poi:

$$A(\omega) = |A(\omega)|e^{j\Phi(\omega)} \quad (4.16)$$

$$\frac{dA(\omega)}{d\omega} = e^{j\Phi(\omega)} \left(\frac{d|A(\omega)|}{d\omega} + j|A(\omega)| \frac{d\Phi}{d\omega} \right) \quad (4.17)$$

si vede che:

$$\tau = \sqrt{\frac{\int_{-\infty}^{+\infty} \left[\left| \frac{d|A(\omega)|}{d\omega} \right|^2 + |A(\omega)|^2 \left(\frac{d\Phi}{d\omega} \right)^2 \right] d\omega}{\int_{-\infty}^{+\infty} |A(\omega)|^2 d\omega}} \quad (4.18)$$

A pari caratteristica di ampiezza, il valore minimo di τ si raggiunge per fase zero, quando il secondo addendo (positivo o nullo) nell'integrale è nullo. Ancora, quando la fase non è nulla, la fase totale sarà la somma delle fasi contribute da ogni zero (si ricorda che la dimostrazione è valida anche per segnali campionati). Portando uno zero da fase minima a fase massima la caratteristica di ampiezza non cambia ma il termine $|d\Phi/d\omega|$ cresce, come è facile verificare nel caso semplice di uno zero reale. Quindi, la condizione di fase minima porta anche a ridurre al minimo il secondo addendo dell'integrale e infine τ .

4.5 Poli

E' interessante studiare sequenze aventi trasformata z inversa di quella corrispondenti agli zeri, e cioè del tipo

$$C(z) = \frac{1}{B(z)} = \frac{1}{1 - z_0 z^{-1}}; \quad B(z) = 1 - z_0 z^{-1}; \quad (4.19)$$

Convolvere una sequenza qualsiasi prima con la sequenza di trasformata $B(z)$ (lo zero) e poi con quella di trasformata $C(z)$ è come non fare nulla o fare una convoluzione con un impulso nell'origine, o moltiplicarne la trasformata per una trasformata z unitaria. Abbiamo quindi introdotto una sequenza $\{c_n\}$ [l'antitrasformata di $C(z)$] "inversa" alla sequenza $\{b_n\} = \{1, -z_0\}$ corrispondente allo zero $B(z)$; infatti convolvere per $\{c_n\}$ *compensa* le modifiche apportate alla funzione di trasferimento dalla convoluzione con la sequenza $\{b_n\}$ [corrispondente a $B(z)$]. Le sequenze "inverse" delle sequenze elementari di due campioni (zeri) sono indicate con il nome di **poli**.

Poiché però la funzione $C(z)$ non è un polinomio di grado finito, la sequenza $\{c_n\}$ non è univocamente determinata.

Infatti vedremo ora che l'operazione di antitrasformazione di $C(z)$ non è immediata; questa operazione è univoca solo se la funzione è un polinomio di ordine finito nella variabile z^{-1} o magari z se vi sono anticipi invece che ritardi; in questo caso, basta la diretta identificazione del coefficiente dei termini di grado i -esimo per trovare il termine i -esimo della sequenza.

Il problema è sostanziale e non solo formale: infatti esistono infinite sequenze, di durata illimitata, che convolute con una sequenza di due campioni danno come risultato l'impulso unitario nell'origine e cioè hanno una trasformata z inversa di quella di uno zero. La verifica è semplice: si consideri la sequenza di due campioni $b_0 = 1$; $b_1 = -z_0$ e si cerchi la sequenza $\{c_n\}$ tale che:

$$\{b_n\} * \{c_n\} = \delta(n); \quad (4.20)$$

$$\begin{aligned} \delta(n) = 1; \quad n = 0; \quad \delta(n) = 0; \quad n \neq 0 \\ c_n - z_0 c_{n-1} = \delta(n) \end{aligned} \quad (4.21)$$

Questa è un'equazione alle differenze finite che si può risolvere per sostituzione. Supponiamo ora che z_0 sia uno zero a minima fase e cioè $|z_0| < 1$. Poniamo $c_{-1} = 0$ e risolviamo per indici diversi da -1 ; risulta:

$$c_n = 0; \quad n < 0; \quad c_n = z_0^n; \quad n \geq 0 \quad (4.22)$$

Peraltro, se avessimo posto $c_1 = 0$ avremmo trovato:

$$c_n = 0; \quad n \geq 0; \quad c_n = -z_0^n; \quad n < 0 \quad (4.23)$$

CAPITOLO 4 LA TRASFORMATA Z

Quindi abbiamo trovato per la sequenza $\{c_n\}$ almeno due diverse soluzioni, ed anzi infinite perchè infinite combinazioni lineari di queste due soluzioni sono ancora soluzioni del problema. Nel caso quindi di uno zero a minima fase, la sequenza causale (quella nulla per $n < 0$) è convergente e quella anticausale è divergente. Si osserva invece che con $|z_0| > 1$ e cioè con uno zero a massima fase, è vero il contrario e cioè la sequenza anticausale è convergente e quella causale è divergente. Allo stesso risultato si arriva

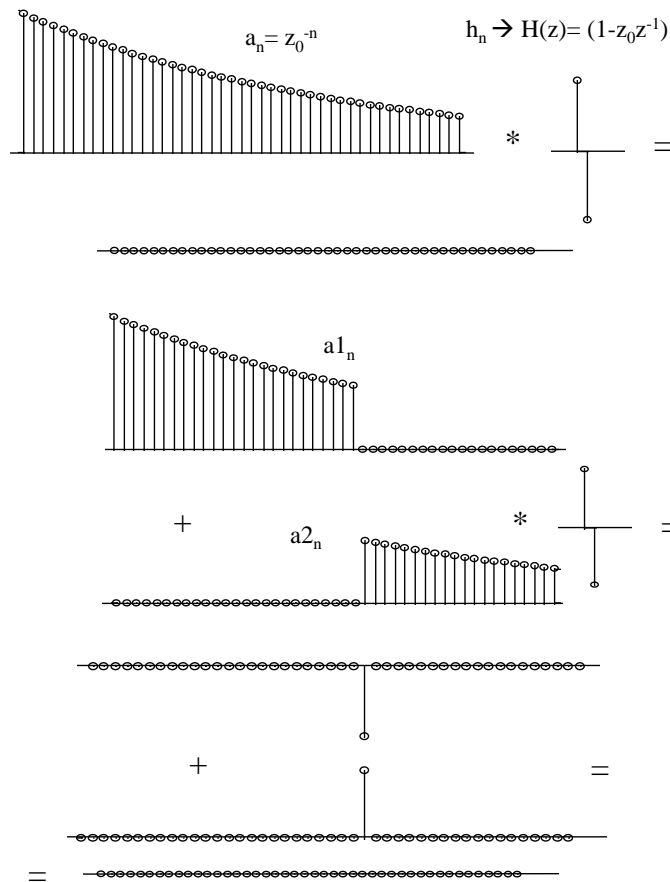


Figura 4.7: Rappresentazione delle sequenze "inverse" degli zeri dal dominio delle trasformate utilizzando la serie geometrica; è possibile

scrivere che:

$$\frac{1}{1 - z_0 z^{-1}} = 1 + (z_0 z^{-1}) + (z_0 z^{-1})^2 + (z_0 z^{-1})^3 + \dots \rightarrow 1, z_0, z_0^2, z_0^3, \dots \quad (4.24)$$

sequenza causale e convergente se lo zero è a minima fase; ma anche

$$\begin{aligned} \frac{1}{1 - z_0 z^{-1}} &= -\frac{1}{z_0} \frac{z}{1 - z_0^{-1} z} = \\ &-\frac{z}{z_0} \left(1 + \frac{z}{z_0} + \left(\frac{z}{z_0}\right)^2 + \left(\frac{z}{z_0}\right)^3 + \dots \right) \end{aligned} \quad (4.25)$$

sequenza anticausale e convergente se lo zero è a massima fase.

Da un altro punto di vista, si può osservare che lo zero a fase minima o a fase massima annulla una sinusoida generalizzata che decresce o cresce con l'indice. Si divida ora la sinusoida generalizzata in due sottosequenze in corrispondenza dell'origine, e cioè $\{z_0^n; n \geq 0\}$ e $\{z_0^n; n < 0\}$; il risultato della convoluzione di ognuno dei due tronconi con lo zero sarà sempre nullo tranne che nell'origine. Ancora, poiché sommando i due tronconi si ritrova la sinusoida generalizzata, si conclude che i due impulsi che si trovano convolvendo i due tronconi destro o sinistro sono di segno opposto (fig. 4.7).

La sequenza corrispondente ad un polo può essere sia l'uno che l'altro dei due tronconi (a meno del segno); più in generale una qualsiasi combinazione lineare dei due tronconi; in un modo o nell'altro si trova una sequenza che convoluta con uno zero dà un impulso; è importante osservare che sempre solo uno dei due tronconi è convergente; l'altro tende ad infinito.

Da quanto osservato, si conclude che vi sono infiniti modi per *compensare* l'effetto di uno zero. Quindi, ad un polo corrispondono infinite sequenze; il grado di libertà che ne consegue può essere utilizzato in vario modo: ad esempio imponendo che la sequenza inversa assuma un valore assegnato ad un certo istante. **Peraltro, tra tutte le sequenze inverse possibili, una sola sarà convergente**, ed in particolare sarà causale nel caso di poli a fase minima (inversi di zeri a fase minima) e anticausale nel caso di poli a fase massima.

La sequenza inversa è unica qualora il problema esige la convergenza (valore assegnato pari a zero a tempo infinito), o qualora esige la causalità (valore assegnato pari a zero per $n = -1$), o infine se si impone che la risposta all'impulso abbia ad un tempo qualsiasi un valore qualsiasi.

4.6 Equazioni alle differenze finite

Si può fare la convoluzione di una sequenza causale corrispondente ad un polo (di lunghezza infinita), con un segnale numerico qualsiasi al costo molto basso di una sola moltiplicazione ed una addizione per campione, come se si trattasse di uno zero. Infatti sia

$$Y(z) = \frac{X(z)}{1 - z_0 z^{-1}} \quad (4.26)$$

Antitrasformando, se ne ricava l'equazione alle differenze finite

$$y_n = x_n + z_0 y_{n-1} \quad (4.27)$$

da cui si calcola la sequenza in uscita e pertanto il risultato della convoluzione cercato.

La condizione di causalità è applicata implicitamente se si fa il calcolo per indici crescenti iniziando p.e. da y_0 assegnando ad y_{-1} un valore qualsiasi; il risultato sarà convergente se il polo è a minima fase.

Se invece il polo fosse a massima fase, per avere un risultato ancora convergente, si dovrebbe risolvere all'incontrario l'equazione di cui sopra andando verso indici decrescenti, assegnando cioè ad y_1 un valore qualsiasi e quindi risolvendo a ritroso:

$$y_{n-1} = \frac{1}{z_0}(y_n - x_n) \quad (4.28)$$

Ovviamente in questo secondo caso, si ha convergenza per $z_0 > 1$.

Sia ora in generale:

$$Y(z) = X(z) \frac{A(z)}{B(z)} \quad (4.29)$$

dove $X(z)$ è la trasformata della sequenza $\{x_n\}$ e $A(z)$, $B(z)$ sono polinomi qualsiasi.

Un sistema di questo tipo ha poli posizionati sia fuori sia dentro il cerchio unitario e anche zeri. In questo caso è necessario prima capire se è la soluzione convergente quella cercata: in questo caso, si deve:

trovare tutte le radici del denominatore (poli) e decomporre il denominatore $B(z)$ nel prodotto di due polinomi, l'uno $B_{max}(z)$ con le radici

4.6 EQUAZIONI ALLE DIFFERENZE FINITE

tutte fuori dal cerchio unitario (poli a fase massima), l'altro $B_{min}(z)$ con le radici tutte dentro il cerchio unitario (poli a fase minima);

$$Y(z) = X(z) \frac{A(z)}{B_{min}(z)B_{max}(z)} = X(z) \frac{A(z)}{B_{min}(z)} \times \frac{1}{B_{max}(z)} \quad (4.30)$$

convolvere con la parte causale e convergente della risposta all'impulso del sistema, e cioè quella corrispondente ai poli a fase minima e agli zeri; ciò si attua utilizzando l'ingresso $\{x_n\}$ come termine forzante dell'equazione alle differenze finite corrispondente a questa parte del sistema. In sostanza si considera il sistema:

$$Y_1(z) = X(z) \frac{A(z)}{B_{min}(z)} \quad (4.31)$$

Sia ora la sequenza $\{y_{1,n}\}$ l'uscita della parte causale del sistema.

Infine, si deve convolvere la sequenza $\{y_{1,n}\}$ con la risposta all'impulso corrispondente ai poli a fase massima, che è convergente e quindi anticausale.

$$Y(z) = \frac{Y_1(z)}{B_{max}(z)} \quad (4.32)$$

Per fare ciò in modo semplice ed a basso costo computazionale si deve invertire l'asse temporale. Formalmente, si può scrivere:

$$Y(1/z) = \frac{Y_1(1/z)}{B_{max}(1/z)} \quad Y_1(1/z) \leftrightarrow \{y_{1,-n}\} \quad (4.33)$$

Infatti la sequenza $Y_1(1/z)$ è la trasformata della sequenza $\{y_{1,n}\}$, dopo che sia stata ribaltata nei tempi.

A sua volta, il polinomio $B_{max}(1/z)$ ha le radici reciproche di quelle del polinomio $B_{max}(z)$ e che quindi sono all'interno del cerchio unitario; l'equazione alle differenze finite descritta dalla precedente equazione è quindi corrispondente ad un sistema causale e convergente che può essere risolto come nel caso precedente. Infine, la sequenza $Y(1/z)$ deve venire ancora ribaltata nel tempo per avere il risultato finale e cioè la sequenza $Y(z)$.

Si osservi che per ribaltare l'asse temporale di una sequenza, ne devono essere disponibili tutti i campioni, presenti e futuri. La cosa non può essere fatta in tempo reale con sistemi fisici che si evolvono nel tempo (sistemi causali).

CAPITOLO 4 LA TRASFORMATA Z

Tuttavia, anche nel caso in cui non si hanno a disposizione tutti i dati, ma si deve operare con un ritardo finito, l'equazione con i poli fuori del cerchio unitario può essere risolta approssimativamente con la tecnica discussa nel paragrafo sulle sequenze geometriche troncate.

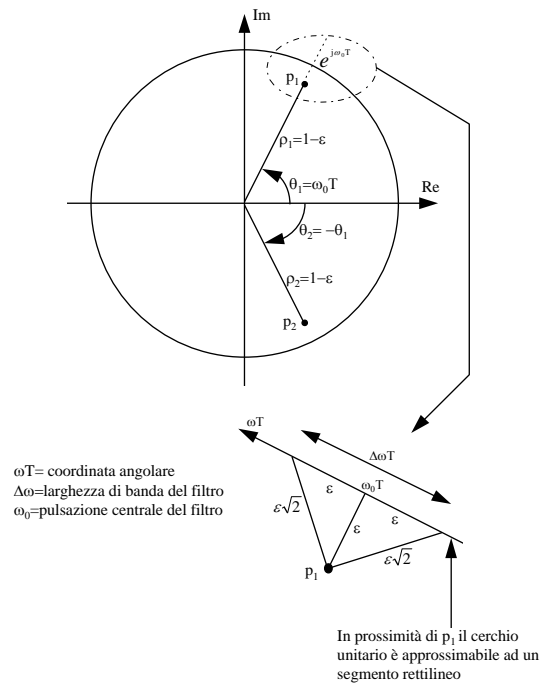


Figura 4.8: Tecnica vettoriale per il progetto di un filtro passa-banda realizzato con due poli complessi coniugati p_1 e p_2 molto vicini al cerchio di raggio unitario

4.7 Filtri passa banda

I poli si prestano per realizzare a basso costo computazionale funzioni di trasferimento a banda molto stretta; infatti le sequenze rappresentate da poli sono lunghe, qualora il polo sia prossimo al cerchio unitario. Ad esempio, supponiamo che il polo sia in $z = (1 - \epsilon)e^{j\phi_0}$ con $\epsilon \ll 1$ (figg.

4.7 FILTRI PASSA BANDA

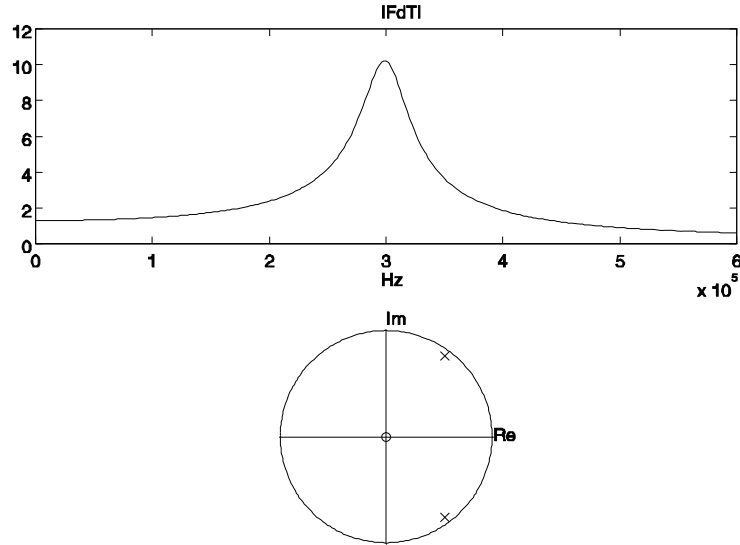


Figura 4.9: Caratteristiche di un filtro passa-banda avente una banda passante di $40k\text{Hz}$ attorno alla frequenza di $300k\text{Hz}$; $f_c = 2\text{MHz}$

4.84.9).

$$A(z) = \frac{1}{1 - z^{-1}(1 - \varepsilon)e^{j\phi_0}}; \quad \phi_0 = \omega_0 T$$

Calcoliamo ora la funzione di trasferimento per valori di ϕ prossimi a ϕ_0 e quindi per:

$$z = e^{j\phi}; \quad \phi = \phi_0 + \gamma; \quad |\gamma| \ll 1$$

$$A(z)|_{z=e^{j\phi}} = \frac{1}{1 - (1 - \varepsilon)e^{-j\gamma}}; \quad \phi = \phi_0 + \gamma; \quad |\gamma| \ll 1$$

e sviluppando in serie l'esponenziale

$$A(z)|_{z=e^{j\phi}} = \frac{1}{1 - (1 - \varepsilon)(1 - j\gamma)} \sim \frac{1}{\varepsilon + j\gamma}$$

Usando la tecnica vettoriale per il calcolo della funzione di trasferimento, si osserva che essa ha un valore massimo di $1/\varepsilon$ per $\omega = \omega_0$ e cala di 3 dB (un fattore pari a $\sqrt{2}$) fuori dell'intervallo $|\gamma| = |(\omega - \omega_0)T| \leq 2\varepsilon$.

Basta scegliere ε piccolo, e risultano filtri a banda molto stretta. Come vedremo poi, discutendo della quantizzazione dei coefficienti, questi

CAPITOLO 4 LA TRASFORMATA Z

filtri sono molto sensibili alle variazioni dei coefficienti causate dalla loro rappresentazione con un limitato numero di cifre binarie.

Di notevole interesse è la realizzazione di filtri arresta banda fatta con una coppia di un polo posto presso il cerchio unitario ed uno zero sul cerchio unitario, isofrequenziale. La funzione di trasferimento si annulla alla frequenza ω_0 ma ritorna 3 dB al di sotto del valore unitario appena fuori dell'intervallo $\Delta\omega T \simeq 2\varepsilon$ (fig. 4.10).

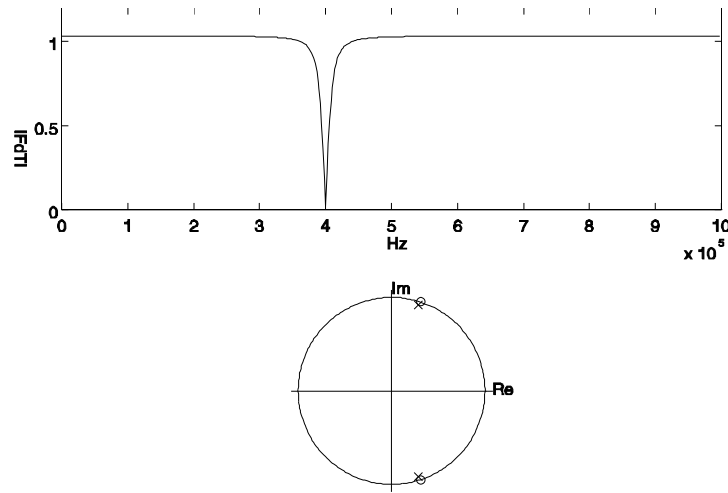


Figura 4.10: Caratteristiche di un filtro arresta banda avente una banda soppressa di $20kHz$ attorno alla frequenza di $400kHz$; $f_c = 2MHz$

4.7.1 Sfasatori puri

Altro caso interessante è quello dello sfasatore puro; qui abbiamo un polo ed uno zero messi in posizione reciproca coniugata rispetto al cerchio unitario; pertanto il modulo della funzione di trasferimento è costante, ma non così la fase che avrà una rapida variazione all'interno del solito intervallo di frequenza $\Delta\omega T \simeq 2\varepsilon$ (fig. 4.10). Si osservi che trasformate z del tipo

$$\frac{\tilde{A}(z)}{A(z)} = Q(z); \quad |Q(z)|_{z=e^{j\omega T}} = 1 \quad (4.34)$$

4.7 FILTRI PASSA BANDA

hanno modulo sempre unitario (sul cerchio unitario, ovviamente) e rappresentano sfasatori puri, causali e convergenti se il polinomio $A(z)$ è a minima fase.

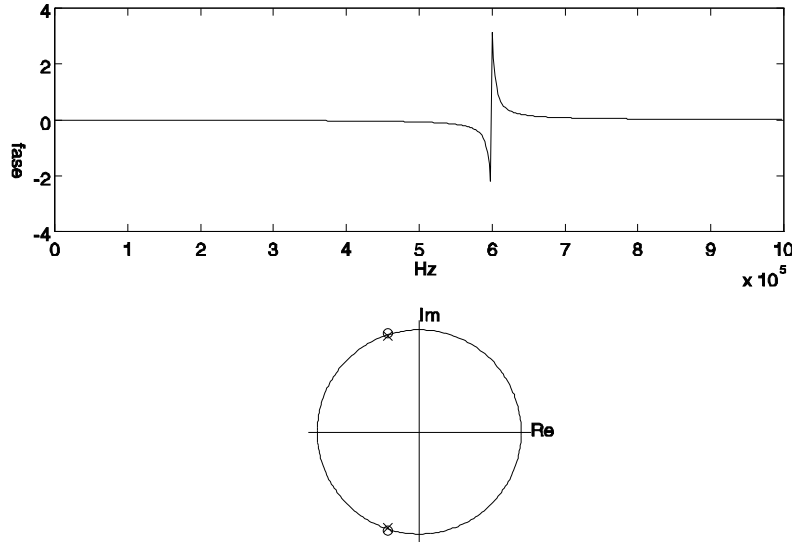


Figura 4.11: Caratteristiche di uno sfasatore puro con sfasamento intorno ai $600kHz$ in un intervallo di $50kHz$; $f_c = 2MHz$

La caratteristica di fase di uno sfasatore puro causale e convergente è sempre monotona decrescente; per dimostrarlo basta considerare il caso di un sistema con un solo zero ed un solo polo reali; casi di maggiore complessità si riporteranno a cascata di sistemi come questo, centrati a differenti frequenze e quindi ruotati intorno al cerchio unitario in vario modo:

$$H(z) = \frac{-a + z^{-1}}{1 - a z^{-1}} = \frac{\tilde{A}(z)}{A(z)}$$

La caratteristica di fase è:

$$\angle H(z) = \phi(\omega) = -\angle A(z) + \angle \tilde{A}(z) = -\omega T - 2\angle A(z)$$

$$\phi(\omega) = -\omega T - 2 \tan^{-1} \frac{a \sin \omega T}{1 - a \cos \omega T}$$

Il ritardo di gruppo dello sfasatore puro $\tau(\omega)$, che corrisponderebbe al ritardo di un segnale passabanda centrato a pulsazione ω , viene ottenuto

Figura 4.12: Andamento del ritardo di gruppo di uno sfasatore puro ($a=0.8$; $T=1$).

differenziando la fase dello sfasatore puro $\phi(\omega)$ rispetto a ω e cambiando segno: esso è necessariamente positivo per i sistemi causali. Infatti:

$$\tau(\omega) = -\frac{d\phi}{d\omega} = T \frac{1 - a^2}{1 + a^2 - 2a \cos \omega T} > 0 \text{ se } a = 1 - \varepsilon < 1 \quad (4.35)$$

Pertanto la fase di uno sfasatore puro causale e stabile è sempre decrescente. Posto poi:

$$\tau_{\max} = \frac{(1 + a)T}{(1 - a)} \sim \frac{2T}{\varepsilon};$$

si vede che il grafico del ritardo di gruppo in funzione della frequenza (fig. 4.12) è una campana di ampiezza τ_{\max} e banda $\Delta\omega \sim 2\varepsilon/T$ centrata a frequenza 0. Inoltre, τ_{\max} può crescere quanto si desidera. In teoria, per un segnale a banda stretta e cioè di banda $\Delta\omega = 2\varepsilon/T$, uno sfasatore puro potrebbe essere utilizzato al posto di una memoria di durata $2T/\varepsilon$. Invece, con $a > 1$, la fase è monotonamente crescente e lo sfasatore puro, per essere stabile, deve essere anticausale.

Nel caso di polo e zero complessi, sostituendo $ae^{j\psi}$ ad a :

$$H_1(z) = \frac{\tilde{A}(z)}{A(z)} = \frac{-ae^{-j\psi} + z^{-1}}{1 - ae^{j\psi} z^{-1}} = e^{-j\psi} \frac{-a + z^{-1}e^{j\psi}}{1 - ae^{j\psi} z^{-1}} = e^{-j\psi} \frac{\tilde{A}(ze^{-j\psi})}{A(ze^{-j\psi})};$$

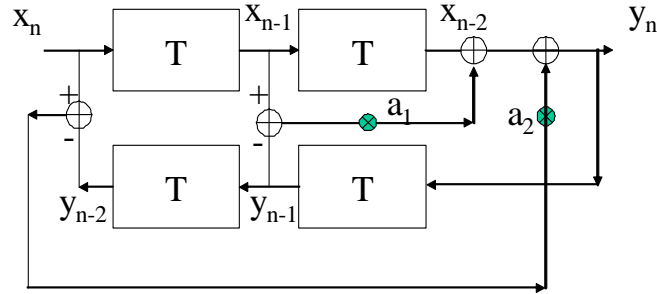


Figura 4.13: Schema a blocchi di uno sfasatore puro con due poli e due zeri realizzato con due moltiplicazioni per campione di ingresso.

e quindi il ritardo di gruppo è ancora quello della (4.35), purchè si sposti la pulsazione centrale della campana di $\Delta\omega = \psi/T$. In figura 4.13, è presentato lo schema a blocchi di uno sfasatore puro

$$H(z) = \frac{\tilde{A}(z)}{A(z)}; \quad A(z) = 1 + a_1 z^{-1} + a_2 z^{-2}$$

caratterizzato da due poli reali e due zeri nella posizione reciproca; esso viene realizzato con due sole moltiplicazioni reali per ogni campione che si presenti all'ingresso. Si osservi che mantiene la sua caratteristica di sfasatore puro e quindi avere caratteristica di ampiezza unitaria, qualsiasi siano i valori di a_1, a_2 . Naturalmente, per avere stabilità, $A(z)$ deve essere a minima fase.

4.7.2 Oscillatori, divani, splines

Anzitutto consideriamo un esempio di equazione alle differenze finite che si riferisce ad un sistema che si evolve nel tempo e pertanto soggetto alla relazione di causalità. Questo è il caso del sistema con due poli in equazione (4.36); il sistema viene qui indicato come oscillatore numerico

CAPITOLO 4 LA TRASFORMATA Z

poichè, per opportuni valori dei parametri, la sua risposta all'impulso può essere una sinusoidale campionata; al sistema corrisponde la seguente equazione alle differenze finite:

$$Y(z) = \frac{X(z)}{1 - 2z^{-1}\rho \cos \phi + \rho^2 z^{-2}}; \quad \phi = \omega_0 T \quad (4.36)$$

corrispondente ad un sistema causale perché riferita ad un oggetto fisico che opera nel tempo e non risponde prima della sua eccitazione. Qualora $X(z) = 1$ la risposta del sistema si riduce alla combinazione delle risposte ai due poli complessi coniugati (posti vicino o sul cerchio unitario se $\rho \leq 1$) e quindi a due sequenze causali complesse coniugate la cui somma è:

$$y_n = \rho^n \frac{\sin[(n+1)\phi]}{\sin \phi}; \quad n \geq 0; \quad (4.37)$$

Si osservi come il primo campione è sempre 1, e al decrescere di ϕ l'oscillazione ha ampiezze crescenti. Se poi $\rho \geq 1$ l'oscillazione diverge ed il sistema andrà incontro a fenomeni di saturazione. Se vogliamo che l'uscita dell'oscillatore sia una cosinusoidale esponenzialmente decrescente, a fase γ preassegnata, si dovrà aggiungere uno zero per ottenere lo sfasamento desiderato alla pulsazione della sinusoidale; p.e. ponendo:

$$Y_1(z) = \frac{1}{2} \left(\frac{e^{j\gamma}}{1 - \rho e^{j\phi} z^{-1}} + \frac{e^{-j\gamma}}{1 - \rho e^{-j\phi} z^{-1}} \right) = \frac{\cos \gamma - z^{-1} \rho \cos(\gamma - \phi)}{1 - 2z^{-1} \rho \cos \phi + \rho^2 z^{-2}}$$

La risposta nei tempi è ottenuta sviluppando le due frazioni in serie geometriche e poi sommando i valori coniugati:

$$y_{1n} = \rho^n \cos(n\phi + \gamma)$$

e si ritrova il caso precedente (a meno di un fattore) ponendo $\gamma = \phi - \pi/2$.

Vediamo ora un'applicazione ad un sistema non causale: un divano a molle. Si supponga di avere un sistema meccanico formato da nodi collegati da molle; il nodo i -esimo è collegato al suolo da una molla verticale di rigidità (rapporto tra forza e deformazione) k_1 ed è collegato ai nodi adiacenti $i - 1$, $i + 1$ da molle di rigidità k_2 ; ogni nodo è poi soggetto alla forza verticale f_i . L'equazione di equilibrio del nodo i -esimo è, indicando con x_i lo spostamento elastico verticale del nodo i :

$$k_2(x_i - x_{i-1}) + k_2(x_i - x_{i+1}) + k_1 x_i = f_i \quad (4.38)$$

4.7 FILTRI PASSA BANDA

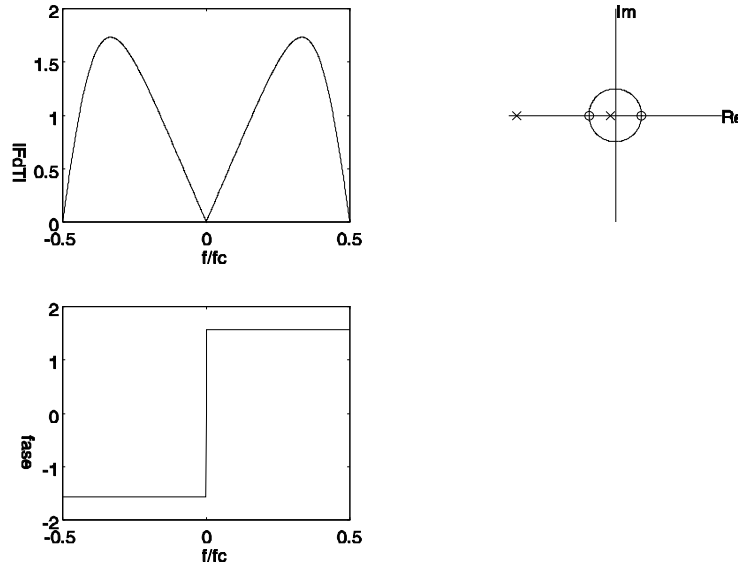


Figura 4.14: Caratteristiche relative alla sequenza delle derivate della risposta impulsiva dello spline in corrispondenza dei campioni

Passando alle trasformate z delle sequenze degli spostamenti $X(z)$ e delle forze $F(z)$ si ha:

$$X(z) [(k_1 + 2k_2) - (z + z^{-1})k_2] = F(z) \quad (4.39)$$

e quindi la risposta all'impulso ($F(z) = f_0$ e cioè una sola forza posizionata sul campione di posto 0) è:

$$X(z) = \frac{f_0}{(k_1 + 2k_2) - (z + z^{-1})k_2} \quad (4.40)$$

La trasformata z della risposta è di nuovo una funzione di trasferimento simile a quella del caso precedente: inoltre, i poli del sistema sono reali, positivi e posti in posizioni reciproche rispetto al cerchio unitario, in quanto il prodotto delle radici del denominatore è 1 ed il discriminante è positivo.

La risposta da attendersi non è però quella causale, come nel caso precedente, e cioè non ci attendiamo spostamento nullo per indici inferiori a 0, e vogliamo una risposta convergente. *Sappiamo che il divano non si deforma p.e. solo a sinistra di chi si siede e non esplose: peraltro,*

CAPITOLO 4 LA TRASFORMATA Z

la deformazione corrispondente alla trazione violenta verso il soffitto di una molla nella parte p.e. sinistra di un divano, potrebbe essere bilanciata da un carico opportuno, annullando così ogni deformazione nel divano, a destra del carico. Si pensi infatti ad un lenzuolo elastico piuttosto che un divano a molle. Si possono avere due situazioni: in un caso p.e. un ragazzo che è disteso sul lenzuolo e lo deforma simmetricamente. Per capire invece l'altra situazione, si pensi ad una mamma che voglia svegliare il figlio addormentato e che tiri verso l'alto il lenzuolo su cui il figlio dorme. P.e. a sinistra del ragazzo, il lenzuolo è tirato verso l'alto con una curva diciamo esponenziale; a destra invece il lenzuolo è orizzontale. È quindi possibile sia una soluzione simmetrica, che una soluzione causale" in cui il fenomeno si sviluppa solamente da un lato. In assenza di causali esterne, sceglieremo la soluzione causale per il polo interno al cerchio e quella anticausale per il polo esterno; sostanzialmente cercando una soluzione che sia simmetrica negli spazi.

Posto:

$$q = \frac{2k_2 + k_1 - \sqrt{k_1^2 + 4k_1k_2}}{2k_2}; \quad 0 \leq q \leq 1 \quad (4.41)$$

$$X(z) = -f_0 \frac{1 + q^2}{(k_1 + 2k_2)(1 - q^2)} \left(\frac{1}{1 - qz} + \frac{1}{1 - qz^{-1}} - 1 \right) \quad (4.42)$$

avremo:

$$x_n = -f_0 \frac{1 + q^2}{(k_1 + 2k_2)(1 - q^2)} q^{|n|} \quad (4.43)$$

In sostanza, se le molle orizzontali sono molto flessibili, il divano si infossa solo localmente ($q \simeq 0$); se sono rigide, il divano si abbassa quasi senza deformarsi ($q \leq 1$).

Come ultimo caso consideriamo l'equazione alle differenze finite che, per una bacchetta elastica, lega le pendenze ai nodi y'_i e gli spostamenti dei nodi y_i ; con la solita convenzione per le trasformate z rispettive, si ha, ricordando quanto osservato nel capitolo II:

$$(z) = 3Y(z) \frac{z - z^{-1}}{z + 4 + z^{-1}} \quad (4.44)$$

I poli del sistema sono reali e negativi in $-2 + \sqrt{3}$, $-2 - \sqrt{3}$ dentro e fuori il cerchio unitario in posizioni reciproche (infatti il prodotto delle radici del denominatore é unitario) (fig. 4.13).

4.8 SERIE GEOMETRICHE TRONCATE

Dobbiamo ancora separare il polo dentro da quello fuori del cerchio e scrivere:

$$Y'(z) = Y(z) \left(\frac{1}{1 + .268z} - \frac{1}{1 + .268z^{-1}} \right) \quad (4.45)$$

nel caso di $Y(z) = 1$ (ingresso impulsivo, solo il campione di posto 0 è non nullo) allora $Y'(z)$ è differenza di due sequenze, l'una causale e l'altra anticausale, simmetriche.

$$y'_n = -(-.268)^{|n|} \text{sign}(n); \quad y'_0 = 0 \quad (4.46)$$

Infatti la soluzione da scegliere in questo caso è ancora quella convergente; una bacchetta elastica la cui deflessione è nulla ad intervalli regolari, se abbassata o sollevata in corrispondenza del campione centrale, convergerà a zero con spostamenti a segni alterni; le derivate nei nodi sono anch'esse decrescenti esponenzialmente a segni alterni.

4.8 Serie geometriche troncate

E' utile conoscere a quale distribuzione di zeri corrisponda la sequenza relativa ad un polo, troncata, e quindi divenuta di lunghezza finita. Allo scopo serve il prodotto notevole:

$$H_1(x) = \frac{1 - x^N}{1 - x} = \sum_{h=0}^{N-1} x^h \quad (4.47)$$

Ponendo $x = az^{-1}$,

$$H(z) = \sum_{h=0}^{N-1} a^h z^{-h} = \frac{1 - a^N z^{-N}}{1 - az^{-1}} = H_1(az^{-1}); \quad (4.48)$$

$$H(z_k) = 0 \rightarrow z_k = ae^{2\pi jk/N}; \quad 0 < k \leq N - 1$$

Lo zero in $z = a$ ($k = 0$) si cancella con lo zero del denominatore e rimangono pertanto $N - 1$ zeri (i campioni della sequenza la cui trasformata z è $H(z)$ sono N) in posizione regolare attorno al cerchio di raggio a , tranne che in corrispondenza del polo della sequenza originaria. In conclusione, una sequenza corrispondente ad un polo, troncata dopo N campioni, corrisponde a $N - 1$ zeri disposti regolarmente tranne quello

CAPITOLO 4 LA TRASFORMATA Z

Figura 4.15: Caratteristiche della FdT $H_1(z)$ caratterizzata da $N = 19$ zeri

cancellato dal polo. Per $N \rightarrow \infty$ ne emerge l'equivalenza tra un polo ed infiniti zeri messi ovunque lungo il cerchio che passa per il polo, tranne che sul polo stesso.

Tuttavia, volendo approssimare un polo con un filtro FIR di ordine N , non sono necessari N moltiplicatori, ma ne bastano circa $\log_2 N$; infatti vale il seguente sviluppo in prodotto infinito:

$$\frac{1}{1 - \rho z^{-1}} = (1 + \rho z^{-1})(1 + \rho^2 z^{-2})(1 + \rho^4 z^{-4}) \cdots = \prod_{i=0}^{\infty} (1 + \rho^{2^i} z^{-2^i}); \quad (4.49)$$

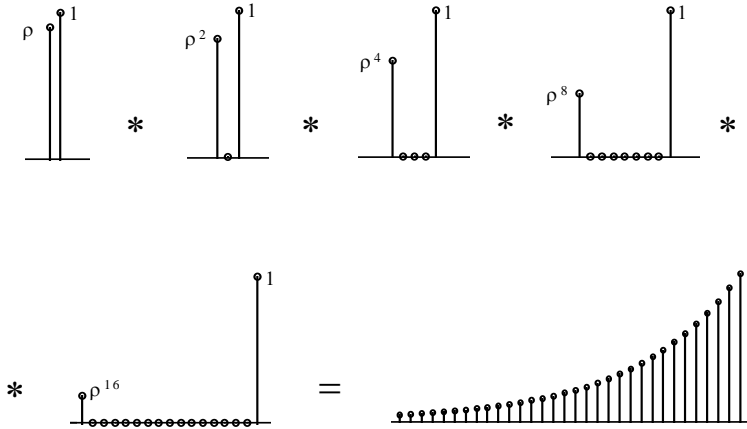
Si vede allora che con la prima moltiplicazione si realizza uno zero, con la seconda due, con la terza quattro e così via.

4.9 APPENDICE: IL CALCOLO DEI RESIDUI

Un'interessante applicazione di questo sviluppo può essere la realizzazione con un filtro FIR a massima fase di un polo anticausale troncato ad N campioni. Nella formula precedente, possiamo sostituire z al posto di z^{-1} e poi aggiungere un ritardo z^{-N} per rendere il tutto causale. Ne risulta la formula che segue:

$$\frac{z^{-N}}{1 - \rho z} \simeq (z^{-1} + \rho)(z^{-2} + \rho^2)(z^{-4} + \rho^4) \dots \quad (4.50)$$

e la conseguente generazione di N zeri a fase massima con $\log_2 N$ moltiplicazioni (fig. ??).



Realizzazione di un polo anticausale stabile ritardato e troncato mediante la convoluzione di zeri a fase massima

4.9 Appendice: Il calcolo dei residui

Esiste un modo semplice per decomporre un rapporto di polinomi in una somma di fratti semplici, utilizzando il calcolo dei residui. Rinviando per la teoria a un qualsiasi libro di analisi complessa o a Wikipedia, qui daremo solo la ricetta per il calcolo dei residui nel caso di poli semplici. Questo calcolo risulta utile per l'antitrasformata z p.e. di funzioni di autocorrelazione ecc.

CAPITOLO 4 LA TRASFORMATA Z

É utilizzabile la formula che segue, dove z_n ; $n = 1..N$, sono le N radici del denominatore $A(z)$; indichiamo poi con $A'(z)$ la derivata rispetto a z del polinomio $A(z)$:

$$\frac{B(z)}{A(z)} = R_0 + \sum_{n=1}^N \frac{R_n}{z - z_n}; R_0 = \lim_{z \rightarrow \infty} \frac{B(z)}{A(z)}; R_n = \frac{B(z_n)}{A'(z_n)} \quad (4.51)$$

Per dare un esempio semplice, nel caso del calcolo dell'autocorrelazione di una risposta all'impulso caratterizzata da un polo solo:

$$H(z) = \frac{1}{1 - az^{-1}}; R(z) = H(z) H^*(1/z)$$

$$R(z) = \frac{B(z)}{A(z)} = \frac{z}{(z - a)(1 - az)} = \frac{z}{A'(a)(z - a)} + \frac{z}{A'(1/a)(z - 1/a)}$$

ma poichè:

$$A'(z) = 1 + a^2 - 2az$$

$$R(z) = \frac{1}{1 - a^2} \left(\frac{az^{-1}}{1 - az^{-1}} + \frac{1}{1 - az} \right)$$

e infatti nel dominio dei tempi, i campioni dell'autocorrelazione sono:

$$r_n = \frac{a^{|n|}}{1 - a^2} \quad (4.52)$$

Un risultato equivalente sarebbe stato trovato per la risposta all'impulso dello spline, di cui alla formula 4.40.

Il caso di poli doppi o multipli è decisamente più complesso: anzitutto si osservi che convolvendo con se stessa una sequenza a^n ; $n \geq 0$ si ottiene la sequenza $(n + 1)a^n$. Infatti:

$$a^n * a^n = (n + 1)a^n; n \geq 0 \quad (4.53)$$

Un polo triplo invece porta alla convoluzione tripla che è:

$$a^n * a^n * a^n = \frac{(n + 1)(n + 2)}{2} a^n; n \geq 0 \quad (4.54)$$

4.9 APPENDICE: IL CALCOLO DEI RESIDUI

Passando alle trasformate z , consideriamo il caso di due poli appena differenti posti in $a_1 \sim a_2$:

$$H(z) = \frac{1}{(1 - a_1 z^{-1})(1 - a_2 z^{-1})} = \frac{z^2}{(z - a_1)(z - a_2)} \quad (4.55)$$

$$= \frac{z}{a_1 - a_2} \left(\frac{z}{z - a_1} - \frac{z}{z - a_2} \right) \quad (4.56)$$

Ma:

$$\frac{z}{z - a_1} - \frac{z}{z - a_2} \sim (a_1 - a_2) \frac{d}{da} \left(\frac{z}{z - a} \right) \quad (4.57)$$

Quindi:

$$H(z) \sim z \frac{d}{da} \left(\frac{z}{z - a} \right) \quad (4.58)$$

Ma se:

$$\frac{z}{z - a} = \sum_{n=0}^{\infty} a^n z^{-n} \quad (4.59)$$

scambiando la derivata e la somma ritroviamo la convoluzione già vista:

$$H(z) \sim z \frac{d}{da} \left(\frac{z}{z - a} \right) = z \frac{d}{da} \sum_{n=0}^{\infty} a^n z^{-n} = \quad (4.60)$$

$$= \sum_{n=1}^{\infty} n a^{n-1} z^{-n+1} = \sum_{n=0}^{\infty} (n+1) a^n z^{-n} \quad (4.61)$$

Quindi, nei casi di poli doppi ed in generale n-upli compaiono le derivate n-esime del denominatore della trasformata z rispetto al polo. Ne conseguono formule complicate che non è utile considerare qui.

Comunque, gli stessi risultati si poteva trovare in modo veloce osservando che:

$$\left(\frac{1}{z - a} \right)^2 = -\frac{d}{dz} \left(\frac{1}{z - a} \right) \quad (4.62)$$

Poi, sfruttando la proprietà della trasformata z della differenziazione in z :

$$-z \frac{d}{dz} X(z) \rightarrow \{n x_n\}; n \geq 0 \quad (4.63)$$

risulta:

$$\left(\frac{1}{z - a} \right)^2 \rightarrow \{(n+1) a^n\} \quad (4.64)$$

Capitolo 5

Filtri numerici

5.1 Introduzione

In questo capitolo discuteremo delle tecniche di progetto di filtri numerici: essi possono essere di due categorie e cioè FIR (Finite Impulse Response) o IIR (Infinite Impulse response) a seconda che la loro trasformata z consista di soli zeri o di poli e zeri. Una buona progettazione di un filtro numerico deve tendere ad ottenere il massimo risultato e cioè la massima aderenza alle specifiche progettuali; queste sono assegnate nel dominio delle frequenze mediante una maschera all'interno della quale deve essere contenuta la funzione di trasferimento del filtro sintetizzato (fig. 5.1). Inoltre, si deve minimizzare il costo e cioè minimizzare il numero di operazioni per campione; come vedremo ciò non corrisponde necessariamente al minor numero di moltiplicazioni; strutture di filtri più complesse possono essere convenienti se usano moltiplicazioni per numeri con pochi bit diversi da zero (potenze di due, ad esempio) o in genere se sono meno sensibili alla quantizzazione dei coefficienti e quindi operano correttamente con moltiplicazioni meno costose.

Tuttavia, per comprendere le basi delle tecniche di sintesi dei filtri FIR è opportuno discutere anzitutto delle tecniche di finestramento (windowing) e cioè gli effetti del troncamento nel tempo delle risposte all'impulso desiderate con finestre (rastremazioni a zero) di vario tipo. **Per semplicità l'analisi degli effetti del finestramento verrà fatta per segnali tempo continui, invece che campionati.** Questo semplifica molto le formule e le rende più leggibili.

CAPITOLO 5 FILTRI NUMERICI

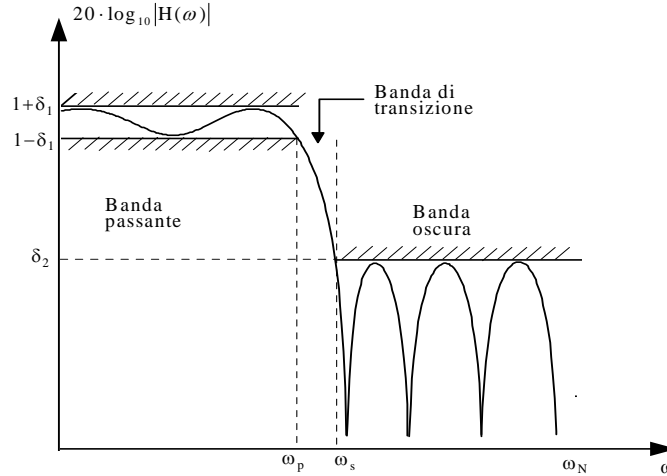


Figura 5.1: Maschera di un filtro passa-basso

5.2 Finestre rettangolari e triangolari

Indichiamo con $D(\omega)$ la risposta all'impulso desiderata per il filtro che si deve sintetizzare. La antitrasformata di Fourier verrà indicata con $d(t)$ ed in generale sarà illimitata nei tempi. Vogliamo vedere cosa accade troncando il segnale $d(t)$ con una finestra rettangolare di lunghezza NT . Nel caso discreto, considereremo la sequenza troncata dopo N campioni (fig. 5.2).

$$d'(t) = d(t) \times w_R(t) \quad (5.1)$$

$$w_R(t) = 1; |t| \leq \frac{NT}{2}; \quad w_R(t) = 0; |t| \geq \frac{NT}{2}$$

Moltiplicare nei tempi equivale a convolvere nelle frequenze e cioè a considerare la

$$D'(\omega) = D(\omega) * W_R(\omega); \quad W_R(\omega) = NT \operatorname{sinc}(fNT) \quad (5.2)$$

La $W_R(\omega)$ tende ad un impulso ideale se N diventa molto grande e questo corrisponde all'ovvia considerazione che finestrare con finestre lunghe

5.2 FINESTRE RETTANGOLARI E TRIANGOLARI

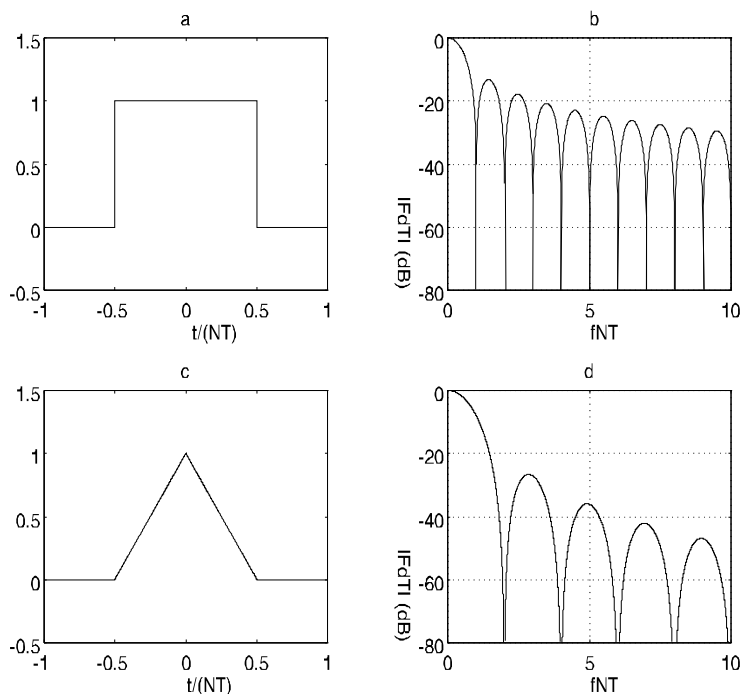


Figura 5.2: a) Finestra rettangolare b) Trasformata di Fourier della finestra rettangolare in dB c) Finestra triangolare d) Trasformata di Fourier della finestra triangolare in dB

porta a piccole distorsioni. Quello che interessa però sono appunto le distorsioni della risposta all'impulso dovute al finestramento: esse sono di due tipi diversi: quelle dovute alla convoluzione per il **lobo principale** del sinc che ha i primi zeri in $f = \pm 1/NT$ e quelle dovute alla convoluzione per i **lobi laterali**, che come è noto decrescono come $1/\Delta f NT$ dove con Δf abbiamo indicato l'intervallo in frequenza tra la posizione del lobo principale e quella del lobo secondario considerato. La convoluzione per il lobo principale smussa l'andamento in frequenza del filtro, e lo rende meno selettivo. I fronti di salita e di discesa, nel dominio della frequenza, non potranno essere più brevi di $\simeq 2/NT$; questa è una diretta conseguenza della limitazione della lunghezza del filtro.

Il brusco troncamento della risposta $d(t)$ porta alla creazione di componenti ad alta frequenza che sarebbero altrimenti state assenti dalla risposta desiderata; questo fatto viene osservato, nel dominio delle fre-

CAPITOLO 5 FILTRI NUMERICI

quenze, con la formazione di componenti ad alta frequenza dovute alla convoluzione con i lobi laterali della finestra. Siamo quindi di fronte a due fenomeni diversi che però interagiscono: consideriamo infatti ora una finestra triangolare della durata ancora di N campioni;

$$w_T(t) = 1 - \frac{2|t|}{NT}; \quad |t| \leq \frac{NT}{2} \quad W_T(\omega) = \frac{NT}{2} \operatorname{sinc}^2 \frac{fNT}{2}; \quad (5.3)$$

Come si vede, per la finestra triangolare, i lobi laterali decrescono come $1/(\Delta f NT)^2$ ma il lobo principale ha larghezza doppia. In generale se $W(\omega)$ e $\{w_n\}$ sono legate da una trasformata di Fourier, allora la larghezza di banda equivalente B_{eq} la durata equivalente T_{eq} sono così legate:

$$\frac{1}{2\pi} \int W(\omega) e^{j\omega t} d\omega = W(0) \times B_{eq} = w(0) = 1; \quad (5.4)$$

$$\int w(t) dt = 1 \times T_{eq} = W(0); \rightarrow B_{eq} \times T_{eq} = 1$$

Pertanto, se per ridurre l'effetto dei lobi laterali si addolciscono le transizioni temporali (si apre e chiude la finestra in modo graduale e quindi si decresce T_{eq} a pari durata NT) nello stesso tempo si allarga l'area sottesa dalla $W(\omega)$ perchè ne allarga il lobo centrale di banda B_{eq} .

5.3 Finestre cosinusoidali

Una forma di finestre particolarmente efficiente è quella delle finestre cosinusoidali (coseno rialzato, Hann, Hamming) che danno un miglior compromesso lobi laterali - lobo principale che non la finestra triangolare. La finestra di Hann, o cosinusoidale è così costituita:

$$w_C(t) = \cos\left(\frac{\pi t}{NT}\right); \quad |t| < \frac{NT}{2} \quad (5.5)$$

La sua trasformata è deducibile in modo sintetico con il seguente ragionamento: poichè la $w_C(t)$ è il prodotto, nei tempi, di una cosinusoide di periodo $2NT$ e di una finestra rettangolare di durata NT , la sua trasformata sarà la convoluzione con il sinc della trasformata della cosinusoide e cioè due impulsi di eguale segno in $f = \pm 1/2NT$. Ma allora, i lobi laterali dei due sinc si sommano dopo una traslazione di frequenza

5.3 FINESTRE COSINUSOIDALI

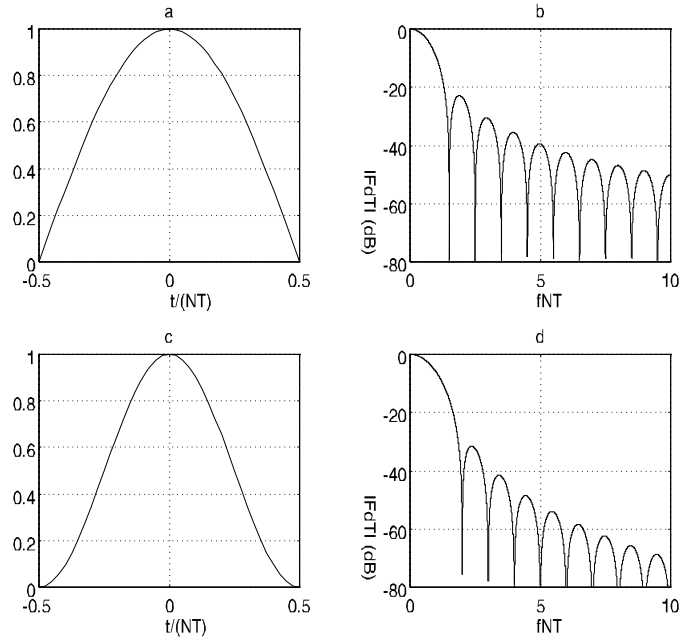


Figura 5.3: a) Finestra cosinusoidale b) Trasformata di Fourier della finestra cosinusoidale in dB c) Finestra a coseno rialzato d) Trasformata di Fourier della finestra a coseno rialzato in dB

pari a $1/NT$ e quindi le loro ampiezze si sottraggono, avendo i due lobi successivi segni opposti.

Pertanto, l' n -esimo lobo laterale, di larghezza $1/NT$ come tutti gli altri lobi laterali, decrescerà con una legge del tipo (fig. 5.3):

$$\frac{1}{n-1} - \frac{1}{n} = \frac{1}{n(n-1)} \quad (5.6)$$

e quindi quadraticamente con n . Peraltro, la larghezza del lobo principale è $3/NT$ e cioè il 25% meno del lobo principale della finestra triangolare. Quindi, la finestra cosinusoidale è meglio della finestra triangolare, per queste applicazioni: il vero vantaggio della finestra triangolare sta nel fatto che lo spettro è sempre positivo, cosa non vera per la finestra cosinusoidale; ma questo vantaggio sarà utile per la stima spettrale. Si possono studiare finestre più complesse, ma fondate sullo stesso principio:

CAPITOLO 5 FILTRI NUMERICI

ad esempio la finestra a coseno rialzato (fig. 5.3):

$$w_R(t) = \frac{1}{2} \left[1 + \cos \left(\frac{2\pi t}{NT} \right) \right]; \quad |t| < \frac{NT}{2} \quad (5.7)$$

Qui, la trasformata si trova come convoluzione del sinc e della somma di un impulso di ampiezza .5 nell'origine e due impulsi di ampiezza .25 alle frequenze $f \pm 1/NT$. In questo caso, i lobi laterali decrescono con legge cubica come risulta da

$$-\frac{1}{4(n-1)} + \frac{1}{2n} - \frac{1}{4(n+1)} = \frac{1}{2n(n^2-1)} \quad (5.8)$$

Il lobo centrale ha larghezza $4/NT$ come nel caso della finestra triangolare: anche questa finestra è piú efficiente della finestra triangolare, perchè a pari lobo centrale, i lobi laterali decrescono con legge cubica invece che quadratica.

5.4 Altri tipi di finestre

In generale, si può pensare di costruire tanti tipi diversi di finestra, che ottimizzano diverse quantità; p.e. il raggio di girazione della risposta all'impulso, e cioè l'area incrementale moltiplicata per il quadrato della distanza dall'origine. Esistono anche famiglie di finestre (Finestre di Kaiser) che assomigliano a delle gaussiane e quindi tendono ad avere eguale aspetto nelle frequenze e nei tempi. La finestra di Kaiser è:

$$w_K(t) = \frac{I_0 \left[k \sqrt{1 - \left(\frac{2t}{NT} \right)^2} \right]}{I_0(k)}; \quad |t| \leq \frac{NT}{2} \quad (5.9)$$

La funzione I_0 è la funzione di Bessel modificata del primo tipo, di ordine 0; Il parametro k serve ad ottenere il compromesso desiderato tra ampiezza dei lobi laterali e larghezza del lobo principale. Una finestra interessante è la finestra di Hamming (fig. 5.4),

$$w_H(t) = .54 + .46 \cos \left(\frac{2\pi t}{NT} \right); \quad |t| < \frac{NT}{2} \quad (5.10)$$

Questa strana finestra ha un piccolo gradino agli estremi: essa quindi ha lobi che discendono come $1/N$ ma con piccola ampiezza perchè lo sbilanciamento dei tre lobi che vengono mediati nella convoluzione è piccolo. Il merito sta nel fatto che il primo lobo laterale ha ampiezza minima.

5.5 PROGETTO CON LA TECNICA MINIMAX

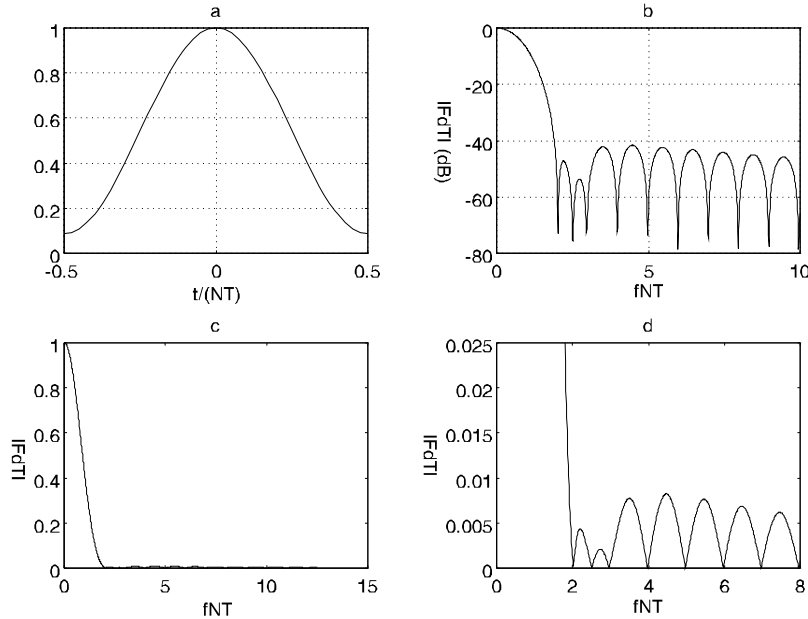


Figura 5.4: a) Finestra di Hamming b) Trasformata della finestra di Hamming in dB c) Trasformata della finestra di Hamming d) Particolare della trasformata della finestra di Hamming

5.5 Progetto con la tecnica MINIMAX

Nei casi studiati finora, per un dato lobo centrale, e quindi per un dato smussamento della risposta in frequenza, i lobi laterali decrescono con la differenza di frequenza; quindi le zone della funzione di trasferimento più lontane dalle transizioni ripide sono meno disturbate dalla finestatura.

Può essere utile progettare un filtro in modo che tutti i lobi laterali abbiano la stessa ampiezza e cioè in modo da minimizzare l'errore massimo dovunque capiti. Questo tipo di progetto, detto MINIMAX, è effettuato in modo iterativo. In questo caso, è essenziale che il numero dei parametri che caratterizza il filtro sia limitato e ci si riferirà ad un segnale a fase zero, cioè simmetrico nei tempi: si dimostra infatti che, a parità di lunghezza del filtro, i filtri a fase zero hanno le transizioni di ampiezza più ripide, nel dominio delle frequenze. Sia $\{h_n\}$; $|n| \leq N$ la sequenza corrispondente al filtro, lungo quindi $2N + 1$ campioni, di cui

CAPITOLO 5 FILTRI NUMERICI

però solo $N + 1$ sono diversi tra loro. La sua trasformata è:

$$H(\omega) = h_0 + \sum_{n=1}^N 2h_n \cos(n\omega T) \quad (5.11)$$

Vogliamo che

$$W(\omega)|H(\omega) - D(\omega)| \leq \delta \quad (5.12)$$

dove $W(\omega)$ è una funzione peso, positiva e la $D(\omega)$ è la funzione di trasferimento desiderata.

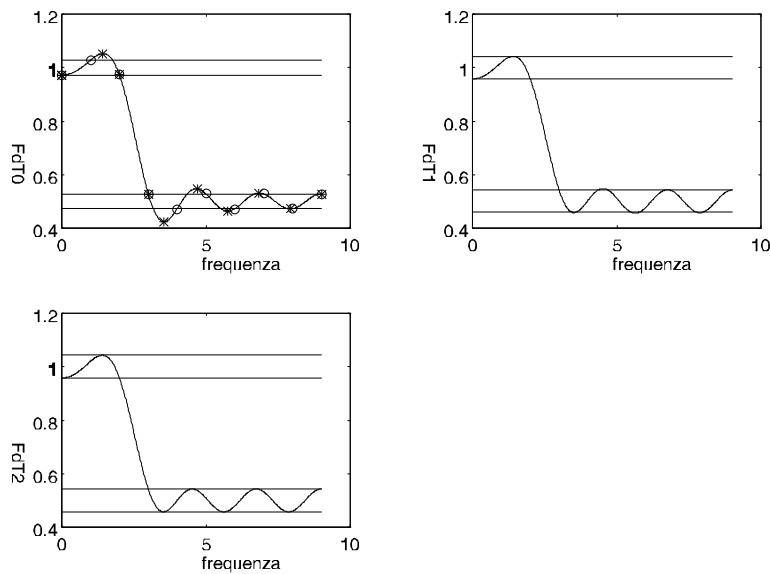


Figura 5.5: Progetto di un filtro FIR passa basso con la tecnica MINIMAX; progressive iterazioni

Esiste un teorema, dovuto a Chebyshev, che afferma che quando un polinomio trigonometrico di ordine N deve approssimare una funzione qualsiasi minimizzando l'errore massimo allora la differenza tra la funzione approssimata e quella approssimante ha $N + 2$ massimi e minimi successivi, di modulo eguale e segno alterno. Tuttavia, non sappiamo a quali frequenze la differenza sia massima; nasce allora una procedura iterativa.

5.6 LUNGHEZZA DEI FILTRI FIR PASSA - BASSO

Supponiamo di partire da un insieme di $N + 2$ frequenze $f_i^{(0)}$; $i = 1, \dots, N + 2$ e **imponiamo** che a queste frequenze le differenze tra funzioni approssimante ed approssimata siano tutte eguali ad un valore δ incognito. Inoltre, in corrispondenza delle discontinuità della maschera la differenza tra la funzione di trasferimento sintetizzata e quella desiderata raggiungerà un minimo od un massimo; pertanto, due frequenze estremali vanno anche posizionate in corrispondenza delle frequenze limite della banda passante e della banda attenuata dove appunto la caratteristica desiderata risulta essere discontinua. Scriveremo poi:

$$W(\omega_i^{(0)}) \left[D(\omega_i^{(0)}) - \left[h_0^{(0)} + \sum_{n=1}^N 2h_n^{(0)} \cos(n\omega_i^{(0)}T) \right] \right] = \quad (5.13)$$

$$= (-1)^i \delta^{(0)}; \quad i = 1, \dots, N + 2; \quad \omega_i^{(0)} = 2\pi f_i^{(0)}$$

Questo è un sistema lineare di $N + 2$ equazioni nelle $N + 2$ incognite $h_n^{(0)}$, $n = 0, \dots, N$; $\delta^{(0)}$, che dà sempre una soluzione. Tuttavia, in generale non sarà vero che le frequenze $f_i^{(0)}$ scelte all'inizio risultino essere le frequenze estremali, e cioè le frequenze alle quali la differenza (approssimante - approssimata) raggiunge i massimi e minimi: le frequenze estremali che risultano in questa prima approssimazione saranno invece altre, che indicheremo con $f_i^{(1)}$; pertanto alle frequenze estremali la differenza non risulta essere costante, come sarebbe richiesto. E' necessario iterare con il secondo insieme di frequenze; si risolve di nuovo il sistema di equazioni (5.13) e così via fino a che $f_i^{(k)} \simeq f_i^{(k+1)}$. A questo punto siamo alla convergenza ed accade appunto che a queste frequenze le differenze tra funzione approssimante ed approssimata hanno i valori (estremi) tutti eguali a $\delta^{(k+1)}$. A questo punto si verifica se $\delta^{(k+1)} \leq \delta$. Se sì, il problema è risolto. Altrimenti si deve ricominciare con N più grande. Naturalmente, se δ è troppo piccolo rispetto a quello richiesto, si può ricominciare con un valore di N minore (fig. 5.5).

5.6 Lunghezza dei filtri FIR passa - basso

È opportuno disporre di una stima del numero N di campioni da cui è costituito un filtro FIR, passa - basso; a priori, è chiaro che N non dipende tanto dalla frequenza di taglio del filtro quanto dalla larghezza Δf della banda di transizione del comportamento del filtro tra passaggio

CAPITOLO 5 FILTRI NUMERICI

e attenuazione; con f_c indichiamo la frequenza di campionamento. La lunghezza illimitata del filtro corrispondente alla risposta all'impulso del tipo *sinc* va delimitata con una finestra $\{w_n\}$ e quindi la risposta in frequenza convoluta con la trasformata della finestra $W(z)$. La larghezza del lobo principale della finestra determina quindi la larghezza della banda di transizione Δf . Utilizzando la tecnica MINIMAX per il calcolo del filtro, è disponibile una formula approssimata che dà la lunghezza del filtro in funzione dell'ondulazione (ripple) (δ_1) della funzione di trasferimento in banda passante e dell'attenuazione (δ_2) nella banda oscura.

$$N \approx .66 \times \frac{f_c}{\Delta f} \times \log_{10} \left(\frac{1}{10 \times \delta_1 \times \delta_2} \right)$$

Ad esempio, per $\delta_1 = \delta_2 = 10^{-4}$, $N \approx 4.6 \times f_c/\Delta f$. Poichè poi la larghezza della banda di transizione va commisurata con la larghezza di banda del filtro (ad esempio, il 20%) ne deriva una relazione tra la frequenza di taglio del filtro f_{\max} e la sua durata:

$$\text{Se: } \Delta f = .2f_{\max} \quad \text{allora } \longrightarrow N \approx 20f_c/f_{\max}$$

È chiaro che filtri passa basso molto selettivi sono molto costosi in termini di numero di moltiplicazioni; diventa interessante allora la tecnica di procedere per stadi e cioè:

- a) filtrare passa basso la sequenza iniziale riducendone la frequenza massima di un fattore ragionevole;
- b) sottocampionare p.e. dello stesso fattore la sequenza così ottenuta;
- c) iterare filtraggio passa basso e sottocampionamento.

In questo modo il risparmio può essere notevole, visto che la frequenza di taglio relativa dei filtri di ognuno degli stadi della sequenza è ben maggiore di quella finale, desiderata. In altre parole, per costruire un filtro a frequenza di taglio $f_c/1000$, si possono utilizzare 3 stadi di riduzione di un fattore 10, in cascata. La lunghezza del filtro necessario per ottenere il valore di Δf desiderato è ottenuta con un segnale fortemente sottocampionato e perciò a basso costo computazionale. Infatti, i filtri degli stadi successivi al primo operano su sequenze sottocampionate e pertanto hanno un costo ridotto di un fattore pari al fattore di sottocampionamento (10, nel caso considerato). Ulteriori analisi di questi filtri verranno fatte al capitolo dedicato a decimazione e interpolazione.

5.7 Filtri IIR

Abbiamo visto che per realizzare filtri molto selettivi, è necessario utilizzare sequenze lunghe: ciò corrisponde a costi computazionali elevati a meno che non si ricorra all'uso di poli, e cioè costruire filtri con equazioni alle differenze finite recursive in cui il termine forzante è il segnale di ingresso e l'uscita è la soluzione. Come già visto, se i filtri sono causali, per assicurare la convergenza, sarà necessario garantire che i poli siano contenuti nel cerchio unitario. Tuttavia, ciò non è necessario se si opera in tempo differito, quando cioè tutti i dati di ingresso sono disponibili prima del trattamento e pertanto si può invertire l'asse temporale per ribaltare all'interno del cerchio unitario i poli esterni. Nel seguito, per semplicità ci riferiremo a filtri causali e pertanto avremo necessità di verificare la convergenza.

In generale, anche a scopo didattico, la progettazione di filtri IIR viene fatta trasponendo nel discreto i filtri analogici realizzati con sistemi di equazioni differenziali a coefficienti costanti. Un sistema lineare elettrico, meccanico, ecc. è rappresentato da un sistema di equazioni differenziali lineari, tempo continue; si desidera trasformare questi sistemi in sistemi alle differenze finite. Per fare ciò è necessario anzitutto definire l'operatore di derivazione per segnali numerici, e poi farne combinazioni lineari per ottenere l'equivalente di operatori di derivazione di ordine superiore.

5.7.1 Derivazione numerica

Se i dati fossero noti tutti, dal primo all'ultimo, e se potessimo operare in tempo differito, ed inoltre se le sequenze a disposizione fossero molto lunghe, non vi sarebbero problemi nel calcolo del derivatore. Infatti, come già visto nel capitolo sulla trasformazione di Fourier discreta, basta trasformare la sequenza, per poi moltiplicarla per $j2\pi k/NT$. Si ha così la trasformata della sequenza delle derivate, **periodicizzata**. Per comodità, ricalcoliamo la risposta all'impulso del derivatore campionato. La risposta del filtro ideale è:

$$x(t) = x = \frac{\sin \pi t/T}{\pi t/T};$$

CAPITOLO 5 FILTRI NUMERICI

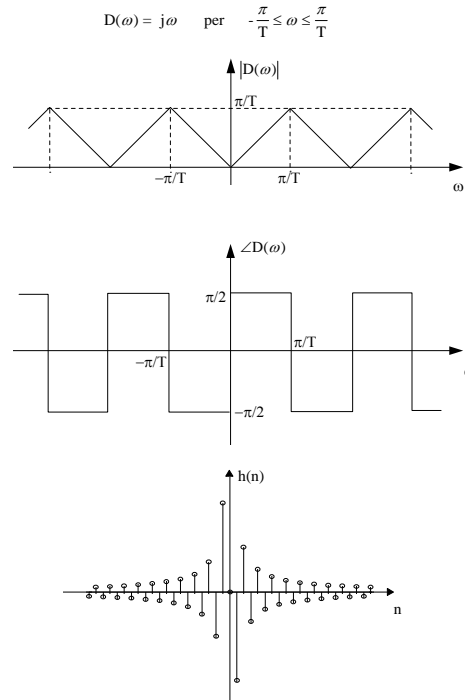


Figura 5.6: Derivatore ideale campionato

La derivata, campionata agli istanti $t = nT$, è:

$$\frac{dx}{dt} \Big|_{t=nT} = \frac{\cos n\pi}{nT} = \frac{(-1)^n}{nT}; n \neq 0; = 0; n = 0$$

Se la sequenza è lunga, la periodicizzazione non dà problemi, tranne che agli estremi (fig. 5.6). Però non è possibile operare in tempo reale, perchè abbiamo bisogno di sequenze lunghe e quindi verranno introdotti notevoli ritardi per attuare l'elaborazione nel dominio delle frequenze. Abbiamo pertanto necessità di tecniche che operino in tempo reale, e quindi dovremo determinare un operatore di derivazione causale; inoltre, vorremo che l'operatore di derivazione sia a fase minima perchè anche il suo inverso sia causale e stabile. Si consideri il sistema caratterizzato da

5.7 FILTRI IIR

un'equazione differenziale di primo ordine:

$$\frac{dy(t)}{dt} = ay(t) + bx(t) \quad (5.14)$$

e trasformiamolo in un sistema discreto. La discretizzazione desider-

$$D_0(\omega) = j \cdot \frac{2}{T} \cdot \sin\left(\frac{\omega \cdot T}{2}\right) \cdot e^{j\omega \frac{T}{2}}$$

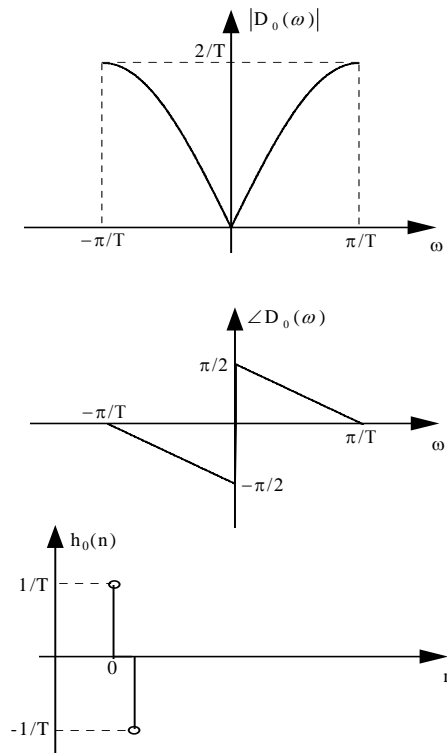


Figura 5.7: Derivatore non ricorsivo

abile corrisponderebbe al derivatore ideale discreto, che è come visto in precedenza:

$$d_f(nT) = \frac{(-1)^n}{nT}; \quad d_f(0) = 0$$

Questa soluzione però è antisimmetrica, e quindi non causale e quindi inaccettabile, perché non la si potrebbe realizzare in pratica con filtri

CAPITOLO 5 FILTRI NUMERICI

senza ritardo. La soluzione causale più semplice è:

$$\frac{y_n - y_{n-1}}{T} = ay_n + bx_n \quad (5.15)$$

in cui si è rappresentata la derivata con la differenza tra il campione attuale e quello precedente, divisa per T ; trasformiamo nelle frequenze per vedere la validità della scelta fatta:

$$D(\omega)Y(\omega) = Y(\omega)\frac{(1 - e^{-j\omega T})}{T} = aY(\omega) + bX(\omega) \quad (5.16)$$

$$D(\omega) = \frac{(1 - e^{-j\omega T})}{T} = \frac{2j}{T} \sin\left(\frac{\omega T}{2}\right) e^{-j\omega T/2} \simeq j\omega; \quad \omega T \ll 2 \quad (5.17)$$

L'operatore di derivazione $D(\omega)$ considerato va abbastanza bene, purchè la frequenza sia piccola rispetto alla frequenza di Nyquist; sia la caratteristica di ampiezza che quella di fase differiscono però da quella di un derivatore esatto fino alla frequenza di Nyquist.

La caratteristica di ampiezza di $D(\omega)$ (fig. 5.7) ha alla frequenza di Nyquist valore $2/T$ invece che π/T ; inoltre nella caratteristica di fase, si osserva un ritardo pari a $T/2$, che fa sì che l'operatore non sia a fase minima, se non al limite; infatti, ha uno zero sul cerchio unitario. Il ritardo di mezzo campione è comprensibile perchè la derivata è riferita all'istante $(n-1/2)T$. E' ragionevole quindi fare come proposto da Crank e Nicholson e cioè riferire anche il secondo membro dell'equazione (5.15) allo stesso istante.

$$\frac{y_n - y_{n-1}}{T} = a\frac{(y_n + y_{n-1})}{2} + b\frac{(x_n + x_{n-1})}{2} \quad (5.18)$$

Nel dominio delle frequenze questo diventa:

$$Y(\omega)\frac{(1 - e^{-j\omega T})}{T} = aY(\omega)\frac{1 + e^{-j\omega T}}{2} + bX(\omega)\frac{1 + e^{-j\omega T}}{2}$$

$$Y(\omega)\frac{2(1 - e^{-j\omega T})}{T(1 + e^{-j\omega T})} = aY(\omega) + bX(\omega) \quad (5.19)$$

L'operatore di derivazione è perciò rappresentato dall'operatore di derivazione ricorsiva:

$$D_r(\omega) = \frac{2(1 - e^{-j\omega T})}{T(1 + e^{-j\omega T})} = j\frac{2}{T} \tan\frac{\omega T}{2} \quad (5.20)$$

5.7 FILTRI IIR

La caratteristica di fase è quella giusta e cioè un anticipo costante di $\pi/2$; la caratteristica di ampiezza, invece, tocca valore infinito alla frequenza di Nyquist. Peraltro, dalla figura 5.8 è evidente come il derivatore $D_r(\omega)$ sia piú preciso dell'operatore $D(\omega)$ fino a frequenze piú elevate. Abbiamo quindi una tecnica per trasformare un sistema differenziale continuo in un sistema discreto: basterà sostituire alla variabile ω_a (la frequenza analogica) la variabile ω_n (la frequenza numerica):

$$\omega_n = \frac{2}{T} \tan^{-1} \left(\frac{\omega_a T}{2} \right) \quad (5.21)$$

Rappresentando ora la sequenza corrispondente alla $D_r(\omega)$ con la trasformata z , scriveremo la seguente corrispondenza:

$$j\omega \Rightarrow D_r(z) = \frac{2}{T} \frac{1 - z^{-1}}{1 + z^{-1}} \quad (5.22)$$

Di nuovo abbiamo un sistema che ha uno zero sul cerchio unitario e quindi rimane da dimostrare che un sistema differenziale stabile viene trasformato in un sistema discreto stabile e cioè poli nel semipiano sinistro nel caso di sistemi tempo continui vengono trasformati in poli all'interno del cerchio unitario con la corrispondenza dell'equazione (5.21). Piuttosto che fare una dimostrazione, faremo una verifica con il sistema RC della fig. 5.9:

La tensione $V_r(\omega)$ è:

$$V_r(\omega) = \frac{E(\omega)}{1 + j\omega\tau} \quad \tau = RC \quad (5.23)$$

Nel caso discreto, il segnale $e(t)$ si trasforma in una sequenza $\{e_n\}$, e l'equazione si trasforma in un'equazione alle differenze finite di cui presentiamo la trasformata

$$\begin{aligned} \frac{V(z)}{E(z)} &= 1 / \left(1 + \tau \frac{2}{T} \frac{1 - z^{-1}}{1 + z^{-1}} \right) \\ &= \frac{1 + z^{-1}}{1 + a + (1 - a)z^{-1}}; \quad a = \frac{2\tau}{T} \end{aligned} \quad (5.24)$$

Antitrasformando, si ha:

$$v_n = \frac{a - 1}{a + 1} v_{n-1} + \frac{1}{1 + a} (e_n + e_{n-1}) \quad (5.25)$$

CAPITOLO 5 FILTRI NUMERICI

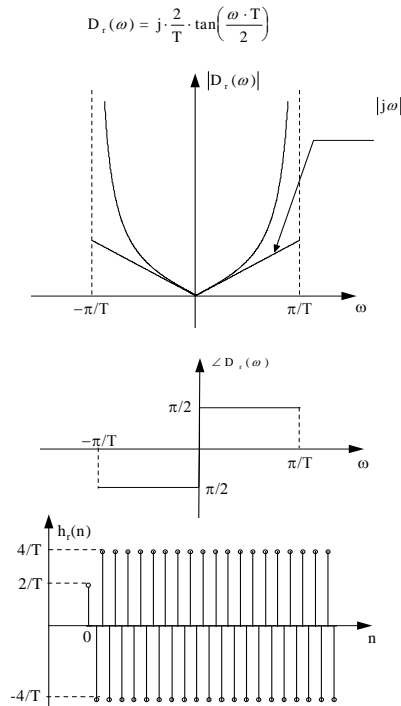


Figura 5.8: Derivatore ricorsivo

Il polo è nel cerchio unitario e cioè:

$$-1 < \frac{a-1}{a+1} < 1 \quad \text{se } \tau > 0 \quad (5.26)$$

e cioè se e solo se la resistenza R è positiva e quindi se il sistema che opera nel tempo continuo è stabile. In conclusione, la derivazione recursiva determinata con l'operatore

$$D_r(z) = \frac{2}{T} \frac{1 - z^{-1}}{1 + z^{-1}} \quad (5.27)$$

trasforma sistemi continui stabili in sistemi discreti stabili: inoltre la trasformazione non altera sostanzialmente l'asse delle frequenze fino a circa $1/3$ della frequenza di Nyquist, come vedremo al prossimo paragrafo.

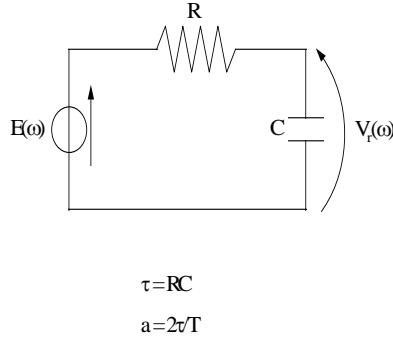


Figura 5.9: Filtro passa - basso RC

5.7.2 La trasformazione bilineare

La corrispondenza tra sistemi tempo continui causali e tempo discreti causali indotta dall'equazione (5.21)

$$j\omega \rightarrow D_r(z) = \frac{2}{T} \frac{1 - z^{-1}}{1 + z^{-1}} = j \frac{2}{T} \tan \frac{\omega T}{2} \quad (5.28)$$

può essere vista come una corrispondenza tra l'asse delle frequenze di un sistema continuo e quello di un sistema discreto (fig. 5.10). Infatti se alla variabile $j\omega$ si sostituisce la variabile $D_r(w)$ si ottiene un sistema discreto al posto di un sistema continuo, come si è visto. Se $H_c(\omega_a)$ è il valore della funzione di trasferimento del sistema tempo continuo ad una pulsazione analogica ω_a assegnata, la funzione di trasferimento del sistema tempo discreto corrispondente $H_d(\omega)$ avrà lo stesso valore alla pulsazione ω_d dove

$$\omega_d = \frac{2}{T} \tan^{-1} \left(\frac{\omega_a T}{2} \right) \quad (5.29)$$

$$\omega_a = \frac{2}{T} \tan \left(\frac{\omega_d T}{2} \right) \quad (5.30)$$

Quindi avviene che i valori assunti dalla funzione di trasferimento analogica nell'intervallo $0 \div \infty$ sono assunti dalla funzione di trasferimento numerica corrispondente nell'intervallo $0 \div f_N$. I filtri passa basso perciò vanno a zero alla frequenza di Nyquist. Ogni tecnica per il progetto di filtri analogici, p.e. i filtri di Butterworth o di Chebyshev può essere

CAPITOLO 5 FILTRI NUMERICI

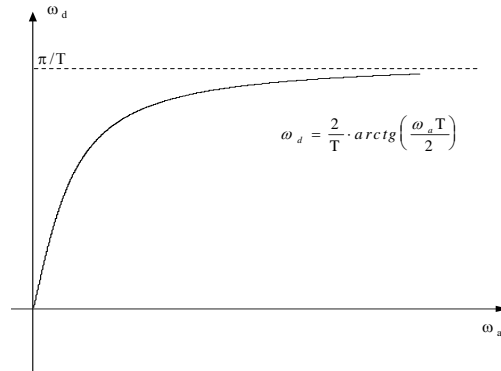


Figura 5.10: Trasformazione bilineare: relazione tra frequenza analogica e frequenza numerica

trasposta in campo numerico facilmente con la trasformazione bilineare, appunto. Si devono però poi applicare delle predistorsioni in maniera da compensare la distorsione non uniforme dell'asse delle frequenze indotta dalla trasformazione bilineare. Ad esempio, si verifica che:

$$\omega_a \simeq \omega_d \left[1 + \frac{\omega_d^2 T^2}{12} \right] + \dots \quad (5.31)$$

L'errore di frequenza commesso è pertanto inferiore al 10% per frequenze $f < \simeq .35 f_N$. Se il filtro da progettare è fortemente selettivo, la caratteristica di frequenza va predistorta. Per esempio, si voglia progettare un filtro passa banda tra $.8 f_N \div .998 f_N$; basterà cercare sulle tabelle il filtro analogico passabanda tra le frequenze:

$$\frac{1}{\pi T} \tan .4\pi \div \frac{1}{\pi T} \tan .499\pi$$

Quindi si applica la trasformazione bilineare per passare dal filtro analogico al filtro numerico, e il risultato è ottenuto. Ritornando all'esempio del circuito RC del paragrafo precedente, si può osservare che il sistema analogico ha come risposta all'impulso un esponenziale con costante di tempo τ ; il sistema numerico corrispondente ha un polo posizionato in $a - 1/a + 1$ ed uno zero alla frequenza di Nyquist. La risposta all'impulso

5.7 FILTRI IIR

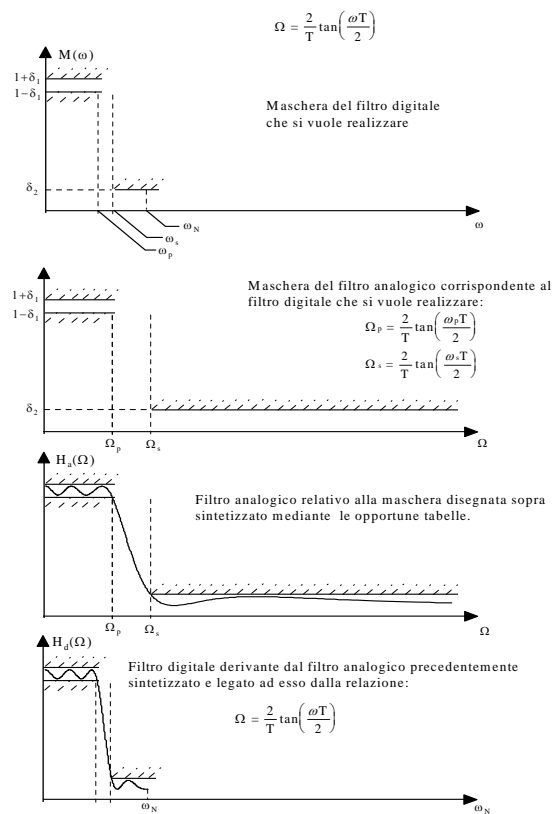


Figura 5.11: Esempio di progettazione di un filtro digitale mediante trasformazione bilineare

CAPITOLO 5 FILTRI NUMERICI

del sistema numerico è:

$$h_n = 0; n < 0; h_0 = \frac{1}{1+a}; h_n = \frac{2a}{a^2-1} \left(\frac{a-1}{a+1} \right)^n \quad (5.32)$$

Il termine esponenziale è approssimabile poi come:

$$\frac{2a}{a^2-1} \left(\frac{a-1}{a+1} \right)^n \simeq \frac{T}{\tau} e^{n \log(a-1/a+1)} \simeq \frac{T}{\tau} e^{-nT/\tau}; \quad \log(1+x) \simeq x \quad (5.33)$$

e quindi approssima (a parte il solito fattore T) la risposta all'impulso del sistema analogico, campionata, ma ne è diversa soprattutto in corrispondenza del primo campione, come vedremo meglio al prossimo paragrafo

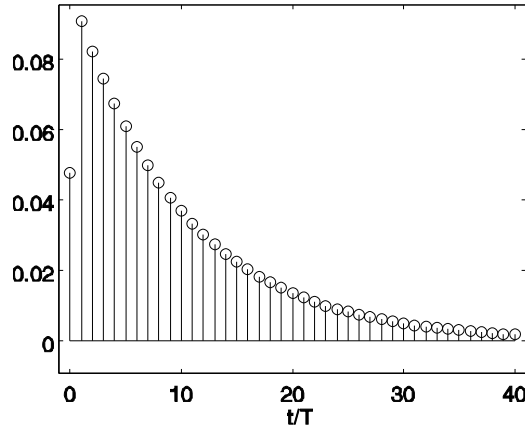


Figura 5.12: Risposta all'impulso del sistema numerico (RC); $a = \frac{2\tau}{T} = 20$

5.8 L'invarianza della risposta impulsiva

Una diversa corrispondenza tra sistemi analogici e sistemi numerici può essere stabilita imponendo che la risposta all'impulso dei secondi sia ottenuta dal campionamento della risposta all'impulso dei primi. Ad esempio alla funzione di trasferimento

$$H(\omega) = \frac{1}{1+j\omega\tau} \rightarrow h(t) = \frac{1}{\tau} e^{-t/\tau}$$

5.9 SINTESI DI FILTRI IIR DALL'AUTOCORRELAZIONE

si farebbe corrispondere la sequenza:

$$h_n = \frac{T}{\tau} e^{-nT/\tau}; \quad n = 0, 1, \dots$$

e quindi la trasformata z

$$H(z) = \frac{T}{\tau} \frac{1}{1 - e^{-T/\tau} z^{-1}}$$

Di nuovo, a sistemi stabili corrisponderebbero sistemi stabili; le risposte all'impulso si somiglierebbero, ma le risposte in frequenza sarebbero piuttosto diverse tranne ovviamente che per frequenze molto basse e per $T \ll \tau$. Infatti, il campionamento, qualsiasi sia l'intervallo T , è **effettuato senza prefiltraggio** e quindi comporta equivocazione. In particolare, alla frequenza di Nyquist, dove vi è il ripiegamento, la funzione di trasferimento desiderata sarebbe sempre riprodotta con un errore del 100%. Nel caso della funzione di trasferimento discussa prima, risultati migliori si avrebbero imponendo (*per continuità*) che il valore assunto al tempo 0 sia $T/2\tau$ invece che T/τ ; così si aggiunge uno zero (non proprio alla frequenza di Nyquist come nel caso della trasformazione bilineare, ma molto vicino se $T \ll \tau$), riducendo il fenomeno dell'equivocazione. Infatti in questo caso risulta:

$$H'(z) = \frac{T}{2\tau} \frac{1 + e^{-T/\tau} z^{-1}}{1 - e^{-T/\tau} z^{-1}}$$

Peraltro, con la tecnica dell'invarianza impulsiva, sarebbe più difficile fare delle trasposizioni da un sistema analogico ad un sistema numerico che non con la trasformazione bilineare (fig. 5.12).

5.9 Sintesi di filtri IIR dall'autocorrelazione

5.9.1 Le equazioni di Yule Walker

Si desideri determinare un filtro tutti poli di ordine N , $1/A_N(z)$, la cui caratteristica di ampiezza approssimi la caratteristica di ampiezza desiderata, $|H(z)|$:

$$\left| \frac{1}{A_N(z)} \right| \sim |H(z)|$$

CAPITOLO 5 FILTRI NUMERICI

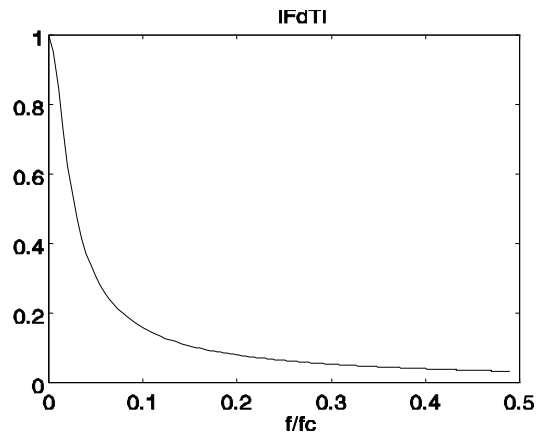


Figura 5.13: Modulo della fdt del sistema RC numerico ($a = \frac{2\pi}{T} = 20$)

Figura 5.14: Corrispondenza tra frequenze analogiche e numeriche nei casi di trasformazione bilineare e invarianza impulsiva ($f_c = 1000Hz$).

5.9 SINTESI DI FILTRI IIR DALL'AUTOCORRELAZIONE

Il filtro ottenuto dovrà essere stabile e quindi a minima fase. Per trovare $A_N(z)$, utilizzeremo la tecnica di Yule Walker. Si osservi anzitutto che assegnare la funzione $|H(z)|$ implica anche il rendere disponibile l'autocorrelazione della risposta all'impulso $\{h_n\}$ del filtro $H(z)$. Infatti, l'antitrasformata di $|H(z)|^2$ non è altro che l'autocorrelazione $\{r_m\}$ di $\{h_n\}$. Infatti, l'autocorrelazione è la convoluzione tra una sequenza e la stessa sequenza ribaltata nei tempi e coniugata. La trasformata dell'autocorrelazione è quindi il prodotto di $H(z)$ e $H^*(1/z)$. Sul cerchio unitario, $H^*(1/z) = (H(z))^*$ e quindi

$$r_m = \sum_{n=0}^{\infty} h_n h_{n-m}^* \Leftrightarrow H(z) H^*(1/z) = |H(z)|^2$$

Utilizzeremo ora i primi N tra i valori assegnati di $\{r_m\}$ per costruire il filtro $A_N(z)$.

Sia $\{x_n\}$ la sequenza in uscita del filtro IIR di ordine N , $1/A_N(z)$ quando questo viene alimentato da una sequenza bianca $\{w_n\}$, e cioè da una sequenza di campioni indipendenti. Senza ledere la generalità, $A_N(z)$ è rappresentato da un polinomio monico, e cioè con il primo coefficiente pari ad 1. Quindi,

$$A_N(z) = 1 + \sum_{h=1}^N a_{h,N} z^{-h}$$

è descritto da N coefficienti $a_{h,N}$, $h = 1, \dots, N$. La sequenza $\{x_n\}$, uscita del filtro, si ottiene pertanto dall'equazione alle differenze finite:

$$x_n = - \sum_{h=1}^N a_{h,N} x_{n-h} + w_n \quad (5.34)$$

e passando alle trasformate z

$$X(z) = \frac{W(z)}{A_N(z)}$$

L'autocorrelazione r_m della sequenza $\{x_n\}$ che, se appunto il filtro A_N è alimentato da una sequenza bianca, è ancora proporzionale all'autocorrelazione della risposta all'impulso del filtro, può essere dedotta moltiplicando la equazione (5.34) per il campione della sequenza x_{n-m}^* (anticipato

CAPITOLO 5 FILTRI NUMERICI

di m campioni e coniugato) e poi considerando il valore atteso del prodotto:

$$r_x(m) = r_m = E[x_n x_{n-m}^*] = -E\left[\sum_{h=1}^N a_{h,N} x_{n-h} x_{n-m}^*\right] + E[w_n x_{n-m}^*] = \quad (5.35)$$

$$r_m = -\sum_{h=1}^N a_{h,N} r_{m-h} + \sigma_w^2 \delta(m); \quad \forall m \quad (5.36)$$

Infatti ogni campione della sequenza $\{w_n\}$, costituita da campioni indipendenti tra loro, è incorrelato con i campioni dell'uscita generati in precedenza x_{n-m} ; $m > 0$.

$$E[w_n x_{n-m}^*] = E\left[w_n \left(-\sum_{h=1}^N a_{h,N} x_{n-h-m} + w_{n-m}\right)^*\right] = 0; \quad m > 0$$

Per $m = 0$, si ha invece:

$$E[w_n x_n^*] = E[w_n w_n^*] = \sigma_w^2$$

e quindi

$$E[w_n x_{n-m}^*] = \sigma_w^2 \delta(m)$$

Le equazioni (5.36) sono note come le equazioni di Yule - Walker e forniscono:

- gli N coefficienti del filtro $A_N(z)$ a partire dai primi N valori della funzione di autocorrelazione $\{r_m\}$, a parte r_0 , ottenuti con l'antitrasformazione del filtro desiderato $|H(z)|^2$.

- tutti i campioni dell'autocorrelazione dell'uscita del filtro, dati i coefficienti del filtro $A_N(z)$ ed il valore quadratico medio del segnale d'ingresso σ_w^2 . I primi N campioni saranno eguali a quelli desiderati, per costruzione. Non così i successivi, a meno che il filtro $H(z)$ non sia stato ricostruito perfettamente e cioè se $H(z)$ è un filtro a N_1 poli, e $N_1 < N$.

Ancora, con la tecnica che sarà vista nel capitolo sulle strutture riverberanti, da $A_N(z)$ si possono determinare gli N coefficienti di riflessione c_1, c_2, \dots, c_N che caratterizzano la struttura riverberante che consta di N celle e che approssima la funzione di trasferimento $H(z)$. Si osservi

5.9 SINTESI DI FILTRI IIR DALL'AUTOCORRELAZIONE

anche che nel caso di segnale e o coefficienti $a_{h,N}$ complessi, le equazione (5.36) vanno scritte tenendo conto che $r_{-m} = r_m^*$.

Per motivi mnemonici, può essere utile osservare che la formula (5.36) assomiglia ad una convoluzione tra le sequenze r_m ed $a_{h,N}$; non è una convoluzione, per la limitazione degli indici. L'equazione corrispondente a $m = 0$ contiene anche la varianza del segnale di ingresso σ_w^2 ; peraltro, il filtro non dipende da questo fattore di scala e si può sempre riscalarlo la funzione di autocorrelazione ponendo $r_0 = 1$ risolvendo poi le equazioni (Y-W) per $m = 1, \dots, N$.

5.9.2 L'autocorrelazione dai coefficienti del filtro

È possibile determinare la funzione di autocorrelazione dai coefficienti del filtro utilizzando le equazioni di Yule Walker. Infatti, ponendo $a_{0,N} = 1$, eliminando per semplicità il secondo indice e supponendo l'autocorrelazione reale, le equazioni diventano, ordinando il vettore delle attuali incognite $r_0 \dots r_N$:

$$\begin{aligned} \sum_{h=0}^N a_h r_{m-h} &= \sum_{n=m}^{m-N} a_{m-n} r_n = \sum_{n=m}^0 a_{m-n} r_n + \sum_{n=-1}^{m-N} a_{m-n} r_n = \\ &= \sum_{n=0}^m a_{m-n} r_n + \sum_{n=1}^{N-m} a_{m+n} r_n = \sum_{n=1}^m a_{m-n} r_n + \sum_{n=0}^{N-m} a_{m+n} r_n = \sigma_w^2 \delta(m) \end{aligned}$$

Queste formule si possono scrivere in modo molto semplice in forma matriciale (le matrici sono di Toeplitz entrambe, ma la prima lungo la diagonale principale e la seconda lungo la diagonale secondaria, a parte la

CAPITOLO 5 FILTRI NUMERICI

prima colonna di zeri nella seconda matrice).

$$\left[\begin{array}{c} \left(\begin{array}{cccccc} 1 & 0 & 0 & .. & 0 & 0 \\ a_1 & 1 & 0 & .. & 0 & 0 \\ a_2 & a_1 & 1 & 0 & .. & 0 \\ .. & .. & .. & .. & .. & .. \\ a_{N-1} & a_{N-2} & a_{N-3} & .. & 1 & 0 \\ a_N & a_{N-1} & a_{N-2} & .. & a_1 & 1 \end{array} \right) \\ \left(\begin{array}{cccccc} 0 & a_1 & a_2 & a_3 & .. & a_N \\ 0 & a_2 & a_3 & .. & a_N & 0 \\ 0 & a_3 & .. & a_N & 0 & 0 \\ .. & .. & .. & .. & .. & .. \\ 0 & a_N & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) \end{array} \right] + \begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ .. \\ r_{N-1} \\ r_N \end{bmatrix} = \begin{bmatrix} \sigma_w^2 \\ 0 \\ 0 \\ .. \\ 0 \\ 0 \end{bmatrix}$$

Trovati i primi N valori $r_1..r_N$ o in generale assegnata qualsiasi N -pla consecutiva di valori di r_n , poi possiamo determinare $A_N(z)$ e quindi i valori di r_n per qualsiasi altro valore di n .

5.9.3 Equazioni di Yule Walker in forma matriciale

Le equazioni di Yule Walker si possono scrivere in forma matriciale indicando i seguenti vettori con i corrispondenti simboli:

$$\underline{\mathbf{a}}_N = \begin{bmatrix} a_{1,N} \\ a_{2,N} \\ a_{3,N} \\ \dots \\ a_{N,N} \end{bmatrix} ; \underline{\boldsymbol{\xi}}_N = \begin{bmatrix} x_{n-1}^* \\ x_{n-2}^* \\ x_{n-3}^* \\ \dots \\ x_{n-N}^* \end{bmatrix} ; \underline{\mathbf{r}}_N = \begin{bmatrix} r_x(1) \\ r_x(2) \\ r_x(3) \\ \dots \\ r_x(N) \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ \dots \\ r_N \end{bmatrix}$$

Allora si ha:

$$x_n = -\underline{\boldsymbol{\xi}}_N^* \underline{\mathbf{a}}_N + w_n \quad (5.37)$$

5.9 SINTESI DI FILTRI IIR DALL'AUTOCORRELAZIONE

Moltiplicando a sinistra per $\underline{\xi}_N$ e considerando il valore atteso, visto che $E \left[\underline{\xi}_N w_n \right] = \underline{\mathbf{0}}$:

$$E \left[\underline{\xi}_N x_n \right] = \underline{\mathbf{r}}_N = E \left[\underline{\xi}_N \left(-\underline{\xi}_N^* \underline{\mathbf{a}}_N + w_n \right) \right] = - E \left[\underline{\xi}_N \underline{\xi}_N^* \right] \underline{\mathbf{a}}_N = -\underline{\mathbf{R}} \underline{\mathbf{a}}_N$$

La matrice dei coefficienti delle equazioni (Y-W) è una matrice di Toeplitz; cioè è simmetrica (o Hermitiana, per sequenze complesse) e tutti gli elementi appartenenti alla stessa diagonale o subdiagonale sono eguali tra loro. La matrice $\underline{\mathbf{R}}$ $[N \times N]$ è pertanto caratterizzata da N numeri. Il vettore dei termini noti $-\underline{\mathbf{r}}_N$ è anch'esso formato dai campioni della funzione di autocorrelazione, r_m ; $m = 1, \dots, N$, col segno cambiato. Riscrivendo in forma matriciale le equazioni di Yule Walker si ottiene:

$$\begin{bmatrix} r_0 & r_{-1} & r_{-2} & \dots & r_{-N} \\ r_1 & r_0 & r_{-1} & \dots & r_{-N+1} \\ r_2 & r_1 & r_0 & \dots & \dots \\ \dots & \dots & \dots & \dots & r_{-1} \\ r_N & r_{N-1} & \dots & r_1 & r_0 \end{bmatrix} \begin{bmatrix} a_{1,N} \\ a_{2,N} \\ a_{3,N} \\ \dots \\ a_{N,N} \end{bmatrix} = - \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ \dots \\ r_N \end{bmatrix} \quad (5.38)$$

ed in equazione vettoriale:

$$\underline{\mathbf{R}}_N \underline{\mathbf{a}}_N = -\underline{\mathbf{r}}_N \quad (5.39)$$

Per completezza, si aggiunge una formulazione completa, facilmente verificabile, in cui compare anche la prima equazione che contiene la varianza della sequenza bianca di ingresso $\sigma_{w,N}^2$.

$$\begin{bmatrix} r_0 & r_{-1} & r_{-2} & \dots & r_{-N} \\ r_1 & r_0 & r_{-1} & \dots & r_{-N+1} \\ r_2 & r_1 & r_0 & \dots & \dots \\ \dots & \dots & \dots & \dots & r_{-1} \\ r_N & r_{N-1} & \dots & r_1 & r_0 \end{bmatrix} \begin{bmatrix} 1 \\ a_{1,N} \\ a_{2,N} \\ \dots \\ a_{N,N} \end{bmatrix} = \begin{bmatrix} \sigma_{w,N}^2 \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix} \quad (5.40)$$

Si ricorda infine che in queste equazioni non sempre le incognite sono contenute nel vettore $\underline{\mathbf{a}}_N$; a volte è la funzione di autocorrelazione a non essere nota, oppure il valore di $\sigma_{w,N}^2$, ecc.

5.9.4 La ricorsione di Levinson

Nel calcolo di $A_N(z)$, si può utilizzare la stessa recursione di Levinson, che incontreremo nello studio delle strutture riverberanti. Infatti, si può risolvere l'equazione (5.39) a partire da $N = 1$, per poi crescere progressivamente. Indicando con:

$$\underline{\mathbf{R}}_{N+1} = \begin{bmatrix} \underline{\mathbf{R}}_N & \tilde{\mathbf{r}}_N \\ \tilde{\mathbf{r}}_N^* & r_0 \end{bmatrix}; \quad \tilde{\mathbf{r}}_N = \begin{bmatrix} r_{-N} \\ r_{-N+1} \\ \dots \\ r_{-1} \end{bmatrix}; \quad \tilde{\mathbf{a}}_N = \begin{bmatrix} a_{N,N}^* \\ \dots \\ a_{2,N}^* \\ a_{1,N}^* \end{bmatrix}$$

i vettori colonna $\tilde{\mathbf{r}}_N, \tilde{\mathbf{a}}_N$ sono i vettori colonna $\mathbf{r}_N, \mathbf{a}_N$ coniugati, ribaltati negli indici e ritardati per avere indici tutti positivi, con l'usuale convenzione che abbiamo visto legare le due sequenze $\tilde{A}(z)$ ed $A(z)$:

$$\tilde{A}_N(z) = z^{-N} A_N^*(1/z)$$

Il vettore riga $\tilde{\mathbf{r}}_N^*$ è il trasposto Hermitiano (e quindi ancora coniugato) del vettore colonna $\tilde{\mathbf{r}}_N$.

$$\begin{bmatrix} \underline{\mathbf{R}}_N & \tilde{\mathbf{r}}_N \\ \tilde{\mathbf{r}}_N^* & r_0 \end{bmatrix} \mathbf{a}_{N+1} = - \begin{bmatrix} \mathbf{r}_N \\ r_{N+1} \end{bmatrix}$$

Si verifica che la soluzione è:

$$\mathbf{a}_{N+1} = \begin{bmatrix} \mathbf{a}_N \\ 0 \end{bmatrix} + c_{N+1} \begin{bmatrix} \tilde{\mathbf{a}}_N \\ 1 \end{bmatrix}$$

e quindi:

$$\begin{bmatrix} \underline{\mathbf{R}}_N & \tilde{\mathbf{r}}_N \\ \tilde{\mathbf{r}}_N^* & r_0 \end{bmatrix} \left\{ \begin{bmatrix} \mathbf{a}_N \\ 0 \end{bmatrix} + c_{N+1} \begin{bmatrix} \tilde{\mathbf{a}}_N \\ 1 \end{bmatrix} \right\} = - \begin{bmatrix} \mathbf{r}_N \\ r_{N+1} \end{bmatrix}$$

Difatti le prime N righe corrispondono alle N equazioni:

$$\underline{\mathbf{R}}_N \mathbf{a}_N + \tilde{\mathbf{r}}_N \times 0 + c_{N+1} (\underline{\mathbf{R}}_N \tilde{\mathbf{a}}_N + \tilde{\mathbf{r}}_N \times 1) = -\mathbf{r}_N \quad (5.41)$$

5.9 SINTESI DI FILTRI IIR DALL'AUTOCORRELAZIONE

che sono verificate poichè:

$$\underline{\mathbf{R}}_N \underline{\mathbf{a}}_N = -\underline{\mathbf{r}}_N ; \underline{\mathbf{R}}_N \tilde{\underline{\mathbf{a}}}_N = -\tilde{\underline{\mathbf{r}}}_N \quad (5.42)$$

L'ultima equazione serve per determinare c_{N+1} ed è:

$$\tilde{\underline{\mathbf{r}}}_N^* (\underline{\mathbf{a}}_N + c_{N+1} \tilde{\underline{\mathbf{a}}}_N) + c_{N+1} r_0 = -r_{N+1}$$

da cui:

$$c_{N+1} = - \frac{r_{N+1} + \tilde{\underline{\mathbf{r}}}_N^* \underline{\mathbf{a}}_N}{r_0 + \tilde{\underline{\mathbf{r}}}_N^* \tilde{\underline{\mathbf{a}}}_N} \quad (5.43)$$

Quindi:

$$\underline{\mathbf{a}}_{N+1} = \begin{bmatrix} a_{1,N} \\ a_{2,N} \\ \dots \\ a_{N,N} \\ 0 \end{bmatrix} + c_{N+1} \begin{bmatrix} a_{N,N}^* \\ \dots \\ a_{2,N}^* \\ a_{1,N}^* \\ 1 \end{bmatrix}$$

e ricordando che

$$A_N(z) = 1 + \sum_{h=1}^N a_{h,N} z^{-h} \quad (5.44)$$

ne consegue la tecnica detta di Levinson per la soluzione ricorsiva delle equazioni di Yule e Walker. Si ha cioè:

$$A_{N+1}(z) = A_N(z) + c_{N+1} z^{-1} \tilde{A}_N(z); \quad \tilde{A}_N(z) = z^{-N} A_N^*(1/z) \quad (5.45)$$

dove c_N risulterà essere il coefficiente di riflessione N -esimo della struttura riverberante che realizza la funzione di trasferimento $1/A_N(z)$ e sarà ottenuto dai valori della funzione di autocorrelazione e dal polinomio $A_{N-1}(z)$ nel modo osservato in (5.43) e come verrà di nuovo dimostrato nel capitolo della predizione lineare. Si dimostra anche che:

$$\sigma_{w,N}^2 = r_0 (1 - |c_1|^2) (1 - |c_2|^2) \dots (1 - |c_N|^2)$$

e quindi:

$$\sigma_{w,N}^2 \leq \sigma_{w,N-1}^2$$

CAPITOLO 5 FILTRI NUMERICI

Questo modo di risolvere le equazioni di Yule e Walker è computazionalmente efficiente. Infatti, per risolvere un sistema di N equazioni in N incognite servono un ordine di grandezza di N^3 moltiplicazioni; usando la ricorsione di Levinson, invece, il costo computazionale di un incremento di indice corrisponde, a parte le somme, al costo di calcolo della (5.43) e quindi $2N$ moltiplicazioni ed una divisione. Il costo totale diventa quindi circa:

$$\sum_{n=1}^N 2n \sim 2 \frac{N^2}{2} = N^2$$

5.9.5 Calcolo dei c_n dagli $a_{h,N}$; la verifica di Schur

Siamo ora in grado di determinare i coefficienti di riflessione c_n ; $n = 1, \dots, N$ a partire dal polinomio $A_N(z)$: basta applicare la ricorsione di Levinson all'incontrario, partendo da N per arrivare ad 1, p.e. supponendo $c_0 = 1$. Si può poi dimostrare (come proposto da Schur e Cohn) che un polinomio qualsiasi, il cui primo coefficiente è 1 come accade per il polinomio $A_N(z)$, è a minima fase se tutti i coefficienti di riflessione sono in modulo minori di 1 (ciò verrà dimostrato nel capitolo sulla predizione lineare).

Si può verificare facilmente se tutti i coefficienti di riflessione c_i sono in modulo minori di 1. Per fare ciò è utile scrivere i coefficienti dei polinomi $A_N(z)$ ed $A_{N-1}(z)$ in termini di vettori colonna. La ricorsione di Levinson (5.45) diventa:

$$\begin{bmatrix} 1 \\ a_{1,N} \\ a_{2,N} \\ \dots \\ a_{N-1,N} \\ a_{N,N} \end{bmatrix} = \begin{bmatrix} 1 \\ a_{1,N-1} \\ a_{2,N-1} \\ \dots \\ a_{N-1,N-1} \\ 0 \end{bmatrix} + c_N \begin{bmatrix} 0 \\ a_{N-1,N-1} \\ a_{N-2,N-1} \\ \dots \\ a_{1,N-1} \\ 1 \end{bmatrix} \quad (5.46)$$

Anzitutto si osserva (5.45) che:

$$c_N = a_{N,N} \quad (5.47)$$

5.9 SINTESI DI FILTRI IIR DALL'AUTOCORRELAZIONE

e cioè che l'ultimo coefficiente del polinomio è un coefficiente di riflessione e pertanto deve essere in modulo inferiore ad 1. Per calcolare c_{N-1} dovremo calcolare $A_{N-1}(z)$. Per calcolare i coefficienti di $A_{N-1}(z)$, ad esempio $a_{h,N-1}$ dovremo risolvere il sistema:

$$a_{h,N} = a_{h,N-1} + c_N a_{N-h,N-1} \quad (5.48)$$

$$a_{N-h,N} = a_{N-h,N-1} + c_N a_{h,N-1} \quad (5.49)$$

da cui risulta :

$$a_{h,N-1} = \frac{a_{h,N} - c_N a_{N-h,N}}{1 - c_N^2} \quad (5.50)$$

Troveremo così $c_{N-1} = a_{N-1,N-1}$ e poi progressivamente tutti gli altri coefficienti di riflessione; se qualcuno di essi ha modulo maggiore di 1, il polinomio di partenza $A_N(z)$ non è a minima fase (la verifica di Schur - Cohn). I coefficienti di riflessione troveranno un uso diretto nel capitolo sulle strutture riverberanti; in particolare, serviranno per costruire dei filtri numerici poco sensibili alla quantizzazione dei coefficienti.

5.9.6 La sintesi del filtro

Riassumendo, se viene assegnata la caratteristica di ampiezza di un filtro da approssimare con una funzione trasferimento tutti poli di ordine N , $A_N(z)$, per effettuarne la sintesi e quindi per determinarne il denominatore:

$$|H(z)| \sim \frac{1}{|A_N(z)|}$$

si opererà così:

1) Si calcola $|H(\omega)|^2$ e, facendone l'antitrasformata z , si trova la funzione di autocorrelazione:

$$r(n); \quad n = 0, 1, \dots, N, \dots, \infty$$

2) Utilizzando le equazioni di Yule Walker, si trova, per ogni valore desiderato di N , il vettore \underline{a}_N ;

3) Si determina la funzione di trasferimento del filtro:

$$\frac{1}{A_N(z)} \sim H(z)$$

È necessario garantire che il filtro $1/A_N(z)$ sia stabile e pertanto che il polinomio $A_N(z)$ sia a minima fase; si dimostrerà che ciò accade **se e solo se** i campioni $r(n)$ che si utilizzano nelle equazioni di Yule Walker sono campioni di una funzione di autocorrelazione (a spettro positivo). Ciò è automaticamente ottenuto se $r(n)$ è antitrasformata di una funzione tutta positiva come $|H(\omega)|^2$. Poichè, per ottenere il filtro di ordine N , non serve **tutta** la funzione di autocorrelazione, ma ne bastano i primi N campioni, lo spettro $|H(\omega)|^2$ può essere smussato, purchè sia convoluto con una funzione tutta positiva come p.e. un sinc².

5.10 Filtri da combinazione di sfasatori puri

Da una semplice costruzione vettoriale ricaveremo ora un interessante classe di filtri, poi molto utilizzati nella pratica. Si considerino due sfasatori puri, e cioè filtri caratterizzati da poli all'interno del cerchio unitario (i denominatori sono funzioni a fase minima) e zeri nelle posizioni reciproche coniugate (i numeratori sono delle funzioni a fase massima) così da avere caratteristica di ampiezza unitaria. Si ricordi che uno sfasatore puro di ordine M comporta solo M moltiplicazioni per campione; inoltre, rimane sempre tale, anche al variare della quantizzazione dei coefficienti. Ad esempio, siano gli sfasatori puri $A_0(z)$, $A_1(z)$ di ordine N_0 , N_1 anche diverso:

$$A_0(z) = \frac{\sum_{n=0}^{N_0} a_{0,N_0-n}^* z^{-n}}{\sum_{n=0}^{N_0} a_{0,n} z^{-n}}; \quad |A_0(z)| = 1, \quad z = e^{j\omega T} \quad (5.51)$$

$$A_1(z) = \frac{\sum_{n=0}^{N_1} a_{1,N_1-n}^* z^{-n}}{\sum_{n=0}^{N_1} a_{1,n} z^{-n}}; \quad |A_1(z)| = 1, \quad z = e^{j\omega T} \quad (5.52)$$

$$H_0(z) = \frac{A_0(z) + A_1(z)}{2} \quad (5.53)$$

$$H_1(z) = \frac{A_0(z) - A_1(z)}{2} \quad (5.54)$$

Con la costruzione vettoriale di fig. 5.15 si vede che $H_0(\omega)$, $H_1(\omega)$ sono semidiagonali di un rombo che ha per lati $|A_0(\omega)| = |A_1(\omega)| = 1$; pertanto:

$$|H_0(z)|^2 + |H_1(z)|^2 = 1; \quad z = e^{j\omega T} \quad (5.55)$$

5.10 FILTRI DA COMBINAZIONE DI SFASATORI PURI

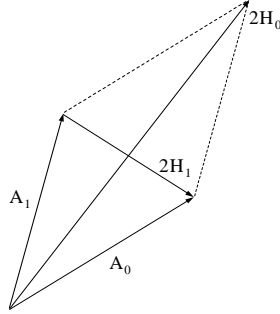


Figura 5.15: Filtri IIR ottenuti come somma e differenza di due sfasatori puri

e cioè i due filtri $H_0(\omega)$, $H_1(\omega)$ sono complementari in potenza; alle frequenze a cui è passa basso l'uno sarà passa alto l'altro e viceversa. Inoltre i due filtri sono anche complementari in ampiezza in quanto la loro somma è di ampiezza costante:

$$|H_0(z) + H_1(z)| = |A_0(z)| = 1; \quad z = e^{j\omega T} \quad (5.56)$$

I due filtri $H_0(\omega)$, $H_1(\omega)$ sono pertanto detti *doppiamente complementari*, in potenza ed ampiezza cioè. Ancora vi è un'altra utile caratteristica dei due filtri $H_0(\omega)$, $H_1(\omega)$; le sequenze causali corrispondenti ai numeratori delle due funzioni di trasferimento sono l'una simmetrica e l'altra antisimmetrica. Infatti:

$$A_0(\omega) = \frac{\tilde{D}_0(\omega)}{D_0(\omega)}, \quad A_1(\omega) = \frac{\tilde{D}_1(\omega)}{D_1(\omega)}; \quad (5.57)$$

$$H_{0,1}(\omega) = \frac{\tilde{D}_0(\omega)D_1(\omega) \pm \tilde{D}_1(\omega)D_0(\omega)}{2D_0(\omega)D_1(\omega)} = \frac{N(\omega) \pm \tilde{N}(\omega)}{2D_0(\omega)D_1(\omega)} \quad (5.58)$$

ma anche:

$$N(\omega) = \tilde{D}_0(\omega)D_1(\omega); \quad \tilde{N}(\omega) = \tilde{D}_1(\omega)D_0(\omega) \quad (5.59)$$

Si verifica che le famiglie di filtri di Butterworth, Chebyshev, ellittici sono appunto decomponibili nella semisomma o semidifferenza di due sfasatori puri.

CAPITOLO 5 FILTRI NUMERICI

Come trovare i due sfasatori puri date le funzioni di trasferimento ? Nel caso interessante dei filtri ellittici (nonchè Butterworth e Chebyshev, che ne sono casi limite), la cosa è semplice perchè si verifica che i poli di $A_0(\omega)$, $A_1(\omega)$ che sono anche poli di $H_0(\omega)$, $H_1(\omega)$ sono alternati nelle frequenze; in altre parole i poli di uno sfasatore puro si alternano con quelli dell'altro. Pertanto per fare la decomposizione, basta assegnare i poli a frequenze crescenti alternativamente ai due sfasatori. Gli zeri degli sfasatori sono dedotti dai poli. Nel caso di funzioni di trasferimento di ordine pari, l'osservazione è tuttora valida, ma i coefficienti degli sfasatori puri sono complessi [Vedi: Vaidyanathan, P. P., Multirate Systems and Filter Banks, Prentice Hall, 1993].

Elemento chiave per la progettazione dei filtri è il rapporto tra $|H_0(\omega)|$, $|H_1(\omega)|$ e cioè la funzione

$$K(\omega) = \frac{|N(\omega) + \tilde{N}(\omega)|}{|N(\omega) - \tilde{N}(\omega)|} \quad (5.60)$$

che ripartisce la potenza tra i due filtri, al variare della frequenza.

5.11 Filtri di Butterworth, Chebyshev, ellittici

Queste famiglie di filtri numerici, derivano dalle famiglie omologhe di filtri utilizzati per segnali analogici attraverso la trasformazione bilineare. Pertanto ne faremo solo un brevissimo cenno, limitato ai filtri di Butterworth, rinviando gli interessati alla letteratura specializzata ed inoltre al software *MATLAB*TM.

Ad esempio, il filtro passabasso di Butterworth di ordine n e frequenza di taglio ω_0 ha la funzione di trasferimento

$$|B_n(\omega)|^2 = \frac{1}{1 + (\frac{\omega}{\omega_0})^{2n}} \quad (5.61)$$

I poli della funzione di trasferimento analogica sono su di un semicerchio del piano s di raggio ω_0 e di fasi:

$$\frac{\pi}{2n} \times (n + 1 + 2k); \quad 0 \leq k \leq n - 1 \quad (5.62)$$

5.11 FILTRI DI BUTTERWORTH, CHEBYSHEV, ELLITTICI

Dalla posizione dei poli del filtro analogico, si ricava la funzione di trasferimento del filtro analogico e poi di quello numerico, usando sempre la relazione:

$$j\omega \leftrightarrow \frac{2}{T} \frac{1 - z^{-1}}{1 + z^{-1}} \quad (5.63)$$

La funzione di ripartizione della potenza in questo caso è

$$K(\omega) = |B_n(\omega_d)|^2 = \left[\frac{2}{\omega_0 T} \frac{|1 - z^{-1}|}{|1 + z^{-1}|} \right]^{2n} \quad (5.64)$$

Se $K(\omega) \rightarrow \infty$, la potenza del segnale è deviata attraverso H_0 ; attraverso H_1 , se invece $K(\omega) = 0$.

Nella fig.5.16 sono mostrate funzioni di trasferimento, risposte all'impulso e disposizione di poli e zeri di quattro filtri passa basso IIR di ordine 10. Si osservi come il filtro ellittico ha la massima ripidità di transizione e quindi transitorio di lunghezza maggiore mentre il filtro di Butterworth ha il transitorio più breve. Si osservi come l'ottimizzazione della distribuzione delle singolarità (*calcolando la funzione di trasferimento con il metodo grafico si comprende che le singolarità devono raggrupparsi vicino al cerchio unitario in corrispondenza delle transizioni ripide*) conduca a funzioni di trasferimento più vicine alla maschera di progetto. Dalla figura è poi facile comprendere la posizione delle singolarità dei due sfasatori puri in cui si potrebbe decomporre il filtro.

Dai grafici in figura si comprende come l'ottimizzazione della posizione di poli e zeri, possibile sistematicamente nei casi semplici di filtri che approssimino i filtri ideali, dia risultati incomparabilmente migliori di quelli che si ottengono sia dai filtri FIR ottenuti con la tecnica delle finestre, sia dai filtri IIR ottenuti con la tecnica di Yule Walker. Infatti, in entrambi questi casi, si fa tutto il possibile per riprodurre la funzione di autocorrelazione del filtro desiderato nei primi N campioni, non considerando affatto quelli successivi. Infatti questi sono nulli nel caso FIR e derivati con le equazioni di Yule Walker dai primi N , nel caso IIR. Invece, nei casi di filtri di Butterworth, Chebyshev, ecc. l'approssimazione è distribuita lungo tutto l'asse dei tempi.

CAPITOLO 5 FILTRI NUMERICI

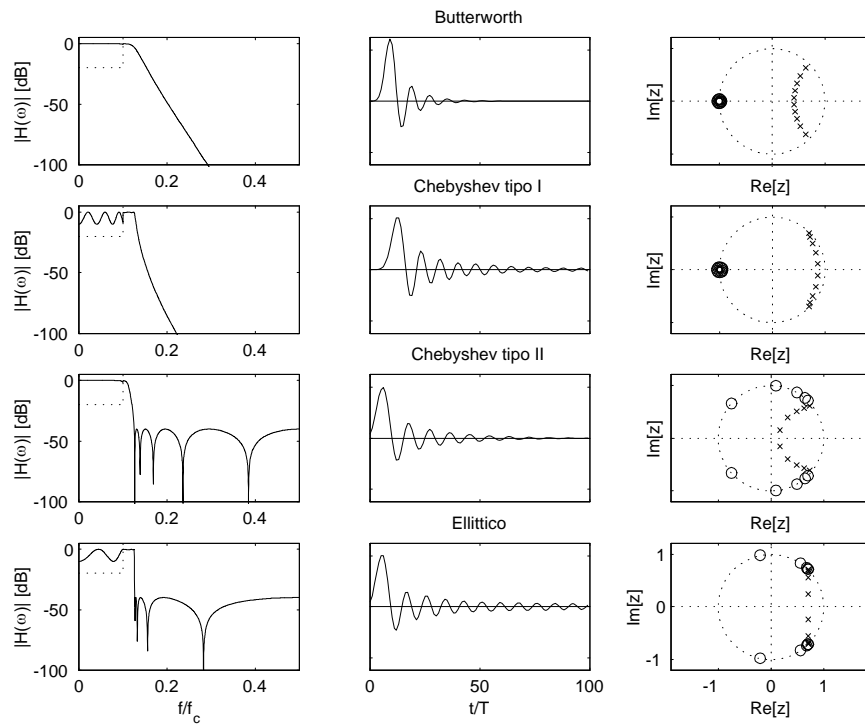


Figura 5.16: Filtri passa-basso IIR di ordine 10 [specifiche progetto: $f_p = f_c/4$, $\delta_1 = 0.5dB$, $\delta_2 = -40dB$]. La caratteristica della risposta in frequenza nella banda passante è amplificata (zona tratteggiata).

Capitolo 6

Campionamento in frequenza

6.1 Introduzione

Una classe di filtri molto utilizzata e di facile progettazione é quella dei filtri a campionamento in frequenza. Si tratta di filtri che sono teoricamente dei filtri FIR, anche se con dei poli, in quanto questi poli sono cancellati da opportuni zeri. Nei fatti, visto che la cancellazione é solo parziale, risultano essere dei filtri IIR.

Si consideri un oscillatore numerico, come quello studiato al capitolo 3. Se questo oscillatore é eccitato da un impulso, la sua risposta é una sinusoidale, di ampiezza proporzionale a quella dell'impulso. Se si applicano piú impulsi nel tempo, la risposta sará la somma delle risposte. Supponiamo ora che la sinusoidale generata compia un numero intero k di cicli, nell'intervallo di tempo NT . Se allora inviamo all'oscillatore due impulsi di eguale ampiezza e segno opposto a distanza di NT la risposta globale sará una sinusoidale di frequenza k/NT troncata dopo k periodi (fig. 6.1). Quindi la risposta all'impulso del filtro é finita ed il filtro é FIR, anche se realizzato con oscillatori numerici e cioé con strutture IIR. La realizzazione é presentata in fig. 6.4.

Essa é ottenuta combinando un oscillatore numerico ed un blocco di ritardo di N campioni realizzato con una memoria di N elementi, per creare sistematicamente l'impulso di spegnimento N campioni dopo l'impulso di attivazione dell'oscillatore. In linea di principio, il costo computazionale é basso, visto che si richiede una moltiplicazione per campione nell'oscillatore ed un'altra per stabilire il guadagno.

CAPITOLO 6 CAMPIONAMENTO IN FREQUENZA

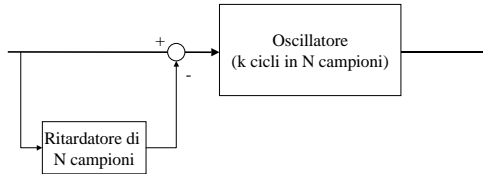


Figura 6.1: Funzionamento di un oscillatore a campionamento di frequenza

La risposta in frequenza è quella di una senoide a frequenza f_0 finestrata con una finestra rettangolare di durata NT ; quindi del tipo

$$H_c(\omega) = \text{sinc}[(f - f_0)NT] \quad (6.1)$$

A parte le repliche spettrali dovute al campionamento nei tempi, lo spettro corrisponde ad un *sinc* con il primo zero ad un intervallo di frequenza pari a $1/NT$ centrato sulla frequenza dell'oscillatore. Quindi, la selettività non è legata al costo computazionale, ma solo alla dimensione della memoria, che determina la durata della risposta all'impulso.

Poli e zeri

Il sistema di fig. 6.4 è caratterizzato da N zeri dovuti al blocco di ritardo e i poli dovuti all'oscillatore (uno se corrispondono a esponenziali complessi o due se a frequenze reali); la trasformata z della risposta è in questo caso (fig. 6.2):

$$H(z) = \frac{1 - z^{-N}}{1 - 2 \cos(2\pi k/N)z^{-1} + z^{-2}}; \quad k \neq 0, N/2 \quad (6.2)$$

$$H(z) = \frac{1 - z^{-N}}{1 \mp z^{-1}} \quad k = 0, N/2. \quad (6.3)$$

6.1 INTRODUZIONE

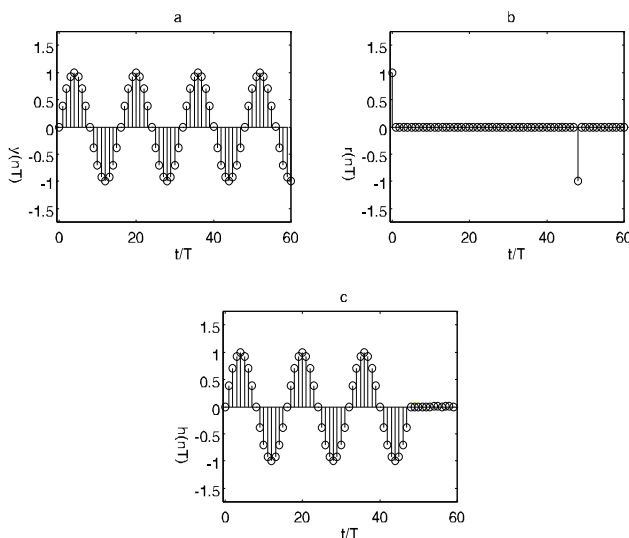


Figura 6.2: a) Risposta impulsiva dell'oscillatore b) Risposta impulsiva del ritardatore c) Risposta all'impulso del filtro a campionamento di frequenza

Gli zeri del primo blocco corrispondono alle N radici N -esime dell'unit  e quindi sono disposti regolarmente lungo il cerchio unitario a distanza angolare pari a $2\pi/N$ con il primo zero nel punto $z = 1$ (la componente continua   soppressa dal primo blocco). I poli sono posizionati in corrispondenza di qualche zero e ne sono cancellati, **se la frequenza dell'oscillatore   esatta**. Quindi il sistema   tutti zeri, FIR.

In generale, come rappresentato in figura 6.6 il blocco di ritardo pu  alimentare fino a $N/2 + 1$ oscillatori, alle frequenze $f_k = k/NT$; $k = 0..N/2$; in questo modo si potrebbero costruire complesse caratteristiche di ampiezza senza per  avere vantaggi computazionali, perch  con $N/2 + 1$ oscillatori il numero di moltiplicazioni per campione diventa pari a N come con un qualsiasi filtro FIR di pari lunghezza. Quindi i filtri a campionamento in frequenza hanno interesse dal punto di vista del costo computazionale solo se il numero di oscillatori alimentati   piccolo rispetto ad N e quindi se hanno delle caratteristiche di tipo passa banda piuttosto pronunciate.

Un altro motivo per cui i filtri a campionamento in frequenza sono interessanti, a parte quello computazionale visto in precedenza, sta nella

CAPITOLO 6 CAMPIONAMENTO IN FREQUENZA

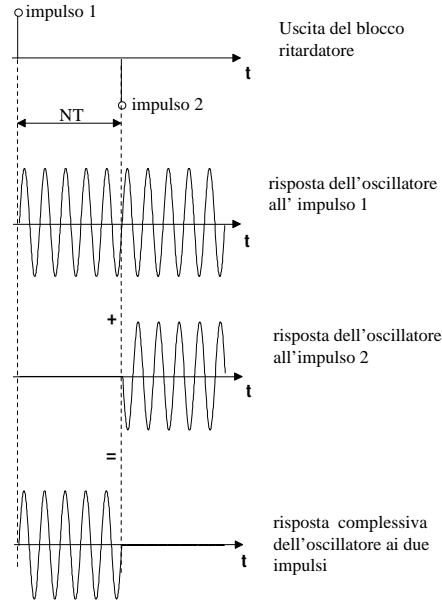


Figura 6.3: Modo alternativo per vedere il funzionamento di un filtro a campionamento di frequenza

modalità di progetto, assai semplice.

Infatti si consideri la caratteristica del filtro desiderato $D(\omega)$; (**attenzione** : se il filtro approssimante è causale, anche il filtro approssimando deve essere causale! Quindi $D(\omega)$ deve tenere conto di un ritardo pari circa a $N/2$ campioni) si campioni questa caratteristica alle $N/2 + 1$ pulsazioni $\omega_k = 2\pi k/NT$; $k = 0, \dots, N/2$ e siano D_k i valori (in generale complessi) di questi campioni. Ora, si interpolino questi campioni lungo l'asse delle frequenze con una funzione sinc che ha appunto il primo zero a intervallo $1/NT$; la funzione risultante $H(\omega)$ è appunto quella del sistema a campionamento in frequenza presentato nella figura 6.5, purché i guadagni H_k siano pari a D_k . Per ritrovare valori reali dei guadagni, si ricordi di accoppiare i rami corrispondenti alle frequenze positive con quelli delle frequenze negative.

La dimostrazione di quanto osservato segue dal fatto che stiamo facendo un'interpolazione (con un filtro ideale nel dominio delle frequenze) di un segnale costituito da una sequenza di impulsi (le sinusoidi alle frequen-

6.1 INTRODUZIONE

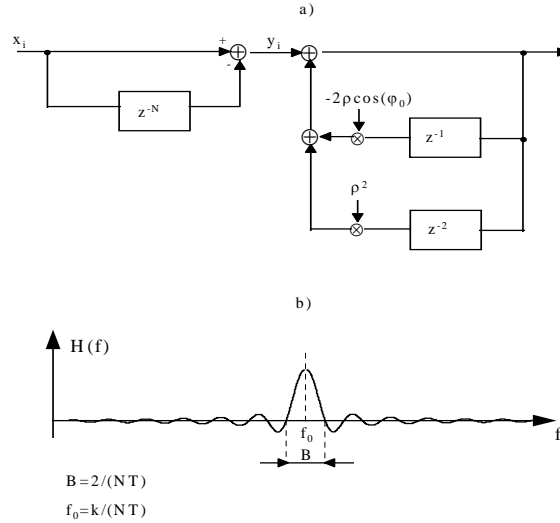


Figura 6.4: Schema a blocchi (a) e funzione di trasferimento (b) di un filtro a campionamento in frequenza

ze $f_k = k/NT$); nel capitolo 2 un'interpolazione simile era stata fatta nel dominio dei tempi. La convoluzione con il *sinc* nel dominio delle frequenze corrisponde alla finestratura rettangolare fatta nel dominio dei tempi. Peraltro si ricordi che sia ha a che fare con segnali campionati e pertanto periodici nel dominio delle frequenze; la trasformata del *sinc* discreto è presentata nel Capitolo 4.

Al crescere di N , cresce la qualità dell'interpolazione e quindi decresce la differenza tra approssimante ed approssimata. Per filtri passa banda, come si è detto, c'è poi anche un vantaggio computazionale.

Si osservi che la somma in parallelo dei vari rami nel sistema di figura 6.6 comporta la formazione di nuovi zeri accanto alla cancellazione dei poli. Altrimenti si arriverebbe al paradosso che la risposta finale ha solo un campione.

CAPITOLO 6 CAMPIONAMENTO IN FREQUENZA

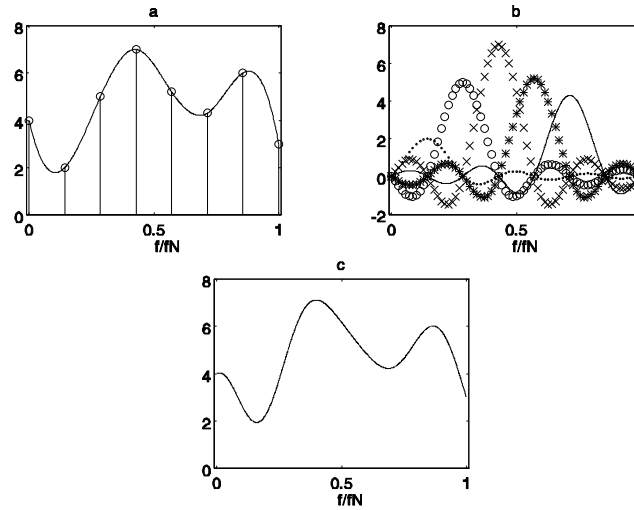


Figura 6.5: a) Campionamento della F.d.T. da sintetizzare b) Interpolazione con funzioni *sinc* c) Risultato dell' interpolazione

6.2 Progetto

Il progetto di filtri FIR utilizzando i campioni della trasformata di Fourier desiderata presi ad intervallo regolare $1/NT$ può essere ricondotto alla progettazione di filtri assegnando i valori dei campioni della DFT $\{H_k\}_{k=0}^{N-1}$. L'unica accortezza è che il filtro ricavato per questa via ha una risposta temporale periodica che deve essere troncata su un periodo per ottenere un filtro a risposta temporale limitata (FIR). Orbene, **il troncamento non può essere arbitrario**. Vengono presentati di seguito i passi principali:

a) Antitrasformazione discreta di Fourier:

$$\{H_k\}_{k=0}^{N-1} \Leftrightarrow \{\tilde{h}_n\}_{n=0}^{N-1}$$

.Si ricordi che la sequenza è periodica (ovvero $\tilde{h}_n = \tilde{h}_{n+N}$).

b) Troncamento su 1 periodo della sequenza periodica (sequenza a-periodica): $h_n = \tilde{h}_n$ per $n = 0, \dots, N - 1$ e zero altrove.

6.2 PROGETTO

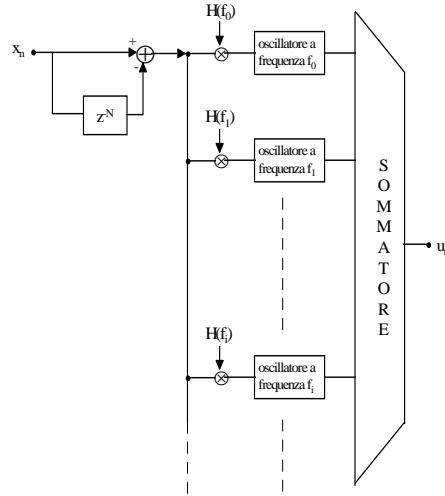


Figura 6.6: Schema a blocchi di un filtro a campionamento in frequenza i cui campioni assegnati siano $H(f_i)$

c) Trasformata z della sequenza troncata:

$$\begin{aligned}
 H(z) &= \sum_{n=0}^{N-1} h_n z^{-n} = \sum_{n=0}^{N-1} \underbrace{\left[\frac{1}{N} \sum_{k=0}^{N-1} H_k W_N^{kn} \right]}_{\text{IDFT}} z^{-n} = \\
 &= \frac{1}{N} \sum_{k=0}^{N-1} H_k \left[\sum_{n=0}^{N-1} W_N^{kn} z^{-n} \right] = \\
 &= \frac{1}{N} \sum_{k=0}^{N-1} H_k \left[\sum_{n=0}^{N-1} (W_N^k z^{-1})^n \right] \\
 &= \frac{1}{N} \sum_{k=0}^{N-1} H_k \frac{1 - W_N^{kN} z^{-N}}{1 - W_N^k z^{-1}} =
 \end{aligned}$$

CAPITOLO 6 CAMPIONAMENTO IN FREQUENZA

$$= \frac{1 - z^{-N}}{N} \sum_{k=0}^{N-1} \frac{H_k}{1 - W_N^k z^{-1}} =$$

$$\frac{1 - z^{-N}}{N} \sum_{k=0}^{N-1} \frac{H_k}{1 - \exp(j\frac{2\pi}{N}k)z^{-1}}$$

ovvero la risposta all'impulso del filtro può essere realizzata da un insieme di oscillatori che generano sinusoidi complesse di ampiezza $|H_k|$ e fase $\angle H_k$. Gli oscillatori sono "spenti" dopo N campioni e avere quindi compiuto un numero intero k di cicli.

d) Trasformata di Fourier della sequenza di durata limitata $\{h_n\}$:

$$H(z = \exp(j\omega T)) = \frac{1 - e^{-j\omega NT}}{N} \sum_{k=0}^{N-1} \frac{H_k}{1 - e^{j(\frac{2\pi}{N}k - \omega T)}}$$

e indicando con $W(\omega)$ la trasformata di una finestra rettangolare causale (ritardata in tempi di $(N - 1)/2$ campioni):

$$W(\omega) = \frac{1}{N} \frac{1 - e^{-j\omega NT}}{1 - e^{j(\frac{2\pi}{N}k - \omega T)}} = \frac{1}{N} \frac{e^{j\omega NT/2} \sin(\omega NT/2)}{e^{j\omega T/2} \sin(\omega T/2)} =$$

$$= \frac{\exp(-j\omega T(N - 1)/2)}{N} \cdot \frac{\sin(\omega NT/2)}{\sin(\omega T/2)}$$

sostituendo in ciascuno dei termini nella sommatoria la trasformata diventa:

$$H(z = \exp(j\omega T)) = \sum_{k=0}^{N-1} H_k \cdot W(\omega T - \frac{2\pi}{N}k)$$

dove la $W(\omega)$ compare ancora modulata (spostata lungo l'asse delle frequenze o ruotata attorno al cerchio unitario) di una pulsazione

$$2\pi f_k = \frac{2\pi k}{NT}$$

La trasformata di Fourier della sequenza h_n può essere quindi ricavata da H_k per (convoluzione) interpolazione con la funzione $W(\omega)$ (nota che $H(2\pi k/N) = H_k$), oppure osservando che la sequenza periodica è troncata da una finestra rettangolare (causale) di N campioni:

$$W(\omega) \longleftrightarrow w_n = \begin{cases} 1 & n = 0, 1, \dots, N - 1 \\ 0 & \text{altrove} \end{cases}$$

6.3 CAMPIONAMENTO IN FREQUENZA E FINESTRATURA

Si osservi infine che anche se solo una coppia degli H_k, H_{-k} è diversa da zero, pur tuttavia, per avere il valore desiderato della funzione di trasferimento, si forma uno zero reale. Se infatti $H_{-k} = H_k^*, W_N = \exp(j2\pi/N)$:

$$H(z) = 2 \frac{1 - z^{-N}}{N} \left(\frac{\operatorname{Re}(H_k) - z^{-1} \operatorname{Re}(H_k W_N^{-k})}{1 + z^{-2} - 2z^{-1} \operatorname{Re}(W_N^k)} \right)$$

6.3 Campionamento in frequenza e finestatura

Si può comprendere la differenza tra campionamento in frequenza e finestatura: in entrambi i casi, la sequenza $\{h_{d,n}\}; n = -\infty, \dots, \infty$, anti-trasformata della funzione desiderata $H_d(\omega)$ viene limitata ad N campioni. Nel caso della finestatura, la limitazione avviene con un troncamento e quindi la funzione desiderata $H_d(\omega)$ viene filtrata passa basso nel dominio delle frequenze con la convoluzione la trasformata della finestra, p.e. un sinc.

$$\{h_n\}; n = N/2, \dots, N/2 \rightarrow H(\omega) = H_d(\omega) * \frac{\sin \frac{\omega NT}{2}}{\frac{\omega NT}{2}} \quad (6.4)$$

Nel caso del campionamento in frequenza, la limitazione temporale viene ottenuta con la replicazione nei tempi a distanza di N campioni, e pertanto con un campionamento in frequenza **seguito** dalla convoluzione con lo stesso sinc.

Nel secondo caso abbiamo anche degli errori di equivocazione, dovuti al campionamento **non** preceduto dal filtraggio, che non avremmo avuto nel primo caso. In conclusione, il campionamento in frequenza è *peggio* che non la finestatura; in compenso, nel caso di filtri passa banda a banda stretta, realizzare un filtro lungo N , può costare molto meno di N moltiplicazioni per campione.

6.4 Effetti del ritardo

Il controllo della funzione di trasferimento del filtro a campionamento in frequenza è limitato poichè essa è progettata solo sulla base dei cam-

CAPITOLO 6 CAMPIONAMENTO IN FREQUENZA

pioni in frequenza H_k . L'andamento effettivo della funzione di trasferimento tra i campioni è controllato dalla funzione $W(\omega)$. Il termine di ritardo $\exp(-j\omega T(N-1)/2)$ introduce una oscillazione nella risposta in frequenza del filtro che può essere eliminata solo se si introduce una traslazione antitetica opportuna (o ritardo) della risposta all'impulso periodica. Riscrivendo la $H(\omega)$:

$$H(\omega) = \exp(-j\omega T(N-1)/2) \sum_{k=0}^{N-1} H_k \underbrace{\exp(j \frac{2\pi}{NT} k \times \frac{(N-1)T}{2})}_{\text{anticipo seq. periodica di } (N-1)/2 \text{ campioni}} \cdot C$$

$$C = \frac{\sin((\omega T - \frac{2\pi}{N}k)N/2)}{N \sin((\omega T - \frac{2\pi}{N}k)/2)}$$

si evidenziano tre contributi (nell'ordine): il ritardo di $(N-1)/2$ campioni necessario per avere un filtro causale, i coefficienti del filtro con uno sfasamento corrispondente ad un anticipo (il termine $\exp(j \frac{2\pi}{N}k \times \frac{(N-1)T}{2})$) e il seno cardinale periodicizzato che effettua l'interpolazione. Per ovviare ai termini indesiderati indotti dall'interpolazione è necessario scegliere i coefficienti del filtro H_k in modo che questi contengano un ritardo di $(N-1)/2$ campioni tale da compensare l'anticipo visto prima. In altri termini, a causa della periodicità di \tilde{h}_n il troncamento da applicare per ottenere h_n non è arbitrario. In pratica è **sempre necessario** ritardare di $(N-1)/2$ campioni la sequenza per ovviare ad un troncamento troppo brusco all'inizio e alla fine della risposta h_n . Questo troncamento altrimenti introdurrebbe termini ad alta frequenza indesiderati e zeri in posizioni difficilmente prevedibili. In fase di progetto il ritardo di $(N-1)/2$ campioni si realizza sfasando opportunamente i campioni in frequenza ovvero scegliendo opportunamente i campioni in frequenza \bar{H}_k , che sono:

$$H_k = \bar{H}_k \exp(-j\pi k \frac{N-1}{N}) = [\text{per } N \text{ grande}] \simeq (-1)^k \bar{H}_k,$$

per segnali reali : $H_k = H_{N-k}^*$

Questo giustifica l'inversione di segno necessaria quando si vuole realizzare un filtro passa banda combinando due risposte all'impulso a frequenze adiacenti. A titolo di esempio in figura si mostra la risposta all'impulso

6.4 EFFETTI DEL RITARDO

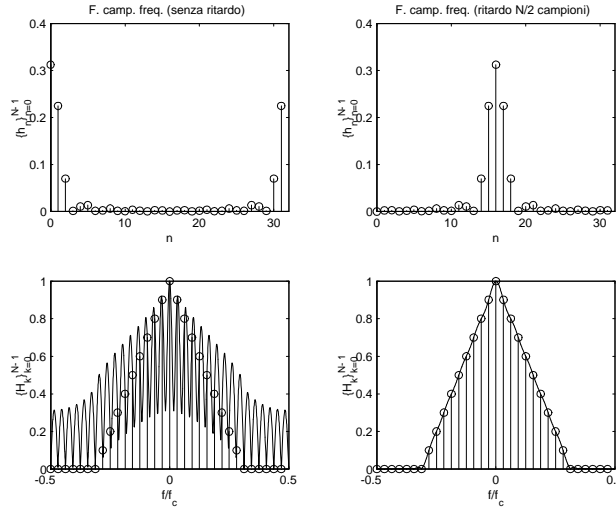


Figura 6.7: Risposta all'impulso (parte superiore) e trasformata di Fourier (parte inferiore) di un filtro passa basso a campionamento in frequenza in cui sono assegnati i campioni della fdt $\{H_k\}_{k=0}^{N-1}$ (pallini vuoti).

e la fdt di un filtro passa basso a campionamento in frequenza senza applicazione di ritardo e con un ritardo di $N/2$ campioni (per semplicità qui si è scelto di ritardare di $N/2$ campioni in quanto $N = 32 \gg 1$). In entrambi i casi la funzione di trasferimento coincide, per $\omega T = 2\pi k/N$, con i campioni della risposta in frequenza del filtro definiti in fase di progetto. Nel secondo caso la funzione di trasferimento appare più “ragionevole” per l'assenza di termini indesiderati ad alta velocità di oscillazione (introdotti dal termine $\exp[-j\omega T(N-1)/2]$ nella $W(\omega)$). La posizione degli zeri per i due casi è indicata nella figura.

Osservazioni:

1) Se la sequenza è reale allora $H_k = H_{N-k}^*$ e quindi la trasformata z diventa (separando il campione alla continua, $k = 0$, e al Nyquist, $k = N/2$, quando quest'ultimo è presente e cioè se N è pari):

$$H(z) = \frac{1 - z^{-N}}{N} \left[\frac{H_0}{1 - z^{-1}} + \frac{H_{N/2}}{1 + z^{-1}} + A \right]$$

$$A = \sum_{k=1}^{N/2-1} \left(\frac{H_k}{1 - \exp(j\frac{2\pi}{N}k)z^{-1}} + \frac{H_k^*}{1 - \exp(-j\frac{2\pi}{N}k)z^{-1}} \right)$$

CAPITOLO 6 CAMPIONAMENTO IN FREQUENZA

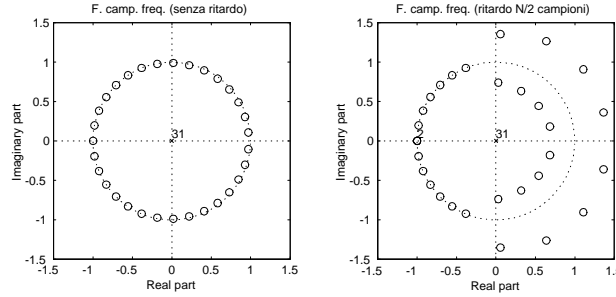


Figura 6.8: Poli/zeri filtro campionamento in frequenza senza e con ritardo di $N/2$ campioni. In entrambi i casi, i valori dei campioni in frequenza sono eguali!

$$= \frac{1 - z^{-N}}{N} \left[\frac{H_0}{1 - z^{-1}} + \frac{H_{N/2}}{1 + z^{-1}} + B \right]$$

$$B = 2 \sum_{k=1}^{N/2-1} \left(\frac{\operatorname{Re}(H_k) - \operatorname{Re}(H_k \exp(-j\frac{2\pi}{N}k)) z^{-1}}{1 - 2 \cos(\frac{2\pi}{N}k) z^{-1} + z^{-2}} \right)$$

L'ultimo termine è una somma di oscillazioni sinusoidali **reali** opportunamente sfasate secondo il termine di fase indicato dallo zero ($\operatorname{Re}(H_k) - \operatorname{Re}(H_k \exp(-j\frac{2\pi}{N}k)) z^{-1}$). È opportuno osservare che per H_k reali la risposta all'impulso di ogni oscillatore è una cosinusoidale.

2) Il ritardo di $(N - 1)/2$ campioni rappresenta un numero intero di campioni nel caso di N dispari mentre corrisponde ad un'interpolazione nei tempi nel caso di N pari (vedi esempio in figura). In quest'ultimo caso si potrebbe adottare un ritardo di $N/2$ a meno di un "piccolo" errore (l'errore è trascurabile solo per risposta all'impulso di durata limitata).

3) Ogni oscillatore ha (nel caso di risposta all'impulso reale) un costo di 3 moltiplicazioni/campione. Complessivamente ci sono $N/2 - 1$ oscillatori con un costo di (circa) $3N/2$ moltiplicazioni/campione. Se confrontato con un'implementazione diretta di un filtro FIR con risposta h_n (N moltiplicazioni/campione) il filtro a campionamento in frequenza ha un vantaggio computazionale solo se il numero di coefficienti H_k è piccolo (se il numero di oscillatori è limitato).

6.5 Quantizzazione dei coefficienti

Un importante problema realizzativo affligge i filtri a campionamento in frequenza: mentre le posizioni degli zeri della funzione di trasferimento sono corrispondenti proprio alle radici dell'unità, le posizioni dei poli dipendono dalla precisione con cui vengono realizzati i fattori $-2 \cos(2\pi k/N)$; se uno di questi fattori è rappresentato con un numero binario finito, e quindi imprecisamente, la perfetta cancellazione polo-zero diventa impossibile. La sinusoide lanciata da un impulso non fa un numero intero di cicli in N campioni, e l'impulso di segno opposto che arriva dopo N campioni lancia quindi una sinusoide che non sarà in perfetta opposizione di fase con la precedente. La somma delle due sinusoidi non si annulla e quindi rimane una sinusoide, residuo della cancellazione imperfetta. La risposta all'impulso del filtro diventa IIR, *non convergente a zero*; il sistema diventa quindi instabile, come vedremo. Un oscillatore con funzione di trasferimento

$$H(z) = \frac{1}{1 + az^{-1} + z^{-2}}; \quad |a| \leq 2 \quad (6.5)$$

genera una sinusoide campionata $\{y_n\}$ con pulsazione

$$\omega T = \varphi = \cos^{-1}\left(\frac{a}{2}\right) \quad (6.6)$$

Noi desidereremmo avere esattamente:

$$\omega = \frac{2\pi k}{N} \rightarrow a = -2 \cos(2\pi k/N)$$

Ma le piccole variazioni di a dovute alla quantizzazione dei coefficienti portano ad una variazione $\Delta\omega$ della pulsazione ω :

$$\Delta\omega = \frac{\Delta a}{2T \sin \varphi} \quad (6.7)$$

La differenza tra la sinusoide desiderata e quella generata dopo N campioni é:

$$\sin \alpha_1 - \sin \alpha_2 = 2 \cos \frac{\alpha_1 + \alpha_2}{2} \sin \frac{\alpha_1 - \alpha_2}{2} \quad (6.8)$$

$$\alpha_1 = (2\pi k/NT + \Delta\omega)(n + N)T; \quad \alpha_2 = (2\pi k/NT + \Delta\omega)nT \quad (6.9)$$

CAPITOLO 6 CAMPIONAMENTO IN FREQUENZA

L'ampiezza della sinusoide residua è (supponendo unitaria l'ampiezza della sinusoide generata):

$$y_{r,n} = 2 \sin \frac{\Delta\omega NT}{2} \cos\left[(\omega + \Delta\omega)nT + \frac{\Delta\omega NT}{2}\right] \simeq \frac{N\Delta a}{4 \sin \varphi} \cos\left(n\varphi + \frac{N\Delta a}{4 \sin \varphi}\right) \quad (6.10)$$

Come si vede, rimane una sinusoide residua di ampiezza anche piccola

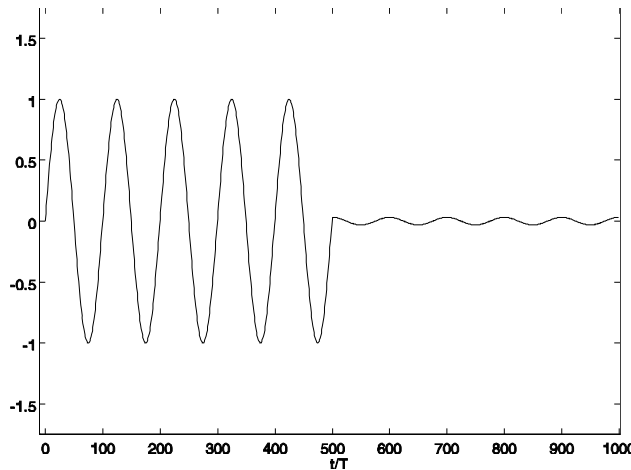


Figura 6.9: Risposta impulsiva di un filtro a campionamento di frequenza reale

se Δa è piccolo, ma di durata illimitata. La situazione è inaccettabile, perché ogni campione in ingresso poi lascerebbe un residuo sinusoidale di durata illimitata; dopo n campioni di ingresso casuali, di eguale varianza σ ed incorrelati tra loro, la varianza del segnale residuo, somma di n sinusoidi di ampiezza e fase casuale, cresce proporzionalmente a $\sigma\sqrt{n}$. Quindi, dopo qualche tempo, il filtro sarebbe comunque totalmente inutilizzabile, per buona che fosse la quantizzazione. Il sistema è diventato instabile.

Per stabilizzarlo, non rimane che riportare i poli degli oscillatori all'interno del cerchio unitario ($\rho < 1$), in modo da limitare in pratica la durata delle sinusoidi residue, che decadranno esponenzialmente a zero ($\rho^N \ll 1$). Se la quantizzazione dei coefficienti è rozza (fatta con pochi

6.6 GENERALIZZAZIONI

bit), sarà opportuno allontanare vieppiù i poli dal cerchio unitario e quindi scegliere un minor valore per ρ . Questo comporta un maggiore degrado della caratteristica di ampiezza del filtro rispetto al caso ideale. Inoltre, per mantenere la cancellazione approssimata, devono rientrare all'interno del cerchio unitario anche gli zeri; questo si ottiene in modo naturale sottraendo dopo N campioni l'impulso di ingresso moltiplicato per ρ^N . In sostanza, la funzione di trasferimento di un filtro passa banda, robusto rispetto al problema della quantizzazione dei coefficienti, deve essere:

$$H(z) = \frac{1 - z^{-N}\rho^N}{1 - 2\rho \cos(2\pi k/N)z^{-1} + \rho^2 z^{-2}}; \quad k \neq 0, N/2; \quad \rho < 1$$

dove ρ è tanto più piccolo quanto più la quantizzazione dei coefficienti è rozza. Si osservi anche che il costo di questo filtro cresce da 1 a 3 moltiplicazioni per campione.

Si osservi che l'arretramento ε dei poli per evitare ondulazioni troppo forti della caratteristica di ampiezza può essere molto maggiore della variazione della posizione del polo Δp dovuta alla quantizzazione. Per esempio, per limitare l'ondulazione a 0.01 dB si deve avere

$$\left| \frac{\Delta p}{\varepsilon} \right| < 1 - 10^{-0.01/10} = 0.002 \rightarrow \varepsilon > 500\Delta p \quad (6.11)$$

e quindi per avere arretramenti ragionevoli p.e. $\varepsilon = 0.001 \rightarrow \Delta p < 2 * 10^{-6}$ e quindi quantizzazioni dei coefficienti con almeno una ventina di bit.

6.6 Generalizzazioni

I filtri che abbiamo visto nei paragrafi precedenti sono piuttosto semplici e poco costosi; può essere interessante renderli più complicati se i benefici superano i costi: ad esempio, per avere una struttura di lobi laterali più rapidamente decrescente a zero di quanto accada con il *sinc*, è possibile disporre in cascata due filtri a campionamento in frequenza identici o magari con due diversi valori per N : N_1 , N_2 . La caratteristica di ampiezza sarà il prodotto delle due e la decrescita dei lobi con la frequenza avverrà più rapidamente; in sostanza, in questo caso, l'involuppo della sinusoidale generata da ogni oscillatore diventa un trapezio (convoluzione di due rettangoli di larghezze N_1 , N_2) invece che un rettangolo.

CAPITOLO 6 CAMPIONAMENTO IN FREQUENZA

È interessante notare che nel caso di sinusoidi reali, i due *sinc* centrati alle frequenze $\pm k/NT$ hanno una struttura di lobi laterali tale che i lobi positivi di un *sinc* si sovrappongono ai lobi negativi dell'altro comportando quindi una decrescita della funzione di trasferimento più veloce con la frequenza ($1/n^2$; dove n è il numero dei lobi del *sinc*); infatti, il troncamento della sinusoidi avviene in occasione di un passaggio per lo zero e quindi viene conservata "la continuità" della funzione ma non della sua pendenza.

Ancora, si può incrementare la scelta delle frequenze degli oscillatori considerando il caso in cui il secondo impulso ha segno positivo. Si abbia quindi:

$$H(z) = \frac{1 + z^{-N}}{N} \frac{1}{1 - 2z^{-1} \cos \phi_k + z^{-2}}$$

In questo caso gli zeri del numeratore sono $z_k = \exp(j\phi_k)$:

$$\phi_k = \frac{(2k + 1)\pi}{N}$$

Essi sono N , distribuiti a distanza angolare $2\pi/N$ attorno al cerchio unitario. Ora, lo zero a frequenza più bassa cade a π/N e quindi non vi sarà mai uno zero alla frequenza 0; ve ne sarà uno alla frequenza di Nyquist solo per N dispari. In sostanza, la sinusoidi lanciata dal primo campione effettua nell'intervallo di tempo NT un numero dispari $(2k + 1)$ di mezzi giri, cosicchè il secondo impulso rilasciato dalla memoria dopo N campioni, eguale al primo, lancia una sinusoidi di segno opposto a quella corrente e avviene la cancellazione. La struttura dei lobi laterali è ancora favorevole, visto che il troncamento avviene in corrispondenza di un passaggio per lo zero della sinusoidi.

Ancora, si può generalizzare ulteriormente il comportamento dei filtri a campionamento in frequenza al caso in cui la frequenza f generata dall'oscillatore sia proprio qualsiasi. Allora, si possono spillare dalla memoria i campioni ritardati non solo N ma anche $N - 1$ campioni. Si può alimentare l'oscillatore in modo tale da troncane le sinusoidi (eventualmente smorzate) che hanno frequenza f lanciando, dopo N campioni, delle sinusoidi sfasate dell'angolo opportuno, che può essere qualsiasi. Il polinomio a numeratore del filtro sarà quindi:

$$H(z) = \frac{B(z)}{1 - 2\rho \cos(2\pi f)z^{-1} + \rho^2 z^{-2}}$$

6.6 GENERALIZZAZIONI

$$B(z) = 1 + az^{-(N-1)} + bz^{-N} \text{ invece che } 1 \pm z^{-N}$$

Indicando la pulsazione normalizzata della sinusoide con $\phi = 2\pi fT$, risulta:

$$a = -\frac{\sin N\phi}{\sin \phi}; \quad \phi = 2\pi fT$$
$$b = \frac{\sin (N-1)\phi}{\sin \phi}$$

La frequenza centrale f del filtro passabanda è quindi arbitraria e, per costruzione, due degli N zeri di $B(z)$ cancellano i due poli dell'oscillatore; gli altri zeri di $B(z)$ non saranno più disposti regolarmente attorno al cerchio unitario, ma se ne distanzieranno in modo abbastanza casuale. Infatti, il troncamento della sinusoide **non** in corrispondenza di un passaggio per il valore 0, comporta necessariamente una decrescenza più lenta dei lobi laterali (la situazione peggiore si ha quando $|N\phi - k\pi| = \pi/2$).

Capitolo 7

Strutture canoniche e riverberanti

7.1 Strutture canoniche di filtri numerici

La realizzazione pratica dei filtri con moltiplicatori e sommatore binari con un numero limitato di cifre ha notevoli effetti sul loro comportamento, come già osservato alla fine del precedente capitolo nel caso di filtri a campionamento in frequenza. In questo capitolo illustreremo anzitutto le forme con cui si possono realizzare efficientemente dei filtri numerici. Poi vedremo che la realizzazione efficiente che minimizza il numero delle moltiplicazioni non è sempre la soluzione più valida. Un qualsiasi filtro FIR o IIR è rappresentato dall'equazione alle differenze finite:

$$y_i = \sum_{n=0}^N b_n x_{i-n} + \sum_{m=1}^M a_m y_{i-m} \quad (7.1)$$

Una realizzazione pratica è presentata nella fig. 7.1 dove il blocco che realizza gli zeri viene prima del blocco che realizza i poli; se invece si fa precedere il blocco dei poli, (fig. 7.2) si vede che il contenuto delle memorie è identico nei due blocchi; evitando la duplicazione, si arriva allo schema di figura 7.3 che realizza l'equazione (7.1) con il minimo di blocchi di memoria, di moltiplicazioni e di interconnessioni. Non è necessario fare precedere i poli; è possibile fare precedere gli zeri ritardando i segnali moltiplicati invece che moltiplicare i segnali ritardati (fig. 7.4). In sostanza, è possibile scambiare il senso delle frecce nella figura 7.3 e

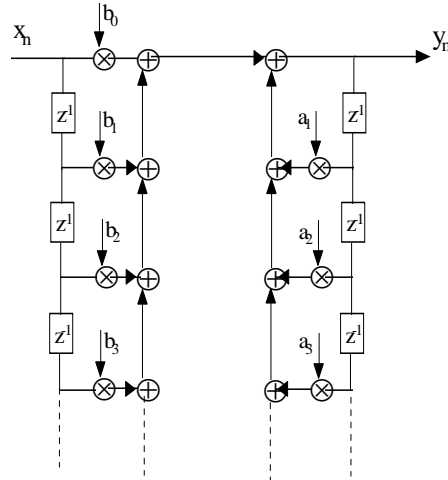


Figura 7.1: Realizzazione di un filtro FIR o IIR con zeri attuati prima dei poli

si ritrova la stessa funzione di trasferimento purchè si scambino i sensi di propagazione nelle memorie e i coefficienti a_i con i b_i .

Il sistema poi si semplifica quando il filtro FIR da realizzare ha fase zero (simmetria dei coefficienti rispetto al tempo 0) ma anche un ritardo per permetterne la realizzazione causale; esso ha quindi fase lineare e pertanto c'è simmetria dei coefficienti (fig. 7.5) rispetto ad un istante centrale. Se il numero dei campioni del filtro è dispari, vi è anche il campione centrale (fig. 7.5b). E' possibile ridurre il numero delle moltiplicazioni alla metà (o quasi, per N dispari) con la struttura di figura 7.5a in cui si calcola prima la somma in parentesi e poi il prodotto.

$$y'_i = \sum_{n=0}^{N/2-1} b_n(x_{i-n} + x_{i-N+1+n}) \quad (7.2)$$

7.2 Quantizzazione dei coefficienti

Se le moltiplicazioni che compaiono nello schema dei filtri canonici vengono realizzate con moltiplicatori binari con un numero finito di cifre la posizione dei poli e degli zeri delle funzioni di trasferimento cambia. La

7.2 QUANTIZZAZIONE DEI COEFFICIENTI

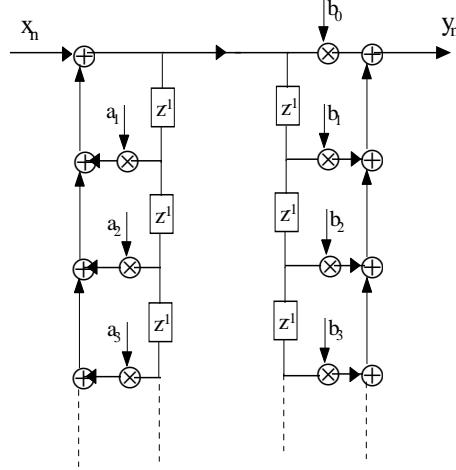


Figura 7.2: Realizzazione di un filtro FIR e IIR con poli attuati prima degli zeri

dipendenza delle radici di un polinomio qualsiasi dalla variazione dei coefficienti (sensibilità) è calcolata come segue. Supponiamo per semplicità di avere un polinomio nella variabile z e non z^{-1} :

$$A(z) = a_0 + \sum_{m=1}^N a_{N-m} z^m = a_N \prod_{n=1}^N (z - z_n); \quad a_0 = a_N \prod_{n=1}^N (-z_n) \quad (7.3)$$

e di alterarne i coefficienti a causa della quantizzazione; vogliamo calcolare le sensibilità come derivate parziali $\partial z_m / \partial a_k$ e cioè la variazione della posizione dello zero m -esimo in funzione della variazione del coefficiente k -esimo. Per fare ciò, possiamo osservare che:

$$\frac{\partial A(z)}{\partial z_h} = -a_N \prod_{n \neq h} (z - z_n); \quad \frac{\partial A(z)}{\partial a_k} = z^{N-k}; \quad (7.4)$$

Inoltre,

$$\frac{\partial A(z)}{\partial a_k} = z^{N-k} = \sum_{m=1}^N \frac{\partial A(z)}{\partial z_m} \frac{\partial z_m}{\partial a_k} = -a_N \sum_{m=1}^N \prod_{n \neq m} (z - z_n) \frac{\partial z_m}{\partial a_k} \quad (7.5)$$

CAPITOLO 7 STRUTTURE CANONICHE E RIVERBERANTI

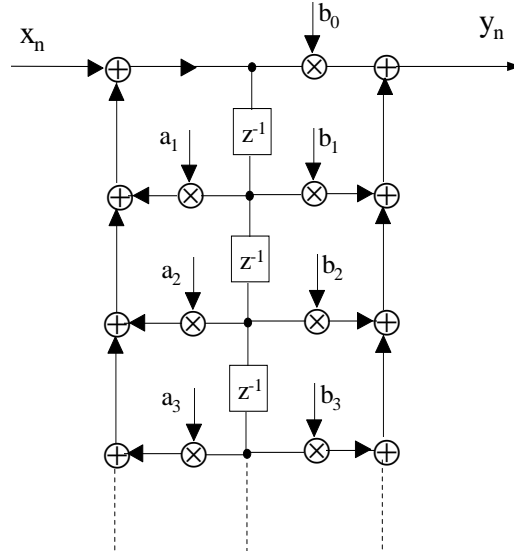


Figura 7.3: Realizzazione usuale dei filtri a struttura canonica

Ponendo $z = z_h$, tutti i termini della sommatoria si annullano tranne quello che non ha tra i fattori $(z - z_h)$,

$$z_h^{N-k} = -a_N \frac{\partial z_h}{\partial a_k} \prod_{n \neq h}^N (z_h - z_n) \quad (7.6)$$

e quindi:

$$\left| \frac{\partial z_h}{\partial a_k} \right| = \left| \frac{z_h^{N-k}}{a_N \prod_{n \neq h}^N (z_h - z_n)} \right| \quad (7.7)$$

Si vede pertanto che la sensitività e cioè il fattore di proporzionalità che lega lo spostamento di una radice alla variazione di un coefficiente è tanto maggiore quanto maggiore è il grado del polinomio e quanto più vicine sono le altre radici alla radice considerata. Quindi, per avere la minima sensitività sarà opportuno realizzare blocchi che contengano solo polinomi di grado minimo e cioè due per evitare di utilizzare aritmetica complessa. Quindi sistemi complicati verranno realizzati come cascate di forme biquadratiche e cioè funzioni di trasferimento che realizzano una coppia di poli ed una coppia di zeri. Inoltre, l'alternanza poli zeri

7.2 QUANTIZZAZIONE DEI COEFFICIENTI

Figura 7.4: Struttura di filtro canonica con ritardo dei risultati dopo le moltiplicazioni

impedisce di cadere in underflow o overflow visto che in generale i poli amplificano e gli zeri attenuano il segnale.

Peraltro, come vedremo in un prossimo paragrafo, sarà opportuno trovare delle strutture filtranti dove sia possibile garantire che i poli siano confinati all'interno del cerchio unitario, per rozza che sia la quantizzazione dei coefficienti. Ciò è possibile, con filtri fondati su analogie fisiche con le strutture riverberanti, per le quali si può garantire una stabilità intrinseca.

7.2.1 Sfasatori puri

E' importante osservare che se si ha uno sfasatore puro, caratterizzato cioè da una funzione di trasferimento del tipo:

$$A(z) = \frac{\tilde{D}(z)}{D(z)} \quad (7.8)$$

numeratore e denominatore sono eguali a parte l'ordine dei coefficienti. Anzitutto, è facile identificare delle strutture con cui, per realizzare uno

CAPITOLO 7 STRUTTURE CANONICHE E RIVERBERANTI

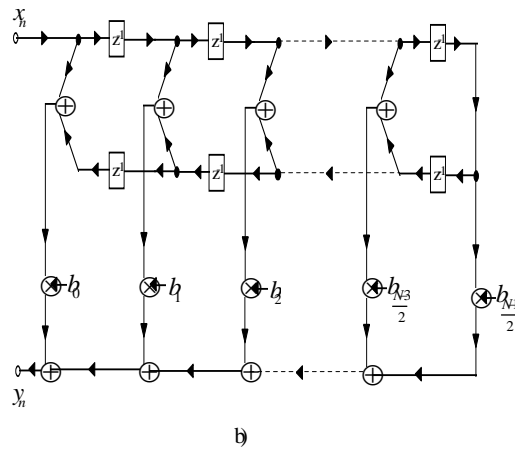
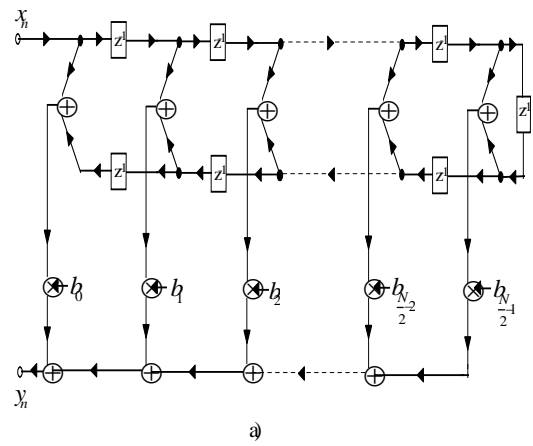


Figura 7.5: a) Realizzazione di un filtro FIR di ordine pari con fase lineare; b) Realizzazione di un filtro FIR di ordine dispari con fase lineare

7.3 STRUTTURE RIVERBERANTI: INTRODUZIONE

sfasatore puro con N zeri ed N poli, bastano N moltiplicazioni per campione, invece di $2N$. Infatti, indicando con $X(z), Y(z)$ le trasformate z delle sequenze di ingresso e di uscita dello sfasatore puro, basta scriverne l'equazione alle differenze finite (supponendo $d_0 = 1$):

$$Y(z) = \frac{\tilde{D}(z)}{D(z)} X(z)$$

$$D(z)Y(z) = \tilde{D}(z)X(z)$$

$$\begin{aligned} \sum_{m=0}^N d_m y_{n-m} &= \sum_{m=0}^N d_{N-m} x_{n-m} \rightarrow \sum_{m=0}^N d_m (y_{n-m} - x_{n-N+m}) = 0 \rightarrow \\ &\rightarrow y_n = x_{n-N} - \sum_{m=1}^N d_m (y_{n-m} - x_{n-N+m}) \end{aligned}$$

Inoltre, per qualsiasi quantizzazione dei coefficienti, lo sfasatore puro rimarrà sempre tale, in quanto la sequenza a numeratore $\tilde{D}(z)$ è comunque la ribaltata della sequenza a denominatore $D(z)$; bisogna però verificare che i poli non escano dal cerchio unitario. Se si ricorda che abbiamo a suo tempo osservato come si possano ottenere filtri passa basso e passa alto complementari in potenza da somme e differenze di due sfasatori puri, si comprende che si possano così ottenere strutture filtranti robuste rispetto alla quantizzazione dei coefficienti. Nella sezione sulle strutture riverberanti, vedremo sistemi per realizzare sfasatori puri certamente stabili.

7.3 Strutture riverberanti: introduzione

Abbiamo visto nel paragrafo sulla quantizzazione dei coefficienti come piccole variazioni dei coefficienti di filtri IIR anche stabili possano portare ad instabilità; vedremo anche la possibilità di formazione di cicli limite in filtri IIR con poli vicini al cerchio unitario. Entrambi questi fenomeni sono poco naturali; non ci attendiamo che un sistema fisico possa, con lievi alterazioni dei parametri, diventare da stabile instabile e generare energia invece di assorbirne. La questione è quindi che i filtri numerici, quali studiati finora, non corrispondono direttamente a delle strutture fisiche da cui possano mutuare pregevoli doti di stabilità.

CAPITOLO 7 STRUTTURE CANONICHE E RIVERBERANTI

Questa é appunto la corrispondenza che vogliamo studiare ora: in particolare verranno introdotti i sistemi numerici chiamati **filtri a traliccio ricorsivi**, derivanti da sistemi fisici indicati come "strutture riverberanti"; mediante questi sistemi é possibile realizzare qualsiasi filtro causale IIR: i loro parametri però avranno un diretto significato fisico ed in particolare saranno i coefficienti di riflessione per onde trasmesse su cascate di tronchi di linee di trasmissione. Per motivi energetici, i coefficienti di riflessione saranno sempre inferiori all'unitá.

Consideriamo una linea di trasmissione illimitata, uniforme; questa linea può essere elettrica (p.e. un cavo coassiale), acustica (p.e. un trombone), o meccanica (p.e. una corda tesa). In questa linea si propagano onde (di tensione e corrente nel caso della linea elettrica, di velocità e pressione nel caso della linea acustica, di velocità e tensione nel caso della corda tesa) nelle due direzioni che indicheremo con S (sinistra) e D (destra) (fig. 7.6). Supponiamo ora di avere due tronchi di linea illimitati,

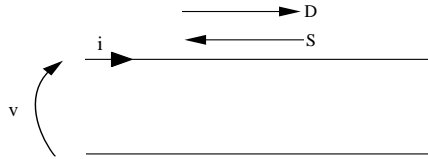


Figura 7.6: Convenzione sulla propagazione delle onde in una linea

diversi, ma entrambi uniformi e collegati insieme (fig. 7.7):

l'onda A_i che arriva alla interfaccia da uno dei due tronchi viene parzialmente riflessa ed in parte viene trasmessa nel secondo tronco; supponendo unitaria l'ampiezza dell'onda incidente A_i indicheremo con c e t le ampiezze delle onde rispettivamente riflesse e trasmesse. Per la continuità delle grandezze fisiche all'interfaccia deve essere:

$$A_i + cA_i = tA_i ; \quad \rightarrow \quad 1 + c = t \quad (7.9)$$

Se ora indichiamo con Y_i ; $i = 1, 2$ le ammettenze dei due tronchi di linea e cioè dei parametri che legano l'energia trasportata dall'onda al modulo quadrato dell'ampiezza dell'onda stessa (ammettenze vere e proprie se

7.3 STRUTTURE RIVERBERANTI: INTRODUZIONE

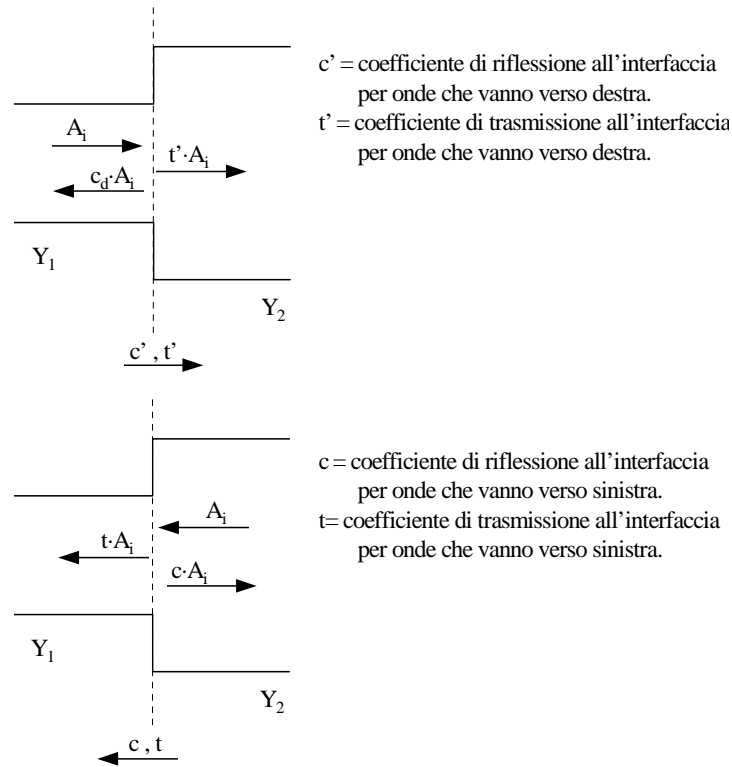


Figura 7.7: Connessione tra tronchi di linee illimitati

si tratta di tensioni, impedenze se si tratta di correnti, altri parametri meccanici in altri tipi di linee ecc.) risulta che l'energia trasportata da un'onda di ampiezza A_i è E_i :

$$E_i = Y_i \times A_i^2 \quad (7.10)$$

Il bilancio energetico impone che la somma delle energie trasportate dalle onde riflessa e trasmessa sia eguale all'energia dell'onda incidente:

$$Y_2 \times 1^2 = Y_2 \times c^2 + Y_1 \times t^2 \quad (7.11)$$

Sostituendo nella relazione 7.9 risulta:

$$c = \frac{Y_2 - Y_1}{Y_2 + Y_1} \quad (7.12)$$

Poiché, per motivi energetici, $Y_i > 0$, ne risulta che i coefficienti di riflessione sono inferiori ad 1; $|c_i| < 1$; inoltre il coefficiente di trasmissione $0 < t < 2$.

Ancora, scambiando i due mezzi, il coefficiente di riflessione cambia segno $c' = -c$; $t' = 1 + c' = 1 - c$. Non deve stupire che il coefficiente di trasmissione possa essere maggiore di 1; basta pensare alle onde del mare che, arrivando vicino alla spiaggia, diventano piú alte, senza che la loro energia sia amplificata.

7.4 Relazione con i sistemi numerici

Fino ad ora, niente appare collegare i sistemi numerici e le strutture riverberanti; pensiamo però ad una cascata di tronchi di linea, diversi tra loro, e che terminano in tronchi di linea illimitati, dalle due parti. Supponiamo anche che tutti i tronchi, sia pure diversi tra loro, abbiano un eguale ritardo di propagazione pari a $T/2$.

Lanciamo ora su questa linea composita un'onda impulsiva da uno dei due tronchi illimitati; l'impulso si propagherà nei vari tronchi e lancerà altre onde impulsive nelle due direzioni. Tuttavia, sia l'onda che finalmente esce dalla struttura, sia l'onda riflessa saranno sequenze di infiniti impulsi intervallati di T ; infatti ad ogni interfaccia sono generate nuove e nuove onde che però si combinano in qualsiasi altra interfaccia sempre dopo intervalli di tempo multipli di T , e cioè dopo un numero intero di percorsi avanti e indietro nelle varie celle.

Per rendersi conto della situazione consideriamo una cascata di tre tronchi di linea (fig. 7.8); indichiamo con Y_i ; $i = 1, 2, 3$ le tre ammettenze dei tre tronchi e c_i , $i = 1, 2$ i coefficienti di riflessione alle due interfacce, questa volta, attenzione, per onde che vadano verso destra (D); per le onde S, che vanno verso sinistra, i coefficienti di riflessione avranno segno opposto.

L'onda impulsiva di ampiezza unitaria arriva al tempo 0 alla prima interfaccia e si dipartono due onde impulsive nelle due direzioni di ampiezze c_1 e t_1 ; l'onda riflessa nel primo mezzo si propaga all'indietro indisturbata, dopo un tempo $T/2$ (ritardo che rappresenteremo con $z^{-1/2}$); l'onda trasmessa arriva alla seconda interfaccia e qui partono due nuove onde,

7.4 RELAZIONE CON I SISTEMI NUMERICI

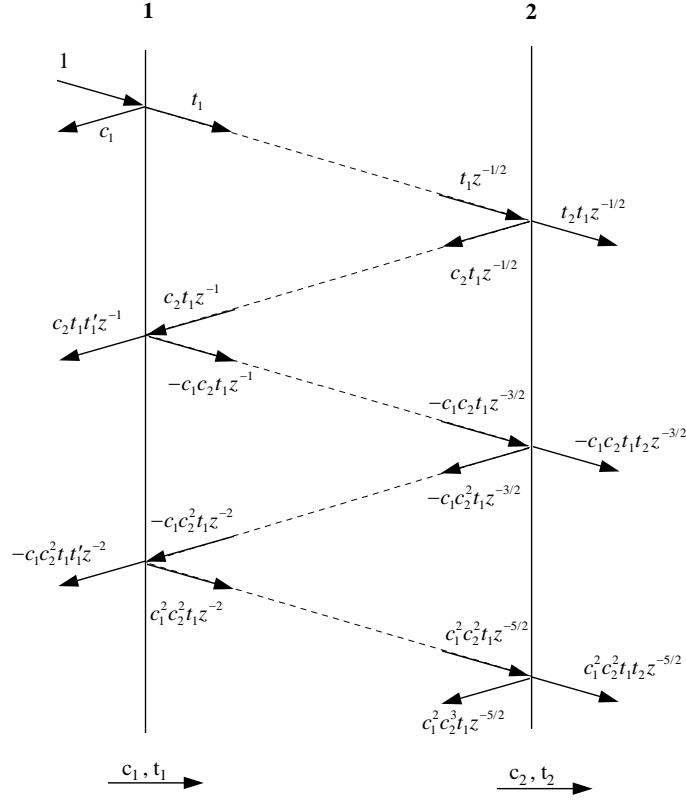


Figura 7.8: Riverberazioni in una cascata di tre tronchi di linea

di ampiezza e ritardo $c_2t_1z^{-1/2}t$ e $t_2t_1z^{-1/2}$; poi, la prima si riflette ancora sulla prima interfaccia, parte rifluendo nel primo mezzo e parte nel secondo, e così via.

L'onda trasmessa nel terzo mezzo é :

$$\begin{aligned} T(z) &= t_1t_2z^{-1/2}(1 - c_2c_1z^{-1} + (c_2c_1z^{-1})^2 - (c_2c_1z^{-1})^3 \dots = \\ &= \frac{t_1t_2z^{-1/2}}{1 + c_1c_2z^{-1}} \end{aligned} \quad (7.13)$$

L'onda riflessa nel primo mezzo é:

$$R(z) = c_1 + (t_1c_2t_1'z^{-1})(1 - c_2c_1z^{-1} + (c_2c_1z^{-1})^2 - (c_2c_1z^{-1})^3 + \dots = \quad (7.14)$$

$$= c_1 + \frac{t_1c_2t_1'z^{-1}}{1 + c_1c_2z^{-1}} = \frac{c_1 + c_2z^{-1}}{1 + c_1c_2z^{-1}} \quad (7.15)$$

La sequenza corrispondente all'onda trasmessa è un polo, a parte il ritardo inessenziale di $1/2$ campione. L'onda riflessa invece ha un polo ed uno zero, e corrisponde ad uno sfasatore puro nel caso che $|c_2| = 1$: infatti, in questo caso l'energia è tutta riflessa e quindi è ragionevole che, sfasamenti a parte, la funzione di trasferimento abbia modulo unitario in tutta la banda. Quando un coefficiente di riflessione è unitario, diremo che la struttura è totalmente disadattata ad un estremo.

7.5 Cascata di tronchi di linea

Consideriamo ora una cascata di tronchi di linea e determiniamo le relazioni tra i segnali ai due ingressi ed alle due uscite: consideriamo le due interfacce tra i tronchi $n-1; n$ ed $n, n+1$; indichiamo con c_n, t_n e $c'_n = -c_n, t'_n = 1 - c_n$ i coefficienti di riflessione e trasmissione per i segnali che si propagano verso sinistra (S) e destra (D) rispettivamente. Indichiamo poi con D_n, S_n i segnali nel tronco n , che si propagano verso destra e sinistra rispettivamente in corrispondenza dell'interfaccia verso il blocco $n-1$ e D'_n, S'_n i segnali (sempre nel tronco n) in corrispondenza dell'interfaccia verso il blocco $n+1$ (vedi figura 7.9).

La relazione tra le ampiezze delle onde ai due lati dell'interfaccia è pertanto:

$$\begin{bmatrix} S'_n(z) \\ D_{n+1}(z) \end{bmatrix} = \begin{bmatrix} t_n & c'_n \\ c_n & t'_n \end{bmatrix} \times \begin{bmatrix} S_{n+1}(z) \\ D'_n(z) \end{bmatrix} \quad (7.16)$$

Ancora possiamo scrivere che:

$$S_n(z) = z^{-1/2} S'_n(z)$$

$$D_n(z) = z^{+1/2} D'_n(z)$$

e combinando le precedenti equazioni troviamo che:

$$\begin{bmatrix} S_{n+1}(z) \\ D_{n+1}(z) \end{bmatrix} = \frac{z^{1/2}}{t_n} \begin{bmatrix} 1 & c_n z^{-1} \\ c_n & z^{-1} \end{bmatrix} \times \begin{bmatrix} S_n(z) \\ D_n(z) \end{bmatrix} \quad (7.17)$$

7.6 IL FILTRO A TRALICCIO RICORSIVO

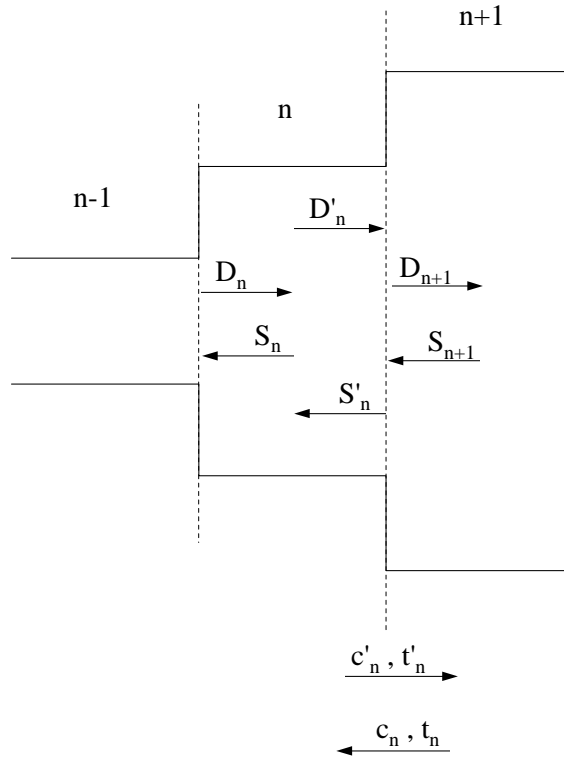


Figura 7.9: Coefficiente di riflessione e trasmissione all'($n+1$)-esima interfaccia

7.6 Il filtro a traliccio ricorsivo

Le strutture riverberanti si trasformano in sistemi numerici in modo semplice: infatti basta costruire lo schema riportato in figura 7.10 corrispondente all'equazione (7.17).

In questo schema compaiono i blocchi:

$$Q_n = \frac{z^{1/2}}{t_n}; \quad \frac{1}{Q_n} = \frac{t_n}{z^{1/2}} \quad (7.18)$$

che hanno funzioni di trasferimento reciproche e si possono elidere contemporaneamente, senza cambiare le funzioni di trasferimento esterne al blocco. Aggiungendo altre celle ed anche una cella iniziale disadatta-

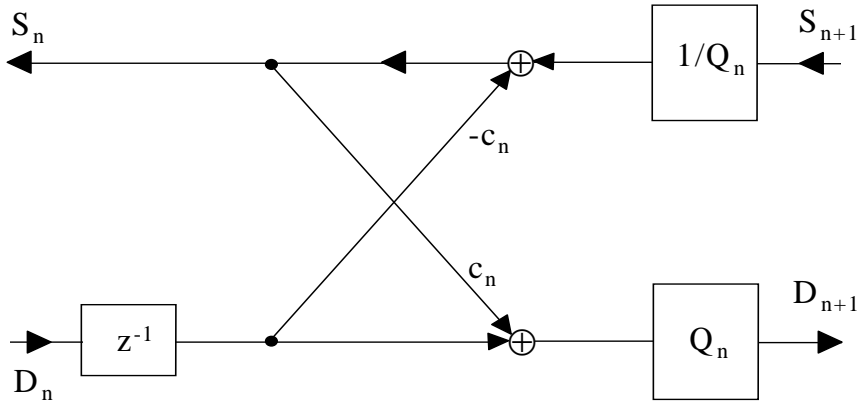


Figura 7.10: Schema della struttura numerica equivalente ad un tronco di linea

ta corrispondente al coefficiente $|c_0| = 1$, (figura 7.11) vedremo che è possibile costruire:

- un filtro tutti poli di ordine N ;
- un'uscita ritardata di N campioni;
- uno sfasatore puro caratterizzato dagli stessi poli.

Si osservi infatti che i sistemi numerici di figura 7.12 o di figura 7.11 hanno due ingressi possibili: l'uno, $I(z)$, dalla parte della cella a coefficiente di riflessione unitario (e la sua funzione di trasferimento verso l'uscita in $C(z)$ sarà caratterizzata da N poli e un ritardo di N campioni). L'altro ingresso possibile, $E(z)$ si trova dalla parte adattata (e la sua funzione di trasferimento verso $S_1(z)$ non ha ritardo tra ingresso ed uscita). Due possibili uscite possono essere prelevate in C ed in S_1 , ma potremmo prelevare il segnale anche altrove. Quindi, è disponibile tra le altre anche la funzione di trasferimento tra l'ingresso E e l'uscita C , che ha sempre gli stessi poli eguali per tutto il sistema, ma che ha anche gli zeri corrispondenti ad uno sfasatore puro, per i motivi fisici già discussi in precedenza e che saranno rivisti ancora.

Le equazioni che governano una cella del sistema numerico senza i

7.6 IL FILTRO A TRALICCIO RICORSIVO

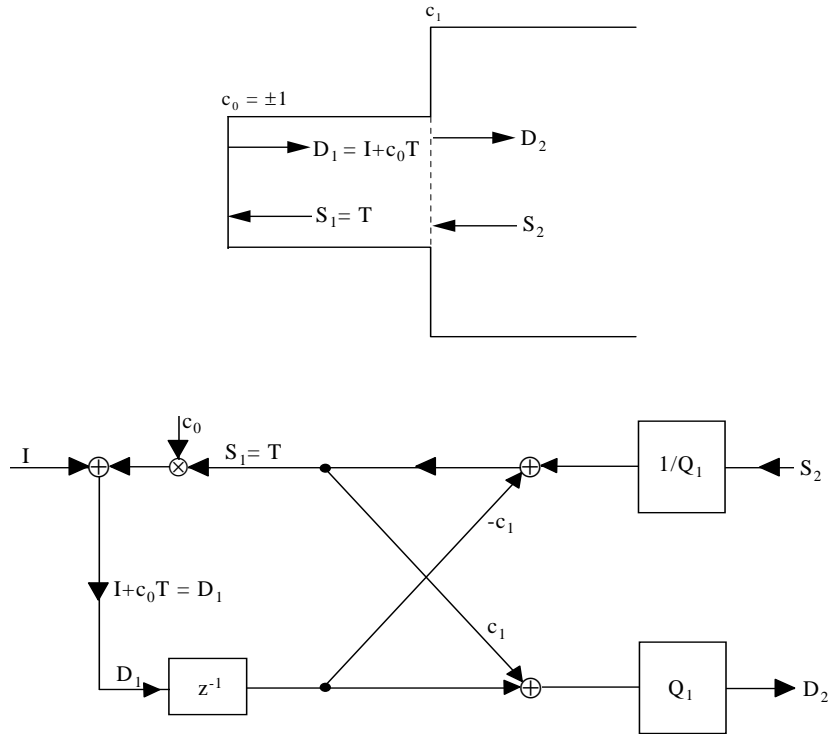


Figura 7.11: Cella iniziale corrispondente al disadattamento e suo corrispondente schema numerico

blocchi Q_n , $1/Q_n$ sono facilmente deducibili dallo schema a blocchi di figura 7.10:

$$\begin{bmatrix} S_{n+1}(z) \\ D_{n+1}(z) \end{bmatrix} = \begin{bmatrix} 1 & c_n z^{-1} \\ c_n & z^{-1} \end{bmatrix} \times \begin{bmatrix} S_n(z) \\ D_n(z) \end{bmatrix} \quad (7.19)$$

Questi filtri vengono chiamati **filtri a traliccio ricorsivi**; se sono ad N celle e quindi di ordine N , (pari a 3 in figura 7.12) sono necessarie $2N$ moltiplicazioni e $2N$ addizioni per ogni campione in ingresso. Se da una parte essi permettono di realizzare sfasatori puri e funzioni di trasfer-

imento tutti poli (con e senza ritardo), peraltro a pari ordine costano di più in termini di operazioni da compiere. Tuttavia, le moltiplicazioni sono tutte per fattori inferiori ad 1 (i coefficienti di riflessione c_n). **La natura fisica della struttura evita così sia cicli limite che instabilità** purchè, anche per valori dei c_n molto prossimi all'unità, il segnale riflesso sia sempre inferiore (in modulo) al segnale incidente.

In questo modo è possibile garantire l'incondizionata stabilità del sistema numerico, per rozza che sia la quantizzazione di coefficienti e segnali. Ancora, potremo rendere questi sistemi tempo varianti e potremo garantirne la stabilità in ogni istante, sempre che naturalmente $|c_n| \leq 1$. Ancora sarà possibile, imponendo $c_N = \pm 1$, ottenere filtri IIR (riverberazioni senza perdite di energia) che si comportano da oscillatori puri, non divergenti se $|c_n| \leq 1$, $n < N$; si è poi visto nel capitolo 5 che si possono ottenere filtri IIR di tipo Butterworth, Chebyshev o ellittici come somme e differenze di sfasatori puri.

7.7 La ricorsione di Levinson

Nel paragrafo precedente, abbiamo visto un esempio di un filtro numerico derivante dalla cascata di più celle corrispondenti a tronchi di linea, chiuso da un lato con una cella corrispondente ad un coefficiente di riflessione unitario (disadattamento da un lato). Le funzioni di trasferimento di questo sistema numerico, a seconda dell'ingresso scelto, sono tutti poli (con o senza un ritardo di N campioni) o uno sfasatore puro; esse sono determinabili in modo semplice a partire dai coefficienti di riflessione delle varie celle, attraverso una formula ricorsiva nota come **ricorsione di Levinson**. Questa formula verrà poi rincontrata nello studio della predizione lineare, a parte essere stata vista nel caso della soluzione delle equazioni di Yule Walker per la sintesi di filtri IIR.

Consideriamo quindi il caso in figura 7.12 in cui la struttura riverberante è chiusa da un lato da un coefficiente di riflessione unitario: numereremo i tronchi di linea a partire dal coefficiente unitario, che diventa quindi il coefficiente $c_0 = \pm 1$; in questo modo possiamo allungare la struttura riverberante senza dovere rinumerare i coefficienti. I sistemi meccanici corrispondenti a questo sistema numerico sono molti: nel caso acustico, questo sistema corrisponde ad uno strumento musicale a fiato. La riflessione totale su c_0 corrisponde a quella delle vibrazioni acustiche

7.7 LA RICORSIONE DI LEVINSON

sull'ancia o sulle labbra del suonatore che chiudono lo strumento (trombone, clarino, oboe, sassofono ecc.). La bocca dello strumento invece corrisponde alla cella adattata ed è svasata appunto per creare adattamento con la propagazione nell'atmosfera. La variabilità delle riverberazioni e quindi dei suoni generati è ottenuta alterando i coefficienti di riflessione (otturando i fori del flauto, allungando il trombone, ecc.). Nel caso del tratto vocale che vedremo al capitolo 15, la struttura riverberante è aperta sullo spazio libero dal lato della bocca. E' questa l'interfaccia caratterizzata da un coefficiente di riflessione pressochè unitario, e quindi, nel caso del tratto vocale numereremo i coefficienti di riflessione a partire dalle labbra, andando verso la gola. Un terzo esempio è quello delle onde sismiche che si propagano nel terreno; la superficie libera del terreno costituisce praticamente un riflettore ideale per le onde che, provenienti dal basso, sono riflesse di nuovo verso il basso.

Il segno del coefficiente di riflessione della cella disadattata iniziale, di modulo unitario, dipende dal tipo di onda che si riflette: nel caso del tratto vocale, avremo un minimo di pressione e quindi le onde di pressione vedono un coefficiente di riflessione $c_0 = -1$; avremo invece un massimo di velocità e per queste onde il coefficiente di riflessione sarà $+1$.

Passando ora al sistema numerico corrispondente realizzato con una cascata di celle senza i blocchi Q_n , $1/Q_n$. Indichiamo con $I(z)$ l'ingresso dalla parte disadattata (fig. 7.11) e con $E(z)$ e $C(z)$ rispettivamente l'ingresso e l'uscita dalla parte adattata. La funzione di trasferimento tra ingresso dalla parte disadattata $I(z)$ ed uscita dalla parte adattata $C(z)$ avrà solo poli, a parte un ritardo di N campioni ed un fattore di scala; la funzione di trasferimento tra $E(z)$ (ingresso dalla parte adattata) e $C(z)$ è invece uno sfasatore puro, perchè tutta l'energia viene riflessa all'indietro dal coefficiente di riflessione a modulo unitario della cella iniziale; infine, la funzione di trasferimento tra l'ingresso $E(z)$ e l'uscita prelevata dalla parte disadattata in $S_1(z)$ è sempre con gli stessi N poli, ma senza ritardo.

Sempre riferendosi alla figura 7.12 ed alle equazioni (7.19) supponiamo che c_2 e c_3 siano nulli. Le funzione di trasferimento $S_1(z)/S_2(z)$, $D_2(z)/S_2(z)$, $D_2(z)/I(z)$ sono rispettivamente:

$$D_1(z) = I(z) + c_0 S_1(z)$$

CAPITOLO 7 STRUTTURE CANONICHE E RIVERBERANTI

Se $S_2(z) = 0$:

$$D_2(z) = I(z) \times \frac{(1 - c_1^2)z^{-1}}{A_1(z)}$$

e se $I(z) = 0$:

$$S_1(z) = S_2(z) \times \frac{1}{A_1(z)}$$

$$D_2(z) = S_2(z) \times \frac{\tilde{A}_1(z)}{A_1(z)}$$

dove

$$A_1(z) = 1 + c_0 c_1 z^{-1}$$

e la sequenza

$$\tilde{A}_N(z) = z^{-N} A_N(z^{-1}) \quad (7.20)$$

è la sequenza $A_N(z)$, ribaltata nei tempi e ritardata di N campioni; in altre parole è la sequenza che si ottiene scambiando l'ordine dei coefficienti. Inoltre, $c_0 = 1$ è il coefficiente di riflessione corrispondente al disadattamento all'ingresso. Supponiamo ora che c_2 sia diverso da 0; si verifica (utilizzando la (7.17)) che:

$$S_3(z) = S_2(z) \times \frac{A_2(z)}{A_1(z)} = S_1(z) A_1(z) \times \frac{A_2(z)}{A_1(z)} = S_1(z) A_2(z)$$

$$D_3(z) = S_3(z) \times \frac{\tilde{A}_2(z)}{A_2(z)}$$

$$D_3(z) = I(z) \times \frac{(1 - c_1^2)(1 - c_2^2)z^{-2}}{A_2(z)}$$

a seconda che siano diversi da 0, $I(z)$ oppure $S_3(z)$. Inoltre,

$$A_2(z) = A_1(z) + c_0 c_2 z^{-1} \tilde{A}_1(z)$$

Proseguendo ricorsivamente si trova che se:

$$D_N(z) = S_N(z) \times c_0 \times \frac{\tilde{A}_{N-1}(z)}{A_{N-1}(z)} \quad (7.21)$$

$$\begin{bmatrix} S_{N+1}(z) \\ D_{N+1}(z) \end{bmatrix} = \begin{bmatrix} 1 & c_N z^{-1} \\ c_N & z^{-1} \end{bmatrix} \times \begin{bmatrix} S_N(z) \\ c_0 S_N(z) \tilde{A}_{N-1}(z) / A_{N-1}(z) \end{bmatrix}$$

7.7 LA RICORSIONE DI LEVINSON

ed eliminando $S_N(z)$, si trova:

$$D_{N+1}(z) = S_{N+1}(z) \times c_0 \frac{\tilde{A}_N(z)}{A_N(z)}$$

dove:

$$A_N(z) = A_{N-1}(z) + z^{-1}c_0c_N\tilde{A}_{N-1}(z) = 1 + \dots + z^{-N}c_0c_N \quad (7.22)$$

Inoltre:

$$S_1(z) = S_{N+1}(z) \times \frac{1}{A_N(z)}$$

$$D_{N+1}(z) = I(z) \times \frac{z^{-N} \prod_{n=1}^N (1 - c_n^2)}{A_N(z)} \quad (7.23)$$

L'equazione (7.22) è già nota come la ricorsione di Levinson e collega la funzione di trasferimento di un sistema riverberante a N celle disadattato da un lato con quello che si ottiene aggiungendo una cella con coefficiente di riflessione c_{N+1} (per le onde che vanno verso il disadattamento). Se il segno di c_0 è negativo, nella formula compare il segno negativo. I **polinomi** $A_N(z)$ **sono a minima fase**, per motivi fisici (*le trombe non esplodono*); procedendo matematicamente si ha:

$$A_{N+1}(z) = A_N(z) \left[1 + c_0c_{N+1}z^{-1} \frac{\tilde{A}_N(z)}{A_N(z)} \right] = A_N(z) \times [1 + \mu_{N+1}(z)] \quad (7.24)$$

Se $A_N(z)$ è a minima fase dimostriamo che anche $A_{N+1}(z)$ lo sarà; infatti è prodotto di una funzione a minima fase per una funzione $[1 + \mu_{N+1}(z)]$ anche essa a fase minima; infatti purchè sia:

$$|\mu_{N+1}(z)| = |c_{N+1}| \leq 1 \quad (7.25)$$

e la fase del vettore $[1 + \mu_{N+1}(z)]$, somma del vettore unitario e di un vettore più piccolo, è compresa tra $\pm \pi/2$ e quindi non supera π ed è minima.

Un modo semplice per ricordare la ricorsione di Levinson si ottiene dallo schema a blocchi in figura 7.13. Questo schema descrive solo parzialmente le strutture riverberanti, in quanto è unifilare e quindi mancano le propagazioni nei due sensi che caratterizzano quelle. Tuttavia, la figura è utile per una ricostruzione mnemonica della ricorsione: entrando dal

CAPITOLO 7 STRUTTURE CANONICHE E RIVERBERANTI

punto A_1 che rappresenta l'ingresso S_{N+1} , ed uscendo dal punto A_2 che rappresenta l'uscita D_{N+1} , la struttura riverberante viene vista come la cascata di una funzione di trasferimento tutti poli $1/A_N(z)$ con la funzione di trasferimento $\tilde{A}_N(z)$, in modo che la funzione totale D_{N+1}/S_{N+1} corrisponde allo sfasatore puro $\tilde{A}_N(z)/A_N(z)$, che corrisponde appunto alla funzione di trasferimento totale della struttura riverberante vista dall'ingresso S_{N+1} guardando verso la cella disadattata di ingresso, per uscire poi in D_{N+1} . Aggiungendo la cella $N + 1$ esima, l'uscita in A_2 e quindi D_{N+1} , si riflette all'indietro entrando in S_{N+1} , dopo essere moltiplicata per $-c_{N+1}$ e ritardata di z^{-1} . La nuova funzione di trasferimento globale, come riportato nella figura è appunto $1/A_{N+1}(z)$. Tuttavia, la rappresentazione non è perfetta, in quanto l'uscita in D_{N+1} non deriva dall'ingresso I con un funzione di trasferimento tutti poli, ma avrebbe anche un ritardo. Peraltro, l'unica funzione di trasferimento tutti poli e senza ritardo è quella tra S_{N+1} (ingresso) e S_1 (uscita) e non D_{N+1} . Quindi, non si pretenda di potere ricostruire tutti gli elementi della struttura riverberante bifilare dalla figura 7.13.

7.7.1 Aggiunta di tronchi di linea

L'equazione (7.22) è importante perchè ci permette di capire che il comportamento passa basso o passa alto del sistema è condizionato dal coefficiente di riflessione più vicino al disadattamento (*per generare un suono più basso (uuuu) o più acuto (iiiiii), si forma una cavità più grande o più piccola vicino ai denti*):

$$|A_{N+1}(z)|^2 = |A_N(z)|^2(1 + c_{N+1}^2) + 2Re[c_{N+1}z^{N+1}A_N(z)^2] \quad (7.26)$$

Aggiungendo il coefficiente di riflessione c_{N+1} , a meno del fattore di scala $1 + c_{N+1}^2$, lo spettro rimane immutato, con l'aggiunta del termine:

$$2Re[c_{N+1}z^{N+1}A_N(z)^2] \quad (7.27)$$

che, a causa del fattore $Re(z^{N+1} \dots)$ ha circa $(N + 1)/2$ ondulazioni tra la frequenza 0 e la frequenza di Nyquist f_N . Quindi, il primo coefficiente di riflessione caratterizza l'andamento globale con la frequenza; i successivi coefficienti provocano una sempre maggiore *ondulazione* dell'andamento dello spettro.

7.8 Calcolo dei coefficienti di riflessione

Siamo ora in grado di determinare i coefficienti di riflessione c_n ; $n = 1, \dots, N$ di una struttura riverberante disadattata da un lato in modo che i poli del sistema siano le radici di un polinomio $A_N(z)$: basta applicare la ricorsione di Levinson all'incontrario, partendo da N per arrivare ad 1, p.e. supponendo $c_0 = 1$. Si può quindi verificare (come proposto da Schur e Cohn) se un polinomio qualsiasi, il cui primo coefficiente è 1 come accade per il polinomio $A_N(z)$, sia a minima fase. Infatti, basta verificare che tutti i coefficienti di riflessione c_i siano in modulo minori di 1. Anzitutto si osservi (7.22) che:

$$c_N = a_{N,N} \quad (7.28)$$

e cioè che l'ultimo coefficiente del polinomio è un coefficiente di riflessione e pertanto deve essere in modulo inferiore ad 1.

Per calcolare c_{N-1} dovremo calcolare $A_{N-1}(z)$. Per fare ciò è conveniente scrivere i coefficienti dei polinomi $A_N(z)$ ed $A_{N-1}(z)$ in termini di vettori colonna. La ricorsione di Levinson (7.22) diventa:

$$\begin{bmatrix} 1 \\ a_{1,N} \\ a_{2,N} \\ \dots \\ a_{N-1,N} \\ a_{N,N} \end{bmatrix} = \begin{bmatrix} 1 \\ a_{1,N-1} \\ a_{2,N-1} \\ \dots \\ a_{N-1,N-1} \\ 0 \end{bmatrix} + c_N \begin{bmatrix} 0 \\ a_{N-1,N-1} \\ a_{N-2,N-1} \\ \dots \\ a_{1,N-1} \\ 1 \end{bmatrix} \quad (7.29)$$

Per calcolare i coefficienti di $A_{N-1}(z)$, ad esempio $a_{h,N-1}$ dovremo risolvere il sistema:

$$a_{h,N} = a_{h,N-1} + c_N a_{N-h,N-1} \quad (7.30)$$

$$a_{N-h,N} = a_{N-h,N-1} + c_N a_{h,N-1} \quad (7.31)$$

da cui risulta :

$$a_{h,N-1} = \frac{a_{h,N} - c_N a_{N-h,N}}{1 - c_N^2} \quad (7.32)$$

Troveremo così $c_{N-1} = a_{N-1,N-1}$ e poi progressivamente tutti i coefficienti di riflessione; se qualcuno di essi ha modulo maggiore di 1, il

CAPITOLO 7 STRUTTURE CANONICHE E RIVERBERANTI

polinomio di partenza $A_N(z)$ non è a minima fase (la verifica di Schur - Cohn).

7.8 CALCOLO DEI COEFFICIENTI DI RIFLESSIONE

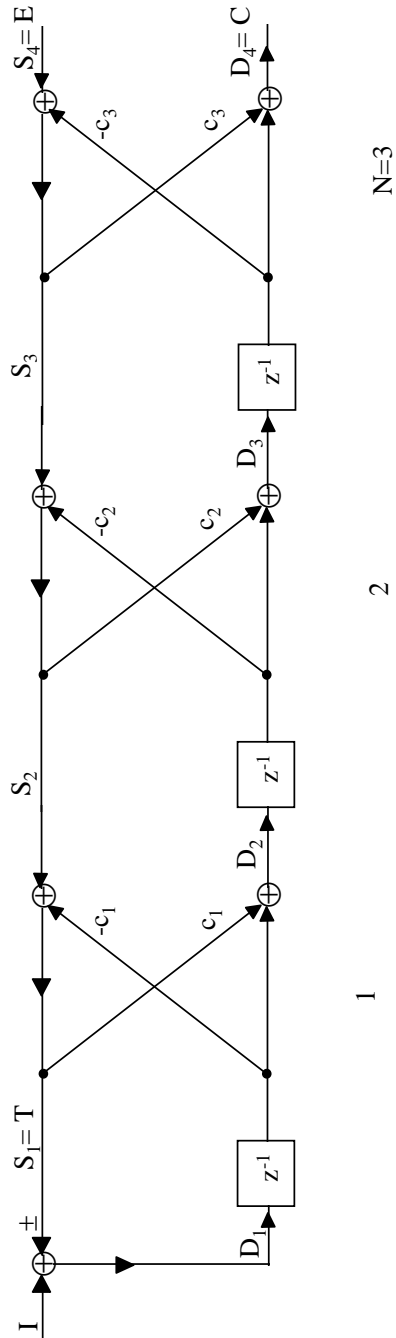


Figura 7.12: Realizzazione di filtri tutti poli con e senza ritardo e di sfasatori puri come cascata di celle elementari

CAPITOLO 7 STRUTTURE CANONICHE E RIVERBERANTI

Figura 7.13: Aggiunta di una cella ad una struttura riverberante: schema a blocchi

Capitolo 8

Filtri polifase

8.1 Filtri per decimazione o interpolazione

E' spesso utile decomporre una sequenza di N campioni in un numero di sottosequenze $M \leq N$ mediante un banco di M filtri passa banda le cui frequenze centrali siano opportunamente sfalsate e le cui bande passanti (dette sottobande) non siano necessariamente eguali tra loro. Le applicazioni di questa decomposizione sono numerose; anzitutto è possibile **trattare in parallelo** le varie sottobande, utilizzando eventualmente tecniche di filtraggio adattativo, come si vedrà nel prossimo capitolo.

La decomposizione in sottobande (subband decomposition) permette di avvicinarsi alla situazione in cui lo spettro del segnale è uniforme in ogni sottobanda. Dal banco di filtri si ottengono infatti M sottosequenze che occupano bande diverse; se queste sottosequenze sono **convertite a frequenza zero** si ottengono M sottosequenze filtrate passabasso ed occupanti $1 : M$ -esimo della banda iniziale. Se poi queste sottosequenze sono massimamente decimate e cioè sono sottocampionate con rapporto $1 : M$ il loro spettro si estenderà fino alla nuova frequenza di Nyquist, $1 : M$ -esimo della frequenza di Nyquist iniziale. Le sottosequenze massimamente decimate tendono quindi ad essere bianche, **anche se la sequenza di partenza è colorata**, poiché nel segnale originale ricoprivano un intervallo di frequenza ridotto, in cui lo spettro può essere considerato costante.

Vi è allora il vantaggio, che sarà particolarmente apprezzabile dopo lo studio dei filtri adattativi, che i campioni successivi delle sequenze

CAPITOLO 8 FILTRI POLIFASE

filtrate passa banda, convertite a frequenza zero, e infine massimamente decimate, sono poco correlati.

Tra le varie applicazioni della decomposizione in sottobande annoveriamo:

a) la realizzazione di filtri passabanda per l'estrazione di canali dallo spettro a radio frequenza;

b) i cosiddetti **transmultiplexer**, e cioè sistemi che trasformano un segnale multiplo a diversità di frequenza in un segnale multiplo a diversità di tempo o di spazio, e viceversa;

c) la divisione in sottobande del segnale musicale, utile perchè le componenti di un segnale musicale che occupano frequenze superiori a quella di una componente forte, sono percepite con minore accuratezza. Questo effetto, detto **mascheramento**, può essere sfruttato per opportune tecniche di codifica;

d) la realizzazione numerica di circuiti utili per esempio in tecniche audio numeriche, per separare i segnali da avviare agli altoparlanti di bassa e di alta frequenza (woofer - tweeter), ecc.

Un requisito molto importante nella decomposizione in sottobande è quello della perfetta ricomposizione e cioè che un sistema di decomposizione messo in cascata con un sistema di ricomposizione dia il segnale di partenza:

i) senza buchi o picchi nella funzione di trasferimento;

ii) con la completa cancellazione del disturbo di equivocazione, generato dal sottocampionamento delle sottobande.

Infatti, la massima decimazione (sottocampionamento $1 : M$ dei segnali delle M sottobande) è necessaria se si vuole che il segnale complessivo delle M sottosequenze e cioè i segnali delle M sottobande abbia ancora lo stesso numero di campioni della sequenza iniziale.

La trasformazione di Fourier discreta è un caso estremo di questa decomposizione, in cui abbiamo il massimo numero di sottosequenze ($M = N$, tante quante i campioni) e la minima durata delle sequenze decimate (1 campione complesso per sottosequenza decimata, che è poi l'ampiezza di una sinusoidale complessa). La tecnica della trasformazione di Fourier discreta permette infatti la decomposizione di una sequenza di N campioni complessi in altrettante sinusoidi complesse campionate e cioè segnali il cui spettro discreto è impulsivo e che, quando siano decimate, si riducono ad un solo campione. All'estremo opposto, c'è il caso in cui $M = 2$; la sequenza viene quindi divisa in due sottosequenze di

8.2 DECIMAZIONE E INTERPOLAZIONE

banda metà (l'una passa basso e l'altra passa alto) da sottocampionare $1 : 2$ e che quindi hanno $N/2$ campioni ognuna e così via. Una tecnica per operare è quella denominata *filtri in quadratura a specchio* (Quadrature Mirror Filters).

E' importante determinare **strutture efficienti** per questa decomposizione (i **filtri polifase**) ed inoltre studiare il problema della **equivocazione** che si formerà nella decimazione; infatti le bande passanti dei filtri devono sovrapporsi parzialmente, per evitare buchi spettrali dovuti alla non idealità della funzione di trasferimento; quindi, sembrerebbe inevitabile avere dell'equivocazione, a meno di sovracampionare e quindi non avere massima decimazione. Si **riesce invece a cancellare totalmente tale equivocazione** e pertanto a decomporre e ricomporre perfettamente il segnale. Purtroppo, tale cancellazione avviene solo se i vari canali non sono modificati prima di ricombinare il segnale in uscita. Altrimenti, perché la cancellazione poi avvenga effettivamente, il banco deve essere riprogettato in modo opportuno.

Contribuirà all'efficienza computazionale delle strutture filtranti il tenere conto del fatto che le uscite dei filtri passabanda vanno decimate; infatti, **non devono essere calcolati tutti i campioni delle uscite** ma solo quelli che non sarebbero eliminati. Ancora, costruendo il segnale multiplo corrispondente alla ricombinazione delle varie bande e quindi **interpolando** ogni sottosequenza ad un intervallo di campionamento inferiore, bisognerà evitare di fare calcoli inutili nella costruzione del segnale multiplo. E' inutile interpolare indipendentemente ogni sequenza e poi sommarle; si deve invece costruire il segnale combinato usando solo i campioni decimati.

8.2 Decimazione e interpolazione

Supponiamo di avere una sequenza reale campionata ad intervallo T , ma che occupa solo una frazione $1/M$ della banda $0 \div f_c$ e cioè una delle r sottobande $r = 0, \dots, M-1$ che hanno frequenze centrali rf_c/M . Per semplicità considereremo $r = 0$ e quindi un segnale che occupa la banda base $\pm f_c/2M$. Per sottocampionare questa sequenza, è sufficiente prenderne 1 campione ogni M (fig. 8.1). Volendo invece partire da una sequenza in una qualsiasi delle r sottobande, essa deve prima essere filtrata con un filtro passa banda $\pm f_c/2M$ (consideriamo la r -esima sottobanda con

CAPITOLO 8 FILTRI POLIFASE

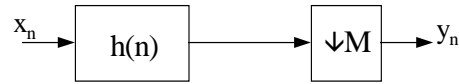


Figura 8.1: Decimazione di sequenze

$r = 0$) e quindi sottocampionata; alternativamente, si può convertire a frequenza zero tale segnale moltiplicando per la sinusoida complessa

$$\exp\left(-j\frac{2\pi}{M}r f_c n T\right) = W_M^{-nr}$$

e poi operare come per la sequenza in banda base. Se il filtro è lungo L campioni, è inefficiente fare altrettante moltiplicazioni per campione di ingresso, come in una convoluzione qualsiasi; infatti è inutile calcolare anche i campioni del segnale che saranno soppressi. In sostanza basta fare non più di $\simeq L/M$ moltiplicazioni per campione di ingresso, calcolando cioè solo i contributi ai campioni di uscita (fig. 8.2 e fig.8.3). Nello

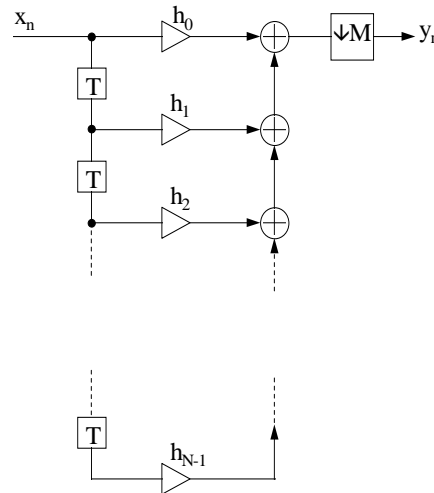


Figura 8.2: Decimazione di sequenze

schema a blocchi di figura, si traspongono le operazioni di moltiplicazione per i coefficienti $\{h_i\}$ e di decimazione per il fattore M , risparmiando così

8.2 DECIMAZIONE E INTERPOLAZIONE

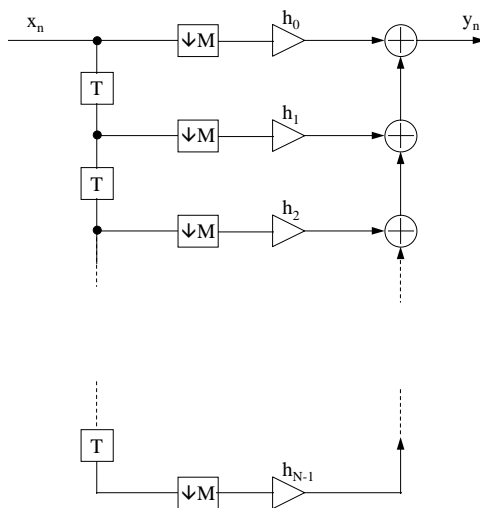


Figura 8.3: Schema efficiente per decimazione

molte operazioni. Lo schema equivalente per l'interpolazione è riportato in fig. 8.6. Si osservi anche che per avere un buon filtro di banda $1/MT$

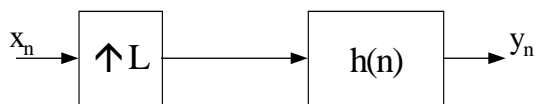


Figura 8.4: Schema per l'interpolazione

sono necessari ben più di M campioni e pertanto il numero di moltiplicazioni per campione di ingresso può essere di diverse unità. Vediamo come questo risultato modifica le figure 8.9, 8.10. Indichiamo con $\{h_i\}$ la risposta all'impulso del filtro campionata ad intervallo T ; supponiamo sia quella di un filtro FIR di $L = KM$ coefficienti, con K intero, per fare in modo che ogni campione dell'uscita sia combinazione lineare di un egual numero di campioni di ingresso. Il segnale filtrato (campionamento

CAPITOLO 8 FILTRI POLIFASE

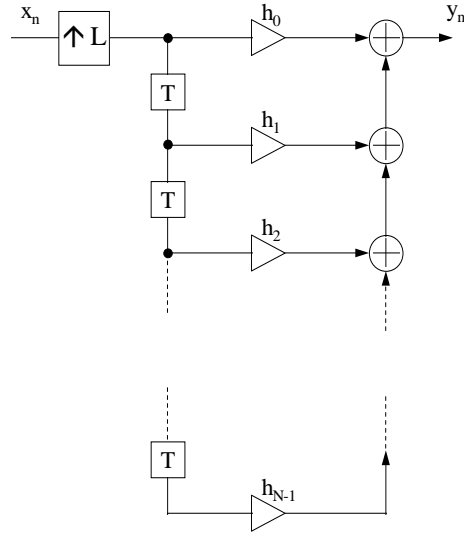


Figura 8.5: Schema di interpolazione

pieno) è:

$$y_i = \sum_{l=0}^{L-1} h_l x_{i-l} \leftrightarrow H(z)X(z) \quad (8.1)$$

Per fissare le idee, poniamo $M = 2$; supponiamo che il filtro $\{h_l\}$ sia lungo $L = 4$ campioni e rimuoviamo dall'uscita i campioni di indice dispari. Al campione uscente di indice $2i$ contribuiscono i campioni di ingresso:

$$y_{2i} = x_{2i}h_0 + x_{2i-1}h_1 + x_{2i-2}h_2 + x_{2i-3}h_3 \quad (8.2)$$

Si vede subito che i campioni di ingresso pari sono convoluti con un "sottofiltro" realizzato con i campioni di indice pari del filtro passa basso iniziale, mentre i campioni di ingresso di indice dispari sono convoluti con il "sottofiltro" realizzato con i campioni di indice dispari. In generale, sottocampionando $1 : M$, suddivideremo l'ingresso in M sottosequenze ognuna delle quali è convoluta con un sottofiltro diverso lungo $K = L/M$ campioni realizzato sottocampionando $1 : M$ il filtro passa basso iniziale, con M pettini sfalsati tra loro di un campione. La situazione è descritta

8.2 DECIMAZIONE E INTERPOLAZIONE

Figura 8.6: Interpolatore efficiente: a sinistra è presentata la soluzione inefficiente, dove ora, differentemente dalla figura 8.5, le moltiplicazioni precedono i ritardi e quindi sono affiancate al blocco di inserzione di zeri. A destra la soluzione efficiente con la trasposizione (detta identità mobile) tra blocco di interpolazione e moltiplicatori. La figura è identica alla figura 8.3, a parte il verso.

CAPITOLO 8 FILTRI POLIFASE

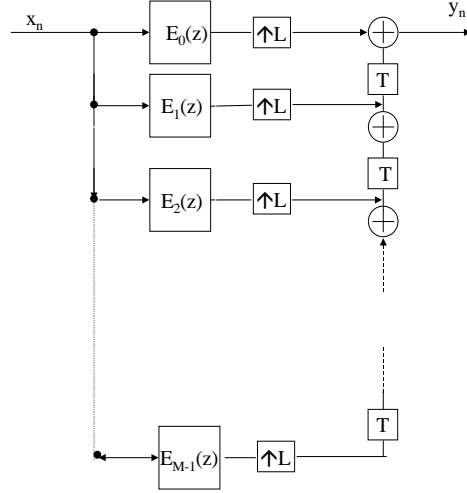


Figura 8.7: Schema efficiente per l'interpolazione con uso di filtri polifase. In sostanza, il segnale interpolato non è altro che la sequenza originaria del segnale, miscelata alla stessa sequenza, ma ritardata di $1/M, 2/M..M - 1/M$ passi di campionamento.

dalla formula che segue:

$$H(z) = \sum_{l=0}^{L-1} h_l z^{-l} = \sum_{k=0}^{K-1} \sum_{m=0}^{M-1} h_{kM+m} z^{-(kM+m)} = \quad (8.3)$$

$$\sum_{m=0}^{M-1} z^{-m} \sum_{k=0}^{K-1} h_{kM+m} (z^M)^{-k} = \sum_{m=0}^{M-1} z^{-m} E_m(z^M)$$

Indicando con:

$$E_m(z) = \sum_{k=0}^{K-1} e_{k,m} z^{-k}; \quad e_{k,m} = h_{kM+m} \quad (8.4)$$

In altre parole, con la risposta del filtro $\{h_l\}$, che come si ricorderà è lunga KM campioni, si formano M sottosequenze di K campioni, indicati come $\{e_{k,m}\}$, chiamati filtri polifase. Queste M sottosequenze sono ottenute dai campioni del filtro passa basso, decimati $1 : M$, con le M diverse

8.2 DECIMAZIONE E INTERPOLAZIONE

posizioni del pettine di sottocampionamento. In sostanza, se il filtro fosse un sinc troncato passa basso di banda $\pm f_c/2M$ si avrebbe:

$$h_n = \frac{\sin \frac{\pi n}{M}}{\frac{\pi n}{M}}; \quad -\frac{KM}{2} \leq n \leq \frac{KM}{2} \quad (8.5)$$

Questo sinc troncato ha K lobi lunghi M campioni, ed i filtri sottocampionati prelevano un campione per ogni lobo. Qualcuno dei coefficienti potrà quindi essere nullo. La struttura del filtro è indicata in fig. 8.8

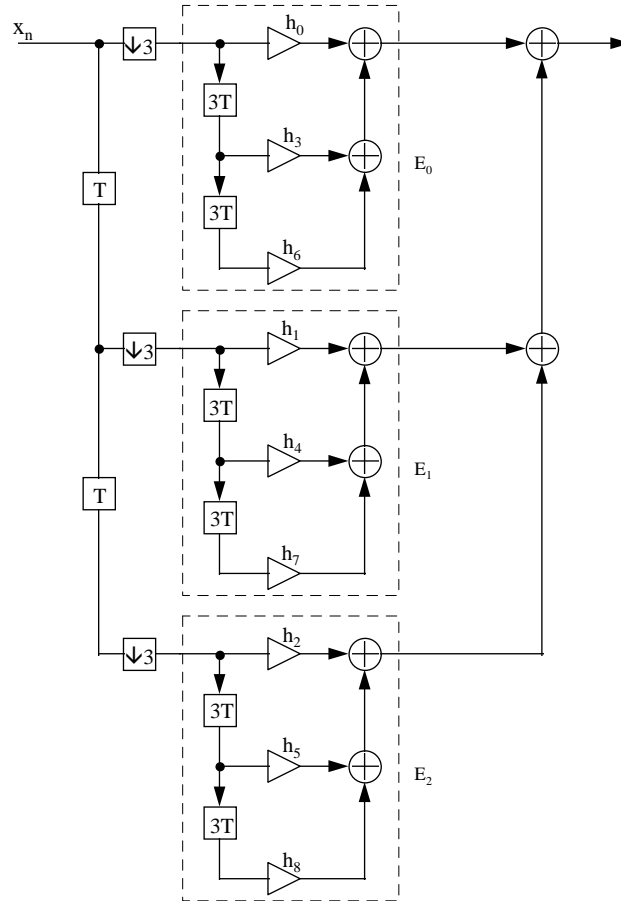


Figura 8.8: Filtro polifase per decimazione; il numero dei blocchi di memoria è pari a L

CAPITOLO 8 FILTRI POLIFASE

(decimazione $1 : 3$, filtro di lunghezza 9 campioni) e può essere vista come un commutatore a spazzola con M posizioni diverse, che alimenta con ritardo progressivo gli M sottofiltri (a intervallo di campionamento MT) indicati come $E_m(z)$. L'uscita filtrata e decimata è la somma delle uscite degli M sottofiltri. Una struttura equivalente che minimizza il numero dei blocchi di memoria è presentata in fig. 8.9 (decimazione $1 : 2$, filtro di lunghezza 6 campioni). In particolare si osserva che servono $M - 1$ blocchi di ritardo unitario (intervallo di campionamento del segnale da decimare) per realizzare il "commutatore a spazzola" di ingresso e poi ancora $K - 1$ ritardi di M campioni (se ci riferisce all'intervallo di tempo da decimare, ma in realtà unitari se intesi come l'intervallo di campionamento del segnale decimato). Similmente, nella fig. 8.10 viene riportata la struttura per il filtro per effettuare un'interpolazione $1 : 3$.

8.2 DECIMAZIONE E INTERPOLAZIONE

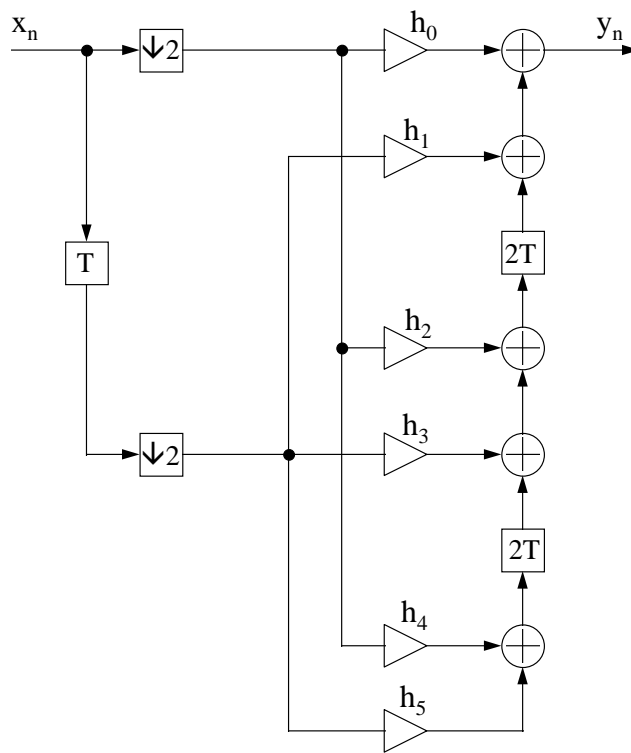


Figura 8.9: Filtro polifase per decimazione 1:2 (filtro lungo 6 campioni); ritardando i campioni dopo la moltiplicazione, il numero dei blocchi di memoria si riduce da $M - 1 + M(K - 1) = MK - 1$ come nel caso della figura precedente, a $M - 1 + K - 1$.

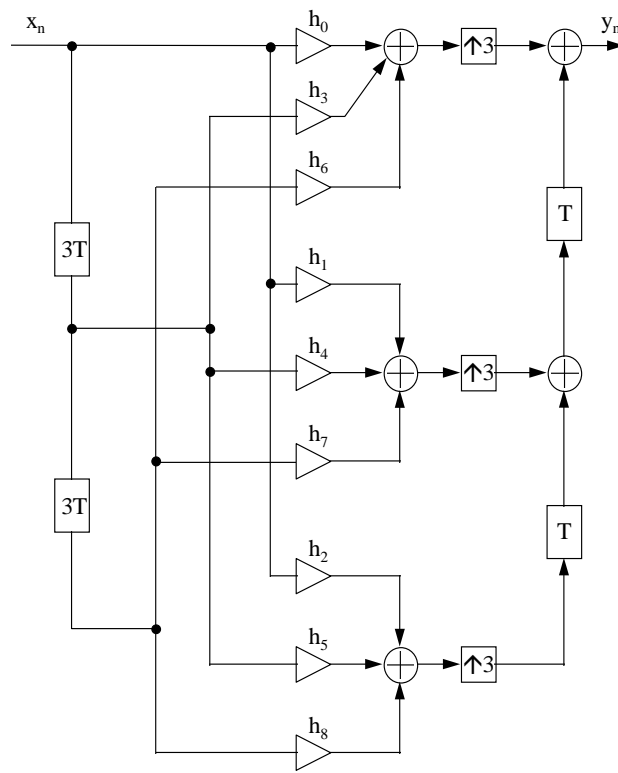


Figura 8.10: Interpolatore polifase a memoria minima; il segnale interpolando ha ritardi fino a $K - 1 = 2$ campioni; il segnale interpolato ha ritardi fino a $M - 1 = 2$ campioni (il commutatore a spazzola).

8.2.1 I filtri polifase

Vediamo ora il perchè del nome polifase, assegnato a questi filtri dall'inventore, M. Bellanger. Se il filtro originale $H(z)$ è un filtro passabasso di banda f_N/M , il sottocampionamento $1 : M$ provoca una replicazione della funzione di trasferimento fino ad ottenere in pratica, con le sovrapposizioni delle repliche, un ricoprimento con ampiezza pressochè uniforme di tutto l'intervallo di frequenze fino alla frequenza di Nyquist f_N (fig. 8.11); quindi, i filtri sottocampionati $E_m(z^M)$ sono filtri passatutto, e

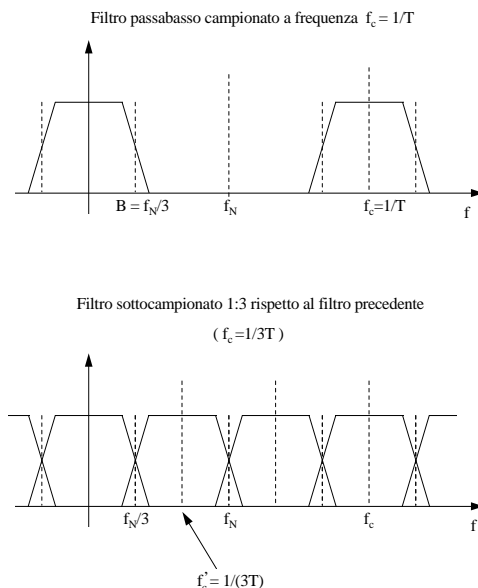


Figura 8.11: Sottocampionamento di un segnale

pertanto sfasatori puri. Ma anche i sottofiltri $E_m(z)$ hanno spettro circa costante e sono quindi filtri passatutto, nella banda $0 \div 2\pi/MT$. Si osservi che in questa ultima espressione di trasformata z compare z e non z^M ; le sequenze $E_m(z)$ corrispondono ai sottofiltri **dopo** la decimazione.

Se il filtro $H(z)$ è approssimativamente un filtro passa basso ideale, l'effetto di uno dei filtri sottocampionati $z^{-m}E_m(z^M)$ è, a parte il filtraggio passa basso seguito dalla decimazione, un ritardo della sequenza da decimare di m campioni al ritmo originale sulla sequenza decimanda;

CAPITOLO 8 FILTRI POLIFASE

peraltro, a causa della decimazione, il filtro lascia anche passare le componenti spettrali al di fuori della banda $0 \div 2\pi/MT$. Lo stesso sottofiltro, dopo l'eliminazione dei campioni nulli, e quindi dopo il sottocampionamento, ha $E_m(z)$ come trasformata z e fa ancora la stessa operazione, e quindi si comporta ancora come un ritardo temporale che è ora, al ritmo ridotto $1 : M$, di una frazione di campioni $m/M \leq 1$; in particolare, si deduce che se l'ingresso è costante, il risultato del filtraggio con uno qualsiasi dei filtri $E_m(z)$, è ancora costante (per una costante, un ritardo è irrilevante).

Allo stesso risultato si può arrivare in modo intuitivo considerando il caso dell'interpolazione $M : 1$ di una sequenza. Infatti, la sequenza interpolata può essere ottenuta generando, a partire dalla sequenza di partenza a bassa frequenza di campionamento $f_c = 1/T$, $M - 1$ sequenze ad intervallo T , ritardate progressivamente di $kT/(M - 1)$, con $0 < k < M$. La sequenza interpolata è poi ottenuta dall'interallacciamento delle M sequenze con i vari ritardi (l'azione del commutatore a spazzola). I filtri polifase sono pertanto dei ritardatori puri di $kT/(M - 1)$, come già osservato.

8.3 Cascata di interpolazione e decimazione

Si osservi che si possono mettere in cascata un interpolatore ed un decimatore, senza dovere replicare il filtraggio. Per evitare equivocazione, è obbligatorio fare prima l'interpolazione, p.e. $1 : L$ e solo dopo decimare l'uscita p.e. $M : 1$ come desiderato. In questo modo si possono realizzare a basso costo ricampionamenti per un fattore razionale L/M prima interpolando di un fattore L e poi decimando di un fattore M . Il costo in termini di moltiplicazioni per campione sarebbe, usando un filtro di durata KL , di KL moltiplicazioni per campione, alla frequenza iniziale. Poiché però serve solo 1 ogni M campioni generati, risparmiando il calcolo dei campioni da eliminare, le moltiplicazioni per campione di ingresso sono pari a KL/M . Se quindi si vuole aggiustare con precisione la frequenza di campionamento ad un valore desiderato e pertanto L, M sono numeri grandi e prossimi, il costo è circa K moltiplicazioni per campione.

Peraltro, se si usasse una tabella in cui sono trascritti i valori della funzione sinc con un piccolissimo intervallo di campionamento, sarebbe pos-

sibile calcolare direttamente i campioni di uscita ad un istante qualsiasi, appunto al costo di K moltiplicazioni per campione.

8.4 Banchi di filtri

8.4.1 La matrice di modulazione

Si può capire ora come avviene la formazione di un intero banco di M filtri; utilizzeremo la sovrapposizione degli effetti supponendo di avere un segnale il cui spettro è diverso da zero solo nella sottobanda r -esima; se esso viene filtrato con i vari filtri $z^{-m}E_m(z^M)$; $m = 0, \dots, M-1$, l'uscita totale deve essere nulla, visto che il filtro complessivo $H(z)$ è passa basso. Infatti accade proprio questo, e ce ne rendiamo conto nel seguente modo: il segnale alla frequenza centrale dell' r -esima sottobanda è la sinusoide $\exp(j\frac{2\pi}{M}nr)$; le m sottosequenze (gli m segnali provenienti dal sottocampionamento $1 : M$) sono delle costanti pari a $\exp(j\frac{2\pi}{M}mr) = W_M^{mr}$, utilizzando la notazione usata per lo studio della Trasformazione Discreta di Fourier; nel dominio sottocampionato, l'uscita totale sarà la somma delle uscite dei sottofiltri $E_m(z)$ ai cui ingressi sono applicate le costanti W_M^{mr} e cioè:

$$U = \sum_{m=0}^{M-1} W_M^{mr} E_m(z=1) = 0 \text{ se } r \neq 0$$

Se però moltiplicassimo i coefficienti dei vari sottofiltri $E_m(z)$ per il fattore W_M^{-mr} compenseremmo le rotazioni W_M^{mr} e la somma degli M sottofiltri ci darebbe una costante; in sostanza otterremmo la r -esima sottobanda, convertita a frequenza zero e decimata. Nasce allora la struttura di fig. (8.12) combinazione di un filtro polifase e di una DFT di M campioni, che viene realizzata ad ogni campione decimato. Le M uscite della FFT sono le M sottobande, già convertite a frequenza 0 e decimate massimamente.

Un altro modo di vedere la stessa cosa è il seguente: possiamo facilmente ottenere la conversione a frequenza zero della r -esima sottobanda; basta moltiplicare il segnale x_n per una sinusoide campionata posta alla frequenza centrale della sottobanda e cioè per il fattore W_M^{-nr} ma anche, tenendo conto delle periodicità, per il fattore W_M^{-mr} dove l'indice

CAPITOLO 8 FILTRI POLIFASE

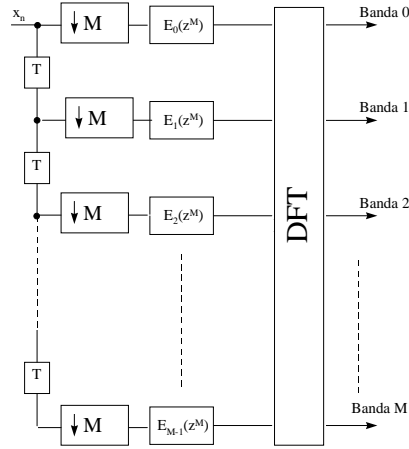


Figura 8.12: Banco di M filtri come combinazione di un filtro polifase e FFT

m varia tra 0 a $M - 1$ e poi ancora tra 0 a $M - 1$ ecc. ($W_M^{Mr} = 1$); per ottenere tutte le sottobande, pertanto, si devono moltiplicare i campioni della M -pla x_{n-m} ; $m = 0, \dots, M - 1$, per la matrice di Fourier che funge da **Matrice di Modulazione**. C'è da osservare che con questo modo di vedere, la conversione a frequenza zero dovrebbe essere fatta **prima** del filtraggio; in realtà, le due operazioni di conversione e filtraggio possono essere trasposte poichè la sottosequenza m -esima che affluisce allo stesso filtro polifase $z^{-m} E_m(z^M)$ per contribuire alla sottobanda r -esima deve venire moltiplicata sempre per lo stesso numero complesso W_M^{-mr} ; in altre parole, la posizione del commutatore a spazzola che alimenta il filtro polifase m -esimo corrisponde sempre allo stesso fattore complesso W_M^{-mr} . La moltiplicazione può quindi essere eseguita **dopo** il filtraggio e pertanto si arriva di nuovo allo schema di fig. (8.12).

Procedendo analiticamente si può partire dalla trasformata z delle M sequenze periodiche di periodo M , $C_{M,\lambda}(n)$ costituite da un campione unitario posizionato al campione λ , e da $M - 1$ zeri. Esse possono essere viste come antitrasformate di Fourier di un ritardo di λ campioni, per

8.4 BANCHI DI FILTRI

sequenze periodiche su M campioni, e cioè:

$$C_{M,\lambda}(n) = \frac{1}{M} \sum_{m=0}^{M-1} W_M^{-\lambda m} W_M^{nm}$$

La moltiplicazione di una sequenza $x(n)$ per la sequenza $C_{M,\lambda}(n)$ (e quindi la decimazione di una sequenza ritardata di λ campioni) porta alla seguente trasformata z :

$$\begin{aligned} x(n)C_{M,\lambda}(n) &\rightarrow \sum_n x(n)C_{M,\lambda}(n)z^{-n} = \frac{1}{M} \sum_n \sum_{r=0}^{M-1} W_M^{-\lambda r} W_M^{nr} x(n)z^{-n} = \\ &= \frac{1}{M} \sum_{r=0}^{M-1} W_M^{-\lambda r} \sum_n x(n)(zW_M^{-r})^{-n} = \frac{1}{M} \sum_{r=0}^{M-1} W_M^{-\lambda r} X(zW_M^{-r}) \end{aligned}$$

e quindi:

- la trasformata z di una sequenza decimata $1 : M$ è la somma di M repliche della trasformata z della sequenza originale;

- la r -esima replica della trasformata z è ruotata nel piano z di un angolo W_M^{-r} e sfasata di un angolo $-2\pi\lambda r/M$, dove λ è il ritardo della sequenza decimata. Un risultato corrispondente era stato ottenuto nel capitolo 1, equazione 1.6. Adesso possiamo arrivare alla decomposizione polifase di un banco di filtri, utilizzando l'equazione dei filtri polifase che riscriviamo per comodità:

$$H(z) = \sum_{n=0}^{L-1} h_n z^{-n} = \sum_{k=0}^{K-1} \sum_{m=0}^{M-1} h_{kM+m} z^{-(kM+m)} = \quad (8.6)$$

$$\sum_{m=0}^{M-1} z^{-m} \sum_{k=0}^{K-1} h_{kM+m} (z^M)^{-k} = \sum_{m=0}^{M-1} z^{-m} E_m(z^M)$$

$$E_m(z) = \sum_{k=0}^{K-1} e_{k,m} z^{-k}; \quad e_{k,m} = h_{kM+m} \quad (8.7)$$

Trasformare il filtro $H(z)$ da passa basso in passa r -esima sottobanda coincide con il ruotare la trasformata z del filtro di un angolo $2\pi r/M$;

CAPITOLO 8 FILTRI POLIFASE

ruotiamo quindi $H(z)$ di un angolo W_M^{-r} nel piano z ; questo equivale a sostituire z con zW_M^{-r} , oppure a moltiplicare la risposta all'impulso del filtro per la sequenza W_M^{-nr} (una sinusoidale campionata che attua la conversione di frequenza). Il risultato è un filtro passa banda che lascia passare la sottobanda r -esima

$$\begin{aligned} H_r(z) &= H(zW_M^{-r}) = \sum_{m=0}^{M-1} z^{-m} W_M^{mr} E_m(z^M W_M^{-rM}) = \\ &= \sum_{m=0}^{M-1} (zW_M^{-r})^{-m} E_m(z^M) \end{aligned}$$

dove abbiamo utilizzato il fatto che $W_M^{-rM} = 1$ e cioè che la sinusoidale modulante sottocampionata $1 : M$ diventa una costante complessa. In conclusione, il filtro che passa la banda r -esima si ottiene combinando le uscite dei filtri polifase $m = 1, \dots, M$ con i coefficienti W_M^{mr} . Per ottenere tutte le bande assieme si utilizza la moltiplicazione per la matrice della trasformazione di Fourier discreta in base M .

La realizzazione efficiente della moltiplicazione per la matrice di Fourier (e pertanto la trasformazione di Fourier) attraverso la FFT, permette poi di realizzare il banco di filtri in modo efficiente con K moltiplicazioni per campione di ingresso a cui vanno sommate $\frac{1}{2} \log_2 M$ moltiplicazioni per campione di uscita; questi, contando tutte le sottobande decimate, sono tanti quanti i campioni di ingresso. In totale pertanto vanno effettuate $K + \frac{1}{2} \log_2 M$ moltiplicazioni per campione della sequenza (decomposta o no in sottobande), dove M è il numero delle sottobande. Per la ricomposizione, vi è una struttura perfettamente simmetrica (fig. 8.13).

8.5 La ricostruzione perfetta

Fino a questo momento, non abbiamo discusso della realizzazione del filtro passabasso $H(z)$, se non per dire che deve essere lungo KM campioni. Vogliamo ora definirne le regole di progetto. I requisiti sono di due tipi diversi: da un lato si desidera un ottimo comportamento filtrante e fase lineare, cosicchè la decomposizione in sottobande sia accurata e non vi siano sovrapposizioni o buchi spettrali; ancora però si desidera che le componenti degli spettri di ingresso, equivocate da una sottobanda all'altra per la massima decimazione, siano poi ricostruite correttamente.

8.5 LA RICOSTRUZIONE PERFETTA

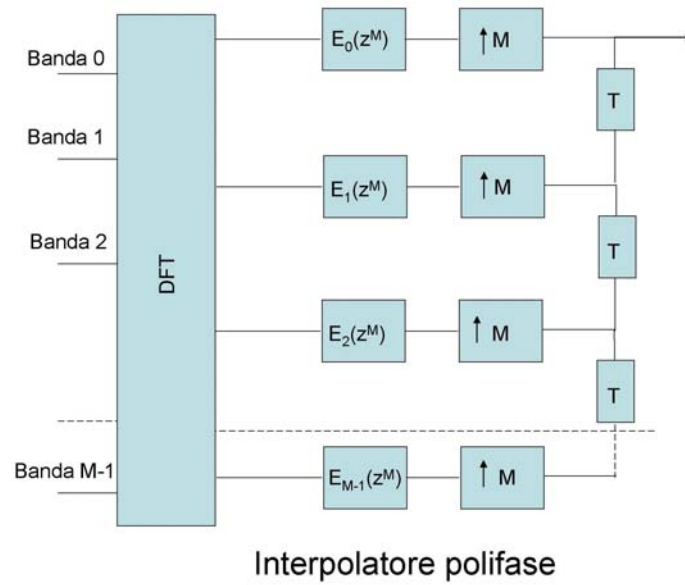


Figura 8.13: Interpolatore Polifase

CAPITOLO 8 FILTRI POLIFASE

In altre parole, pur con una buona separazione delle bande, si desidera una ricombinazione corretta della sequenza di uscita e cioè:

nessuna equivocazione pur con decimazione $1 : M$;

guadagno globale costante;

fase lineare.

E' ovvio che questi requisiti (la cosiddetta ricostruzione perfetta) sono assai pesanti per il filtro ma è possibile dimostrare che sono ottenibili, sia pure a costi crescenti via via che si approssima la situazione ideale. Si vedrà che la ricostruzione perfetta corrisponde alla compensazione del rumore di equivocazione dovuto alla decimazione $1 : M$ e lasciato passare dal filtraggio, necessariamente non ideale, che viene effettuato prima della decimazione. Si osservi però che, non appena il contenuto p.e. della banda r -esima viene modificato o magari semplicemente amplificato in modo differente da quello delle bande adiacenti, proprio per motivare in qualche modo la decisione di effettuare la decomposizione, tale compensazione scompare e quindi la ricostruzione non è più perfetta. In altre parole, non appena la decomposizione viene utilizzata per fare qualcosa di utile, la proprietà di ricostruzione perfetta scompare. Bisognerà addirittura riprogettare il sistema multibanda in queste nuove condizioni (*rifare tutto daccapo*). Se ne evince che è piuttosto difficile (anche se non impossibile) costruire filtri adattativi multibanda che garantiscono la proprietà della ricostruzione perfetta, sia pure in modo solo asintotico. Su questo punto si ritornerà nelle sezioni finali del capitolo.

8.5.1 Banche di due filtri

Vediamo anzitutto come sia possibile realizzare un banco di due filtri, $H_L(\omega)$ e $H_H(\omega)$ rispettivamente passa basso e passa alto che permetta la ricostruzione corretta anche con la massima decimazione. Questi due filtri sono chiamati filtri a specchio in quadratura (Quadrature Mirror Filter QMF) per i motivi che vedremo. Lo schema a blocchi del sistema è presentato in figura 8.14. Indichiamo rispettivamente con $X(\omega)$, $X_L(\omega)$, $X_H(\omega)$ le trasformate di Fourier del segnale e delle sue due componenti passa basso e passa alto.

$$X_L(\omega) = X(\omega)H_L(\omega); \quad X_H(\omega) = X(\omega)H_H(\omega) \quad (8.8)$$

8.5 LA RICOSTRUZIONE PERFETTA

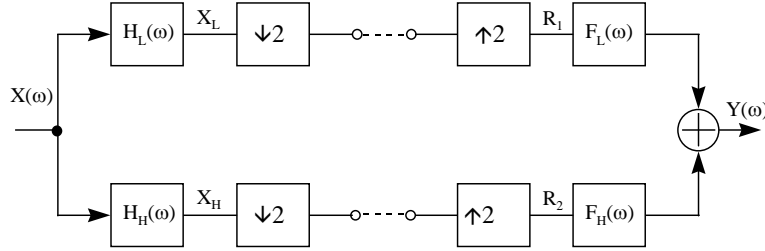


Figura 8.14: Schema a blocchi di un filtro a specchio in quadratura

I segnali dopo decimazione ed interpolazione corrispondono alla replica in frequenza di $X_L(\omega)$, $X_H(\omega)$ a distanza $\omega_0 = \frac{\omega_c}{2} = \pi/T$:

$$R_1(\omega) = X_L(\omega) + X_L(\omega - \omega_0) \quad (8.9)$$

$$R_2(\omega) = X_H(\omega) + X_H(\omega - \omega_0) \quad (8.10)$$

La replica in frequenza è infatti dovuta al sottocampionamento 1 : 2 ed all'inserzione successiva di zeri. Il risultato della ricombinazione con due filtri $F_L(\omega)$, $F_H(\omega)$ è:

$$\begin{aligned} \hat{X}(\omega) &= R_1(\omega)F_L(\omega) + R_2(\omega)F_H(\omega) = & (8.11) \\ &= X(\omega)[H_L(\omega)F_L(\omega) + H_H(\omega)F_H(\omega)] + \\ &+ X(\omega - \omega_0)[F_L(\omega)H_L(\omega - \omega_0) + F_H(\omega)H_H(\omega - \omega_0)] = \\ &X(\omega)T(\omega) + X(\omega - \omega_0)A(\omega) \end{aligned}$$

Vorremmo pertanto che:

$$T(\omega) = H_L(\omega)F_L(\omega) + H_H(\omega)F_H(\omega) = e^{-j\omega kT} \quad (8.12)$$

per avere guadagno costante e fase lineare; inoltre vorremmo:

$$A(\omega) = F_L(\omega)H_L(\omega - \omega_0) + F_H(\omega)H_H(\omega - \omega_0) = 0 \quad (8.13)$$

per non avere equivocazione. In teoria, ad ogni frequenza, abbiamo solo due equazioni per determinare quattro incognite (i valori delle funzioni di trasferimento dei quattro filtri) e pertanto abbiamo ampie possibilità di soluzione; tuttavia, si cerca una soluzione pratica.

8.5.2 I filtri in quadratura a specchio (QMF)

Imponiamo ora dei requisiti sui filtri $H_L(\omega)$, $H_H(\omega)$; anzitutto imponiamo che siano FIR, reali, lunghi N campioni e simmetrici e cioè a fase zero a parte un ritardo di $(N - 1)/2$ campioni. Per ogni zero della funzione di trasferimento, vi sarà sia lo zero coniugato perchè la risposta all'impulso è reale, sia lo zero reciproco, perchè essa è simmetrica. Imponendo quindi la simmetria coniugata rispetto all'asse reale del piano z si ha:

$$H_L(z) = H_L^*(z^*); \quad H_H(z) = H_H^*(z^*) \quad (8.14)$$

Imponiamo poi che i due filtri passa basso e passa alto siano ottenibili l'uno dall'altro con una semplice traslazione lungo l'asse delle frequenze e cioè:

$$H_H(\omega) = H_L(\omega - \omega_0); \quad \rightarrow \quad h_{H,n} = (-1)^n h_{L,n} \quad (8.15)$$

In termini di trasformata z si può scrivere che:

$$H_H(z) = H_L(ze^{-j\omega_0 T}) = H_L(-z) \quad (8.16)$$

e quindi la trasformata z di $H_H(z)$ è la riflessione speculare della trasformata z di $H_L(z)$ rispetto all'asse immaginario del piano z (in fig. 8.14 8.15) sono riportate la caratteristica di ampiezza e gli zeri di un filtro simmetrico di 32 campioni ottimizzato come indicato in [Chen, C. K., *Signal Processing*, Dicembre 1995, pp. 275]). Il nome di filtri a specchio in quadratura (QMF) nasce appunto dalla doppia simmetria rispetto alle rette del piano z : $Re(z) = 0$; $Im(z) = 0$. Per risparmiare sui coefficienti del filtro potremmo ancora imporre che ogni filtro sia **specularmente simmetrico** rispetto alla frequenza $f_c/4$;

$$H_L(\omega) + H_L(\omega_0 - \omega) = H_L(0); \quad H_H(\omega) + H_H(\omega_0 - \omega) = H_H(0); \quad (8.17)$$

Infatti, questo implica che la funzione di trasferimento sommata con la sua replica spettrale è una costante. Avevamo già visto questa condizione discutendo della ricostruibilità di costanti sovrapponendo le risposte all'impulso degli interpolatori di vario ordine; imporre che la somma della trasformata e della sue n repliche spettrali sia costante implica che nel dominio coniugato (nei tempi in questo caso) l'antitrasformata si annulli in corrispondenza del pettine di campionamento relativo alla replicazione,

8.5 LA RICOSTRUZIONE PERFETTA

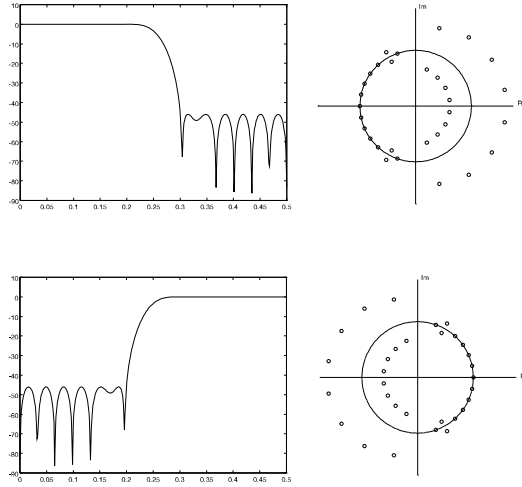


Figura 8.15: Filtri in quadratura a specchio

e quindi un campione ogni n . Se $n = 2$, circa metà dei campioni della risposta all'impulso (i campioni di posto pari tranne quello nell'origine) sono nulli con vantaggi di costo notevoli. I filtri con questa simmetria sono anche chiamati filtri di Nyquist e sono spesso utilizzati per la trasmissione numerica. Essi garantiscono infatti limitata o nulla interferenza intersimbolica, quando i simboli sono trasmessi in corrispondenza dei regolari passaggi per lo zero della risposta all'impulso.

8.5.3 Equivocazione e ricostruzione perfetta

Vediamo ora le conseguenze delle scelte fatte nel paragrafo precedente per quanto riguarda la perfetta ricostruzione; si decida inoltre che anche per i filtri $F_L(\omega)$, $F_H(\omega)$ valgano le stesse condizioni che per i filtri $H_L(\omega)$; $H_H(\omega)$, con la differenza che ora

$$F_H(\omega) = -F_L(\omega - \omega_0) \quad (8.18)$$

Scegliendo infine $F_L(\omega) = H_L(\omega)$ si ha:

$$T(\omega) = H_L^2(\omega) - H_L^2(\omega - \omega_0) = e^{-j\omega kT} \quad (8.19)$$

$$A(\omega) = H_L(\omega - \omega_0)F_L(\omega) + H_H(\omega - \omega_0)F_H(\omega) =$$

CAPITOLO 8 FILTRI POLIFASE

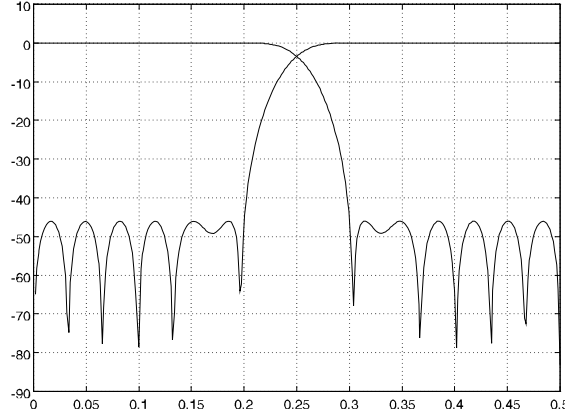


Figura 8.16: Simmetria delle funzioni di trasferimento (lo specchio).

$$= H_L(\omega)H_L(\omega - \omega_0) - H_L(\omega - \omega_0)H_L(\omega) = 0$$

Si vede cioè che, sistematicamente, la ricostruzione non contiene rumore di equivocazione. Bisognerebbe ancora imporre che la funzione di trasferimento $T(\omega)$ sia unitaria, perchè la ricostruzione sia perfetta; si dimostra però che una soluzione esatta che assicuri valore perfettamente unitario per $|T(\omega)|$, porta a filtri $H_L(\omega)$ poco selettivi; il lettore, per esempio, consideri il caso:

$$H_L(z) = 0.5(1 + z^{-1}) \rightarrow H_H(z) = .5(1 - z^{-1}); T(z) = z^{-1}; A(z) = 0$$

Nelle zone in cui $H_L(\omega) \simeq 1 \rightarrow H_H(\omega) \simeq 0$; $|T(z)|$ è quindi circa unitario come desiderato; si tratta solo di ottenere un buon comportamento nella banda di transizione e questo è possibile con ottima approssimazione. Si osserva poi che, poichè $H_L(\omega)$ è FIR, lungo N campioni, a fase lineare $\angle H_L(\omega) = -\omega T(N - 1)/2$ è necessario che sia N **pari** per evitare che $T(\omega)$ sia sistematicamente nulla alla frequenza $f_N/2$. Infatti, usando la (8.14) e la (8.15):

$$T(\omega) = [|H_L(\omega)|^2 - (-1)^{N-1}|H_L(\omega - \omega_0)|^2]e^{-j\omega T(N-1)} \quad (8.20)$$

Si deve imporre quindi che N sia pari e poi si deve anche imporre che:

$$[|H_L(\omega)|^2 + |H_L(\omega - \omega_0)|^2] = |T(\omega)| \simeq 1 \quad (8.21)$$

8.5 LA RICOSTRUZIONE PERFETTA

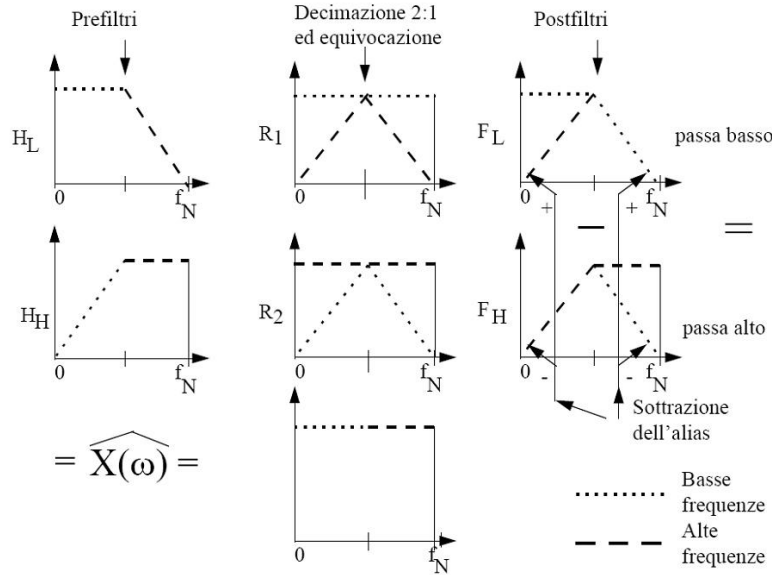


Figura 8.17: Ricostruzione perfetta (sottrazione delle componenti equivocate)

Quest'ultima condizione, insieme al requisito di una buona attenuazione in banda oscura, che implica un buon comportamento in banda passante per le relazioni tra comportamento in banda passante e in banda oscura stabilite in precedenza, si realizza appunto con un'opportuna ottimizzazione dei coefficienti dei filtri.

Si può ora capire come si ottiene l'eliminazione del rumore di equivocazione (alias) (fig. 8.17) nella ricostruzione di $\hat{X}(\omega)$, utilizzando i due filtri $F_L(\omega)$; $F_H(\omega)$; gli spettri dei due segnali decimati e riportati al ritmo di campionamento iniziale con l'aggiunta di zeri, $R_1(\omega)$, $R_2(\omega)$ sono sempre visibili in fig. 8.17. Nella ricombinazione, ad esempio l'effetto del filtro $F_H(\omega)$ su $R_2(\omega)$ sarà quello di lasciare passare ulteriore equivocazione che **cancela** l'equivocazione lasciata in precedenza passare dal filtro $H_L(\omega)$. Infatti, a causa della rotazione di fase dovuta al ritardo, se si **somma** l'uscita del canale passa alto con quella del canale passa basso, si **sottrae** allo stesso tempo il segnale equivocato che ha viaggiato attraverso il filtro ad un'altra frequenza e quindi con un diverso sfasamento

CAPITOLO 8 FILTRI POLIFASE

dovuto al ritardo. Quindi, i rumori di equivocazione lasciati passare dal filtro di ricostruzione, per le simmetrie, **sono cancellati** dai rumori di equivocazione dell'altro canale lasciati passare dai filtri di decimazione, ed alla fine risulta $A(\omega) = 0$. Si comprende meglio quanto osservato in una sezione precedente, e cioè che se poi i due canali fossero amplificati in modo diseguale, tanto per motivare in qualche modo la decomposizione, tale compensazione non avverrebbe più.

La realizzazione polifase del filtro $H_L(\omega)$

Il filtro $H_L(\omega)$ ed anche il corrispondente filtro passa alto $H_H(\omega)$ possono venire realizzati con la tecnica polifase. Avremo allora:

$$H_L(z) = H_0(z^2) + z^{-1}H_1(z^2) \quad (8.22)$$

dove $H_i(z)$; $i = 0, 1$ sono i sottofiltri ottenuti dal sottocampionamento di $H_L(z)$. Si osservi che, nel caso si faccia la scelta di un progetto di filtro simmetrico rispetto a $f_c/4$, a causa dei numerosi passaggi per lo zero della forma d'onda dovuti alla simmetria appunto rispetto alla frequenza $f_c/4$, il filtro $H_1(z^2)$ viene ad avere solo un campione diverso da zero. Di nuovo, si trovano i campioni del filtro $H_0(z^2)$ con una procedura di ottimizzazione, imponendo un buon comportamento filtrante. Ancora, l'equivocazione viene cancellata sistematicamente per l'adozione della (8.19).

8.6 Banchi di M filtri

Abbiamo visto, nel caso di un banco di due filtri, che non è solo possibile ottenere sistematicamente la rimozione del rumore di equivocazione ma anche si può avere una funzione di trasferimento unitaria. Per estendere queste proprietà a banchi di più filtri si può operare in vario modo:

- formazione di alberi binari di filtri, ripetendo l'operazione di divisione in due bande di ogni banda;
- tecniche approssimate (filtri a quasi perfetta ricostruzione);
- tecniche esatte di perfetta ricostruzione.

la prima soluzione è di facile ottenimento; la sequenza iniziale è prima divisa in due sottosequenze ma poi viene ulteriormente frazionata in

8.7 LE TRASFORMATE LAPPED: LA MDCT

quattro e così via fino a raggiungere il numero desiderato di filtri, eventualmente anche con larghezze di banda diverse tra filtro e filtro. Questa tecnica ha il difetto di mettere in cascata numerosi filtri FIR (quelli di tutte le decomposizioni e ricomposizioni) e quindi può portare a ritardi globali inaccettabili; ovviamente il ritardo temporale maggiore è quello dovuto ai filtri di banda più stretta, relativi alla massima decimazione in assoluto; ma poi si deve tenere conto anche degli altri filtri in cascata con risultati spesso inaccettabili.

Piuttosto, si utilizzano tecniche per sintetizzare banchi di filtri in cui la rimozione del rumore di equivocazione è solo approssimata; si suppone cioè che solo il rumore prodotto dalla decimazione delle bande adiacenti sia significativo, mentre i segnali che transitano per filtri più lontani in frequenza non danno contributo di rumore. In questo modo il progetto dei filtri si semplifica.

Infine, la tecnica di perfetta ricostruzione si può estendere a banchi di M filtri. Tuttavia, i filtri così sintetizzati sono più complessi, a parità di selettività. Se poi si tiene conto del fatto che a causa della quantizzazione di coefficienti e segnali si ha sempre del rumore dovuto al filtraggio, si comprende come finora siano utilizzati preferenzialmente i filtri a quasi perfetta ricostruzione. Si osservi poi che il meccanismo di ricostruzione perfetta deve essere modificato in modo da tenere conto del fatto che il rumore di equivocazione da cancellare cade in finestre adiacenti e quindi non si può utilizzare immediatamente la DFT.

8.7 Le trasformate lapped: la MDCT

La stessa metodologia che porta alla costruzione di banchi di filtri non ideali, ma parzialmente sovrapposti in frequenza, pur senza creare equivocazione delle frequenze malgrado il campionamento, può essere applicata dualmente nel dominio delle frequenze, scambiando gli assi di tempo e frequenza. Nel caso del banco di filtri, si filtra il segnale p.e. passa banda con filtri non ideali. Pertanto, si finestra (in frequenza) con delle finestre che sono parzialmente sovrapposte (nelle frequenze: figure 8.16, 8.17). Quando si sottocampiona il segnale proveniente da una sottobanda nasce del rumore di equivocazione dovuto appunto al fatto che le finestre in frequenza non sono ideali, ma sono parzialmente sovrapposte. I contenuti delle due bande di frequenza p.e. quelli al di qua ed al di là

CAPITOLO 8 FILTRI POLIFASE

dell'ascissa 0.25 nella figura 8.16 si ripiegano dopo il sottocampionamento e si sovrappongono. Abbiamo visto che, nella fase di ricostruzione del segnale, questo rumore viene cancellato da un rumore di equivocazione identico e di segno opposto.

La stessa cosa si può fare scambiando gli assi di frequenza e di tempo. L'equivocazione diventa ora nei tempi piuttosto che nelle frequenze. In sostanza, si consideri un segnale lungo l'asse dei tempi. Lo si segmenti con delle finestre temporali non rettangolari, e quindi parzialmente sovrapposte (si veda p.e. la tecnica WOSA nel capitolo del periodogramma). In questo modo i lobi laterali delle trasformate delle finestre saranno molto attenuati. Si facciano ora le trasformate p.e. simili alla Discrete Fourier del contenuto di ognuna delle finestre temporali. La periodicità in tempi creata dal campionamento in frequenza fa nascere dell'equivocazione temporale (si ricordi l'asse dei tempi circolare della DFT: il segnale nei tempi *si morde la coda*). Questa equivocazione temporale verrà appunto rimossa nella ricostruzione da un rumore di equivocazione eguale ed opposto. Si ottengono finestre parzialmente sovrapposte nei tempi (e quindi *morbide*) senza aggiungere i costi della duplicazione del trattamento dei dati. Le tecniche di questo tipo sono note come *lapped transforms*. Una di queste tecniche è nota come Modified Discrete Cosine Transform (MDCT). In essa, si usa una particolare versione della DFT tutta reale (la Discrete Cosine Transform), in cui si combinano le frequenze positive e quelle negative e il segnale invece di *mordersi la coda* a causa del campionamento in frequenza, si ripiega dalla stessa parte (Fig. 8.18, da Princen et al., Analysis/Synthesis Filter Bank Design Based on Time Domain Aliasing Cancellation, IEEE Trans. ASSP, 34, 5, Oct, 1986; Tran et al., Lapped Transform via Time-Domain Pre- and Post-Filtering, IEEE Trans. ASSP, 51, 6, June 2003). In particolare, si usano blocchi lunghi $2N$ campioni, rastremati dolcemente a zero dalle due parti, e che si susseguono ad intervalli di N campioni, e pertanto sono parzialmente sovrapposti (vedi fig. 8.18). Quando se ne calcola lo spettro, per farne la codifica efficiente utilizzando l'effetto del mascheramento, si utilizzano trasformate solo su N campioni. Quindi, si genera una sistematica equivocazione temporale, che però viene cancellata dalla combinazione con i blocchi successivi. Questa tecnica è usata nella codifica AAC, Advanced Audio Coding adottata nella Sony Play Station e in iTunes.

8.7 LE TRASFORMATE LAPPED: LA MDCT

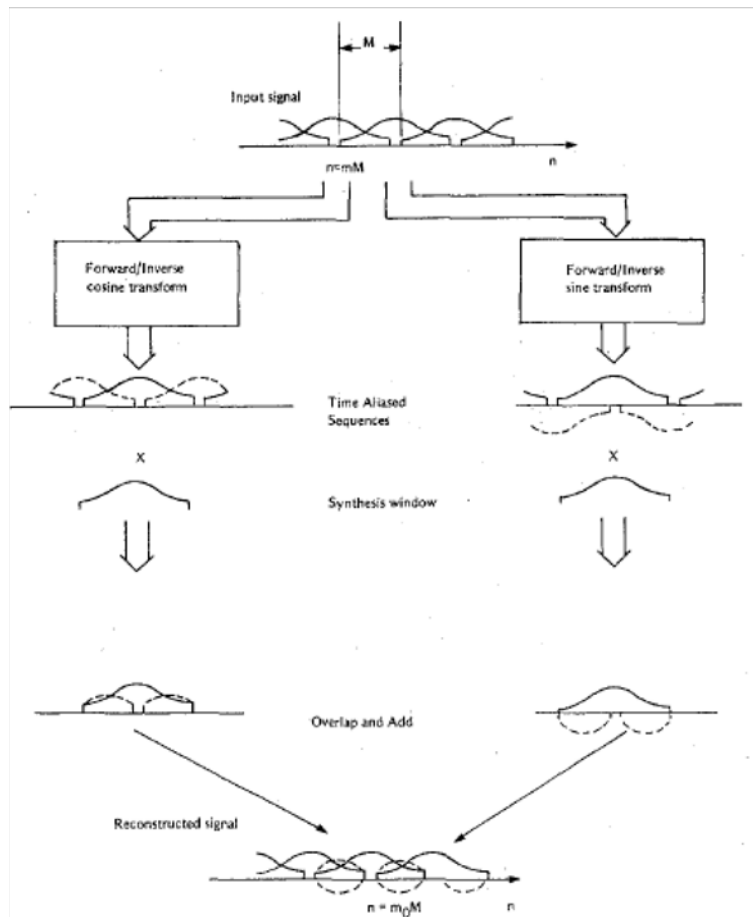


Fig. 5. The mechanism of aliasing cancellation.

Figura 8.18:

Capitolo 9

Quantizzazione dei segnali

9.1 La quantizzazione dei segnali

Il risultato di una moltiplicazione tra due numeri rispettivamente di B_c (il coefficiente del filtro) e B_s (il segnale) cifre binarie, ha un numero di cifre binarie pari a $B_c + B_s$; è ragionevole ricondurre a B_s il numero di cifre binarie che rappresenta un segnale, o dopo il filtraggio, o, come accade ben piú spesso, dopo ogni moltiplicazione di un segnale con un coefficiente. Ciò implica un arrotondamento od un troncamento della cifra binaria, in funzione dell'aritmetica usata; con arrotondamento, indichiamo l'operazione di approssimazione con il numero quantizzato (maggiore o minore) piú vicino al numero dato. Invece, si ha troncamento quando il modulo del numero approssimante, quantizzato, è sempre inferiore al modulo del numero approssimato. Ancora, può a volte essere necessario saturare l'uscita per evitare trabocchi, se si supera il numero massimo che può essere rappresentato con l'aritmetica binaria utilizzata.

In questo paragrafo vogliamo identificare le implicazioni di questa operazione di quantizzazione: sono da considerare due casi di natura molto diversa tra loro:

a) il rumore dovuto alla quantizzazione è piccolo rispetto al segnale e può essere modellato come un disturbo indipendente dal segnale stesso, anche se questa indipendenza è di per sé falsa, visto che ad eguale segnale corrisponderà sempre eguale rumore di quantizzazione.

b) il rumore non può essere considerato indipendente dal segnale. Questo accade se il segnale è periodico o costante, e quindi il rumore di

quantizzazione ha anch'esso delle periodicità. Ancora, può accadere che il rumore di quantizzazione, circolando attraverso il filtro numerico, viene amplificato, crea altro rumore ecc. fino alla formazione di un disturbo di ampiezza anche grande che può essere assai dannoso per il funzionamento del filtro stesso (ciclo limite).

9.1.1 Rumore di quantizzazione dei segnali

Se il numero di cifre binarie con cui si rappresenta un segnale viene riportato a B_s dopo ogni moltiplicazione, si aggiunge al segnale un disturbo che, se il segnale è casuale e con una dinamica grande rispetto al rumore, è sostanzialmente indipendente dal segnale stesso.

Se il segnale ha una variazione di ampiezza tra picco positivo e picco negativo pari a V_{pp} , ed il rumore di quantizzazione nasce da un arrotondamento del segnale, esso ha una distribuzione di probabilità pressoché uniforme; l'escursione di ampiezza del rumore di quantizzazione, tra picco e picco, è $a = V_{pp}/2^{B_s}$; il valore medio è nullo (fig. 9.1). La varianza è:

$$\sigma_n^2 = \frac{1}{12} \frac{V_{pp}^2}{2^{2B_s}} = \frac{a^2}{12} \quad (9.1)$$

La situazione è diversa se il disturbo di quantizzazione nasce da un troncamento; allora il valore medio è nullo solo se il segnale ha valore medio nullo, e la varianza quadruplica. È comune l'ipotesi che i campioni di questo rumore siano indipendenti tra loro; ancora, ciò non è vero in ogni caso come vedremo tra poco, ma è un'ipotesi valida nelle normali condizioni di lavoro di un filtro numerico.

Quindi, si può pensare che accanto ad ogni moltiplicatore (e sommatore) vi sia un iniettore di disturbo; il disturbo poi contribuisce all'uscita, dopo essere transitato per il filtro, attraverso una sua propria funzione di trasferimento, che è in generale diversa da quella del segnale. La potenza totale di disturbo in uscita si otterrà sommando le potenze di disturbo dovute alle varie sorgenti. Bisogna stare assai attenti all'amplificazione del segnale dovuta ai poli, che amplifica poi anche il disturbo.

Lo spettro continuo di potenza bilatero del rumore di quantizzazione iniettato, $|S_q(\omega)|^2$, è costante per l'ipotesi che i campioni successivi del rumore siano indipendenti e quindi incorrelati e quindi la loro funzione di autocorrelazione sia impulsiva (per una discussione di questi concetti, si rinvia al capitolo sull'analisi spettrale). Il suo valore è determinabile

9.1 LA QUANTIZZAZIONE DEI SEGNALI

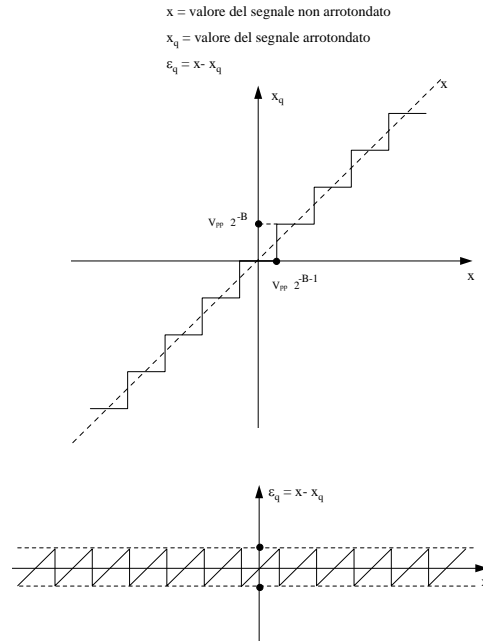


Figura 9.1: Rumore di quantizzazione: arrotondamento del segnale ed escursione di ampiezza del rumore

con il teorema di Parseval; integrato fino alla frequenza di Nyquist, deve dare il valore quadratico medio del rumore.

$$\begin{aligned}
 E[n_q^2] &= \sigma_n^2 = \int_{-\pi/T}^{\pi/T} |S_q(\omega)|^2 \frac{d\omega}{2\pi} = \frac{a^2}{12}; \\
 \Leftrightarrow |S_q(\omega)|^2 &= \frac{a^2 T}{12};
 \end{aligned} \tag{9.2}$$

Se avessimo utilizzato la definizione di trasformazione discreta di Fourier del Cap. 4 avremmo trovato un diverso fattore di scala, come risulta nel capitolo dedicato al periodogramma. La sequenza dei campioni del rumore in uscita n_u e cioè $n_{u,i}$ corrisponde al rumore generato, amplificato dalla funzione di trasferimento $H(\omega)$.

$$E[n_u^2] = \sigma_u^2 = \int_{-\pi/T}^{\pi/T} |S_q(\omega)|^2 |H(\omega)|^2 \frac{d\omega}{2\pi} \tag{9.3}$$

CAPITOLO 9 QUANTIZZAZIONE DEI SEGNALI

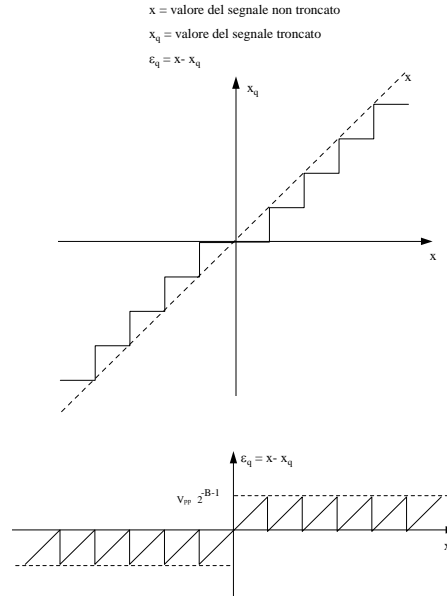


Figura 9.2: Rumore di troncamento: arrotondamento del segnale e ampiezza dell'errore

A ulteriore titolo di esempio, per un filtro del secondo ordine, ricorsivo, riporteremo la potenza di disturbo in uscita σ_o^2 dovuta al rumore di quantizzazione (fig. 9.3); le sorgenti di rumore sono due e possono essere trasferite al nodo sommatore di ingresso. La funzione di trasferimento tra ingresso ed uscita è:

$$H(z) = \frac{1}{1 - 2\rho \cos \theta z^{-1} + \rho^2 z^{-2}} \quad (9.4)$$

Per il teorema di Parseval:

$$E[n_o^2] = \sigma_o^2 = \int_{-\pi/T}^{\pi/T} |S_q(\omega)|^2 |H(\omega)|^2 \frac{d\omega}{2\pi} = 2 \times \frac{a^2 T}{12} \int_{-\pi/T}^{\pi/T} |H(\omega)|^2 \frac{d\omega}{2\pi} = \quad (9.5)$$

$$= \frac{a^2}{6} \frac{1 + \rho^2}{1 - \rho^2} \frac{1}{1 - 2\rho^2 \cos 2\theta + \rho^4} \quad (9.6)$$

9.1 LA QUANTIZZAZIONE DEI SEGNALI

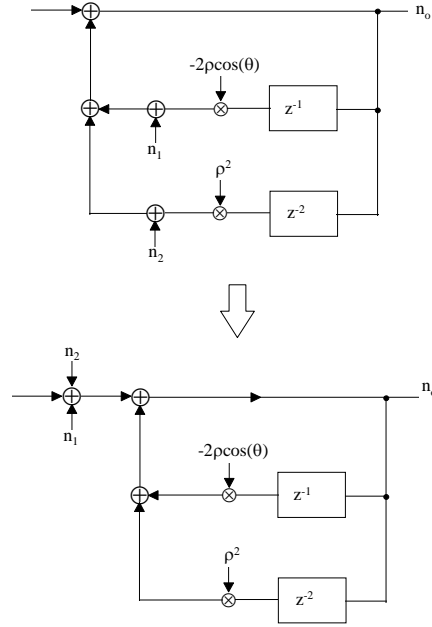


Figura 9.3: Filtro ricorsivo del secondo ordine: calcolo della potenza in uscita del rumore dovuto all'arrotondamento.

$$\text{Se } \rho = 1 - \varepsilon; \lim_{\varepsilon \rightarrow 0} \sigma_o^2 = \frac{a^2}{6} \times \frac{1}{4\varepsilon \sin^2 \theta}$$

Quest'ultima formula è facilmente ottenibile considerando i due poli complessi coniugati distanti di ε dal cerchio unitario nel piano z ; in corrispondenza delle pulsazioni di guadagno massimo

$$\pm\omega_0 = \pm\theta/T$$

la funzione di trasferimento è deducibile dai diagrammi vettoriali:

$$|H(\omega)| = \frac{1}{2 \sin \theta} \times \frac{1}{\sqrt{\varepsilon^2 + (\omega - \omega_0)^2 T^2}}$$

Allora, tenendo conto dei due contributi alle pulsazioni $\pm\omega_0$:

$$\sigma_o^2 = 2 \times \frac{a^2 T}{12} \int_{-\pi/T}^{\pi/T} |H(\omega)|^2 \frac{d\omega}{2\pi} \simeq \frac{a^2}{24\pi\varepsilon \sin^2 \theta} \int_{-\infty}^{\infty} \frac{1}{1+x^2} dx =$$

$$= \frac{a^2}{24\varepsilon \sin^2 \theta}; \quad x = \frac{\omega T}{\varepsilon}$$

In sostanza, la potenza di rumore dovuta all'arrotondamento viene amplificata di un fattore proporzionale a $1/\varepsilon$.

Si osserva infine che il rumore dovuto alla quantizzazione dei segnali in una cella biquadratica cambia con il modo di realizzazione e cioè è diverso se si usi la struttura di figura (9.3) o la struttura con i sensi di propagazione e i coefficienti scambiati. Nel paragrafo successivo troveremo una tecnica alternativa per il calcolo della potenza in uscita del rumore di quantizzazione.

9.1.2 Retroazione dell'errore di quantizzazione

Sono di interesse delle strutture di filtri (considereremo solo sistemi di primo ordine) in cui il prodotto è fatto su B bit, ma in cui i B bit meno significativi delle operazioni effettuate (il resto) vengono opportunamente retroazionati, per cancellare al meglio l'effetto della quantizzazione. Si consideri la figura 9.4. In essa vediamo un sistema di primo ordine in cui il prodotto per il coefficiente ρ viene effettuato con B bit. Il prodotto ρy_{i-1} (che conta $2B$ bit) è poi sommato sempre con la precisione di $2B$ bit sia all'ingresso x_i che con un numero $-w_{i-1}$ che verrà ora definito. Il risultato z_i della somma

$$\rho y_{i-1} + x_i - w_{i-1} = z_i = z_{i,q} + w_i$$

viene ora quantizzato con B bit $z_{i,q}$, ma ne viene calcolato anche il resto w_i , e cioè i B bit meno significativi della somma $\rho y_{i-1} + x_i - w_{i-1}$ calcolata con $2B$ bit. Questo numero, riiniettato nel sistema **cambiato di segno** (vedi figura 9.4) e con un ritardo di 1 campione, dà appunto w_{i-1} . Vediamo ora l'effetto di questa operazione, nel caso in cui $\rho = 1 - \varepsilon$ è vicino a 1. Se supponiamo che il rumore di quantizzazione w_i sia indipendente dal segnale che lo ha generato, vediamo che iniettiamo nel sistema $w_i - w_{i-1}$, invece che w_i . L'uscita dovuta a questo ingresso supponendo di:

- applicare la sovrapposizione degli effetti e di non considerare l'ingresso x_i che pure ha causato l'esistenza di w_i
- non considerare di nuovo l'effetto della quantizzazione;

9.1 LA QUANTIZZAZIONE DEI SEGNALI

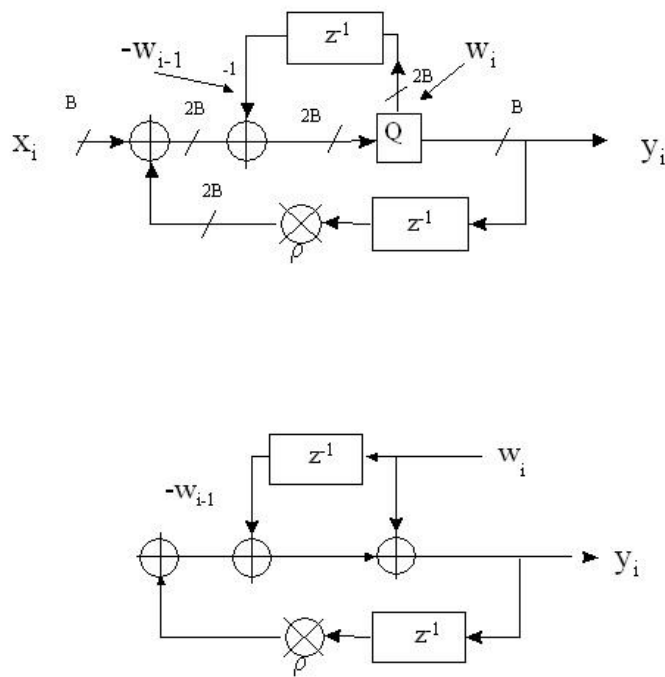


Figura 9.4: Riniezione del rumore di quantizzazione in un sistema di primo ordine. In alto: lo schema a blocchi del sistema utilizzato. In basso: lo schema a blocchi per il calcolo della funzione di trasferimento tra il rumore di quantizzazione generato w_i e l'uscita.

CAPITOLO 9 QUANTIZZAZIONE DEI SEGNALI

è rappresentata nella figura 9.4 in basso.

Si ha allora:

$$y_i = w_i - w_{i-1} + \rho y_{i-1}$$

$$Y(z) = W(z) \frac{1 - z^{-1}}{1 - \rho z^{-1}}$$

Lo zero a frequenza zero che è stato inserito **quasi** cancella l'effetto del polo; lo cancellerebbe perfettamente se si retroazionasse ρw_i invece di w_i : ma questo costerebbe un'altra moltiplicazione e gli ulteriori vantaggi (molto piccoli) non giustificano i maggiori costi. La potenza del rumore in uscita sarà ora calcolata dalla funzione di autocorrelazione del rumore in uscita. La funzione di autocorrelazione di un sistema ad uno o più poli è ottenibile dalle equazioni di Yule Walker viste nel capitolo 5; si ottiene così la formula presentata senza dimostrazione in equazione 9.6.

$$r_m = - \sum_{h=1}^N a_{h,N} r_{m-h} + \sigma_w^2 \delta(m)$$

Nel caso di un solo polo:

$$r_m = \rho r_{m-1} + \sigma_w^2 \delta(m) \rightarrow r_1 = \rho r_0 ; r_0 = \rho r_1 + \sigma_w^2$$

$$r_0 = \frac{\sigma_w^2}{1 - \rho^2}; r_m = r_0 \rho^m$$

$$R(z) = \frac{\sigma_w^2}{1 - \rho^2} [1 + \rho(z + z^{-1}) + \rho^2(z^2 + z^{-2}) + \dots]$$

Ora, per tenere conto dello zero, si deve convolvere la $R(z)$ appena trovata con la funzione di autocorrelazione dovuta allo zero:

$$R_{zero}(z) = (1 - z^{-1})(1 - z) = 2 - z - z^{-1}$$

La funzione di autocorrelazione dell'uscita dovuta al disturbo di quantizzazione è pertanto

$$R_{totale}(z) = R_{zero}(z) R(z) =$$

$$= (2 - z - z^{-1}) \frac{\sigma_w^2}{1 - \rho^2} (1 + \rho(z + z^{-1}) + \rho^2(z^2 + z^{-2}) + \dots)$$

9.1 LA QUANTIZZAZIONE DEI SEGNALI

A noi interessa il valore quadratico medio e cioè:

$$R_{totale}(z=0) = \frac{\sigma_w^2}{1-\rho^2} (2 - \rho - \rho) = \frac{2\sigma_w^2}{1+\rho} \sim \sigma_w^2 \ll \frac{\sigma_w^2}{1-\rho^2}$$

invece di

$$\frac{\sigma_w^2}{1-\rho^2}$$

come sarebbe accaduto se non avessimo reazionato il rumore di quantizzazione. In totale, un significativa riduzione (di un fattore $\sim 2\varepsilon$). Nel caso di sistemi di secondo ordine (vedi figura 9.5) si fanno entrambe le moltiplicazioni a B bit con uscita quindi a $2B$ bit; si fa la somma a $2B$ bit e si reaziona il rumore di quantizzazione con un ritardo unitario (con un fattore β_1) e con ritardo doppio (con un fattore β_2). Volendo una cancellazione perfetta si dovrebbe avere:

$$\beta_1 = -2\rho \cos \phi; \quad \beta_2 = \rho^2;$$

Di nuovo, per evitare delle moltiplicazioni, si usa per ϕ piccolo:

$$\beta_1 = -2; \quad \beta_2 = 1;$$

Per $\phi \sim \pi$, invece si deve usare $\beta_1 = 2$. La moltiplicazione della funzione di trasferimento del rumore per un doppio zero in $z = 1$ attenua fortemente l'effetto di amplificazione del rumore dovuto ai poli, se questi sono vicini.

9.1.3 Effetti di trabocco (overflow)

Si prenderanno ora in considerazione gli effetti di trabocco, dovuti al fatto che i segnali descritti con cifre binarie (in virgola fissa) corrispondono a segnali a dinamica limitata (i valori massimi negativi e positivi sono finiti). Il valore massimo del modulo del segnale non deve superare troppo spesso la condizione di trabocco. Se il segnale è a densità di probabilità gaussiana, la probabilità che il suo valore di picco sia p.e. 3 volte superiore al valore efficace è pari a 2.7×10^{-3} . Per rendere trascurabile l'ulteriore contributo di rumore dovuto al trabocco, bisognerà fare in modo che in tutti i punti del filtro il valore efficace del segnale sia almeno 12 dB al di sotto del valore di V_p , condizione di trabocco del filtro; da

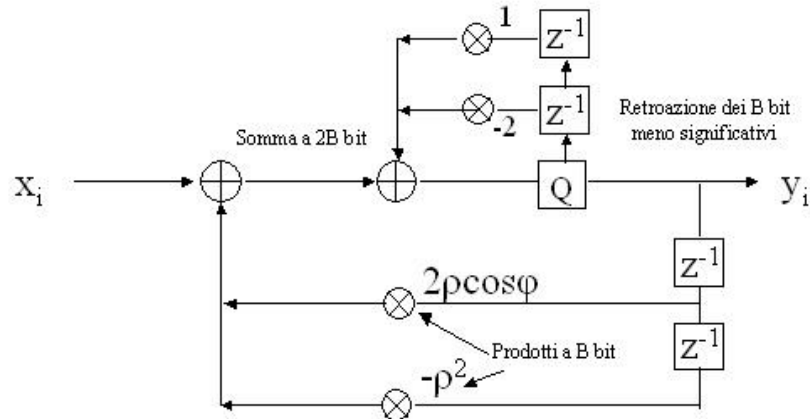


Figura 9.5: Sistema di secondo ordine e riiniezione del rumore di quantizzazione.

questa osservazione e dal numero di bit che rappresentano i segnali si può determinare il rapporto segnale - rumore di quantizzazione risultante. Se poi si ricorda che i poli in generale tendono ad amplificare il segnale e gli zeri ad attenuarlo, sarà opportuno costruire i filtri numerici alternando poli e zeri e realizzando assieme con coppie biquadratiche i poli e gli zeri posizionati nelle stesse zone del piano z , così da mantenere abbastanza uniforme il livello del segnale.

Ancora, è opportuno che il trabocco non generi segnali di segno opposto e ampiezza vicina alla massima, come potrebbe accadere con aritmetica binaria senza controlli; infatti, ne potrebbero nascere ulteriori pericolose oscillazioni (cicli limite).

9.2 Quantizzazione e DFT

Una possibile fonte di disturbo nel trattamento dei segnali con la trasformazione di Fourier discreta deriva dalla quantizzazione dei coefficienti

9.3 I CICLI LIMITE

della trasformazione e dei segnali. Indicando con:

$$X(k) = \sum_{n=0}^{N-1} x_n W^{-nk}; \quad W = e^{j2\pi/N} \quad (9.7)$$

e supponendo che sia segnali che coefficienti siano quantizzati con errori di quantizzazione pari a δ_n e w_n rispettivamente:

$$X'(k) = X(k) + \Delta_k = \sum_{n=0}^{N-1} (x_n + \delta_n)(W^{-nk} + w_n);$$

$$E[w_n] = E[\delta_n] = E[\Delta_k] = 0;$$

$$E[|\Delta_k|^2] = N(E[|w_n|^2] \times \sigma_x^2 + E[\delta_n^2] \times 1 + E[\delta_n^2] \times E[w_n^2]) \simeq 2kN\sigma_x^2 2^{-2B} \quad (9.8)$$

se:

$$k = \frac{V_{pp}^2}{12\sigma^2}; \quad E[w_n^2] = k2^{-2B} \times \sigma_W^2 = 2^{-2B}; \quad E[\delta_n^2] = k2^{-2B}\sigma_x^2 \quad (9.9)$$

e quindi si determina il rumore che inevitabilmente è sovrapposto alla trasformata.

9.3 I cicli limite

Un altro tipo di disturbo é dovuto ai cicli limite: per comprendere il fenomeno, consideriamo un filtro ad un polo, situato assai vicino al cerchio unitario, ad esempio in $z = 1 - \varepsilon$ (fig. 9.6). Supponiamo di alimentare il filtro con un impulso di ampiezza A_0 e di lasciare che il sistema evolva; tuttavia l'uscita del moltiplicatore per $z = 1 - \varepsilon$ é quantizzata con B_s cifre binarie. La moltiplicazione per $z = 1 - \varepsilon$ seguita da una quantizzazione lascia il campione immutato non appena:

$$y_n = [(1 - \varepsilon)y_{n-1}]_{\text{quantizzato}} \quad (9.10)$$

coincide con y_{n-1} e cioè quando:

$$\varepsilon y_{n-1} \leq \frac{\Delta}{2} \rightarrow y_{n-1} \leq \frac{\Delta}{2\varepsilon} = A_{lim} \quad \Delta = \frac{V_{pp}}{2^{B_s}} \quad (9.11)$$

CAPITOLO 9 QUANTIZZAZIONE DEI SEGNALI

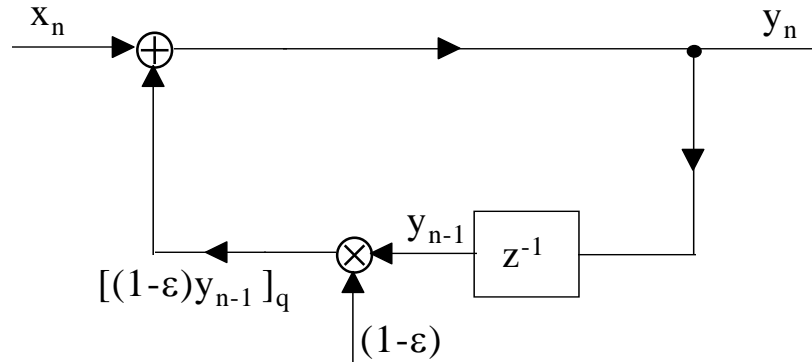


Figura 9.6: Effetti di arrotondamento: il ciclo limite

da questo momento in poi, è come se il polo si fosse portato sul cerchio unitario, e ricircola il valore A_{lim} senza decrescere più. In altre parole, il sistema è diventato "appiccicoso" (sticky) e l'uscita non tende più a zero, ma rimane al valore A_{lim} . Si osservi che se ϵ è piccolo, A_{lim} può essere anche notevole (fig. 9.7).

Una cura per questo semplice caso di ciclo limite può essere un sistema simile a quello di scuotere una ciotola per fare scivolare sul fondo una pallina rimasta attaccata alle pareti appiccicose. In questo caso, si aggiunge all'ingresso del sistema (per poi sottrarlo all'uscita) una sequenza di dither (setaccio). Il dither è una sequenza pseudo casuale e cioè una sequenza periodica, nota, e sufficientemente ampia da spezzare il ciclo limite; la presenza del dither riporta il sistema nelle condizioni per cui il segnale è casuale ed ha ampia dinamica. All'uscita del filtro, si sottrae l'effetto del disturbo pseudo casuale, facilmente calcolabile nel caso di sistema privo di rumore di quantizzazione.

Nel caso del semplice filtro di figura (fig. 9.6) è stato assai facile calcolare il ciclo limite; ciò è assai più difficile nel caso di filtri numerici complessi. Non esistono regole sicure allora, se non quella della simulazione, oppure il ricorrere a strutture intrinsecamente senza cicli limite come le strutture riverberanti che vedremo in un prossimo capitolo.

9.3 I CICLI LIMITE

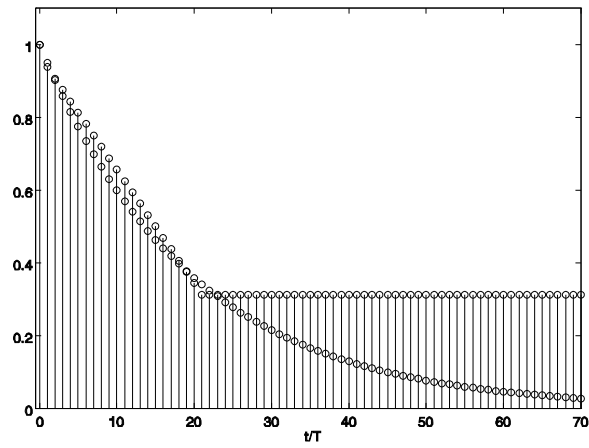


Figura 9.7: Risposta impulsiva con e senza l'effetto del ciclo limite

Capitolo 10

Le trasformate di Hilbert e *chirp* z

10.1 Trasformate di sequenze causali

Esiste una relazione tra la parte reale e la parte immaginaria della trasformata di Fourier di una sequenza causale; la stessa relazione intercorre tra modulo e fase della trasformata di una sequenza a fase minima; ancora, è la stessa relazione che vi è tra una funzione e quella che si ottiene ruotando di $\pi/2$ tutte le componenti della sua trasformata di Fourier: questa relazione è di tipo convolutivo e viene indicata come trasformazione di Hilbert, $\mathcal{H}(\cdot)$. In sostanza si tratta della convoluzione con una funzione del tipo $-1/(\pi t)$. A causa di campionamenti sia nei tempi e/o nelle frequenze, la funzione con cui si fa la convoluzione assume forme diverse, che vedremo. Tuttavia, la ripetuta applicazione della trasformazione di Hilbert porterà sempre alla funzione di partenza con il segno cambiato.

Consideriamo una sequenza **reale e causale** e cioè sia:

$$x_i = 0; \quad i < 0 \quad (10.1)$$

e si decomponga la sequenza nelle sue parti pari e dispari (fig. 10.1). Sia cioè:

$$x_{p,i} = x_i/2; \quad x_{p,-i} = x_i/2; \quad (10.2)$$

$$x_{d,i} = x_i/2; \quad x_{d,-i} = -x_i/2; \quad (10.3)$$

$$x_i = x_{p,i} + x_{d,i} \quad (10.4)$$

CAPITOLO 10 LE TRASFORMATE DI HILBERT E *CHIRP Z*

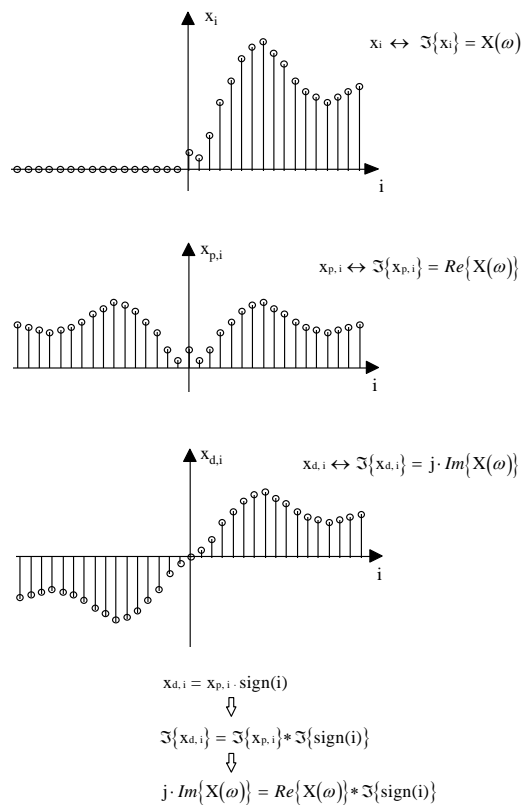


Figura 10.1: Decomposizione di una sequenza nelle componenti pari $x_{p,i}$ e dispari $x_{d,i}$

10.1 TRASFORMATE DI SEQUENZE CAUSALI

Si ha cioè che, definita la sequenza $sign(i)$ in modo che:

$$sign(i) = \frac{i}{|i|}; \quad i \neq 0; \quad sign(0) = 0; \quad (10.5)$$

allora risulta :

$$x_{d,i} = x_{p,i} sign(i) \quad (10.6)$$

e quindi, dette $R(\omega)$, $jIm(\omega)$ le parti reali ed immaginarie della trasformata di Fourier della sequenza x_i , risulta che se:

$$X(\omega) = R(\omega) + jIm(\omega) \quad (10.7)$$

la parte immaginaria è la **Trasformata di Hilbert** della parte reale:

$$jIm(\omega) = \mathcal{H}[Re(\omega)] = Re(\omega) * F(sign(i)) \quad (10.8)$$

$$F[sign(i)] = \sum_{i=1}^{\infty} (z^{-i} - z^i) = \frac{1}{1-z^{-1}} - \frac{1}{1-z} = \frac{z+1}{z-1} = -j \cot(\omega T/2) \quad (10.9)$$

(sul cerchio unitario). Pertanto:

$$Im(\omega) = \mathcal{H}\{Re(\omega)\} = Re(\omega) * [-\cot(\omega T/2)] \quad (10.10)$$

Il risultato ottenuto è facilmente ricordato se si pensa al fatto che la "derivata numerica" della sequenza $sign$ è un impulso. Allora le trasformate di Fourier della sequenza $sign$ e dell'operatore derivata devono essere l'una proporzionale all'inverso dell'altra.

Infatti richiamiamo l'espressione dell'operatore di derivazione causale, recursiva:

$$D_r(z) = \frac{2}{T} \frac{1-z^{-1}}{1+z^{-1}} = \frac{2}{T} \frac{z-1}{z+1} \quad (10.11)$$

Ora, vediamo che le due espressioni ($D_r(z)$, $F[sign(i)]$) corrispondono l'una all'inversa dell'altra. Tuttavia, non ci si stupisca se, convolvendo le due risposte in tempi NON si ottiene un impulso: in un caso (derivatore ricorsivo) il polo è visto come causale; nel caso invece della funzione $sign$ il polo è sia causale che anticausale.

Abbiamo dimostrato che le parti reale ed immaginaria della trasformata di una sequenza causale sono legate dalla convoluzione con la funzione periodica $-\cot(\omega T/2)$, che è la trasformata di Fourier della sequenza $sign(i)$; $-\infty < i < \infty$; essa rappresenta una versione discreta del cosiddetto nocciolo di Hilbert (fig. 10.2).

CAPITOLO 10 LE TRASFORMATE DI HILBERT E *CHIRP Z*

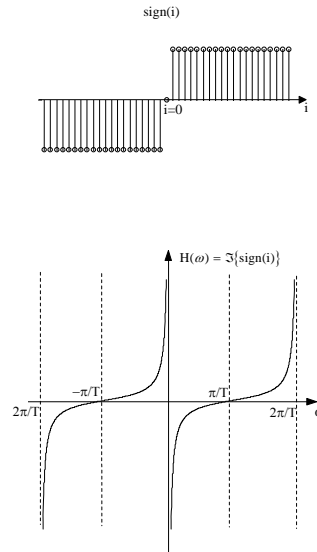


Figura 10.2: Nocciolo di Hilbert illimitato nel dominio coniugato

La trasformazione descritta dalle precedenti equazioni è perciò chiamata **trasformazione di Hilbert discreta**, e corrisponde al caso in cui il segnale è campionato ed illimitato in un dominio (in questo caso quello dei tempi) e periodico e continuo nell'altro (in questo caso quello delle frequenze).

10.2 Sfasatori puri di $\pi/2$

Come osservato, esistono varie versioni del nocciolo di Hilbert, a seconda che il segnale sia continuo nei due domini, campionato e illimitato nell'uno e periodico e continuo nell'altro, o infine campionato e periodico in entrambi.

Consideriamo il caso che il campionamento sia nel dominio coniugato a quello dove si calcola la funzione $\text{sign}(i)$ che quindi viene generalizzata diventando una funzione continua e periodica, invece che discreta ed illimitata.

Il problema di riferimento in questo caso è il seguente: data una sequenza reale x_i se ne consideri la trasformata di Fourier $X(\omega)$ (si ricordi

10.2 SFASATORI PURI DI $\pi/2$

che questa è una funzione periodica!) e si ruotino di $\pi/2$ tutte le sue componenti di Fourier; in altre parole, si moltiplichino la trasformata per la funzione $j\text{sign}(\omega)$ periodicizzata sull'intervallo $-\pi/T \div \pi/T$. L'antitrasformata è una nuova sequenza $\xi_i = \mathcal{H}\{x_i\}$, che anche viene identificata come la trasformata di Hilbert della sequenza x_i . La situazione però è differente da quella vista prima perchè la sequenza equivalente al nocciolo di Hilbert ha trasformata periodica e continua (fig. 10.3).

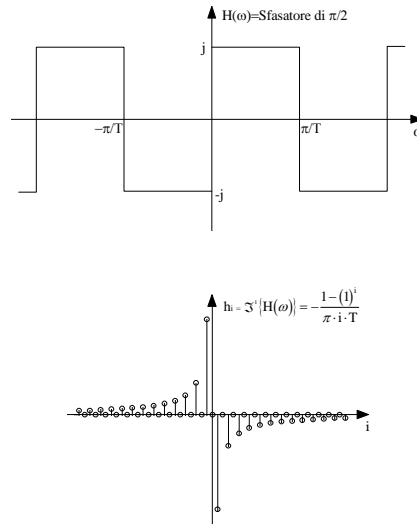


Figura 10.3: Nocciolo di Hilbert periodicizzato nel dominio coniugato

Si tratta quindi di calcolare l'antitrasformata della funzione periodicizzata $j\text{sign}(\omega)$; $|\omega| \leq \pi/T$. Per fare ciò senza fatica, basta calcolare prima il segnale continuo antitrasformata della funzione $j\text{sign}(\omega)$ annullata fuori del periodo principale e poi campionare la funzione del tempo così ottenuta. Il risultato è:

$$h(t) = -\frac{1 - \cos \pi t/T}{\pi t/T}; \quad h(iT) = h_i = -\frac{1 - (-1)^i}{\pi i} \quad -\infty < i < \infty \quad (10.12)$$

Esso assomiglia all'andamento della cotangente con il segno cambiato, ma con differenze notevoli se si considera che la sequenza h_i non è un segnale continuo, è nulla a campioni alterni e non è periodica. La sequenza ξ_i , convoluzione della x_i e della h_i , è detta ancora trasformata di Hilbert

della sequenza x_i .

$$\xi_i = \mathcal{H}\{x_i\} = x_i * h_i$$

Si noti che se una sequenza è una sinusoidale (campionata), la sua trasformata di Hilbert è una cosinusoidale etc.

10.3 Sequenze periodiche e segnali continui

Una terza versione del nocciolo di Hilbert si ha quando la trasformata di Fourier sia periodica e campionata e cioè quando la sequenza di partenza è campionata e periodica. E' evidente che in questo caso il concetto di causalità va riformulato, accettando di definire causale una sequenza nulla per la seconda metà dei campioni. Allora se ne deduce una sequenza *sign* periodicizzata di cui si cerca la trasformata.

$$X(k) = \sum_{n=-N/2+1}^{N/2} x_n W^{-nk};$$

$$W = \exp(2\pi j/N);$$

$$x_n = 1; \quad 0 < n < N/2$$

$$x_n = -1; \quad -N/2 < n < 0; \quad x_0 = x_{N/2} = 0 \quad (10.13)$$

Risulta che:

$$X(k) = -j[1 - (-1)^k] \cot k\pi/N \quad (10.14)$$

e perciò si trova tanto l'effetto della discontinuità alla frequenza di Nyquist (la nullità a elementi alterni della trasformata) quanto l'effetto della periodicizzazione (la cotangente) (fig. 10.4).

L'ultimo caso è quello della continuità nei due domini: il nocciolo di Hilbert è allora:

$$h(t) = -\frac{1}{\pi t} \quad (10.15)$$

Possiamo ora tabellare le varie forme assunte dalla funzione che appare nella trasformazione di Hilbert; essa è sempre la trasformata di Fourier della funzione *sign*, che però può essere continua/discreta e illimitata/periodica, nel dominio dei tempi o delle frequenze ed inoltre, si faccia attenzione, a volte compare con un segno positivo ed a volte con quello negativo.

10.3 SEQUENZE PERIODICHE E SEGNALI CONTINUI

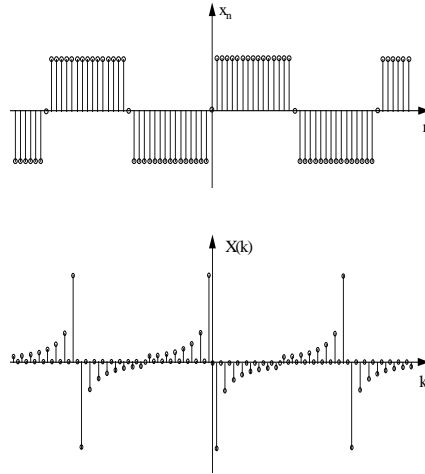


Figura 10.4: Sequenza $sign(n)$ periodicizzata e sua trasformata

Funzione	<i>continua</i>	<i>continua</i>	<i>discreta</i>	<i>discreta</i>
Segnale	<i>illimitato</i>	<i>periodico</i>	<i>illimitato</i>	<i>periodico</i>
Espressione	$-\frac{1}{\pi t}$	$-\frac{1-(-1)^i}{\pi i}$	$-\cot \frac{\omega T}{2}$	$-\left[1 - (-1)^k\right] \cot \frac{k\pi}{N}$

10.3.1 Singolarità nella trasformazione di Hilbert

Apparentemente, la trasformazione di Hilbert non dovrebbe creare singolarità perchè corrisponde a pure rotazioni di fase, senza amplificazioni dello spettro. Quindi il teorema di Parseval ci dovrebbe permettere di concludere che non sono da attendersi singolarità. Tuttavia, la trasformata di Hilbert della sequenza $sign(i)$ ha una singolarità di tipo logaritmico, che nasce dalla somma di una serie armonica (fig. 10.5).

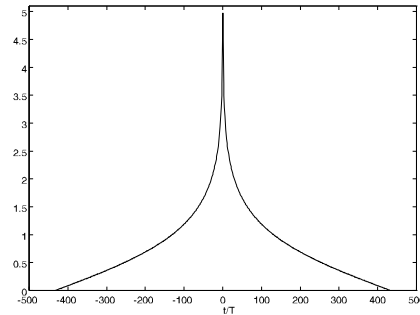


Figura 10.5: Trasformata di Hilbert della sequenza $sign(n)$.

Questo fatto è connesso a quanto osservato nel capitolo sull'interpolazione di sequenze illimitate con il filtro ideale. I segnali continui che interpolano sequenze con transitori bruschi e permanenti hanno picchi che crescono con il logaritmo del numero dei campioni della sequenza e quindi infiniti nel caso di sequenze per le quali il transitorio è illimitato.

10.4 Funzioni a fase minima

Supponiamo di avere una funzione di trasferimento di cui è nota solo la caratteristica di ampiezza. E' chiaro che possiamo assegnarne la caratteristica di fase con vari criteri: ad esempio possiamo imporre che la funzione di trasferimento sia a fase zero, ed allora la risposta all'impulso sarà non causale e simmetrica nel tempo; ancora possiamo imporre che la risposta all'impulso sia causale, aggiungendo un ritardo e quindi una fase che cresce linearmente con la frequenza.

E' importante il caso in cui vogliamo che la risposta sia sì causale, ma anche desideriamo che la sua durata sia minima e che cioè il ritardo della risposta sia il più breve possibile. Similmente, si può desiderare che la funzione di trasferimento abbia una funzione di trasferimento inversa causale e quindi sia a fase minima, come visto nel capitolo sulla Trasformata z . E' perciò un problema interessante quello di determinare la caratteristica di fase minima che si accompagna ad una caratteristica di ampiezza assegnata. Dimosteremo che vi è un facile legame tra caratteristica di ampiezza e caratteristica di fase minima, che corrisponde ad una trasformazione di Hilbert.

10.4 FUNZIONI A FASE MINIMA

Si abbia infatti una sequenza a fase minima e indichiamo la sua trasformata z con $F(z)$; dimostriamo anzitutto che $G(z)$, logaritmo complesso di $F(z)$, è la trasformata z di una sequenza causale. Infatti, se $F(z)$ è a fase minima:

$$G(z) = \log F(z) = \log(|F(z)|e^{j\angle F(z)}) = \log |F(z)| + j\angle F(z) \quad (10.16)$$

La $F(z)$, trasformata z di una sequenza a fase minima, ha poli e zeri all'interno del cerchio unitario. Il suo logaritmo complesso $G(z)$ ha tutti i poli all'interno del cerchio unitario (i valori infiniti della funzione logaritmo sono raggiunti dove l'argomento del logaritmo ha valore zero o infinito e quindi dove la $F(z)$ ha zeri o poli). La funzione $G(z)$ si può

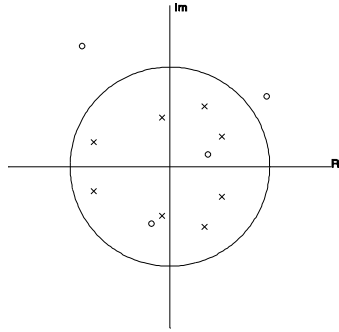


Figura 10.6: Poli e zeri del logaritmo della trasformata di una sequenza a fase minima

annullare (in corrispondenza degli zeri della $G(z)$) in qualsiasi posto del piano complesso, là dove $F(z) = 1$ (fig. 10.6)). Quindi la sequenza antitrasformata della $G(z)$ è una sequenza causale. Perciò le parti reale e parte immaginaria di $G(z)$ lungo il cerchio unitario sono legate dalla trasformazione di Hilbert e si può ottenere l'una dall'altra mediante la convoluzione con la funzione $-\cot(\omega T/2)$.

Ricordando poi che la parte immaginaria di $G(z)$ è la fase di $F(z)$ risulta:

$$\angle F(z) = \text{Im}[G(z)] = \text{Im}(\log F(z)) \quad (10.17)$$

$$\begin{aligned} \angle F(z)|_{z=e^{j\omega T}} &= \text{Im}(\log F(z)|_{z=e^{j\omega T}}) = & (10.18) \\ &= \mathcal{H}\{\text{Re}(\log F(z)|_{z=e^{j\omega T}})\} = \mathcal{H}\{\log |F(z)|_{z=e^{j\omega T}}\}. \end{aligned}$$

In conclusione, per ottenere la caratteristica di fase di una funzione di trasferimento a fase minima basta fare la trasformata di Hilbert del logaritmo naturale del modulo della sua caratteristica di ampiezza. Nel capitolo sulla predizione lineare troveremo un'importante applicazione di questo risultato.

Come esempio, si può considerare un filtro ideale di banda inferiore alla frequenza di Nyquist. La sequenza corrispondente è un *sinc* che poi deve essere opportunamente finestrato (p.e. con un coseno rialzato) per limitarne la durata. Il *sinc* finestrato è presentato in fig. 10.7; il logaritmo della caratteristica di ampiezza è presentato in fig. 10.8; la trasformata di Hilbert (la caratteristica di fase da associare alla caratteristica di ampiezza per avere fase minima) è in fig. 10.9. Infine la sequenza che ha le caratteristiche di ampiezza e di fase mostrate è presentata in fig. 10.10. E' evidente che la durata della sequenza in fig. 10.10 è ben inferiore a quella della sequenza in fig. 10.7.

10.5 Il segnale analitico

Consideriamo una sequenza reale x_i e aggiungiamogli come parte immaginaria una sequenza $-j\xi_i$, ottenuta dalla sequenza x_i mediante trasformazione di Hilbert e moltiplicazione per $-j$. La trasformata di Fourier di questa nuova sequenza complessa:

$$y_i = x_i - j\xi_i \quad (10.19)$$

che viene definita "il segnale analitico" corrispondente alla sequenza x_i ha la parte negativa dello spettro nulla. Infatti, la trasformata della sequenza $j\xi_i$ è identica a quella della sequenza x_i a meno della moltiplicazione per la funzione *sign*. Ancora, la trasformata della sequenza ξ_i è eguale a quella della sequenza x_i a meno di una rotazione di $-\pi/2$ di tutte le componenti. Infine, la trasformata della sequenza y_i è "causale" nelle frequenze e pertanto ha spettro nullo per $\omega < 0$. E' utile osservare che:

$$\begin{aligned} |y_i| &= \sqrt{x_i^2 + \xi_i^2} \\ \angle y_i &= \arctan(\xi_i/x_i) \end{aligned} \quad (10.20)$$

Si ricordi che se la sequenza x_i corrisponde ad una sinusoidale, la sequenza ξ_i corrisponde alla stessa sinusoidale sfasata di $-\pi/2$ e pertanto $|y_i|$ e $\angle y_i$

10.5 IL SEGNALE ANALITICO

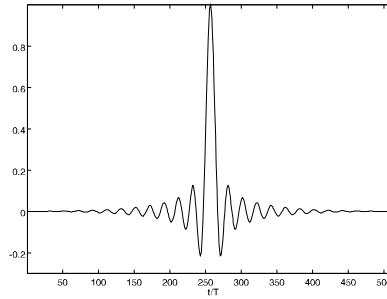


Figura 10.7: Sinc finestrato

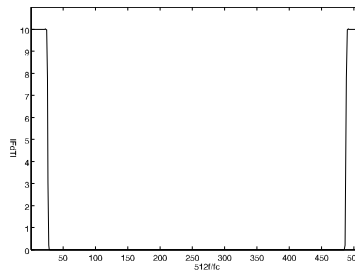


Figura 10.8: Caratteristica d'ampiezza

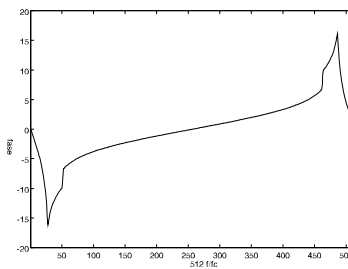


Figura 10.9: Caratteristica di fase

CAPITOLO 10 LE TRASFORMATE DI HILBERT E *CHIRP Z*

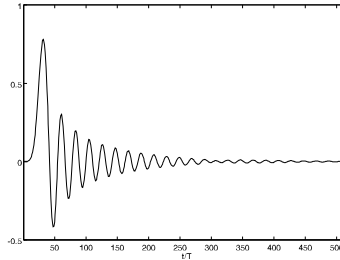


Figura 10.10: Sequenza a fase minima corrispondente ad un filtro passa basso ideale

sono delle costanti e niente altro che l'ampiezza e la fase della sinusoide stessa. Pertanto con il segnale analitico è possibile ottenere un'estensione dei concetti di fase e di ampiezza di una sinusoide, e costruire inviluppo e fase istantanea di un segnale qualsiasi (fig. 10.11); naturalmente, è anche possibile definire la frequenza istantanea di un segnale qualsiasi, come derivata della fase istantanea, fatta rispetto al tempo.

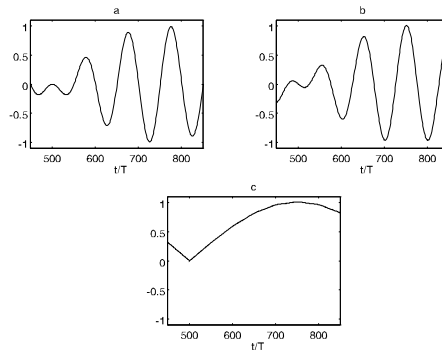


Figura 10.11: a) Segnale modulato in ampiezza x_i ; b) Trasformata di Hilbert del segnale modulato in ampiezza c) Inviluppo del segnale modulato

10.6 La trasformata *chirp z*

La trasformata di Fourier non è altro che la valutazione della trasformata z di una sequenza qualsiasi di N campioni in K punti disposti regolarmente lungo il cerchio unitario a partire dal punto $z = 1$. La *chirp transform* (o *chirp z transform*) consiste nel calcolo della trasformata z della sequenza in K punti non distribuiti uniformemente lungo il cerchio unitario, ma disposti lungo un arco di spirale nel piano z . I campioni della trasformata *chirp z* sono:

$$C_k = \sum_{n=0}^{N-1} x_n z_k^{-n}; k = 0, \dots, K-1 \quad (10.21)$$

dove ora

$$z_k = PQ^k \quad (10.22)$$

e P, Q sono due numeri complessi qualsiasi. Ovviamente, non è affatto detto che la trasformata *chirp* abbia un'inversa, e cioè che si possa ritornare alla sequenza di partenza dai C_k . Si ritrova invece la trasformata discreta di Fourier discreta qualora:

$$|P| = |Q| = 1; \angle P = 0; \angle Q = \frac{2\pi}{N}; K = N \quad (10.23)$$

Tuttavia, al variare dei parametri, il luogo descritto dai punti z_k può essere appunto un arco di cerchio o di spirale (se $|Q| \neq 1$) comunque ruotati nel piano z . Il calcolo dei C_k è reso poco costoso dalla seguente osservazione (Bluestein):

$$nk = \frac{n^2 + k^2 - (k-n)^2}{2} \quad (10.24)$$

Allora,

$$\begin{aligned} C_k &= \sum_{n=0}^{N-1} x_n P^{-n} Q^{-nk} = \sum_{n=0}^{N-1} x_n P^{-n} Q^{-\frac{n^2 + k^2 - (k-n)^2}{2}} = \\ &= Q^{-\frac{k^2}{2}} \sum_{n=0}^{N-1} x_n P^{-n} Q^{-\frac{n^2}{2}} Q^{\frac{(k-n)^2}{2}} \end{aligned} \quad (10.25)$$

e ponendo:

$$h_n = Q^{\frac{n^2}{2}}; \quad n = 0, \dots, \text{Max} (K, N) \quad (10.26)$$

$$y_n = \frac{x_n}{P^n h_n}; \quad D_k = C_k h_k$$

risulta:

$$D_k = \sum_{n=0}^{N-1} y_n Q^{\frac{(k-n)^2}{2}} = y_n * h_n \quad (10.27)$$

Quindi, si ottiene la sequenza D_k come una convoluzione tra le sequenze y_n e h_n , realizzabile con una trasformazione ed una antitrasformazione di Fourier. Poi dalla sequenza D_k si trova la sequenza $C_k = D_k/h_k$.

La sequenza h_k è chiamata sequenza *chirp*, perchè, nel caso semplice in cui $|Q| = 1$ corrisponde ad una senoide la cui frequenza cresce progressivamente nel tempo, come un cinguettio. *Chirp* significa in inglese appunto cinguettio d'uccello. Finora, non si comprende un motivo che giustifichi la trasformata chirp; due applicazioni si vedranno nelle prossime sezioni.

10.7 Il segnale chirp

Il segnale *chirp* trova applicazioni nel telerilevamento e nel trattamento dei segnali radar. È utile quando si desidera studiare la propagazione di un segnale in un mezzo di trasmissione: si desidera trasmettere un segnale ad elevata energia, per ottenere un buon rapporto segnale rumore in ricezione, ma con bassa potenza di picco per evitare di avere trasmettitori troppo costosi. La sequenza h_n ha questa proprietà perchè corrisponde ad una senoide la cui frequenza istantanea cresce linearmente con il tempo: perchè l'energia in essa contenuta sia sufficientemente elevata, basta costruire una sequenza abbastanza lunga allungando così anche il tempo di permanenza in ogni intervallo spettrale. Per esempio, per avere una sequenza campionata a T che *spazzola* l'intero intervallo di frequenze $-\frac{1}{2T} \div \frac{1}{2T}$ nell'intervallo di tempo $T_{\text{tot}} = NT$ basta porre:

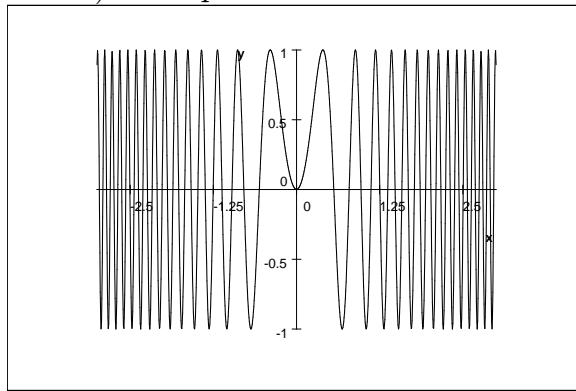
$$h_n = \exp\left(j \frac{\pi n^2}{N}\right); \quad -\frac{N}{2} < n < \frac{N}{2};$$

10.7 IL SEGNALE CHIRP

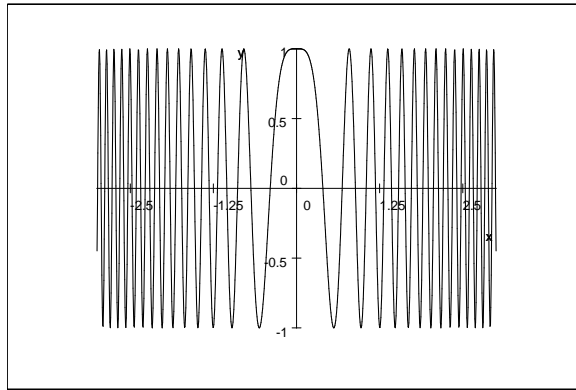
Infatti per $n = \pm N/2$; $\angle h_n - \angle h_{n-1} = \pm\pi$. La costante N prende anche il nome di prodotto *Banda - tempo* del chirp:

$$N = \frac{T_{\text{tot}}}{T} = BT_{\text{tot}}$$

Se si vuole ridurre di un fattore α l'escursione di frequenza ($\alpha < 1$ o si cade nell'equivocazione) basta porre $\alpha\pi$.



Parte immaginaria di una sequenza *chirp* h_n ; $N = 1000$; $\alpha = 0.125$



Parte reale della stessa sequenza h_n

La sequenza $\{h_n\}$ ha ampiezza costante nel tempo e quindi l'energia totale è proporzionale al prodotto banda tempo N . La sua funzione di autocorrelazione ha un modulo che assomiglia al *sinc* che ha banda corrispondente e infatti il suo spettro è approssimativamente uniforme nell'intervallo di frequenza *spazzolato* dalla sinusoidale. La trasformata del segnale chirp tempo continuo è facile da ottenersi, ricordando che gaussiane nei tempi corrispondono a gaussiane nelle frequenze e che:

$$\exp(-\pi t^2) \leftrightarrow \exp(-\pi f^2)$$

e scalando il tempo di un fattore \sqrt{jk}

$$c(t) = \exp(-j\pi kt^2) \leftrightarrow C(f) = \frac{1}{\sqrt{k}} \exp\left(-j\pi \frac{f^2}{k} - j\frac{\pi}{4}\right)$$

La finestrazione nei tempi del segnale chirp in un intervallo $\pm T_{\text{tot}}/2$ ne modifica lo spettro, a causa della convoluzione nelle frequenze con il *sinc* :

$$c_w(t) = c(t); |t| < \frac{T_{\text{tot}}}{2}; \text{ zero altrove}$$

$$C_w(f) = C(f) * W(f); W(f) = \frac{\sin \pi f T_{\text{tot}}}{\pi f T_{\text{tot}}}$$

Si osservi che la convoluzione azzerà lo spettro di $C_w(f)$ quando nel lobo principale del sinc di durata $2/T_{\text{tot}}$ la funzione $C(f)$ ha una rotazione di fase superiore a 2π :

$$C_w(f) \sim 0; \text{ se } \frac{d}{df} \left(\pi \frac{f^2}{k} \right) \frac{2}{T_{\text{tot}}} > 2\pi \rightarrow |f| > \frac{kT_{\text{tot}}}{2}$$

Congruentemente, la frequenza istantanea del chirp è:

$$f = \frac{1}{2\pi} \frac{d}{dt} (\pi kt^2) = kt; |t| < \frac{T_{\text{tot}}}{2}; \rightarrow |f| < \frac{kT_{\text{tot}}}{2}$$

Ha particolare interesse la decomposizione polifase di una sequenza chirp: infatti se dividessimo la sequenza originaria lunga N campioni in M sottobande, otterremmo M sottosequenze lunghe N/M campioni ognuna. In realtà queste sottosequenze sono M volte ancora più corte, poiché il segnale nella sua *spazzolata* si trattiene in ogni sottobanda solo per un M -esimo del tempo totale. Quindi la convoluzione effettuata con il filtro polifase costerà solo N/M moltiplicazioni per campione di ingresso, a parte il costo del banco di filtri. La cosa può essere utile quando si deve avere a che fare con segnali *chirp* tempo varianti quali si incontrano nella focalizzazione dei radar ad apertura sintetica, quando non si riesce ad adottare la tecnica della DFT.

10.8 Interpolazione con fattore qualsiasi

Nei casi visti nel capitolo sulla decimazione ed interpolazione, abbiamo interpolato (o decimato) delle sequenze sempre usando fattori interi. Peraltro, usando una cascata di interpolazione e decimazione, si può arrivare

10.8 INTERPOLAZIONE CON FATTORE QUALSIASI

ad un qualsiasi fattore **razionale** m/n di interpolazione. È però possibile effettuare delle interpolazioni con fattori di scala anche non razionali, e quindi con maggiore libertà [Munson, D. C., Arikan, O., Interpolation and the chirp transform: DSP meets optics, Proc. ICASSP 1999, pp. 2099 - 2102]. Allo scopo si utilizza la trasformata *chirp* z che abbiamo visto nella precedente sezione. Iniziamo l'analisi supponendo di avere un asse delle frequenze continuo. Interpolare in un dominio corrisponde ad una scalatura inversa dell'asse dell'altro dominio. Indichiamo con T_1, T_2 i due intervalli di campionamento delle due sequenze $\{x_{1,n}\}$, $\{x_{2,m}\}$ che corrispondono allo stesso segnale continuo, una volta fatto un filtraggio ideale nella banda di frequenza normalizzata comune. Supponendo di avere la sequenza $\{x_{1,n}\}$ e di volere calcolare la sequenza $\{x_{2,n}\}$ si comprende che se:

$$T_2 < T_1$$

il segnale interpolato $\{x_{2,n}\}$ potrà essere identico al vecchio perchè la banda disponibile cresce. Quindi Ω_2 , la banda normalizzata della sequenza $\{x_{2,n}\}$, sarà inferiore a π . In caso contrario, tutta la banda Ω_2 sarà occupata e la sequenza $\{x_{2,n}\}$ non potrà descrivere perfettamente la sequenza $\{x_{1,n}\}$. Si ha quindi:

$$\Omega_2 = \pm \text{Min} \left(\pi \frac{T_2}{T_1}, \pi \right)$$

ed inoltre indicando con $X_1(\omega)$, $X_2(\omega)$:

$$X_1(\omega) = \sum_{n=0}^{N-1} x_{1,n} e^{-j\omega n}$$

$$X_2(\omega) = \sum_{n=0}^{M-1} x_{2,m} e^{-j\omega m}$$

risulta:

$$X_2(\omega) = \frac{T_1}{T_2} X_1 \left(\omega \frac{T_1}{T_2} \right); \quad |\omega| < \Omega_2 \quad (10.28)$$

Come calcolare facilmente $\{x_{2,m}\}$ da $\{x_{1,n}\}$? Intanto osserviamo che:

$$x_{2,m} = \frac{T_1}{T_2} \int_{-\Omega_2}^{\Omega_2} X_2(\omega) e^{jm\omega} \frac{d\omega}{2\pi} = \frac{T_1}{T_2} \int_{-\Omega_2}^{\Omega_2} \sum_{n=0}^{N-1} x_{1,n} e^{-j\omega \frac{T_1}{T_2} n} e^{jm\omega} \frac{d\omega}{2\pi} \quad (10.29)$$

CAPITOLO 10 LE TRASFORMATE DI HILBERT E *CHIRP Z*

Poi approssimiamo l'integrale con una sommatoria (sostituiamo alla variabile continua ω la variabile discreta $k\Delta$) per avere:

$$x_{2,m} \sim \frac{\Delta}{2\pi} \frac{T_1}{T_2} \sum_{k=0}^{K-1} \sum_{n=0}^{N-1} x_{1,n} e^{-jk\Delta \frac{T_1}{T_2} n} e^{jmk\Delta} \quad (10.30)$$

Il calcolo della $\{x_{2,m}\}$ apparirebbe molto costoso computazionalmente, ma lo si farà in modo efficiente in due passi, usando appunto la trasformata *chirp*. Indicando con:

$$g_n = x_{1,n} w_n e^{-j \frac{\Delta \cdot T_1}{2T_2} n^2}; \quad h_n = e^{j \frac{\Delta \cdot T_1}{2T_2} n^2} \quad (10.31)$$

e poi calcolando la sequenza:

$$f_k = \{g_n\} * \{h_n\} = \sum_{n=0}^{N-1} x_{1,n} w_n e^{-j \frac{\Delta \cdot T_1}{2T_2} n^2} e^{j \frac{\Delta \cdot T_1}{2T_2} (n-k)^2} \quad (10.32)$$

si ottengono i campioni di uscita $\{x_{2,m}\}$ come:

$$x_{2,m} \sim \frac{\Delta}{2\pi} \frac{T_1}{T_2} \sum_{k=0}^{K-1} f_k \alpha_k e^{-j \frac{\Delta \cdot T_1}{2T_2} k^2} e^{jkm\Delta} \quad (10.33)$$

Sono state aggiunte due finestre α_k, w_n rispettivamente nei domini delle frequenze e dei tempi, per ridurre i lobi laterali provenienti dai troncamenti nei due domini. Per la verifica, basta sostituire nella 10.33 la 10.32, utilizzando la 10.24 per trovare la 10.30. Abbiamo quindi una ricetta poco costosa (due moltiplicazioni per campione ed una convoluzione) per fare un'interpolazione con un fattore di scala qualsiasi.

Si osservi che questa tecnica di utilizzare la trasformata *chirp* per fare un'interpolazione equivale in tutto e per tutto all'uso di una lente di ingrandimento. Infatti, una lente ottica effettua appunto una trasformazione *chirp* sul segnale elettromagnetico incidente, e funziona così da trasformatore di Fourier.

Capitolo 11

Processi casuali: generalità

11.1 Processi casuali discreti

Consideriamo una sequenza di variabili complesse $\{x_n\}$, $n = 0 \dots N$ e supponiamo che la densità di probabilità congiunta delle variabili (parti reali ed immaginarie) sia gaussiana. Indichiamo con $\underline{\mathbf{x}}$ un vettore colonna le cui N componenti sono gli elementi della sequenza $\{x_n\}$ e con $\underline{\mathbf{R}}$ la matrice di covarianza, di dimensioni $[N \times N]$, del vettore (e della sequenza) e cioè:

$$\underline{\mathbf{R}} = E[\underline{\mathbf{x}}\underline{\mathbf{x}}^*] \quad (11.1)$$

L'asterisco indica la trasposizione e, nel caso di vettori complessi, anche il coniugio. La matrice $\underline{\mathbf{R}}$ ha simmetria hermitiana e cioè, indicando con asterisco la trasposizione ed il coniugio si ha:

$$\underline{\mathbf{R}}^* = \underline{\mathbf{R}} \quad (11.2)$$

Dimostriamo ora, in casi semplici, che la densità di probabilità del vettore casuale $\underline{\mathbf{x}}$ è:

$$p(\underline{\mathbf{x}}) = \frac{1}{(2\pi)^{N/2} |\underline{\mathbf{R}}|^{1/2}} \exp\left(-\frac{\underline{\mathbf{x}}^* \underline{\mathbf{R}}^{-1} \underline{\mathbf{x}}}{2}\right) \quad (11.3)$$

Se la sequenza $\{x_n\}$ è bianca, e quindi

$$E[x_m x_{m-i}^*] = \sigma_x^2 \delta_i$$

la matrice di covarianza $\underline{\mathbf{R}}_x$ è diagonale:

$$\underline{\mathbf{R}}_x = \sigma_x^2 \underline{\mathbf{I}} \quad (11.4)$$

CAPITOLO 11 PROCESSI CASUALI: GENERALITÀ

dove con \mathbf{I} abbiamo indicato la matrice identica e con σ_x^2 il valore q.m. degli elementi della sequenza (dei moduli, se la sequenza è complessa). Se le variabili casuali sono a v. m. nullo, reali, gaussiane e i.i.d. e cioè indipendenti e identicamente distribuite, la 11.3 risulta immediatamente poichè la densità di probabilità di un vettore le cui componenti sono delle variabili indipendenti è il prodotto delle loro densità.

$$\begin{aligned} p(\underline{\mathbf{x}}) &= \prod_{i=1}^N p(x_i) = \frac{1}{(2\pi)^{N/2} \sigma_x^N} \exp\left(-\frac{1}{2\sigma_x^2} \sum x_i^2\right) = \quad (11.5) \\ &= \frac{1}{(2\pi)^{N/2} |\underline{\mathbf{R}}_x|^{1/2}} \exp\left(-\frac{\underline{\mathbf{x}}^* \underline{\mathbf{R}}_x^{-1} \underline{\mathbf{x}}}{2}\right) \end{aligned}$$

Passiamo ora a un caso più generale: facciamo una trasformazione di coordinate dal vettore $\underline{\mathbf{x}}$ ad un vettore $\underline{\mathbf{y}}$ e quindi moltiplichiamo il vettore $\underline{\mathbf{x}}$ per una matrice $\underline{\mathbf{A}}$ quadrata ed invertibile (dimentichiamo le singularità):

$$\underline{\mathbf{y}} = \underline{\mathbf{A}}\underline{\mathbf{x}} \rightarrow \underline{\mathbf{x}} = \underline{\mathbf{A}}^{-1}\underline{\mathbf{y}}$$

La matrice di covarianza di $\underline{\mathbf{y}}$ è:

$$E[\underline{\mathbf{y}}\underline{\mathbf{y}}^*] = \underline{\mathbf{R}}_y = \underline{\mathbf{A}}E[\underline{\mathbf{x}}\underline{\mathbf{x}}^*]\underline{\mathbf{A}}^* = \sigma_x^2 \underline{\mathbf{A}}\underline{\mathbf{A}}^*$$

ed è anche vero che:

$$\underline{\mathbf{x}}^* \underline{\mathbf{R}}_x^{-1} \underline{\mathbf{x}} = \frac{1}{\sigma_x^2} \underline{\mathbf{y}}^* (\underline{\mathbf{A}}^{-1})^* \underline{\mathbf{A}}^{-1} \underline{\mathbf{y}} = \underline{\mathbf{y}}^* \underline{\mathbf{R}}_y^{-1} \underline{\mathbf{y}}$$

e sostituendo nella 11.5 si ottiene la densità di probabilità 11.3. Per normalizzare ad 1 l'integrale della probabilità su tutto il volume infinito N -dimensionale dopo la trasformazione di coordinate, si deve dividere per il determinante della matrice Jacobiana di trasformazione delle coordinate da $\underline{\mathbf{x}}$ a $\underline{\mathbf{y}}$:

$$\sigma_x^{2N} |\underline{\mathbf{J}}| = \sigma_x^{2N} |\underline{\mathbf{A}}| = \sigma_x^{2N} \sqrt{|\underline{\mathbf{A}}|^2} = |\underline{\mathbf{R}}_y|^{1/2}$$

e quindi risulta la 11.3:

$$p(\underline{\mathbf{y}}) = \frac{1}{(2\pi)^{N/2} |\underline{\mathbf{R}}_y|^{1/2}} \exp\left(-\frac{\underline{\mathbf{y}}^* \underline{\mathbf{R}}_y^{-1} \underline{\mathbf{y}}}{2}\right)$$

11.2 MATRICI DI COVARIANZA DI SEQUENZE FILTRATE

La dimostrazione del fatto che dopo una trasformazione di coordinate, gli elementi di volume sono scalati per il determinante della matrice di trasformazione, è semplice in casi semplici. Partiamo dalla constatazione che una matrice $\underline{\mathbf{A}}$ può venire diagonalizzata con lo sviluppo in autovettori - autovalori. Si ha cioè:

$$\underline{\mathbf{A}} = \underline{\mathbf{U}}\underline{\mathbf{Q}}\underline{\mathbf{U}}^*$$

dove $\underline{\mathbf{U}}$ è la matrice degli autovettori di $\underline{\mathbf{A}}$, ortogonale, e quindi con determinante unitario. Quindi la trasformazione indotta da $\underline{\mathbf{A}}$ è decomponibile nella cascata di una rotazione N -dimensionale (la matrice $\underline{\mathbf{U}}$), uno scalamento di coordinate diverso da coordinata a coordinata (la matrice diagonale $\underline{\mathbf{Q}}$ che ha come elementi della diagonale principale gli autovalori di $\underline{\mathbf{A}}$, q_i) ed una rotazione di coordinate opposta alla precedente (la matrice $\underline{\mathbf{U}}^* = \underline{\mathbf{U}}^{-1}$). Lo scalamento dei volumi è ovviamente pari al prodotto dei moduli dei fattori di scala lungo le coordinate, e quindi al prodotto dei moduli degli autovalori, che è appunto il modulo del determinante. Si osservi che nella formula compare appunto:

$$|\underline{\mathbf{R}}_y|^{1/2} = \sigma_x^N \sqrt{|\underline{\mathbf{U}}| |\underline{\mathbf{Q}}|^2 |\underline{\mathbf{U}}^*|} = \sigma_x^N \prod_{n=1}^N |q_i|$$

Nel caso di autovalori nulli, la trasformazione è singolare, la nuova densità di probabilità non *occupa* tutto lo spazio, e infatti lungo alcune dimensioni corrisponderà ad un impulso ideale posto sullo zero (uscita sistematicamente nulla).

La generalizzazione della dimostrazione della 11.3 al caso di vettori casuali complessi è formalmente immediata, ma piuttosto laboriosa e verrà omessa. Nel paragrafo successivo si analizza il caso di vettori casuali corrispondenti a sequenze stazionarie, troncate.

11.2 Matrici di covarianza di sequenze filtrate

Consideriamo una sequenza di campioni $\{y_1, y_2, \dots, y_N\}$ ordinata in un vettore $\underline{\mathbf{y}}$. La sequenza $\{y_1, y_2, \dots, y_N\}$ è generata dalla convoluzione tra una

CAPITOLO 11 PROCESSI CASUALI: GENERALITÀ

sequenza bianca (ordinata in un vettore $\underline{\mathbf{x}}$) ed un filtro che ha risposta all'impulso lunga M campioni, $f_m = \{f_1, f_2, \dots, f_M\}$. La convoluzione è troncata dopo N campioni e sia anche $N \gg M$. Ponendo $r_\tau = E[y_n y_{n-\tau}^*]$, la matrice di covarianza del vettore $\underline{\mathbf{R}}_y$ sarà:

$$\underline{\mathbf{R}}_y = \begin{bmatrix} r_0 & r_{-1} & r_{-2} & \dots & r_{1-N} \\ r_1 & r_0 & r_{-1} & \dots & \dots \\ r_2 & r_1 & r_0 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ r_{N-1} & \dots & \dots & \dots & r_0 \end{bmatrix} \quad (11.6)$$

$$y_n = \sum_{m=1}^M x_{n-m} f_m; \quad r_\tau = E[y_n y_{n-\tau}^*] = \sigma_x^2 \sum_{m=1}^M f_m f_{m-\tau}^* = \sigma_x^2 r_f(\tau) \quad (11.7)$$

In sostanza, tutte le righe di $\underline{\mathbf{R}}_y$ sono eguali (a parte effetti di bordo) e contengono la replica di r_τ funzione di autocorrelazione della sequenza y_n sfalsata progressivamente di un passo ogni riga. $\underline{\mathbf{R}}_y$ avrà cioè struttura di Toeplitz, con gli elementi della diagonale principale tutti eguali tra loro ed eguali ad r_0 , gli elementi delle subdiagonali ancora eguali tra loro ma eguali a r_1, r_2 ecc. Si osservi, che nel caso di variabili casuali complesse, $r_{-n} = r_n^*$ e quindi la matrice di covarianza risulta a simmetria hermitiana. Ancora, se la correlazione si riduce a zero per $\tau = M$ ben prima dell' N -esimo campione, e cioè per:

$$r_M = 0; \quad M \ll N \quad (11.8)$$

gli elementi della matrice $\underline{\mathbf{R}}_y$ saranno nulli al di fuori di una banda di diametro $2M$ attorno alla diagonale principale (banded matrix, in inglese).

É utile avere un'idea della struttura della matrice $\underline{\mathbf{R}}_y^{-1}$. Indichiamo con $r_g(\tau)$ l'autocorrelazione di un filtro g_m che sbianca il filtro f_m . Si abbia cioè, indicando con la doppia freccia la corrispondenza indotta dalla trasformazione di Fourier,

$$\{f_m\} \Leftrightarrow F(\omega); \quad \{g_m\} \Leftrightarrow G(\omega);$$

$$|G(\omega)| = \frac{1}{|F(\omega)|}$$

11.2 MATRICI DI COVARIANZA DI SEQUENZE FILTRATE

$$r_f(\tau) \Leftrightarrow |F(\omega)|^2; r_g(\tau) \Leftrightarrow |G(\omega)|^2 = \frac{1}{|F(\omega)|^2}$$

$$r_f(\tau) * r_g(\tau) \sim \delta(\tau) \quad (11.9)$$

Si osservi che la caratteristica di fase del filtro g_m è qualsiasi: basta che il **modulo** della sua trasformata sia inverso del **modulo** della trasformata di f_m . La matrice \mathbf{R}_y^{-1} è ancora una matrice di Toeplitz, ma ogni sua riga contiene $r_g(\tau)/\sigma_x^2$ invece che $\sigma_x^2 r_f(\tau)$. Infatti, facendo il prodotto righe per colonne delle due matrici $\mathbf{R}_y \times \mathbf{R}_y^{-1}$, si ottiene una matrice ancora di Toeplitz a banda che contiene in ogni riga la convoluzione di $r_f(\tau)$ e di $r_g(\tau)$ che è un impulso ideale, per costruzione (11.9). Quindi il prodotto delle matrici è una matrice identica, e questo conferma che \mathbf{R}_y^{-1} ha appunto la struttura che abbiamo visto, almeno approssimativamente. Si trascuri il fatto che il filtro inverso (e quindi la sua autocorrelazione) sarà di lunghezza infinita.

Per fare un esempio, che verrà riutilizzato più avanti, consideriamo due matrici $[N \times N]$, $\mathbf{M}_p, \mathbf{M}_z$ l'una inversa dell'altra. La prima, \mathbf{M}_p , corrisponde allo spettro di un polo ed è esattamente di Toeplitz. Una riga è

$$\mathbf{M}_p \rightarrow \sigma_m^2 (\dots, \rho^2, \rho, 1, \rho, \rho^2 \dots)$$

$$\mathbf{M}_p = \sigma_m^2 \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^N \\ \rho & 1 & \rho & \rho^2 & \dots \\ \rho^2 & \rho & 1 & \rho & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \rho^N & \dots & \rho^2 & \rho & 1 \end{bmatrix} \quad (11.10)$$

La seconda matrice, $\mathbf{M}_{p,inv}$, corrisponde allo spettro di uno zero, è con grande approssimazione ancora di Toeplitz, a parte la prima e l'ultima riga: una riga centrale è

$$\mathbf{M}_{p,inv} \rightarrow \frac{1}{\sigma_m^2 (1 - \rho^2)} (\dots, 0, 0, 0, 0, -\rho, 1 + \rho^2, -\rho, 0, 0, 0 \dots);$$

CAPITOLO 11 PROCESSI CASUALI: GENERALITÀ

ma $m_{p,inv,11} = m_{p,inv,NN} = 1$.

$$\underline{\mathbf{M}}_{p,inv} = \frac{1}{\sigma_m^2 (1 - \rho^2)} \begin{bmatrix} 1 & -\rho & 0 & .. & 0 \\ -\rho & 1 + \rho^2 & -\rho & 0 & .. \\ 0 & -\rho & 1 + \rho^2 & -\rho & 0 \\ .. & .. & .. & .. & .. \\ 0 & .. & 0 & -\rho & 1 \end{bmatrix} \quad (11.11)$$

Orbene:

$$\underline{\mathbf{M}}_p \underline{\mathbf{M}}_{p,inv} = \underline{\mathbf{I}} \quad (11.12)$$

Questa accade perchè:

$$H(z) = 1 - \rho z^{-1}; |H(z)|^2 = 1 + \rho^2 - \rho(z + z^{-1}); \quad (11.13)$$

ed anche:

$$\frac{1}{|H(z)|^2} = \frac{1}{1 + \rho^2 - \rho(z + z^{-1})} = \frac{1}{1 - \rho^2} \left(\frac{1}{1 - z^{-1}\rho} + \frac{1}{1 - z\rho} - 1 \right) \quad (11.14)$$

da cui le righe centrali delle due matrici.

11.2.1 Matrici di covarianza di trasformate

Consideriamo ancora la sequenza di variabili indipendenti identicamente distribuite (*bianca*) $\{x_n\}$ convoluta con un filtro $\{f_m\}$ in modo da risultare *colorata*; come appena visto, la matrice di covarianza della sequenza risultante $\{y_n\} = \{x_n\} * \{f_m\}$ sarà $\underline{\mathbf{R}}_y$.

$$\underline{\mathbf{R}}_y = \begin{bmatrix} r_0 & r_{-1} & r_{-2} & \dots & 0 \\ r_1 & r_0 & r_{-1} & \dots & 0 \\ r_2 & r_1 & r_0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & r_0 & r_{-1} \\ 0 & 0 & .. & r_1 & r_0 \end{bmatrix}$$

11.2 MATRICI DI COVARIANZA DI SEQUENZE FILTRATE

Calcoliamo ora la matrice di covarianza della trasformata di Fourier di una sequenza reale e bianca $\{x_n\}$ e di lunghezza limitata N , e cioè la matrice di covarianza della sequenza $\{X_k\}$ dove:

$$X_k = \sum_{n=0}^{N-1} x_n W^{-nk} \quad (11.15)$$

Indicando ora con $\underline{\mathbf{X}}$ il vettore le cui componenti sono i campioni della sequenza $\{X_k\}$ si ha:

$$\underline{\mathbf{X}} = \underline{\mathbf{W}} \underline{\mathbf{x}} \quad (11.16)$$

dove $\underline{\mathbf{W}}$ è una matrice che corrisponde alla trasformazione di Fourier diretta (attenzione! non è una matrice hermitiana!). La matrice di covarianza relativa al vettore $\underline{\mathbf{X}}$ è:

$$\underline{\mathbf{R}}_{\underline{\mathbf{X}}} = E[\underline{\mathbf{X}} \underline{\mathbf{X}}^*] = E[\underline{\mathbf{W}} \underline{\mathbf{x}} \underline{\mathbf{x}}^* \underline{\mathbf{W}}^*] = \underline{\mathbf{W}} E[\underline{\mathbf{x}} \underline{\mathbf{x}}^*] \underline{\mathbf{W}}^* = \underline{\mathbf{W}} \underline{\mathbf{R}} \underline{\mathbf{W}}^* \quad (11.17)$$

Indicando con A_k e B_k la parte reale ed immaginaria di X_k , si può verificare che, poichè la sequenza $\{x_n\}$ è bianca e reale:

$$X_k = A_k + jB_k;$$

$$E[A_k B_h] = 0; \forall k, h; \quad E[A_k A_h] = E[B_k B_h] = 0; \quad |k| \neq |h| \quad (11.18)$$

$$E[A_k^2] = \frac{N}{2} \sigma_x^2 [1 + \delta(k) + \delta(k - N/2)] \quad (11.19)$$

$$E[B_k^2] = E[A_k^2]; \quad k = 1, \dots, N/2 - 1$$

$$E[B_k^2] = 0; \quad k = 0, N/2$$

Infatti, le relazioni precedenti si possono dedurre dal fatto che:

$$\begin{aligned} E[X_k X_h^*] &= \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} E[x_n x_m^*] W^{-nk} W^{mh} = \\ &= \sigma_x^2 \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \delta(n-m) W^{-nk} W^{mh} = \sigma_x^2 \sum_{m=0}^{N-1} W^{m(h-k)} = \\ &= N \sigma_x^2 \times \delta(h-k); \end{aligned}$$

CAPITOLO 11 PROCESSI CASUALI: GENERALITÀ

Relazioni simili si possono scrivere per le parti reali ed immaginarie degli X_k . Se invece la sequenza $\{x_n\}$ fosse bianca e complessa, si avrebbe:

$$E[A_k B_h] = 0; \quad E[A_k A_h] = E[B_k B_h] = 0; \quad k \neq h \quad (11.20)$$

$$E[A_k^2] = E[B_k^2] = N\sigma_x^2/2$$

Si possono ottenere gli stessi risultati in modo diretto, osservando che:

$$\begin{aligned} \underline{\mathbf{X}} &= \underline{\mathbf{W}} \underline{\mathbf{x}}; \quad \underline{\mathbf{x}} = \frac{1}{N} \underline{\mathbf{W}}^* \underline{\mathbf{X}} = \frac{1}{N} \underline{\mathbf{W}}^* \underline{\mathbf{W}} \underline{\mathbf{x}} \rightarrow \\ \frac{1}{N} \underline{\mathbf{W}}^* \underline{\mathbf{W}} &= \frac{1}{N} \underline{\mathbf{W}} \underline{\mathbf{W}}^* = \underline{\mathbf{I}} \rightarrow (\underline{\mathbf{W}}^*)^{-1} = \frac{1}{N} \underline{\mathbf{W}} \\ \underline{\mathbf{R}}_{\mathbf{X}} &= \underline{\mathbf{W}} \underline{\mathbf{R}}_{\mathbf{x}} \underline{\mathbf{W}}^* = N\sigma_x^2 \underline{\mathbf{I}} \end{aligned} \quad (11.21)$$

Quindi la trasformata di Fourier di una sequenza complessa i.i.d. (bianca) è bianca (i campioni sono incorrelati); se la sequenza di partenza è reale, vanno sempre tenute in conto le simmetrie.

Ora, convolviamo **periodicamente** la sequenza $\{x_n\}$ di lunghezza limitata N , con un qualsiasi filtro $\{h_n\}$ di eguale lunghezza N (o diventato tale dopo un'aggiunta di zeri). La matrice di covarianza della sequenza $\{y_n\} = \{x_n\} * \{h_n\}$ è proprio circolante e cioè tutte le sue righe sono esattamente eguali a meno di uno slittamento progressivo verso destra.

$$\underline{\mathbf{R}}_{\mathbf{y}} = \begin{bmatrix} r_0 & r_{-1} & 0 & \dots & r_1 \\ r_1 & r_0 & r_{-1} & \dots & 0 \\ 0 & r_1 & r_0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & r_{-1} \\ r_{-1} & 0 & \dots & r_1 & r_0 \end{bmatrix}$$

La trasformata di Fourier discreta della sequenza $\{x_n\}$, $\{X_k\}$ verrà moltiplicata per $\{H_k\}$, Trasformata di Fourier discreta della sequenza $\{h_n\}$. Si ha cioè, se:

$$\{x_{n+N}\} = \{x_n\}; \quad \{h_{n+N}\} = \{h_n\}; \quad (11.22)$$

$$\{y_n\} = \{h_n\} * \{x_n\} \rightarrow \{y_{n+N}\} = \{y_n\} \quad (11.23)$$

11.2 MATRICI DI COVARIANZA DI SEQUENZE FILTRATE

$$X_k = \sum_{n=0}^{N-1} x_n W^{-nk}; \quad H_k = \sum_{n=0}^{N-1} h_n W^{-nk}; \quad (11.24)$$

$$Y_k = X_k H_k; \quad y_n = \frac{1}{N} \sum_{k=0}^{N-1} Y_k W^{nk} \quad (11.25)$$

Indicando con vettori sia la sequenza $\{X_k\}$ che la $\{Y_k\}$ e con una moltiplicazione matriciale l'effetto dell'applicazione della funzione di trasferimento $\{H_k\}$ risulta:

$$\underline{\mathbf{X}} = \begin{bmatrix} X_0 \\ X_1 \\ \dots \\ X_{N-1} \end{bmatrix}; \quad \underline{\mathbf{Y}} = \begin{bmatrix} Y_0 \\ Y_1 \\ \dots \\ Y_{N-1} \end{bmatrix}; \quad \underline{\mathbf{H}} = \begin{bmatrix} H_0 & 0 & 0 & 0 \\ 0 & H_1 & 0 & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & H_{N-2} \\ 0 & 0 & 0 & \dots & H_{N-1} \end{bmatrix}$$

$$\underline{\mathbf{Y}} = \underline{\mathbf{H}} \underline{\mathbf{X}}$$

Anzitutto,

$$\underline{\mathbf{R}}_{\mathbf{X}} = E[\underline{\mathbf{X}}\underline{\mathbf{X}}^*] = N\sigma_x^2 \mathbf{I} \quad (11.26)$$

Infatti,

$$E[|X_k|^2] = \sum_{m,n=0}^{N-1} E[x_n x_m] W^{-nk} W^{mk} = \sum_{n=0}^{N-1} E[|x_m|^2] = N\sigma_x^2 \quad (11.27)$$

La matrice di covarianza della sequenza $\underline{\mathbf{Y}}$, $\underline{\mathbf{R}}_{\mathbf{Y}}$ sarà anch'essa diagonale, ma gli elementi della diagonale principale avranno ora valore $N\sigma_x^2 |H_k|^2$.

$$\underline{\mathbf{R}}_{\mathbf{Y}} = E[\underline{\mathbf{H}} \underline{\mathbf{X}} \underline{\mathbf{X}}^* \underline{\mathbf{H}}^*] = \underline{\mathbf{H}} \underline{\mathbf{R}}_{\mathbf{X}} \underline{\mathbf{H}}^* =$$

$$= N\sigma_x^2 \begin{bmatrix} |H_0|^2 & 0 & \dots & 0 & 0 \\ 0 & |H_1|^2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & |H_{N-2}|^2 & 0 \\ 0 & 0 & \dots & 0 & |H_{N-1}|^2 \end{bmatrix} \quad (11.28)$$

CAPITOLO 11 PROCESSI CASUALI: GENERALITÀ

Questo implica che le componenti della Trasformata di Fourier discreta di una qualsiasi sequenza colorata (filtrata) $\{y_n\}$ ottenuta convolvendo **periodicamente** una sequenza bianca con una sequenza qualsiasi $\{h_n\}$ di pari lunghezza, sono incorrelate, anche se di diversa varianza.

In altre parole, i campioni di una sequenza **stazionaria** colorata sono correlati ed i campioni della loro Trasformata di Fourier discreta sono incorrelati, ma la loro **varianza cambia** con la frequenza.

Vediamo di nuovo ora la struttura della matrice di covarianza della sequenza $\{y_n\}$ convoluzione periodica delle due sequenze $\{x_n\}$ e $\{h_n\}$, nel caso semplice in cui solo r_0, r_1 siano diversi da 0;

$$\underline{\mathbf{R}}_y = E[\underline{\mathbf{y}} \underline{\mathbf{y}}^*] \tag{11.29}$$

$$\underline{\mathbf{R}}_y = \begin{bmatrix} r_0 & r_{-1} & 0 & \dots & r_1 \\ r_1 & r_0 & r_{-1} & \dots & 0 \\ 0 & r_1 & r_0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & r_{-1} \\ r_{-1} & 0 & 0 & \dots & r_0 \end{bmatrix} \tag{11.30}$$

Come già visto, la matrice di covarianza di una sequenza stazionaria $\underline{\mathbf{R}}_y$ ha struttura **hermitiana** e cioè è simmetrica se è reale o a simmetria coniugata se complessa. Inoltre, ha struttura di **Toeplitz** e cioè gli elementi lungo la diagonale e le subdiagonali sono costanti. Inoltre, a causa del fatto che la convoluzione è stata fatta in modo periodico, la matrice $\underline{\mathbf{R}}_y$ è **circolante** e cioè ogni riga è ottenuta dalla precedente con lo slittamento di un posto e mettendo l'ultimo campione al posto del primo. La m -esima riga della matrice $\underline{\mathbf{R}}_y$ è quindi l'autocorrelazione circolare del segnale $\{y_n\}$, $r_{mn} = E[y_h y_{h+m-n}^*]$; ricordiamo inoltre la periodicità N delle sequenze $\{y_n\}$, $\{h_n\}$.

Se la sequenza $\{h_n\}$ ha appunto solo due campioni diversi da zero, la $\underline{\mathbf{R}}_y$ ha struttura tridiagonale, come mostrato nell'equazione (11.30). Si osserva pertanto che, se la sequenza $\{h_n\}$ è corta rispetto ad N , la periodizzazione cambia solo pochi elementi della matrice di covarianza.

Questo fatto ci porta a concludere che in pratica, convolvendo sequenze lunghe di campioni indipendenti con filtri corti, la matrice di

11.3 MATRICI DI TOEPLITZ: AUTOVALORI

covarianza risultante dalla convoluzione periodica è poco differente da quella risultante dalla convoluzione aperiodica: quindi gli elementi della trasformata di Fourier di una sequenza ottenuta finestrando un segnale complesso stazionario, bianco o colorato che sia, tendono ad essere indipendenti. Se il segnale è reale, ci si deve ricordare delle relazioni di parità e disparità delle parti reali ed immaginarie.

11.3 Matrici di Toeplitz: autovalori

Si consideri una matrice di covarianza $\underline{\mathbf{R}}_{\mathbf{x}}$ di Toeplitz (e cioè con le diagonali parallele alla diagonale principale composte da elementi eguali). La matrice $\underline{\mathbf{R}}_{\mathbf{x}}$ essendo di covarianza e cioè

$$\underline{\mathbf{R}}_{\mathbf{x}} = E[\underline{\mathbf{x}} \underline{\mathbf{x}}^*]$$

è simmetrica se la sequenza di partenza è reale; altrimenti è a simmetria Hermitiana (la trasposta è eguale alla coniugata). E' intuitivamente attraente collegare gli autovalori e gli autovettori di matrici di questo tipo allo spettro di potenza del processo cui appartiene il vettore casuale $\underline{\mathbf{x}}$ che ha appunto $\underline{\mathbf{R}}_{\mathbf{x}}$ come matrice di covarianza. Ricordiamo che una matrice $\underline{\mathbf{R}}_{\mathbf{x}}$ ha autovalori λ_i (che vedremo corrispondere alle energie spettrali) e autovettori colonna $\underline{\mathbf{e}}_i$ (corrispondenti alle sinusoidi di varia frequenza) tali che:

$$\underline{\mathbf{R}}_{\mathbf{x}} \underline{\mathbf{e}}_i = \lambda_i \underline{\mathbf{e}}_i \quad (11.31)$$

In forma matriciale si ha:

$$\underline{\mathbf{E}} \underline{\mathbf{R}}_{\mathbf{x}} \underline{\mathbf{E}}^* = \underline{\mathbf{\Lambda}}; \quad \underline{\mathbf{E}}^{-1} = \underline{\mathbf{E}}^* \quad (11.32)$$

dove la matrice **ortogonale** $\underline{\mathbf{E}}$ corrisponde agli N autovettori colonna $\underline{\mathbf{e}}_i$, la matrice $\underline{\mathbf{E}}^*$ agli autovettori riga $\underline{\mathbf{e}}_i^*$ (è sia l'inversa che la trasposta della precedente, con coniugio se hermitiana) ed infine la matrice $\underline{\mathbf{\Lambda}}$ è diagonale, e contiene lo "spettro" degli autovalori della matrice $\underline{\mathbf{R}}_{\mathbf{x}}$.

$$\underline{\mathbf{\Lambda}} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_{N-1} & 0 \\ 0 & 0 & \dots & 0 & \lambda_N \end{bmatrix}$$

CAPITOLO 11 PROCESSI CASUALI: GENERALITÀ

$$\underline{\mathbf{E}} = \begin{bmatrix} \underline{\mathbf{e}}_1 & \underline{\mathbf{e}}_2 & \dots & \underline{\mathbf{e}}_{N-1} & \underline{\mathbf{e}}_N \end{bmatrix}$$

Dimostriamo ora che $\lambda_i \geq 0$ e quindi che la matrice $\underline{\mathbf{R}}$ è semidefinita positiva. Infatti, sia $\underline{\mathbf{a}}^* \underline{\mathbf{x}}$ una qualsiasi combinazione lineare degli elementi del vettore $\underline{\mathbf{x}}$. Il suo valore quadratico medio è maggiore od eguale a zero:

$$E[\underline{\mathbf{a}}^* \underline{\mathbf{x}} \underline{\mathbf{a}}^* \underline{\mathbf{x}}] = \underline{\mathbf{a}}^* E[\underline{\mathbf{x}} \underline{\mathbf{x}}^*] \underline{\mathbf{a}} = \underline{\mathbf{a}}^* \underline{\mathbf{R}}_x \underline{\mathbf{a}} \geq 0 \quad (11.33)$$

Se $\underline{\mathbf{a}}$ è un autovettore della matrice $\underline{\mathbf{R}}_x$ risulta:

$$\underline{\mathbf{e}}_i^* \underline{\mathbf{R}}_x \underline{\mathbf{e}}_i = \lambda_i \geq 0 \quad (11.34)$$

e quindi tutti gli autovalori di una matrice di covarianza, come le componenti di uno spettro di potenza, sono maggiori o eguali a 0. Inoltre, per la simmetria della matrice, due autovettori $\underline{\mathbf{e}}_i, \underline{\mathbf{e}}_j$ corrispondenti a due autovalori λ_i, λ_j distinti, sono ortogonali, così come due sinusoidi campionate di frequenza diversa, nel senso che la media del loro prodotto è nulla. Infatti:

$$\underline{\mathbf{e}}_i^* \underline{\mathbf{R}}_x \underline{\mathbf{e}}_j = \lambda_j \underline{\mathbf{e}}_i^* \underline{\mathbf{e}}_j = \lambda_i \underline{\mathbf{e}}_i^* \underline{\mathbf{e}}_j \rightarrow \underline{\mathbf{e}}_i^* \underline{\mathbf{e}}_j = 0$$

Ancora consideriamo due variabili casuali z_1, z_2 combinazione lineare degli elementi della sequenza $\{x_n\}$:

$$z_1 = \underline{\mathbf{a}}_1^* \underline{\mathbf{x}}; \quad z_2 = \underline{\mathbf{a}}_2^* \underline{\mathbf{x}}$$

Risulta ora:

$$E[z_1 z_2] = \underline{\mathbf{a}}_1^* \underline{\mathbf{R}}_x \underline{\mathbf{a}}_2$$

e le due variabili casuali z_1, z_2 sono incorrelate se $E[z_1 z_2] = \underline{\mathbf{a}}_1^* \underline{\mathbf{R}}_x \underline{\mathbf{a}}_2 = 0$ e cioè se i due vettori $\underline{\mathbf{a}}_1, \underline{\mathbf{R}}_x \underline{\mathbf{a}}_2$ sono ortogonali; in questo caso i due vettori $\underline{\mathbf{a}}_1, \underline{\mathbf{a}}_2$ sono definiti come ortogonali coniugati e se ne discuterà nel capitolo 20.

Si ricorda ancora che la traccia di una matrice è pari alla somma degli autovalori, ed infine il determinante è pari al loro prodotto.

Supponiamo ora che la funzione di autocorrelazione della sequenza $\{x_n\}$ tenda a 0 per valori di $\tau \ll N$; la circolarizzazione della matrice $\underline{\mathbf{R}}_x$ non cambia la matrice di covarianza in modo significativo (la cambia solo negli angoli).

11.3.1 Autovettori ed autovalori di matrici di covarianza circolanti

Consideriamo ora la relazione tra una sequenza complessa filtrata con una convoluzione periodica e la sua trasformata di Fourier discreta:

$$\underline{\mathbf{Y}} = \underline{\mathbf{W}} \underline{\mathbf{y}} \quad (11.35)$$

dove con $\underline{\mathbf{W}}$ abbiamo indicato la matrice della trasformazione di Fourier diretta. Inoltre:

$$\underline{\mathbf{R}}_{\mathbf{Y}} = E[\underline{\mathbf{Y}} \underline{\mathbf{Y}}^*] = \underline{\mathbf{H}} \underline{\mathbf{R}}_{\mathbf{X}} \underline{\mathbf{H}}^* = \underline{\mathbf{W}} E[\underline{\mathbf{y}} \underline{\mathbf{y}}^*] \underline{\mathbf{W}}^* = \underline{\mathbf{W}} \underline{\mathbf{R}}_{\mathbf{y}} \underline{\mathbf{W}}^* \quad (11.36)$$

dove con $\underline{\mathbf{R}}_{\mathbf{y}}$ abbiamo indicato la matrice diagonale che contiene le ampiezze, moltiplicate per N , del periodogramma della sequenza $\{y_n\}$; infatti abbiamo dimostrato in precedenza che $E[\underline{\mathbf{Y}} \underline{\mathbf{Y}}^*]$ è diagonale. Possiamo anche scrivere che:

$$\underline{\mathbf{R}}_{\mathbf{y}} = \frac{1}{N} \underline{\mathbf{W}}^* \underline{\mathbf{R}}_{\mathbf{Y}} \underline{\mathbf{W}} \frac{1}{N} = \frac{1}{\sqrt{N}} \underline{\mathbf{W}}^* \frac{\underline{\mathbf{R}}_{\mathbf{Y}}}{N} \underline{\mathbf{W}} \frac{1}{\sqrt{N}} \quad (11.37)$$

e cioè che la sequenza $\{y_n\}$, ottenuta con una qualsiasi convoluzione periodica di un filtro e di una sequenza a spettro bianco, ha una matrice di covarianza circolante $\underline{\mathbf{R}}_{\mathbf{y}}$ che è diagonalizzata dalla trasformazione di Fourier; in altre parole, la sequenza $\{Y_k\}$ trasformata di Fourier della sequenza $\{y_n\}$, ha matrice di covarianza $\underline{\mathbf{R}}_{\mathbf{Y}}$ diagonale. In particolare, la matrice

$$\frac{\underline{\mathbf{W}}}{\sqrt{N}}$$

è ortogonale, in quanto è inversa della sua trasposta e coniugata. Infatti, le sinusoidi sono qui normalizzate per il fattore $1/\sqrt{N}$ in modo da avere energia unitaria.

Quindi, gli autovettori della matrice di covarianza $\underline{\mathbf{R}}_{\mathbf{y}}$ di una sequenza ottenuta da una convoluzione periodica di una sequenza bianca ed un filtro qualsiasi sono le sinusoidi discrete di frequenza k/NT ; $0 \leq k \leq N-1$ ed ampiezza $1/\sqrt{N}$; gli autovalori sono le ampiezze del periodogramma.

Nel caso di una matrice di covarianza di un segnale qualsiasi e non di una matrice di covarianza circolante, gli autovalori sono comunque non negativi e gli autovettori sono comunque ortogonali tra loro: non saranno però sinusoidi vere e proprie, ma gli assomiglieranno abbastanza.

CAPITOLO 11 PROCESSI CASUALI: GENERALITÀ

L'ordinamento dello spettro degli autovalori è diverso però da quello dello spettro di frequenza di una sequenza: lo spettro di frequenza di una sequenza è ordinato sulla base della frequenza degli autovettori (l'indice k); lo spettro degli autovalori di una matrice è in genere ordinato per autovalori decrescenti, visto che non è possibile definire una vera e propria frequenza per ogni autovettore.

11.3.2 Ancora la densità di probabilità di un vettore casuale

Possiamo ora riconsiderare la formula 11.3, che riportiamo qui per comodità.

$$p(\underline{\mathbf{y}}) = \frac{1}{(2\pi)^{N/2} |\underline{\mathbf{R}}_y|^{1/2}} \exp \left(-\frac{\underline{\mathbf{y}}^* \underline{\mathbf{R}}_y^{-1} \underline{\mathbf{y}}}{2} \right) \quad (11.38)$$

Ora, nel caso di campionamento delle frequenze e periodicità nei empi, con sequenze filtrate periodicamente e quindi con autocorrelazioni circolanti, utilizzando la rappresentazione di un vettore mediante le componenti con autovettori e cioè ponendo:

$$\underline{\mathbf{y}} = \sum \eta_k \underline{\mathbf{e}}_k; \quad \eta_k = \underline{\mathbf{y}}^* \underline{\mathbf{e}}_k \quad (11.39)$$

si ha:

$$\underline{\mathbf{y}}^* \underline{\mathbf{R}}_y^{-1} \underline{\mathbf{y}} = \sum \eta_k^* \underline{\mathbf{e}}_k^* \underline{\mathbf{R}}_y^{-1} \underline{\mathbf{e}}_k \eta_k \quad (11.40)$$

Ora, gli autovettori sono le sinusoidi a frequenze discrete campionate, le componenti η_k coincidono con le componenti di Fourier discrete Y_k , gli autovalori sono i valori assunti dal quadrato del modulo della funzione di trasferimento $|F_k|^2$, risulta:

$$\underline{\mathbf{y}}^* \underline{\mathbf{R}}_y^{-1} \underline{\mathbf{y}} = \sum_{k=0}^{N-1} \left| \frac{Y_k}{F_k} \right|^2 = \sum_{k=0}^{N-1} |X_k|^2 \quad (11.41)$$

$$p(\underline{\mathbf{y}}) = \frac{1}{(2\pi)^{N/2} |\prod_{k=0}^N |F_k||} \exp \left(- \sum_{k=0}^{N-1} \left| \frac{Y_k}{F_k} \right|^2 \right) \quad (11.42)$$

Quindi si evince da queste formule l'ovvia conseguenza che, alle frequenze dove lo spettro del segnale è basso, le componenti del segnale sono, con grande probabilità, di assai piccola potenza: conseguenza banale di una formula apparentemente arcana.

11.3.3 Un esempio numerico

Si consideri una sequenza di $N = 4$ campioni casuali, a valor medio nullo e varianza unitaria e si indichi con $\underline{\mathbf{x}}$ il vettore corrispondente:

$$\underline{\mathbf{x}} = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix}; \quad E[\underline{\mathbf{x}}\underline{\mathbf{x}}^*] = \underline{\mathbf{R}}_{\mathbf{x}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \underline{\mathbf{I}}$$

Convolviamo periodicamente $\underline{\mathbf{x}}$ con il filtro:

$$\underline{\mathbf{f}} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

ed otteniamo il vettore $\underline{\mathbf{y}}$ le cui componenti sono:

$$y_i = x_i + x_{i+1}; \quad i = 0, \dots, 3; \quad x_4 = x_0 \quad \text{per la ciclicità}$$

In termini matriciali si ha:

$$\underline{\mathbf{y}} = \underline{\mathbf{A}} \underline{\mathbf{x}}$$

dove

$$\underline{\mathbf{A}} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

Calcoliamo ora $\underline{\mathbf{R}}_{\mathbf{y}}$, la matrice di autocovarianza di $\underline{\mathbf{y}}$, circolante per la ciclicità della convoluzione:

$$\underline{\mathbf{R}}_{\mathbf{y}} = E[\underline{\mathbf{y}}\underline{\mathbf{y}}^*] = \underline{\mathbf{A}} \underline{\mathbf{R}}_{\mathbf{x}} \underline{\mathbf{A}}^* = \begin{bmatrix} 2 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 1 & 2 \end{bmatrix}$$

CAPITOLO 11 PROCESSI CASUALI: GENERALITÀ

come si potrebbe ottenere direttamente dalla funzione di autocorrelazione della sequenza $\{y_i\}$, $[1 \ 2 \ 1]$ periodicizzata anch'essa. Il filtraggio passa - basso fatto con la convoluzione con il filtro $\underline{\mathbf{f}}$ comporta uno zero alla frequenza di Nyquist, visto che il numero dei campioni della sequenza è pari. Infatti il vettore $\underline{\mathbf{F}}$ trasformata di Fourier di $\underline{\mathbf{f}}$, è:

$$\underline{\mathbf{F}} = \underline{\mathbf{W}} \underline{\mathbf{f}} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -j & -1 & j \\ 1 & -1 & 1 & -1 \\ 1 & j & -1 & -j \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 1-j \\ 0 \\ 1+j \end{bmatrix}$$

Il vettore $\underline{\mathbf{Y}}$, trasformata di Fourier di $\underline{\mathbf{y}}$ si ottiene moltiplicando il vettore $\underline{\mathbf{X}}$ trasformata di Fourier di $\underline{\mathbf{x}}$ per una matrice $\underline{\Phi}$ che ha lungo la diagonale principale gli elementi di $\underline{\mathbf{F}}$. Si ha cioè:

$$\underline{\mathbf{Y}} = \underline{\mathbf{W}} \underline{\mathbf{y}} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1-j & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1+j \end{bmatrix} \begin{bmatrix} X_0 \\ X_1 \\ X_3 \\ X_4 \end{bmatrix} = \underline{\Phi} \underline{\mathbf{W}} \underline{\mathbf{x}}$$

Si ha ancora:

$$E[\underline{\mathbf{Y}} \underline{\mathbf{Y}}^*] = \underline{\Phi} E[\underline{\mathbf{X}} \underline{\mathbf{X}}^*] \underline{\Phi}^*$$

Peraltro:

$$E[\underline{\mathbf{X}} \underline{\mathbf{X}}^*] = 4\underline{\mathbf{I}}$$

e quindi

$$E[\underline{\mathbf{Y}} \underline{\mathbf{Y}}^*] = \underline{\mathbf{R}}_{\mathbf{Y}} = 4 \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Da un'analisi diretta, è peraltro evidente che la matrice $\underline{\mathbf{R}}_{\mathbf{y}}$ ha un autovalore nullo e l'autovettore corrispondente è la senoide alla frequenza

11.3 MATRICI DI TOEPLITZ: AUTOVALORI

di Nyquist.

$$\underline{\mathbf{R}}_y \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} = 0 \times \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$$

Per avere gli altri autovalori ed autovettori di $\underline{\mathbf{R}}_y$, si puo' anche ricorrere alla trasformazione di Fourier discreta che diagonalizza la matrice, come già osservato. Se

$$\begin{aligned} \underline{\mathbf{Y}} &= \underline{\mathbf{W}}\underline{\mathbf{y}}; \\ E[\underline{\mathbf{Y}}\underline{\mathbf{Y}}^*] &= \underline{\mathbf{W}}\underline{\mathbf{R}}_y\underline{\mathbf{W}}^* = \underline{\mathbf{R}}_Y \\ \frac{\underline{\mathbf{R}}_Y}{N} &= 1 \times \begin{bmatrix} |F_0|^2 & 0 & 0 & 0 \\ 0 & |F_1|^2 & 0 & 0 \\ 0 & 0 & |F_2|^2 & 0 \\ 0 & 0 & 0 & |F_3|^2 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \end{aligned}$$

La matrice $\underline{\mathbf{R}}_y$ è diagonalizzata dalla trasformazione di Fourier poichè:

$$\begin{aligned} \underline{\mathbf{y}} &= \frac{1}{N}\underline{\mathbf{W}}^*\underline{\mathbf{Y}} \rightarrow \underline{\mathbf{W}}^{-1} = \frac{1}{N}\underline{\mathbf{W}}^* \\ \underline{\mathbf{R}}_y &= \frac{1}{N}\underline{\mathbf{W}}^*\underline{\mathbf{R}}_Y\underline{\mathbf{W}}\frac{1}{N} = \frac{1}{\sqrt{N}}\underline{\mathbf{W}}^*\frac{\underline{\mathbf{R}}_Y}{N}\underline{\mathbf{W}}\frac{1}{\sqrt{N}} \end{aligned}$$

Gli autovettori normalizzati della matrice $\underline{\mathbf{R}}_y$ sono appunto le sinusoidi (le righe o le colonne della matrice $\underline{\mathbf{W}}$) la cui ampiezza è $1/\sqrt{N}$ perchè la somma dei quadrati sia unitaria. Pertanto, gli autovalori della matrice $\underline{\mathbf{R}}_y$ sono proprio gli elementi della diagonale principale della matrice $\underline{\mathbf{R}}_Y/N$ (il periodogramma) ed in questo caso:

$$[4, 2, 0, 2]$$

Infatti:

$$\begin{bmatrix} 2 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix} = 4 \times \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ j \\ -1 \\ -j \end{bmatrix} = 2 \times \begin{bmatrix} 1 \\ j \\ -1 \\ -j \end{bmatrix}$$

11.4 Stime di autovalori di matrici di covarianza

La matrice di covarianza si ottiene dall'espressione teorica:

$$\mathbf{R} = E [\mathbf{x} \mathbf{x}^*]$$

Le dimensioni del vettore \mathbf{x}_i sono $[1 \times N]$ e la matrice \mathbf{R} ha quindi dimensioni $[N \times N]$. In realtà poi non è possibile disporre del valore atteso $E[\]$, ma se ne possono solo fare delle stime; in altre parole ci dobbiamo accontentare di una stima non polarizzata $\hat{\mathbf{R}}$ ottenuta combinando M realizzazioni del vettore casuale \mathbf{x}_i ; $i = 1 \dots M$. :

$$\hat{\mathbf{R}} = \frac{1}{M} \sum_{i=1}^M \mathbf{x}_i \mathbf{x}_i^*; \quad \text{Tr}(\hat{\mathbf{R}}) = \sum_{n=1}^N \hat{\lambda}_n = \frac{1}{M} \sum_{i=1}^M \mathbf{x}_i^* \mathbf{x}_i \quad (11.43)$$

$$E[\hat{\mathbf{R}}] = \mathbf{R}; \quad E[\text{Tr}(\hat{\mathbf{R}})] = \frac{M}{M} N \sigma_x^2 = N \sigma_x^2$$

Tutto La stima è ragionevole se:

$$\mu = \frac{N}{M} \lll 1$$

Allora, le stime non sono molto disperse attorno ai valori attesi e il risultato derivante dall'equazione 11.43 è accettabile. In particolare, $\hat{\lambda}_n$, $n = 1..N$, autovalori di $\hat{\mathbf{R}}$, sono delle ragionevoli stime dei λ_n , $n = 1..N$, autovalori di \mathbf{R} , e similmente per gli autovettori corrispondenti. Tuttavia, non sempre è possibile essere in questa situazione e può accadere che invece addirittura sia:

$$\mu > 1$$

11.4 STIME DI AUTOVALORI DI MATRICI DI COVARIANZA

Cosa accade allora? È evidente che $N - M > 0$ autovalori della matrice sono proprio nulli, perchè la matrice $\widehat{\mathbf{R}}$ è somma di M diadi (prodotti $\mathbf{x}_i \mathbf{x}_i^*$) e il suo rango è non superiore a M . Peraltro, la traccia della matrice $\widehat{\mathbf{R}}$ e quindi la somma dei suoi autovalori (contano solo quelli non nulli) $\widehat{\lambda}_n$, ha sempre valore atteso $N\sigma_x^2$ e dispersione ridotta trattandosi di un numero solo risultato della media di MN quadrati di numeri non tanto dipendenti tra loro. Quindi, gli M autovalori non nulli di $\widehat{\mathbf{R}}$ sono *strizzati* ed hanno cioè un valore atteso medio anche molto più grande in quanto si devono spartire tra M valori la potenza totale di N campioni.

$$E[\widehat{\lambda}_n] = \mu\sigma_x^2; n = 1..M; \quad \widehat{\lambda}_n = 0; n = M + 1, \dots, N;$$

In casi intermedi in cui M non è grande come si desidererebbe, la matrice $\widehat{\mathbf{R}}$ tende ad essere comunque deficiente in rango; quindi, gli autovalori stimati più piccoli sono più piccoli del valore corrispondente, mentre il valore atteso della stima dell'autovalore più grande diventa più grande del valore corretto.

$$E[\widehat{\lambda}_{\max}] > \lambda_{\max}; \quad E[\widehat{\lambda}_{\min}] < \lambda_{\min}$$

Per cui, anche se la stima $\widehat{\mathbf{R}}$ della matrice \mathbf{R} non è polarizzata, le stime degli autovalori lo sono. È disponibile il limite della densità di probabilità degli autovalori λ di uno stimatore campionario $\widehat{\mathbf{R}}$ di una matrice di covarianza di un vettore casuale di N dimensioni con M campioni (la distribuzione di Pastur-Marchenko $g(\xi)$, vedi su Wikipedia, ed in figura 11.1); i risultati sono ottenuti nel caso in cui i campioni osservati sono indipendenti a varianza $1/N$ e N, M tendono entrambe ad infinito, con $N/M = \mu$.

$$p_\lambda(\xi) = \left(1 - \frac{1}{\mu}\right) \delta(\xi) + \frac{g(\xi)}{\mu}; \text{ se } \mu > 1 \quad (11.44)$$

$$p_\lambda(\xi) = g(\xi) \quad \text{se } \mu < 1 \quad (11.45)$$

$$g(\xi) = \frac{1}{2\pi\mu\xi} \sqrt{4\mu - [\xi - (1 + \mu)]^2} \quad (11.46)$$

$$(1 - \sqrt{\mu})^2 < \xi < (1 + \sqrt{\mu})^2 \quad (11.47)$$

CAPITOLO 11 PROCESSI CASUALI: GENERALITÀ

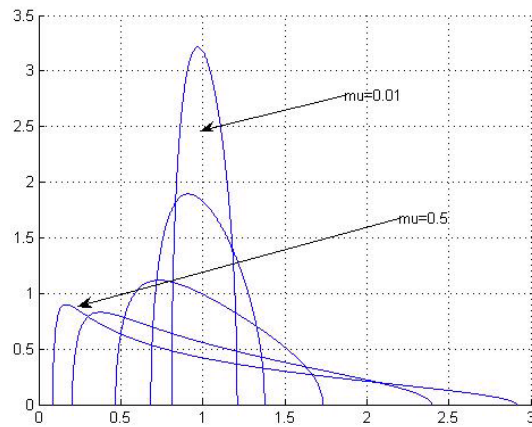


Figura 11.1: Densità di probabilità degli autovalori di matrici casuali per $\mu = 0.01, 0.03, 0.1, 0.3, 0.5$. Se $\mu = 0.01$, p.e., si hanno 100 realizzazioni per ogni variabile casuale, e quindi è ragionevole che la stima della potenza oscilli in un intervallo di circa il 10%. Il caso $\mu = 0.5$ (2 realizzazioni per ogni variabile casuale) corrisponde alla distribuzione delle potenze delle componenti di un periodogramma di una sequenza casuale bianca, con media pari a 1.

11.5 Fluttuazione dei corsi dei titoli di borsa

Il problema della dispersione degli autovalori di matrici causali diventa di notevole interesse nel caso dell'analisi delle fluttuazioni del corso di titoli in borsa, p.e. per la costruzione di portafogli di titoli a rischio minimo. In questi casi, si cerca di verificare se le fluttuazioni del corso di due titoli siano anticorrelate o correlate o incorrelate. Il tipico esempio è quello delle azioni di compagnie di aerolinee e di compagnie petrolifere: se sale il prezzo del petrolio, i titoli delle aerolinee scendono e quelli delle petrolifere salgono, e quindi le fluttuazioni dei corsi dei titoli rispettivi saranno anticorrelate; infatti, ci si attende che i profitti delle une scendano e quelli delle altre salgano. Allora, chi volesse rendersi indipendente dalle fluttuazioni del costo del petrolio, comprerebbe un giardinetto composto da titoli sia delle une che delle altre, in modo da minimizzare quel rischio. Gli autovettori della matrice di covarianza forniscono giardinetti di titoli le cui fluttuazioni sono incorrelate. Poi ci saranno giardinetti a massima fluttuazione (l'autovettore corrispondente all'autovalore massimo) ecc.

Il problema è però che ci sono migliaia di titoli in borsa (N), e pochissimi anni (M) su cui fare delle statistiche, peraltro tempo varianti. Quindi, non si avrebbe comunque nessuna utilità a risalire troppo indietro nel tempo per arricchire le statistiche; ancora, non si ha molta utilità a sovracampionare temporalmente (p.e. rilevando i corsi ogni ora od ogni giorno), perchè poi si ha a che fare con campioni correlati tra loro e scarsamente utili per ridurre la varianza della stima. Gli analisti finanziari sono praticamente sempre nella condizione $M < N$ e quindi devono ricordare bene che le loro stime della matrice di covarianza delle fluttuazioni di migliaia di titoli sono matrici di rango sempre ridottissimo e le stime degli autovalori sono polarizzate. Una soluzione a questo problema è la parametrizzazione e quindi la riduzione della dimensionalità della matrice di covarianza. Si veda p.e. il lavoro: Ledoit, O. and Wolf, M. , *Honey, I Shrunk the Sample Covariance Matrix* (June 2003), UPF Economics and Business Working Paper No. 691). Si possono raggruppare i titoli in comparti (aziende di semiconduttori, supermercati, aerolinee, assicurative, petrolifere, acciai, ecc.) e si studia la matrice di covarianza di dimensioni molto più ridotte che corrisponde alla covarianza delle fluttuazioni medie dei titoli di un comparto rispetto a quelle dei

CAPITOLO 11 PROCESSI CASUALI: GENERALITÀ

titoli di un altro comparto. Poi, all'interno di ogni comparto, si studiano le caratteristiche rispettive delle varie aziende. Se va bene per una, ci si aspetta che vada bene per tutte, e viceversa ... Quindi accade in borsa il fenomeno apparentemente strano per cui se p.e. un'azienda di semiconduttori o di farmaceutici dichiara profitti (o perdite) eccezionali, tutto il comparto ne risente positivamente (o negativamente). Eppure, sarebbe ragionevole indurre che per i meccanismi di concorrenza in un mercato limitato, il valore di un comparto dovrebbe essere bilanciato, per cui se ad un'azienda va meglio, alle altre dovrebbe andare peggio. Ma non è così: solo in momenti successivi sarà compito degli analisti di borsa ricostruire esattamente la posizione concorrenziale di un'azienda rispetto alle altre e ribilanciare i pesi delle aziende nel comparto, stabilendo se un'azienda sia meglio gestita di un'altra, ecc. Peraltro, possono poi nascere compagnie diversificate che operano contemporaneamente in diversi comparti e possono compensare, all'interno dei loro bilanci, profitti e perdite degli uni e degli altri. Questi conglomerati sono essere meno trasparenti e più difficili da studiare e da categorizzare.

11.6 Stima a massima verosimiglianza (Wikipedia)

Il metodo della massima verosimiglianza (in inglese Maximum Likelihood, e in breve ML) in statistica è un procedimento matematico per determinare uno stimatore. Il metodo consiste nel massimizzare la funzione di verosimiglianza, definita in base alla probabilità di osservare una data realizzazione campionaria, condizionatamente ai valori assunti dai parametri oggetto di stima. Il metodo è stato originariamente sviluppato dal genetista e statistico Sir Ronald Fisher, tra il 1912 e il 1922.

11.6.1 Il metodo

Dato un parametro θ , caratterizzato da una distribuzione di probabilità \mathcal{D} , con densità di probabilità \mathcal{L}_D , e dato un campione di dati osservati $\{x_i\}_{i=1}^N$ di dimensione N funzione del parametro θ , si può calcolare la probabilità associata ai dati osservati.

$$P\left(\{x_i\}_{i=1}^N \mid \theta\right) = \mathcal{L}_D\left(\theta \mid \{x_i\}_{i=1}^N\right) \quad (11.48)$$

11.6 STIMA A MASSIMA VEROSIMIGLIANZA (WIKIPEDIA)

Il parametro θ è ignoto, sebbene sia noto che θ è estratto dalla distribuzione \mathcal{D} . Utilizziamo i dati a nostra disposizione, $\{x_i\}_{i=1}^N$, per valutare θ . Il metodo della massima verosimiglianza ricerca il valore più *verosimile* di θ , ossia ricerca, tra tutti i possibili valori di θ , il valore del parametro che *massimizza* la probabilità di aver ottenuto il campione dato. Da un punto di vista matematico, $\mathcal{L}_D(\theta|\{x_i\}_{i=1}^N)$ è detta funzione di verosimiglianza, e lo stimatore ML è ottenuto come:

$$\hat{\theta} = \arg \max_{\theta} \mathcal{L}_D(\theta|x_1, x_2, \dots, x_N) \quad (11.49)$$

Esempio

Si consideri un campione $\{x_i\}_{i=1}^N$ di variabili casuali identicamente e indipendentemente distribuite, con distribuzione gaussiana a media μ e varianza σ^2 , ignote. La funzione di verosimiglianza associata è:

$$\mathcal{L}_D(\mu, \sigma^2|x_1, x_2, \dots, x_N) = \prod_{i=1}^N p(x_i) = \frac{1}{(2\pi)^{N/2} \sigma^N} \exp \left[- \frac{\sum_{i=1}^N (x_i - \mu)^2}{2\sigma^2} \right] \quad (11.50)$$

Massimizzare la funzione di verosimiglianza è come massimizzarne il logaritmo che a meno di costanti è:

$$\log(\mathcal{L}_D(\theta|x_1, x_2, \dots, x_N)) = -N \log \sigma - \frac{\sum_{i=1}^N (x_i - \mu)^2}{2\sigma^2}$$

Le stime dei parametri $\hat{\mu}, \hat{\sigma}$ sono determinate annullando le derivate di $\log(\mathcal{L}_D(\theta|x_1, x_2, \dots, x_N))$ rispetto a $\hat{\mu}, \hat{\sigma}$:

$$\begin{aligned} \frac{\sum_{i=1}^N (x_i - \hat{\mu})}{\hat{\sigma}^2} = 0 &\rightarrow \hat{\mu} = \sum \frac{x_i}{N} \\ -N\hat{\sigma}^2 - \sum_{i=1}^N (x_i - \hat{\mu})^2 = 0 &\rightarrow \hat{\sigma} = \sqrt{\frac{\sum_{i=1}^N (x_i - \sum \frac{x_i}{N})^2}{N}} \end{aligned}$$

Questo esempio illustra alcune proprietà degli stimatori ML. Si osserva che $\hat{\mu}$ è uno stimatore non polarizzato e cioè il suo valore atteso

CAPITOLO 11 PROCESSI CASUALI: GENERALITÀ

coincide con il valore vero.

$$E[\hat{\mu}] = \mu$$

Invece, $\hat{\sigma}$ non gode di tale proprietà. Si verifica che:

$$E[\hat{\sigma}^2] = \frac{N-1}{N}\sigma^2$$

Lo stimatore di massima verosimiglianza è comunque uno stimatore asintoticamente non polarizzato; infatti per $N \rightarrow \infty$:

$$\lim_{N \rightarrow \infty} E[\hat{\sigma}^2] = \sigma^2$$

In particolare, si dimostra che uno stimatore ML è sempre asintoticamente ($N \rightarrow \infty$) non polarizzato e asintoticamente normalmente distribuito.

Consideriamo ora il caso in cui i campioni della sequenza osservata, indicati con y_n , siano correlati, e cioè ottenuti filtrando una sequenza i.i.d. x_n con un filtro f_m . La stima ML della loro varianza sarebbe da effettuarsi usando non direttamente i campioni correlati y_n , ma i campioni *sbiancati* x_n : infatti compare nella funzione di verosimiglianza la seguente funzione dei dati:

$$\underline{\mathbf{y}}^* \underline{\mathbf{R}}_y^{-1} \underline{\mathbf{y}} \sim \underline{\mathbf{x}}^* \underline{\mathbf{x}} \quad (11.51)$$

che abbiamo visto in un paragrafo precedente potersi ottenere filtrando i dati y_n con un filtro g_m la cui caratteristica di ampiezza è inversa a quella del filtro f_m . Incidentalmente, si osservi che la sequenza sbiancata $\underline{\mathbf{x}}$ coincide con quella originaria di ingresso al filtro f_m , ma solo a meno dell'effetto di uno sfasatore puro qualsiasi, che altera la sequenza ma non la sua autocorrelazione. Infine, questo sbiancamento è efficace e perfettamente possibile, ma solo in teoria: nei fatti, la presenza di rumore additivo rende dannoso lo sbiancamento a quelle frequenze dove il livello del segnale attenuato scende sotto il livello del rumore. Lo sbiancamento perfetto è possibile solo se il rumore è assente.

Supponiamo ora che la coloratura del dato sia vista come il risultato di un filtraggio ideale, con riduzione della banda di un fattore $1/L$. Il segnale colorato potrà essere decimato $1:L$, senza perdere informazione, e senza peggiorare la situazione nei riguardi del rumore, almeno in banda passante. La sequenza decimata è una sequenza i.i.d., ma questa volta con N/L campioni, invece di N . Allora, il filtraggio ideale $1:L$ comporta

11.6 STIMA A MASSIMA VEROSIMIGLIANZA (WIKIPEDIA)

la crescita della dispersione della stima di un fattore L , irrecuperabile. Un risultato simile viene trovato nel paragrafo in cui si analizza la dispersione della stima della potenza di una sequenza correlata.

Sarebbe invece molto complessa la trattazione completa ML della stima di una varianza di un segnale, tenendo conto sia del rumore additivo che del filtraggio. Si avrebbero però risultati ottimali, pesando opportunamente il segnale nelle varie bande in funzione della potenza del rumore.

11.6.2 Difficoltà del metodo; stima di una media

Può accadere che il valore dello stimatore ML non appartenga allo spazio dei parametri; lo stimatore può poi non essere unico. Ancora, l'operazione di massimo può essere difficile da effettuare: si pensi al caso di variabili a distribuzione uniforme, o a masse impulsive. Per esempio, si voglia stimare una percentuale di popolazione, p.e. i fumatori. Indicando con p la probabilità di appartenere a tale categoria, considerando una variabile casuale x_i che valga 1 se la persona i -esima fuma e zero altrimenti, la densità di probabilità della variabile x_i è

$$\begin{aligned} p(x) &= (1-p)\delta(0) + p\delta(1); \quad E[x] = p; \quad E[x^2] = p; \\ E[x_i x_j] &= p \text{ se } i = j; \quad E[x_i x_j] = p^2 \text{ se } i \neq j; \end{aligned} \quad (11.52)$$

Una stima non polarizzata di p è facilmente ottenuta come:

$$\hat{p} = \sum \frac{x_i}{N}; \quad E[\hat{p}] = \sum \frac{E[x_i]}{N} = p$$

La varianza di questo stimatore è:

$$\begin{aligned} \text{var}(\hat{p}) &= E[\hat{p}^2] - E[\hat{p}]^2 = \sum_{i,j} \frac{E[x_i x_j]}{N^2} - p^2 = \\ &= \frac{Np + (N^2 - N)p^2}{N^2} - p^2 = \frac{p(1-p)}{N} < \frac{1}{4N} \end{aligned} \quad (11.53)$$

Per avere con una significativa probabilità che l'errore commesso nella valutazione di p sia inferiore p.e. a 1%, si deve ragionevolmente dimensionare N in modo che

$$\frac{1}{\sqrt{4N}} \ll 0.01 \rightarrow N \gg 2500$$

CAPITOLO 11 PROCESSI CASUALI: GENERALITÀ

Poi, si può rendere più precisa questa valutazione almeno dal punto di vista statistico, imponendo la probabilità che l'errore sia inferiore ad un certo valore, come fatto nella sezione 10.2.1. Arrivare a questo semplice risultato dalla stima a massima verosimiglianza sarebbe piuttosto complesso.

11.6.3 La stima lineare MMSE

Un caso molto frequente cui si applicano bene i metodi probabilistici è quello di un modello lineare in presenza di una distribuzione di probabilità a priori dei parametri da stimare. Si abbia quindi che il dato osservato sia il vettore:

$$\underline{\mathbf{y}} = \underline{\mathbf{A}}\underline{\mathbf{x}} + \underline{\mathbf{n}} \quad (11.54)$$

dove $\underline{\mathbf{x}}$, $\underline{\mathbf{n}}$ sono vettori incogniti gaussiani indipendenti (a valor medio nullo, per semplicità) le cui densità di probabilità siano:

$$p(\underline{\mathbf{x}}) = \frac{1}{(2\pi)^{N/2} |\underline{\mathbf{R}}_x|^{1/2}} \exp\left(-\frac{\underline{\mathbf{x}}^* \underline{\mathbf{R}}_x^{-1} \underline{\mathbf{x}}}{2}\right)$$

$$p(\underline{\mathbf{n}}) = \frac{1}{(2\pi)^{N/2} |\underline{\mathbf{R}}_n|^{1/2}} \exp\left(-\frac{\underline{\mathbf{n}}^* \underline{\mathbf{R}}_n^{-1} \underline{\mathbf{n}}}{2}\right)$$

La densità di probabilità di $\underline{\mathbf{x}}$, dato $\underline{\mathbf{y}}$, usando il teorema di Bayes, è:

$$p(\underline{\mathbf{x}}|\underline{\mathbf{y}}) = \frac{p(\underline{\mathbf{y}}|\underline{\mathbf{x}})p(\underline{\mathbf{x}})}{p(\underline{\mathbf{y}})} \quad (11.55)$$

$$p(\underline{\mathbf{x}}|\underline{\mathbf{y}}) = \frac{1}{p(\underline{\mathbf{y}})} \frac{1}{(2\pi)^N (|\underline{\mathbf{R}}_x| |\underline{\mathbf{R}}_n|)^{1/2}} \exp\left[-\frac{(\underline{\mathbf{y}} - \underline{\mathbf{A}}\underline{\mathbf{x}})^* \underline{\mathbf{R}}_n^{-1} (\underline{\mathbf{y}} - \underline{\mathbf{A}}\underline{\mathbf{x}})}{2} - \frac{\underline{\mathbf{x}}^* \underline{\mathbf{R}}_x^{-1} \underline{\mathbf{x}}}{2}\right] \quad (11.56)$$

Cerchiamo la stima che minimizza l'errore di stima quadratico medio (MMSE : Minimum Mean Square Error); si verifica che lo stimatore ottimo da questo punto di vista è la media condizionata, $E(\underline{\mathbf{x}}|\underline{\mathbf{y}})$. Nel caso

11.6 STIMA A MASSIMA VEROSIMIGLIANZA (WIKIPEDIA)

di variabili congiuntamente gaussiane questo valore è quello che massimizza la funzione di probabilità a posteriori $p(\underline{\mathbf{x}}|\underline{\mathbf{y}})$ e quindi troviamo, osservando che $p(\underline{\mathbf{y}})$ non dipende da $\underline{\mathbf{x}}$:

$$\hat{\underline{\mathbf{x}}} = \arg \min_{\underline{\mathbf{x}}} [(\underline{\mathbf{y}} - \underline{\mathbf{A}}\underline{\mathbf{x}})^* \underline{\mathbf{R}}_n^{-1} (\underline{\mathbf{y}} - \underline{\mathbf{A}}\underline{\mathbf{x}}) + \underline{\mathbf{x}}^* \underline{\mathbf{R}}_x^{-1} \underline{\mathbf{x}}] \quad (11.57)$$

Il minimo di questa forma quadratica nel vettore incognito $\underline{\mathbf{x}}$ si trova annullando le derivate vettoriali ed è una funzione lineare di $\underline{\mathbf{y}}$:

$$\begin{aligned} \hat{\underline{\mathbf{x}}} &= \underline{\mathbf{R}}_x \underline{\mathbf{A}}^* (\underline{\mathbf{A}} \underline{\mathbf{R}}_x \underline{\mathbf{A}}^* + \underline{\mathbf{R}}_n)^{-1} \underline{\mathbf{y}} = \\ &= (\underline{\mathbf{A}}^* \underline{\mathbf{R}}_n^{-1} \underline{\mathbf{A}} + \underline{\mathbf{R}}_x^{-1})^{-1} \underline{\mathbf{A}}^* \underline{\mathbf{R}}_n^{-1} \underline{\mathbf{y}} \end{aligned} \quad (11.58)$$

Nel caso particolare in cui:

$$\underline{\mathbf{R}}_x = \sigma_x^2 \underline{\mathbf{I}}, \quad \underline{\mathbf{R}}_n = \sigma_n^2 \underline{\mathbf{I}}$$

si ha:

$$\hat{\underline{\mathbf{x}}} = \underline{\mathbf{A}}^* \left(\underline{\mathbf{A}} \underline{\mathbf{A}}^* + \frac{\sigma_n^2}{\sigma_x^2} \underline{\mathbf{I}} \right)^{-1} \underline{\mathbf{y}} \quad (11.59)$$

congruentemente con quanto si vedrà al capitolo sulle direzioni di arrivo e con i metodi di minimo errore di predizione o di minimo errore quadratico medio. Si osservi che il numero delle componenti del vettore $\underline{\mathbf{x}}$ e del vettore osservato $\underline{\mathbf{y}}$ possono essere diversi e quindi può essere più conveniente l'uso dell'una o dell'altra delle due formulazioni di 11.58. Quindi questa soluzione si presta alla stima di una costante o di un comportamento polinomiale sommati a rumore, o infiniti altri casi.

11.6.4 Stima della potenza di segnali

Campioni indipendenti

Per valutare ora la dispersione della stima del valore quadratico medio di un segnale di cui sono disponibili solo un numero limitato di campioni, eventualmente correlati tra loro, supponiamo di avere una sequenza di M campioni reali x_m che sappiamo essere **indipendenti e gaussiani**, a valor medio nullo e con eguale valore q.m. σ^2 ; la stima a massima verosimiglianza del loro v.q.m é:

$$\hat{S} = \frac{1}{M} \sum_{m=0}^{M-1} x_m^2$$

CAPITOLO 11 PROCESSI CASUALI: GENERALITÀ

\widehat{S} , la stima della varianza della sequenza x_m è una variabile casuale, la cui distribuzione è nota come χ^2 con M gradi di libertà; \widehat{S} è approssimativamente gaussiana, ha valor medio pari a σ^2 ; la dispersione verrà indicata con $\sigma_{\widehat{S}}$. \widehat{S} non può essere gaussiana se non in modo approssimato, perché sempre positiva. Come verrà mostrato dall'equazione (11.64), $\sigma_{\widehat{S}}$ decresce con \sqrt{M} , e quindi lentamente. Calcoleremo ora $\sigma_{\widehat{S}}$ nell'ipotesi che i campioni x_m siano incorrelati;

$$E[\widehat{S}] = \frac{1}{M} E\left[\sum_{m=0}^{M-1} x_m^2\right] = \frac{1}{M} \sum_{m=0}^{M-1} E[x_m^2] = \sigma^2$$

$$\sigma_{\widehat{S}}^2 = E[\widehat{S}^2] - E[\widehat{S}]^2 = \frac{1}{M^2} \sum_{n,m=0}^{M-1} E[x_n^2 x_m^2] - \sigma^4 \quad (11.60)$$

$$\begin{aligned} &= \frac{ME[x_m^4] + (M^2 - M)E_{n \neq m}[x_n^2 x_m^2]}{M^2} - \sigma^4 = \\ &= \frac{3M\sigma^4 + (M^2 - M)\sigma^4}{M^2} - \sigma^4 = \frac{2\sigma^4}{M} \end{aligned} \quad (11.61)$$

dove abbiamo utilizzato la relazione:

$$E[x_1 x_2 x_3 x_4] = E[x_1 x_2]E[x_3 x_4] + E[x_1 x_3]E[x_2 x_4] + E[x_1 x_4]E[x_2 x_3] \quad (11.62)$$

valida per variabili gaussiane reali a v.m. nullo, come sono le $\{x_m\}$. In conclusione

$$E[x_m^4] = 3\sigma^4; E[x_m^2 x_n^2] = \sigma^4 \text{ se } (m \neq n); \quad (11.63)$$

e quindi abbiamo:

$$\sigma_{\widehat{S}} = \sqrt{\frac{2}{M}} \sigma^2 = \sqrt{\frac{2}{M}} E[\widehat{S}] \quad (11.64)$$

Supponiamo ora di voler calcolare il numero M di campioni indipendenti che permette una stima \widehat{S} che, con una probabilità dell' 85%, sia differente dal valore vero σ^2 di una quantità inferiore a $\pm 0.05\sigma^2$. Una variabile casuale gaussiana con varianza σ^2 è contenuta entro un intervallo $\pm k\sigma$ attorno al suo valor medio con una probabilità $\text{erf}(k/\sqrt{2})$; poichè $0.85 = \text{erf}(1.016)$, si ha $k = 1.016 \times \sqrt{2} = 1.43$. Inoltre perchè sia

$$1.43\sigma_{\widehat{S}} = 0.05\sigma^2$$

11.6 STIMA A MASSIMA VEROSIMIGLIANZA (WIKIPEDIA)

è necessario che:

$$\sigma_{\hat{S}} = \sqrt{\frac{2}{M}\sigma^2} = \frac{\sigma^2}{28.6} \rightarrow M \simeq 1635$$

Campioni correlati

Quando i campioni del segnale sono correlati, a meno di sbiancarli, come fatto nello stimatore ML (vedi Appendice 1) la dispersione della stima del loro v.q.m. cresce, perché decresce il numero equivalente di campioni indipendenti che contribuiscono alla stima: *se vogliamo stimare la distribuzione delle età dei cittadini, non è bene raccogliere dati in scuole elementari o case di riposo.* Fare valutazioni quantitative precise comporterebbe molti calcoli; cercheremo qui con un esempio di dare un'idea del problema.

Consideriamo due variabili casuali gaussiane a valor medio nullo ed eguale varianza che siano tra loro correlate, e le indichiamo con x, y . In generale possiamo supporre che sia:

$$y = \mu x + w; \quad E[x^2] = E[y^2] = \sigma^2; \quad (11.65)$$

$$E[w^2] = \sigma^2(1 - \mu^2); \quad E[xw] = 0$$

dove w è una variabile casuale incorrelata con x . Risulta poi, utilizzando la solita relazione

$$\begin{aligned} E[x^4] &= 3\sigma^4 \\ E[xy] &= \mu\sigma^2; \quad E[x^2y^2] = (1 + 2\mu^2)\sigma^4 \end{aligned} \quad (11.66)$$

Consideriamo ora una sequenza $\{x_m\}$ di campioni **correlati** tale che:

$$E[x_n x_m] = \mu\sigma^2; \quad n \neq m; \quad E[x_m^2] = \sigma^2$$

Ritornando ora alla stima del v.q.m. della sequenza:

$$\hat{S} = \frac{1}{M} \sum_{m=0}^{M-1} x_m^2$$

congruentemente con l'equazione (11.60) essa sarà caratterizzato dalla varianza

$$\sigma_{\hat{S}} = \sigma^2 \sqrt{\frac{2}{M}(1 - \mu^2) + 2\mu^2}$$

CAPITOLO 11 PROCESSI CASUALI: GENERALITÀ

e la varianza **NON** decresce a zero anche se M tende ad ∞ . Supponiamo ora che i campioni corrispondano ad una sequenza bianca filtrata con un filtro passabasso con un polo in $z = \rho = 1 - \varepsilon$, e quindi di banda $B \sim \varepsilon/2\pi T$. Nel dominio z :

$$E[|Y|^2] = \frac{E[|X|^2]}{(1 - \rho z^{-1})(1 - \rho z)}$$

L'autocorrelazione dell'uscita è:

$$E[y_m y_{m+k}] = \rho^{|k|} \sigma^2; \quad \rho \leq 1$$

La stima \hat{S} del v.q.m. σ^2 della sequenza è sempre ottenuta come:

$$\hat{S} = \frac{1}{M} \sum_{m=0}^{M-1} x_m^2$$

Dopo qualche calcolo si trova che:

$$E[\hat{S}] = \sigma^2; \quad \sigma_{\hat{S}} \simeq \sigma^2 \sqrt{\frac{2}{M}} \times \sqrt{\frac{1 + \rho^2}{1 - \rho^2}}$$

Avendo quindi da stimare la varianza di una serie di M campioni complessi non indipendenti, ma correlati con un decadimento esponenziale della correlazione di ragione ρ , è quindi come avere a disposizione un numero equivalente inferiore di campioni indipendenti, pari a:

$$M_{eq} \simeq M \frac{1 - \rho^2}{1 + \rho^2}; \text{ se } \rho = 1 - \varepsilon \rightarrow M_{eq} \sim M\varepsilon$$

coerentemente con il fatto che se campionassimo l'uscita del filtro a frequenza $2B = \varepsilon/\pi T$ avremmo, dopo un tempo MT , $2M\varepsilon/\pi$ campioni complessi approssimativamente indipendenti. Nell'Appendice 1 si considera il caso simile in cui i campioni sono filtrati da un filtro passabasso ideale di banda f_c/L e poi decimati $1:L$. In questa approssimazione:

$$L_{eq} = \frac{f_c}{2B} = \frac{1}{2BT} = \frac{\pi}{\varepsilon}; \quad \frac{M}{L_{eq}} = \frac{2}{\pi} M\varepsilon \text{ piuttosto che } M\varepsilon$$

11.7 LE DIREZIONI DI ARRIVO

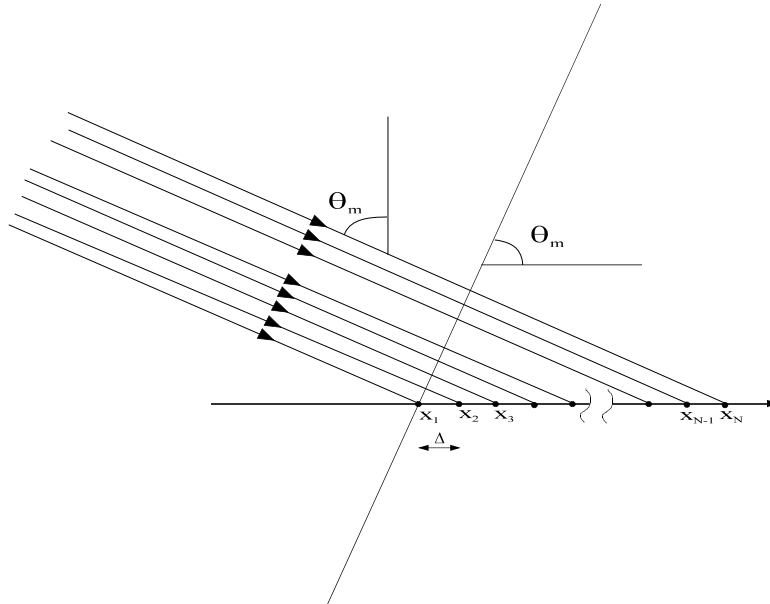


Figura 11.2: Schiera di sensori

11.7 Le direzioni di arrivo

Accanto a sequenze in cui l'indice rappresenta l'asse temporale, possiamo avere sequenze il cui indice si riferisce a posizioni nello spazio. Nel problema delle Direzioni di Arrivo (Directions of Arrival: DOA), p.e. si suppone che segnali temporalmente monofrequenziali (o a spettro molto limitato) illuminano una schiera di N sensori.

Si dispongono N sensori p.e. lungo un cavo, e siano $\{x_{n,k}\}$, $n = 1, \dots, N$, i segnali da essi misurati all'istante k -esimo. La distanza tra il sensore n -esimo e l'inizio del cavo sia $d_n = (n - 1)\Delta$; supponiamo che esistano M sorgenti di radiazioni, monofrequenziali o a banda stretta, a grande distanza dal cavo; l'angolo tra la direzione della sorgente m ; $m = 1 \dots M$ e la normale al cavo (vedi fig. 17.1) è indicato con θ_m ; quindi la radiazione arriva al ricevitore x_n con il ritardo $\tau_n = (n - 1) \times \Delta \sin \theta_m / c$ rispetto al ricevitore $n = 1$. Inoltre, indichiamo con $A_m e^{j\omega k T}$ l'ampiezza sinusoidale della radiazione all'istante k , e con c la velocità di propagazione delle onde. Il segnale misurato dal sensore n all'istante k è

CAPITOLO 11 PROCESSI CASUALI: GENERALITÀ

quindi (per semplicità abbiamo posto $M = 1$):

$$x_{n,k} = A \exp(j\omega kT + j \frac{(n-1)\omega\Delta \sin\theta_m}{c}) = A e^{j\omega kT} \times e^{jk_m(n-1)\Delta} = A e^{j\omega kT} \times e^{ju_m(n-1)}; \quad (11.67)$$

$$k_m = \frac{\omega \sin\theta_m}{c}; \quad u_m = k_m\Delta$$

Il segnale misurato nella stessa posizione al variare del tempo è una sinusoidale complessa di pulsazione ω ; il segnale misurato allo stesso istante kT da tutta la schiera di sensori (al variare della posizione nello spazio) è una **sinusoidale complessa**, nel dominio degli spazi, avente **numero d'onda** k_m (per le sinusoidi spaziali, è la grandezza corrispondente alla pulsazione delle sinusoidi temporali). Misurare la direzione di arrivo equivale a misurare il numero d'onda della sinusoidale complessa registrata. La pulsazione normalizzata $u_m = k_m\Delta$ è la rotazione di fase del segnale m -esimo al passare da un ricevitore della schiera al successivo.

Fare un'analisi dello spettro di frequenza (spaziale) della sequenza $\{x_{n,k}\}$, per un dato istante di tempo k , corrisponde quindi a determinare numero e direzione di arrivo delle radiazioni incidenti sul cavo. La distanza massima tra i sensori per evitare equivocazione spaziale è:

$$|k| = \frac{\omega \sin\theta}{c} \leq k_{Nyquist} = \frac{\pi}{\Delta}; \quad \Delta \leq \frac{c}{2f} = \frac{\lambda}{2} \quad (11.68)$$

Per esempio, con segnali a $20kHz$, se $c = 1500m/sec$, allora $\Delta \leq 0.0375m$. Scegliendo questo valore per Δ avremo che alla direzione $\theta = 0$ corrisponde $k = 0$; alle direzioni $\theta = \pm\pi/2$ corrisponde $k = \pm k_{Nyquist}$

11.7.1 Autocorrelazione di una sequenza spaziale

Il segnale in arrivo sia caratterizzato da un numero discreto M di fronti d'onda, ognuno con la sua ben definita direzione di arrivo. La sequenza $\{x_i\}$ sarà somma di M sinusoidi complesse corrispondenti agli M fronti. Le loro pulsazioni spaziali (i numeri d'onda) sono:

$$k_m = \frac{\omega \sin\theta_m}{c}; \quad u_m = k_m\Delta \quad (11.69)$$

Supponiamo che $M = 2$, l'ampiezza della radiazione A_m è una variabile casuale complessa di varianza $E[|A_m|^2] = \sigma_m^2$ e che sui dati è sovrapposto

del rumore bianco w_n , complesso, di varianza σ_w^2 . Si ha:

$$x_{n,k} = e^{j\omega k T} (A_1 e^{ju_1(n-1)} + A_2 e^{ju_2(n-1)}) + w_n$$

e i campioni della funzione di autocorrelazione spaziale della sequenza $\{x_{n,k}\}$ sono indipendenti da k e sono:

$$E[x_n x_{n-h}^*] = r_h = r_{-h}^* = \sigma_1^2 e^{ju_1 h} + \sigma_2^2 e^{ju_2 h} + \sigma_w^2 \delta(h); 0 \leq h \leq N \quad (11.70)$$

Si osservi che possiamo stimare con precisione i campioni della funzione di autocorrelazione facendo delle operazioni di media **nel tempo**; non abbiamo necessit  di fare delle medie nello spazio. Il numero dei campioni della funzione di autocorrelazione, misurati con la precisione desiderata, corrisponde al numero dei sensori.

11.8 Appendici

11.8.1 Autocorrelazione e spettri per variabili continue

La definizione dell'autocorrelazione nel dominio del tempo continuo per segnali complessi  :

$$r(\tau) = E[x(t)x^*(t-\tau)] = E[x^*(t)x(t+\tau)]; \quad r(-\tau) = r^*(\tau)$$

Per definizione, lo spettro di potenza $S(\omega)$  :

$$S(\omega) = \int_{-\infty}^{\infty} r(\tau) e^{-j\omega\tau} d\tau$$

Infatti, considerando un intervallo di tempo di osservazione pari a T_{tot} :

$$X(\omega) = \int_{-T_{tot}/2}^{T_{tot}/2} x(t) e^{-j\omega t} dt; \quad X^*(\omega) = \int_{-T_{tot}/2}^{T_{tot}/2} x^*(t) e^{j\omega t} dt$$

$$|X(\omega)|^2 = \int_{-T_{tot}/2}^{T_{tot}/2} \int_{-T_{tot}/2}^{T_{tot}/2} x(t) x^*(\xi) e^{j\omega\xi} e^{-j\omega t} dt d\xi$$

Ponendo $t - \xi = \tau$;

$$|X(\omega)|^2 = \int_{-T_{tot}/2}^{T_{tot}/2} d\xi \int_{-T_{tot}/2-\xi}^{T_{tot}/2-\xi} x(\xi + \tau) x^*(\xi) e^{-j\omega\tau}$$

CAPITOLO 11 PROCESSI CASUALI: GENERALITÀ

Passando alla media di insieme:

$$\begin{aligned} E[|X(\omega)|^2] &= \int_{-T_{tot}/2}^{T_{tot}/2} d\xi \int_{-T_{tot}/2-\xi}^{T_{tot}/2-\xi} r(\tau) e^{-j\omega\tau} d\tau \simeq \\ &\simeq S(\omega) \int_{-T_{tot}/2}^{T_{tot}/2} d\xi = S(\omega) \times T_{tot} \end{aligned}$$

perchè si può pensare che $r(\tau)$ si annulli già per valori dell'argomento ben inferiori a T_{tot} . Volendo una maggiore precisione, si può tenere conto della finestatura in tempi (T_{tot}) con una finestatura triangolare della funzione di autocorrelazione (vedi eq. 10.23). Passando al limite per $T_{tot} \rightarrow \infty$:

$$\lim_{T_{tot} \rightarrow \infty} \frac{E[|X(\omega)|^2]}{T_{tot}} = S(\omega)$$

Lo spettro di energia è per il teorema di Parseval e per un tempo T_{tot} limitato:

$$\int_{-\infty}^{\infty} |X(\omega)|^2 \frac{d\omega}{2\pi} = \int_{-T_{tot}/2}^{T_{tot}/2} |x(t)|^2 dt = T_{tot} P_m$$

dove con P_m abbiamo indicato la potenza media (il valore quadratico medio) del segnale $x(t)$. Quindi

$$\int_{-\infty}^{\infty} \frac{|X(\omega)|^2}{T_{tot}} \frac{d\omega}{2\pi} = \int_{-\infty}^{\infty} S(\omega) \frac{d\omega}{2\pi} = \frac{1}{T_{tot}} \int_{-T_{tot}/2}^{T_{tot}/2} |x(t)|^2 dt = P_m$$

e quindi $S(\omega)$ corrisponde allo spettro di potenza del segnale e cioè alla ripartizione della potenza del segnale nelle varie bande spettrali. Si osservi che $S(\omega)$ non è necessariamente simmetrico rispetto ad ω in quanto il segnale $x(t)$ può essere complesso.

11.8.2 Dimensione di variabili continue e discrete

Se i segnali $x(t)$ sono delle tensioni, indicheremo tale dimensionalità con:

$$x(t) \rightarrow [V];$$

Nel discreto

$$x_n = x(nT) \rightarrow [V]$$

Le trasformate continue hanno dimensionalità:

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt \rightarrow [VT]$$

Le trasformate discrete hanno invece la stessa dimensionalità del segnale

$$X_k = \sum_{n=0}^{N-1} x_n W_N^{-nk} \rightarrow [V]$$

e pertanto i campioni della trasformata continua si ottengono da quelli della trasformata discreta moltiplicando per T , sempre se il campionamento è corretto:

$$X(\omega_k) = \int_{-\infty}^{\infty} x(t)e^{-j\omega_k t} dt \simeq T \sum_{n=0}^{N-1} x_n W_N^{-nk} = X_k T$$

La funzione di autocorrelazione ha dimensione $[V^2]$ e cioè è una potenza (trascuriamo il termine dovuto alla resistenza $[\Omega^{-1}]$). Infine lo spettro di potenza di una variabile continua $S(\omega)$

$$S(\omega) = \frac{|X(\omega)|^2}{T_{tot}} \rightarrow \left[\frac{V^2 T^2}{T_{tot}} \right] = [V^2 T]$$

Nel discreto, il periodogramma è collegato allo spettro di potenza dalla relazione:

$$S(\omega_k) = \frac{E[|X(\omega_k)|^2]}{T_{tot}} \simeq \frac{|X_k|^2 T^2}{NT} = \frac{T}{N} |X_k|^2$$

11.8.3 Teorema del limite centrale

Il teorema del limite centrale è una conseguenza immediata del meccanismo delle convoluzioni ripetute; si consideri una variabile casuale x_i , caratterizzata da avere solo due valori possibili ed equiprobabili con probabilità 0.5:

$$x_i = \pm \frac{1}{\sqrt{N}}; \quad p_x(x_i = 1/\sqrt{N}) = p_x(x_i = -1/\sqrt{N}) = .5$$

La trasformata di Fourier della densità di probabilità é:

$$P(u) = \cos\left(\frac{u}{\sqrt{N}}\right)$$

CAPITOLO 11 PROCESSI CASUALI: GENERALITÀ

La variabile casuale a , somma di N valori della x_i , ha media nulla e varianza unitaria; infatti:

$$a = \sum_{i=1}^N x_i; \quad E[a^2] = N \sum E[x_i^2] = 1; \quad E[x_i x_j] = 0; i \neq j$$

La densità di probabilità di a è approssimativamente gaussiana; essa è ottenuta convolvendo N volte con se stessa la densità di probabilità $p_x(x)$. La sua trasformata di Fourier è quindi la potenza N -esima della $P(u)$:

$$P_a(u) = [P(u)]^N = \cos^N \left(\frac{u}{\sqrt{N}} \right)$$

$$\cos^N \left(\frac{u}{\sqrt{N}} \right) \simeq \left(1 - \frac{u^2}{2N} \right)^N \simeq e^{-u^2/2}; \quad u \ll \sqrt{N}$$

L'antitrasformata di Fourier, gaussiana a varianza unitaria è quindi:

$$p_a(\xi) = \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2}$$

Poichè la variabile x assume solo due valori discreti, anche la variabile a è discreta, a passo $1/\sqrt{N}$; infatti, la sua trasformata $P_a(u)$, essendo trasformata di una variabile campionata, è periodica di periodo $2\pi\sqrt{N}$ e quindi il periodo cresce con N . Peraltro, l'approssimazione fatta ponendo $u \ll \sqrt{N}$, implica un filtraggio passa - basso e quindi lo smussamento della distribuzione di a e quindi ci impone di trascurare il fatto che essa assuma solo dei valori discreti ad intervalli $1/\sqrt{N}$.

Capitolo 12

Analisi spettrale non parametrica

12.1 Introduzione

In questa seconda parte del testo, affronteremo il problema dell'analisi spettrale dei segnali e cioè la stima del loro spettro di potenza.

Il problema della stima spettrale é molto vasto ed ha applicazioni molto diverse tra loro. Per introdurre il problema, si può pensare allo studio dello spettro delle radiofrequenze in una grande città; l'uno accanto all'altro si incontrano i vari canali radio e televisivi, i segnali di telefonia cellulare, i segnali radar ecc. Gli intervalli di frequenza, se considerati con sufficiente larghezza di banda, sono occupati da segnali totalmente indipendenti tra loro, con diversa ampiezza e diverse caratteristiche statistiche. Per analizzare lo spettro, appare logico utilizzare un ricevitore selettivo che misura il contenuto energetico in ogni intervallo di frequenza. Cercheremo la stima piú precisa possibile di queste energie nell'intervallo di tempo disponibile senza fare alcuna ipotesi ulteriore, non cercando modelli della generazione del segnale (**analisi spettrale non parametrica**).

Un caso ben diverso si ha ad esempio considerando lo spettro di un segnale acustico dovuto a vibrazioni o ad un segnale vocale. In questo caso, tutto il segnale nel suo complesso ha origini uniche e quindi vi saranno delle relazioni tra i contenuti delle varie bande spettrali. Un caso estremo poi di analisi spettrale é quello della ricerca della frequenza

CAPITOLO 12 ANALISI SPETTRALE NON PARAMETRICA

di un oscillatore o delle direzioni di arrivo di onde piane su una schiera di sensori. In questi casi, si deve anzitutto scegliere un modello per la generazione del segnale e quindi determinare i parametri del modello stesso. In questo modo, si dá la descrizione dello spettro del segnale piú accurata possibile. Ad esempio, si cercheranno i parametri di un filtro lineare che, alimentato da un segnale a spettro uniforme (rumore bianco) produce uno spettro di potenza simile allo spettro sotto analisi. Questa tecnica (**analisi spettrale parametrica**) ha numerosissime applicazioni: studiare le vibrazioni di un sistema, studiare la stabilitá di frequenza di un oscillatore, stimare posizione e numero delle sorgenti di segnale in un campo acustico, stimare i parametri del tratto vocale di un parlatore, e in generale stimare e predire segnali nel tempo o nello spazio. Ovviamente, il successo della tecnica parametrica dipende dalla qualità e dalla correttezza del modello scelto; modelli validi portano ad una descrizione **parsimoniosa** del segnale, caratterizzata cioè dal minimo numero di parametri necessario; come vedremo, ciò porterá ad una migliore stima di questi parametri e quindi a risultati ottimali. Infine, e questo é poi uno dei risultati piú importanti, l'analisi spettrale parametrica porta all'identificazione del modello e questo apre una successiva fase di studio per comprendere e poi eventualmente controllare lo stato e l'evoluzione del sistema sotto osservazione. L'analisi spettrale non parametrica é una questione concettualmente semplice se si fa uso del concetto di media di insieme: infatti, se si hanno abbastanza realizzazioni del segnale, basta calcolarne le trasformate discrete di Fourier e mediane le potenze, componente per componente. Tuttavia, raramente si hanno numerose repliche del segnale; spesso, si ha a disposizione una sola replica, per un intervallo di tempo assegnato la cui lunghezza indicheremo sistematicamente con $T_{tot} = NT$ dove T é l'intervallo di campionamento. Per determinare lo spettro di potenza, si dovrá ricorrere ad ulteriori ipotesi come stazionarietá ed ergodicitá che supporremo sempre valide, tranne quando verranno discussi i sistemi adattativi, in cui la non stazionarietá delle statistiche verrá considerata esplicitamente.

12.2 Analisi spettrale non parametrica

Volendo stimare lo spettro di una sequenza $\{x_n\}$, $n = 0, \dots, N - 1$ senza fare alcuna ipotesi addizionale, é necessario utilizzare un banco di K filtri

12.2 ANALISI SPETTRALE NON PARAMETRICA

passa banda con bande $B_k; k = 1, \dots, K$ e frequenze centrali $f_k; k = 1, \dots, K$ attraverso i quali fare transitare il segnale; si stimano poi le potenze medie delle uscite $y_{n,k}$ dei vari filtri del banco (fig. 10 .1). Emerge il problema di determinare la dispersione della stima spettrale e cioè la dispersione della stima \hat{S}_k della potenza (il valore quadratico medio) σ_k^2 dei segnali $y_{n,k}$ in uscita dei vari filtri.

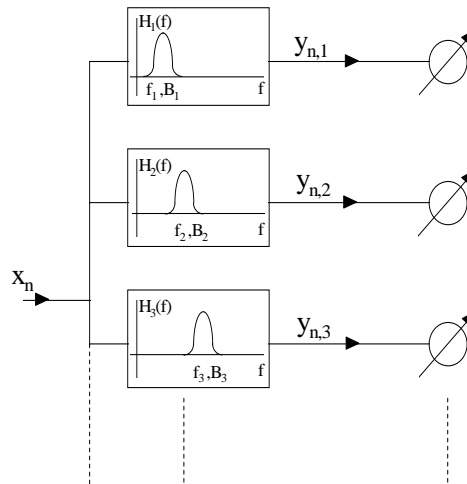


Figura 12.1: Stima dello spettro di un segnale

Questa stima è funzione sia del tempo totale di osservazione concesso $T_{tot} = NT$ sia delle larghezze di banda B_k dei vari filtri (risoluzione spettrale), che non sono necessariamente eguali tra loro. Vedremo che dispersione della stima e risoluzione spettrale sono entrambe limitate dal valore di T_{tot} ed inoltre al migliorare della risoluzione spettrale peggiorerà (crescerà) la dispersione di \hat{S}_k , $\sigma_{\hat{S}_k}$. Infatti, se la banda passante B_k del filtro k -esimo è molto stretta, i campioni (complessi) della sua uscita sono correlati tra loro per un intervallo di tempo lungo $\Delta T_k \simeq 1/B_k$. Quindi, il tempo totale necessario per misurare **all'uscita del filtro** un numero di campioni complessi indipendenti M abbastanza grande (per avere una buona stima del loro valore quadratico medio e quindi della densità spettrale media in quella banda) sarà:

$$T_{tot} = NT = M\Delta T_k = M/B_k$$

e sarà tanto maggiore per quanto più ridotta è B_k e per quanto maggiore è M . A pari T_{tot} invece, dovremo scegliere se fare grande M (stima precisa) o ΔT_k (elevata risoluzione spettrale).

12.3 Risoluzione spettrale e dispersione

12.3.1 La dispersione della stima spettrale

Consideriamo ora il necessario compromesso da raggiungere tra la risoluzione spettrale e la dispersione della stima se la sequenza di cui facciamo la stima spettrale è lunga N campioni e cioè è osservata in un intervallo di tempo limitato $T_{tot} = NT$. Infatti sia B la risoluzione spettrale e quindi B sia la banda passante di ogni filtro da utilizzare per la stima dello spettro di potenza del segnale attorno ad una qualsiasi frequenza assegnata: in altre parole, la risoluzione spettrale non dipende dalla frequenza. Ripetendo in modo leggermente diverso il ragionamento fatto alla fine del paragrafo precedente, convertiamo in banda base il segnale passa banda uscita del filtro k -esimo. Centriamo quindi attorno alla frequenza 0 le componenti alle frequenze positive del segnale reale di partenza, eliminando del tutto quelle alle frequenze negative. Otteniamo così un segnale complesso con frequenza massima $B/2$; se la banda B è piccola, sarà ragionevole ritenere uniforme lo spettro del segnale convertito in questa banda. Campionando la parte reale e la parte immaginaria del segnale complesso con l'intervallo di tempo $\Delta T = 1/B$ e quindi alla frequenza $B = 2 \times B/2$, otteniamo due sequenze di campioni, tutti incorrelati tra loro (lo spettro è uniforme) e di pari v.q.m. Inoltre, poichè il filtro passa banda è a banda stretta e quindi combina linearmente tra loro numerosi campioni del segnale di ingresso campionati ad intervallo $T \ll \Delta T$, il segnale filtrato è gaussiano per il teorema del limite centrale e pertanto l'incorrelazione tra i campioni implica anche la loro indipendenza. In ogni modo, **abbiamo per ogni banda B , $M = 2BT_{tot}$ campioni incorrelati.**

Quindi, utilizzando sia le parti reali che quelle immaginarie per la stima \hat{S} , utilizzando il risultato del capitolo precedente per la dispersione della stima, troviamo che:

$$\frac{\sigma_{\hat{S}}}{E(\hat{S})} = \frac{\sigma_{\hat{S}}}{\sigma^2} = \sqrt{\frac{2}{M}} = \sqrt{\frac{2}{2BT_{tot}}} = \frac{1}{\sqrt{BT_{tot}}} \quad (12.1)$$

12.3 RISOLUZIONE SPETTRALE E DISPERSIONE

Pertanto, se la risoluzione spettrale migliora (B decresce), aumenta anche la dispersione della stima; poiché però dobbiamo finestrare i dati nell'intervallo T_{tot} la risoluzione spettrale non può comunque essere maggiore di $1/T_{tot}$ a causa della finestatura rettangolare che comporta la convoluzione con dei dati con $W_r(\omega) = \text{sinc}(fT_{tot})$.

12.3.2 Polarizzazione della stima spettrale dovuta al troncamento dei dati

Ancora, il troncamento dei dati nell'intervallo di osservazione e quindi la convoluzione della trasformata con la trasformata della finestra comporta **la polarizzazione della stima**. Con questo termine si indica il fatto che:

$$E[\widehat{S}_k] \neq \sigma_k^2 = S(\omega_k)B_k \quad (12.2)$$

e cioè che il valore medio del risultato dello stimatore è diverso dal valore medio della variabile casuale. Infatti, la presenza di componenti spettrali al di fuori della banda osservata B_k (per esempio una sinusoidale di grande ampiezza ad un'altra frequenza), a causa del troncamento con la finestra temporale adottata e quindi della convoluzione del segnale con la sua trasformata $W_r(\omega)$, crea infiltrazioni spurie di energia (leaks) nella banda B_k : risulta infatti che la trasformata di Fourier del segnale finestrato è:

$$X_f(\omega) = X(\omega) * W_r(\omega)$$

$$\begin{aligned} E[|X_f|^2] &= E\left[\int \int X(\psi) W_r(\omega - \psi) X^*(\eta) W_r^*(\eta - \omega) d\psi d\eta\right] = \\ &= \int \int W_r(\omega - \psi) W_r^*(\eta - \omega) E[X(\psi) X^*(\eta)] d\psi d\eta = \end{aligned}$$

e poichè, come visto nel capitolo precedente nel caso di segnali campionati:

$$E[X(\psi) X^*(\eta)] = E[|X(\psi)|^2] \delta(\psi - \eta)$$

$$E[|X_f(\omega)|^2] = \int |W_r(\omega - \psi)|^2 E[|X(\psi)|^2] d\psi$$

risulta che lo spettro del segnale finestrato è la convoluzione dello spettro originale con il modulo quadro della trasformata della finestra

$$S_f(\omega) = S(\omega) * |W_r(\omega)|^2$$

CAPITOLO 12 ANALISI SPETTRALE NON PARAMETRICA

e, nel caso discreto,

$$E[\widehat{S}_k] = B_k \times (S(\omega) * |W_r(\omega)|^2)_{\omega=\omega_k}$$

La risoluzione spettrale peggiora se si usano finestre rastremate a zero, p.e. triangolari o cosinusoidali così da evitare che i lobi laterali della $|W_r(\omega)|^2$, a causa di componenti spettrali esterne all'intervallo di frequenza B_k , vi iniettino energia spuria e creino troppa polarizzazione. Queste finestre rastremate diminuiscono la polarizzazione (i lobi laterali di ampiezza inferiore che con la finestra rettangolare) ma riducono la risoluzione spettrale. Il lobo centrale della trasformata della finestra avrà infatti larghezza maggiore che non quello corrispondente alla finestra rettangolare di eguale durata. Inoltre, la rastremazione della finestra porta a pesare poco dei dati e quindi ad accrescere la dispersione della stima.

Si osservi che la finestra che deve essere applicata ai dati per delimitarli temporalmente può essere qualsiasi, purchè abbia bassi lobi laterali. Vedremo invece più avanti che, quando finestreremo la funzione di autocorrelazione allo scopo di ridurre la dispersione della stima, avremo bisogno di finestre la cui trasformata di Fourier è sempre positiva.

12.4 Il periodogramma

Se abbiamo a disposizione solo un intervallo di osservazione dei dati di durata T_{tot} , abbiamo visto che il più piccolo intervallo di frequenza (che nel seguito indicheremo anche come cella) su cui si può avere una stima spettrale è:

$$B_{min} = \frac{1}{T_{tot}} = \frac{2}{N} f_N$$

Questa risoluzione spettrale (la massima possibile) si ottiene utilizzando una finestra rettangolare prima del filtraggio passa banda del segnale. Dalla formula precedentemente vista, risulta che in questo caso:

$$\frac{\sigma_{\widehat{S}}}{E(\widehat{S})} = \sqrt{\frac{1}{B_{min} T_{tot}}} = 1 \quad (12.3)$$

e cioè la dispersione della stima non dipende dalla lunghezza dei dati; **allungando la serie dei dati e volendo sempre la massima risoluzione**

spettrale, la dispersione della stima non decresce perchè contemporaneamente cresce la risoluzione spettrale. In questo caso, infatti, il numero M di campioni indipendenti usato per la stima dello spettro in ogni intervallo di frequenza è sempre 2. Supponiamo ora che i filtri passa banda vengano realizzati con filtri a campionamento in frequenza, complessi, con frequenze centrali appunto

$$f_k = \frac{k}{T_{tot}}; \quad k = -\frac{N}{2} + 1, \dots, \frac{N}{2}$$

La risposta all'impulso del filtro k -esimo, complesso, è:

$$h_k(n) = e^{j2\pi knT/T_{tot}} = e^{j2\pi kn/N} = W_N^{nk}; \quad n = 0, \dots, N-1 \quad (12.4)$$

$$W_N = e^{j2\pi/N}$$

L'ingresso del filtro è la sequenza $\{x_n\}$, $n = 0, \dots, N-1$; l'uscita del filtro, complessa, è utilizzabile solo all'istante $n = N-1$ quando il transitorio iniziale di caricamento dei dati nel filtro è terminato, ma anche i dati sono terminati. In sostanza, con la sequenza di dati lunga N , per ogni filtro alla frequenza f_k disponiamo di solo un campione in uscita, complesso. Esso è:

$$X_k = \sum_{n=0}^{N-1} x_n W^{-nk} \quad (12.5)$$

$$P_k = \frac{|X_k|^2}{N} \quad (12.6)$$

Abbiamo visto nel capitolo precedente che gli X_k , uscite dei vari filtri passa banda che costituiscono il banco, sono sostanzialmente indipendenti tra loro, a garanzia che la risoluzione spettrale sia effettivamente $1/NT$. Determinare le N uscite X_k del banco degli N filtri passa banda corrisponde poi al calcolo della trasformata di Fourier discreta della sequenza.

L'unico campione a nostra disposizione dell'uscita del k -esimo filtro passa banda, il campione complesso X_k , ha potenza $|X_k|^2$; questa, divisa per N , è la stima P_k della densità spettrale del segnale in quella banda; questa stima spettrale viene indicata con il nome di **periodogramma**.

CAPITOLO 12 ANALISI SPETTRALE NON PARAMETRICA

Con il periodogramma abbiamo raggiunto la massima risoluzione possibile della stima spettrale e anche abbiamo rinunciato totalmente a migliorarne la dispersione: infatti, per fare la stima di energia usiamo solo $M = 2$ dati indipendenti, e cioè le parti reale ed immaginaria di $X(k)$. Il periodogramma perciò ha massima risoluzione spettrale e massima dispersione della stima. Allungando T_{tot} , aumenta la risoluzione spettrale ma non migliora la dispersione. Peraltro, calcolare il periodogramma è molto poco costoso.

In figura 10.4 sono presentati gli spettri di potenza, calcolati in vari modi, corrispondenti ad un processo costituito da una sinusoidale a frequenza 125 Hz ($f_c = 1000Hz$) cui è aggiunta una sequenza di campioni indipendenti (rumore bianco) convoluti con la sequenza $(-0.5, 0, 1, 0, -0.5)$ e quindi caratterizzata dallo spettro:

$$S(\omega) = [1 - \cos(2\omega T)]^2$$

In figura 10.4a è presentato il periodogramma calcolato su 2048 campioni. Si osservi il forte picco causato dalla sinusoidale (tutta la potenza cade in una sola cella spettrale) e la forte dispersione della stima alle frequenze a cui lo spettro del segnale è diverso da zero. Dove invece l'ampiezza spettrale è piccola, anche la dispersione è piccola.

12.5 La stima dell'autocorrelazione

Stimare la densità spettrale equivale a stimare la funzione di autocorrelazione del processo: infatti l'una sarà la trasformata di Fourier dell'altra. Con autocorrelazione definiamo la seguente funzione,

$$r(\tau) = E[x_n^* x_{n+\tau}]; \quad r(-\tau) = r^*(\tau)$$

indipendente dall'indice n ; il processo causale $\{x_n\}$ è a media nulla e può essere complesso. La trasformata di Fourier dell'autocorrelazione è quindi lo spettro di potenza:

$$E \frac{[X(\omega)X^*(\omega)]}{N} \longleftrightarrow \frac{1}{N} \sum_{n=0}^{N-1} E[x_n^* x_{n+\tau}]$$

in quanto è la trasformata del valore atteso della convoluzione tra la sequenza e la stessa sequenza **ribaltata nel tempo e coniugata**.

12.5 LA STIMA DELL'AUTOCORRELAZIONE

E' quindi interessante individuare quale sia la stima dell'autocorrelazione corrispondente al periodogramma. La risposta e' semplice: per definizione il periodogramma e' l'insieme di variabili casuali:

$$P_k = \frac{|X_k|^2}{N} = \frac{1}{N} \left(\sum_{n=0}^{N-1} x_n^* W_N^{nk} \right) \left(\sum_{n=0}^{N-1} x_n W_N^{-nk} \right) \quad (12.7)$$

esso quindi corrisponde, nel dominio della frequenza, alla convoluzione **circolare** della sequenza $\{x_n\}$ con se stessa, ribaltata nei tempi e coniugata, se complessa (il coniugio della trasformata). Nel seguito, supporremo per semplicita' che la sequenza $\{x_n\}$ sia reale, tranne quando specificato. **Il periodogramma e' quindi la trasformata dell'autocorrelazione della sequenza periodicizzata.** Apparentemente, quindi la periodicizzazione ci porterebbe a gravi imprecisioni. Ma, stranamente, non e' cosı: infatti si consideri la sequenza $\{x_n\}$ nell'intervallo T_{tot} (la sequenza sia lunga N campioni) e il ragionevole stimatore **polarizzato** dell'autocorrelazione che vedremo corrispondere al periodogramma:

$$\hat{r}(\tau) = \frac{1}{N} \sum_{n=0}^{N-1} x_n^* x_{n+\tau} \quad x_{n+\tau} = 0; \quad \forall n + \tau > N; \quad \forall n + \tau < 0;$$

Lo stimatore considerato e' polarizzato perch :

$$E[\hat{r}(\tau)] = \frac{N - |\tau|}{N} r(\tau) \neq r(\tau) \quad (12.8)$$

Ricordiamo qui che l'indice τ dell'autocorrelazione viene anche chiamato in inglese *lag* (ritardo). Questa stima dell'autocorrelazione si ottiene bordando con N zeri la sequenza $\{x_n\}$ (fig. 10.5) e poi facendone la correlazione con se stessa, eliminando quindi qualsiasi errore di periodicizzazione.

La Trasformata di Fourier X_k della sequenza x_n di lunghezza N bordata con N zeri e' la Trasformata di Fourier della sequenza originale, interpolata ad intervallo di frequenza met  (e quindi pari a $1/2NT$) con un sinc. Pertanto, i valori che assume alle frequenze $f_k = k/NT$ sono eguali a quelli della Trasformata di Fourier della sequenza originale, **non bordata con zeri**. Quindi a parte un'irrilevante interpolazione su un intervallo di frequenza met  la Trasformata di Fourier della $\hat{r}_{,\tau}$ coincide con il periodogramma, come si era detto; si deve ricordare tuttavia che

CAPITOLO 12 ANALISI SPETTRALE NON PARAMETRICA

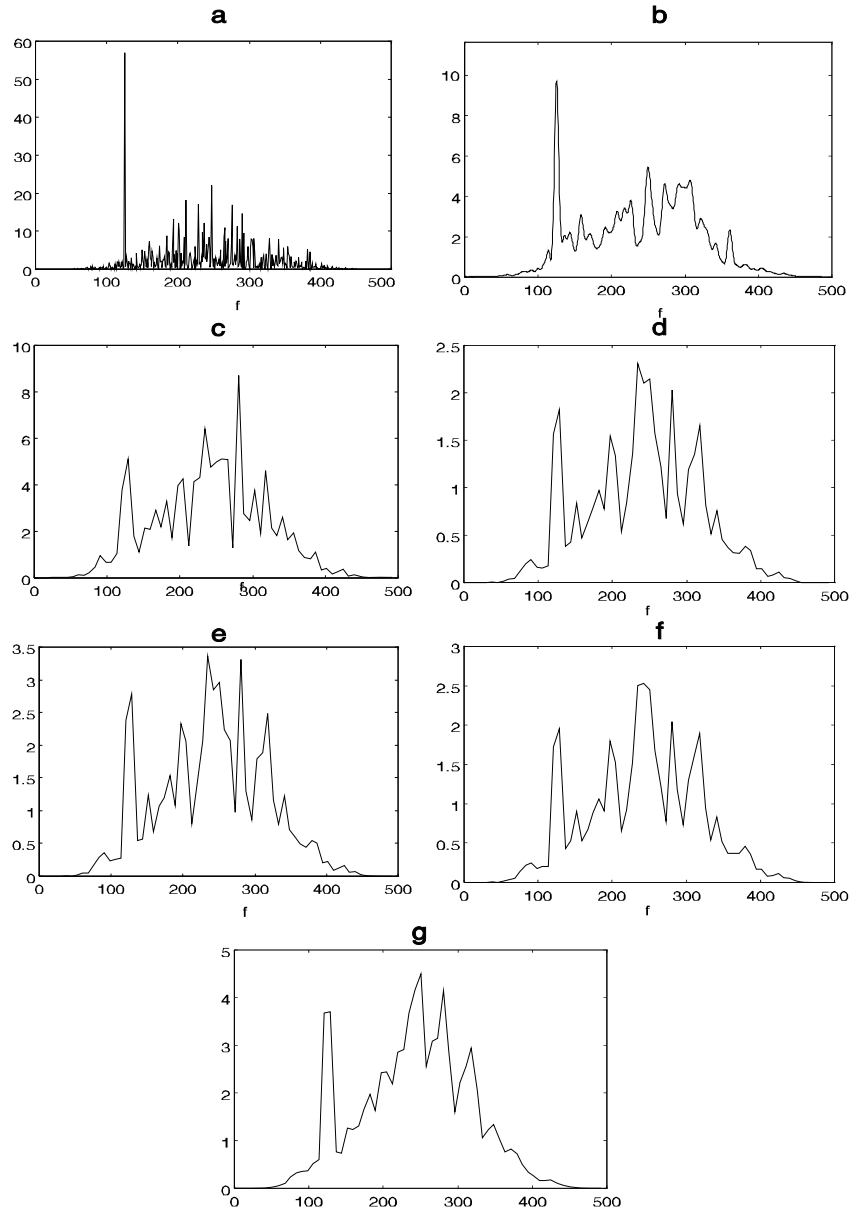


Figura 12.2: a) Spettro di potenza calcolato utilizzando un solo periodogramma su tutti i 2048 campioni; b) periodogramma precedente smussato c) Periodogramma ottenuto suddividendo i dati in 16 sottosequenze con finestre rettangolari e mediando le potenze; d) finestre triangolari e) finestre cosinusoidali f) finestre a coseno rialzato g) finestre a coseno rialzato sovrapposte (WOSA)

12.5 LA STIMA DELL'AUTOCORRELAZIONE

l'interpolazione 2:1 deve essere fatta sulla trasformata X_k e non sul modulo quadrato $|X_k^2|$; in altre parole, **prima si fa l'interpolazione e poi il modulo quadrato e non viceversa**. La periodicizzazione non ha influenza sulla stima spettrale anche se l'autocorrelazione periodicizzata é ben diversa da quella non periodicizzata. La polarizzazione nel do-

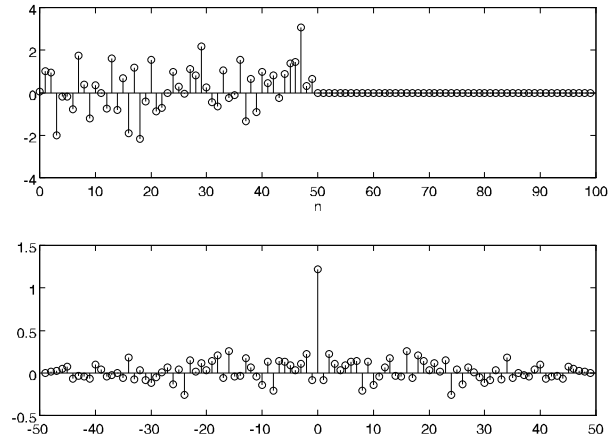


Figura 12.3: Sequenza di N campioni indipendenti bordata con N zeri e sua funzione di autocorrelazione

minio dei tempi che risulta dalla (10 .23) corrisponde alla polarizzazione del periodogramma trovata nel dominio delle frequenze; sono entrambe dovute alla finestratura dei dati con una finestra rettangolare. Infatti, sia $\{z_n\}$ una sequenza illimitata e $\{x_n\}$ la sequenza ottenuta finestrando temporalmente la sequenza $\{z_n\}$ con una finestra rettangolare $w_{r,n}$ (fig. 10 .6):

$$x_n = z_n \times w_{r,n}; \quad w_{r,n} = 1; \quad 0 \leq n \leq N - 1 \quad (12.9)$$

$$E[\hat{r}(\tau)] = E\left[\frac{1}{N} \sum_{n=0}^{N-1} z_n w_{r,n} z_{n+\tau} w_{r,n+\tau}\right] = \frac{1}{N} \sum_{n=0}^{N-1} w_{r,n} w_{r,n+\tau} E[z_n z_{n+\tau}] = \quad (12.10)$$

$$= r_{z,\tau} \frac{1}{N} \sum_{n=0}^{N-1} w_{r,n} w_{r,n+\tau} = r_{z,\tau} r_{w,\tau}$$

CAPITOLO 12 ANALISI SPETTRALE NON PARAMETRICA

dove con $r_{w,\tau}$ e $r_{z,\tau}$ sono state indicate rispettivamente le autocorrelazioni della finestra e della sequenza di ingresso. Quindi, finestrare i dati comporta che il valore atteso della stima dell'autocorrelazione trovato sia il prodotto dell'autocorrelazione della sequenza originale per quella della finestra. Nel caso di finestra rettangolare, si ritrova il risultato che l'autocorrelazione "vera" è moltiplicata per una finestra triangolare che va a zero dopo N campioni.

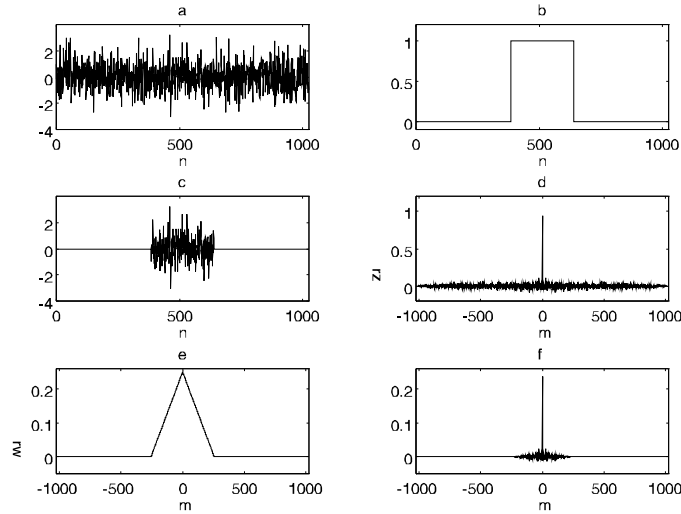


Figura 12.4: a) Sequenza illimitata z_n b) Finestra rettangolare w_n c) Sequenza finestrata x_n d) funzione di autocorrelazione della sequenza z_n e) funzione di autocorrelazione della finestra w_n f) $r_z \times r_w$

12.6 Errori da evitare

Si potrebbe credere di potere fare meglio del periodogramma tentando di rimuoverne la polarizzazione ad esempio:

- a) depolarizzando la $\hat{r}_{\cdot,\tau}$ moltiplicandola per il fattore $N/(N - |\tau|)$
- b) utilizzando altri campioni del processo eventualmente disponibili e utilizzando come stimatore dell'autocorrelazione la funzione:

$$\hat{r}_{\cdot,\tau} = \frac{1}{N} \sum_{n=0}^{N-1} x_n^* x_{n+\tau} \quad x_{n+\tau} \neq 0; \quad \forall n + \tau > N$$

12.7 MODIFICHE DEL PERIODOGRAMMA

Operare così porta a gravi errori anche se apparentemente si fa una cosa utile. Infatti, l'autocorrelazione polarizzata dell'equazione (12.8) ha spettro tutto positivo (il periodogramma appunto), come dovuto, ed il suo valore nell'origine, somma di tutti i termini dello spettro, tutti positivi, è massimo (le due questioni sono evidentemente collegate). Le autocorrelazioni non polarizzate ottenute con gli stimatori a, b non hanno affatto questa garanzia; sia nel caso a) che nel caso b) potrebbe benissimo accadere che $\hat{r}_{,\tau} > \hat{r}_0$ per qualche valore del ritardo τ . Inoltre, nel caso a) la crescente amplificazione della stima dell'autocorrelazione per ritardi crescenti (dovuta alla divisione per $N - |\tau|$ numero sempre piú vicino allo zero) peraltro necessaria per evitare la polarizzazione, comporta un inaccettabile aumento della dispersione.

12.7 Modifiche del periodogramma

12.7.1 Smussamento dello spettro

Usando lo stimatore $\hat{r}(\tau)$ viene fornita una stima dell'autocorrelazione fino a ritardi $\tau \leq N$; è chiaro tuttavia che per ritardi grandi, l'autocorrelazione è stimata utilizzando ben pochi campioni (un solo campione se $\tau = N - 1$) e quindi la stima ha grande dispersione; ciò risulta anche dal periodogramma dove la risoluzione spettrale è $1/NT$ e quindi corrisponde ad un'autocorrelazione stimata fino al ritardo NT . Ciò porta all'elevata dispersione che abbiamo visto nell'equazione 12.3. Appare opportuno modificare il periodogramma in modo da ridurre sia la dispersione che la risoluzione spettrale.

Per ridurre la risoluzione spettrale basta p.e. mediare i valori assunti dalla stima spettrale fornita dal periodogramma in $L = 2N/P$ intervalli di frequenza adiacenti; ciò corrisponde a smussare il periodogramma stesso convolvendolo con un operatore di filtro. Così facendo, si riduce anche la dispersione della stima in quanto si mediano tra loro stime basate su $M = 2L$ variabili casuali indipendenti $\hat{S}_k = |X_k|^2 = A_k^2 + B_k^2$. Per effettuare poi questa riduzione della risoluzione spettrale si può, nell'altro dominio, **moltiplicare** la funzione di autocorrelazione per una **finestra**, che è l'antitrasformata di Fourier del filtro che smussa il periodogramma (fig. 10 .7). Tuttavia, non tutte le finestre sono valide: dovremo moltiplicare l'autocorrelazione per finestre con Trasformata di Fourier

CAPITOLO 12 ANALISI SPETTRALE NON PARAMETRICA

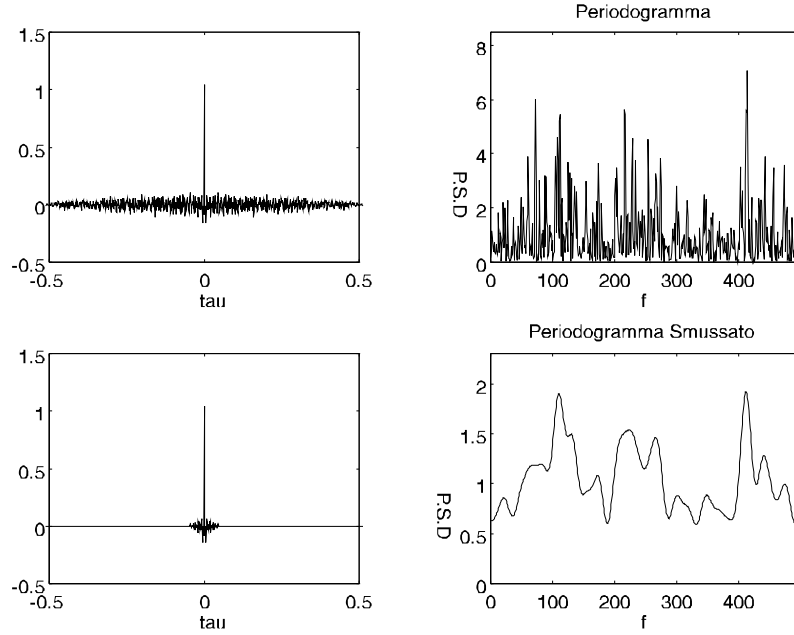


Figura 12.5: Troncamento della funzione di autocorrelazione e smussamento dello spettro

positiva per evitare di generare valori negativi dello spettro smussato, e quindi, in qualche situazione, ridurre pericolosamente \hat{r}_0 . In genere, si userà una finestra triangolare nei tempi di durata corrispondente alla riduzione di risoluzione spettrale desiderata. Ad esempio, se si moltiplica l'autocorrelazione per una finestra triangolare che va a zero per $\tau_{max} = PT \simeq NT/10$, come viene spesso consigliato (metodo di Blackman e Tukey), il lobo principale della trasformata della finestra va a zero per $f = 1/PT \simeq 10/NT$ e quindi combina insieme i contenuti \hat{S}_k di almeno 20 celle spettrali ad intervallo $\Delta f = 1/NT$. Facendo i conti più precisamente risulta però che, agli effetti della varianza della stima della locale densità spettrale, il numero equivalente di celle del periodogramma i cui contenuti sono combinati tra loro è $\simeq 15$ invece che 20. Infatti, indicando con \hat{S}_B la stima spettrale secondo Blackman e Tukey, si ha, posto che:

$$E[\hat{S}_k] = S; \quad E[\hat{S}_k^2] - E^2[\hat{S}_k] = \frac{2}{M}S^2 = S^2$$

12.7 MODIFICHE DEL PERIODOGRAMMA

$$\widehat{S}_B = \sum p_k \widehat{S}_k; \quad \sum p_k = 1; \quad E[\widehat{S}_B] = S$$

dove i pesi p_k sono dati dalla formula (trasformata di un triangolo campionato di lunghezza di base pari a $2P$ periodicizzato su lunghezza N ; P dispari, altrimenti non torna la periodicità su N campioni):

$$p_k = \frac{1}{PN} \left(\frac{\sin(\pi k P/N)}{\sin(\pi k/N)} \right)^2 \quad (12.11)$$

$$p_k = \sum_{l=0}^{N-1} q_l W_N^{-kl}; \quad q_l = \frac{1}{N} \sum_{k=0}^{N-1} p_k W_N^{kl};$$

Si osservi che i pesi p_k hanno un lobo principale che comprende $2N/P$ celle e quindi ci aspetteremmo una dispersione ridotta di un fattore $\sqrt{P/2N}$. Calcolando la dispersione come:

$$\sigma_{\widehat{S}_B}^2 = E \left[\left(\sum p_k \widehat{S}_k \right)^2 \right] - \left(E \left[\sum p_k \widehat{S}_k \right] \right)^2 = S^2 \sum p_k^2$$

risulta invece:

$$\frac{\sigma_{\widehat{S}_B}}{E[\widehat{S}_B]} = \sqrt{\sum p_k^2} \simeq \sqrt{\frac{2P}{3N}} = \sqrt{\frac{4}{3} \frac{P}{2N}} \quad (12.12)$$

Infatti i pesi nel dominio delle frequenze p_k sono trasformati di pesi nel dominio del tempo q_l decrescenti linearmente (a triangolo); si deve imporre che il valore al tempo 0 sia $q_0 = 1/N$ perchè la somma dei pesi nel dominio trasformato p_k sia unitaria.

$$q_l = \frac{1}{N} \left(1 - \frac{|l|}{P} \right); \quad \sum_{l=-P}^P q_l \simeq \frac{P}{N}; \quad \sum_{l=-P}^P q_l^2 \simeq \frac{2P}{3N^2};$$

e, usando la relazione di Parseval:

$$\sum p_k^2 = N \sum_{l=-P}^P q_l^2 \simeq \frac{2P}{3N} \quad (12.13)$$

In conclusione troviamo che la varianza cresce di un fattore $4/3$ appunto perchè lo spettro della finestra (e quindi i pesi p_k) è molto basso per

$f \simeq B$; nel caso considerato di una finestra triangolare che va a zero per $\tau = NT/10$ si ha $P = N/10$ e il numero equivalente di stime spettrali indipendenti \widehat{S}_k combinate (15) è minore di quello totale (20) originalmente contenuto nelle celle spettrali combinate dalla media fatta con i pesi. Tenendo conto delle parti reali ed immaginarie della trasformata, i campioni di potenza indipendenti contenuti nelle 20 celle spettrali adiacenti da cui si stima lo spettro di potenza nell'intervallo cercato sono appunto $M \simeq 30$; le stime indipendenti \widehat{S}_k che sono state mediate si susseguono nel dominio delle frequenze. Smussamenti maggiori si otterrebbero con finestre sull'autocorrelazione che vanno a zero per tempi inferiori ecc. Si osservino gli spettri in figura 10.7.

Un altro risultato si sarebbe ottenuto con pesi tutti eguali, e cioè convolvendo il periodogramma con una finestra rettangolare. In questo caso, con maggiore efficienza, la riduzione di risoluzione spettrale di un fattore 20 avrebbe comportato la riduzione della dispersione di un fattore $\sqrt{20}$, invece che $\sqrt{15}$. Peraltro, sarebbe stato equivalente moltiplicare l'autocorrelazione per una finestra a seno cardinale, che non si annulla nemmeno a tempi grandi. In pratica, si usa spesso la tecnica di Blackman e Tukey, anche se è leggermente inefficiente.

12.7.2 Media di periodogrammi di sottosequenze

Un altro modo per diminuire la dispersione è quello di suddividere l'intervallo NT di osservazione dei dati in L sottointervalli disgiunti e quindi dividere la sequenza totale in L sottosequenze. Poi vengono mediate tra loro i periodogrammi delle sottosequenze ed il numero totale di campioni indipendenti combinati tra loro per ogni cella sarà $M = 2L$. Se p.e. $L = 15$, $M = 30$ e dal punto di vista della dispersione statistica i risultati sono equivalenti a quelli ottenibili con una finestatura triangolare dell'autocorrelazione che vada a zero per $\tau_{max} \simeq NT/10$. Quindi, con un periodogramma unico si ottiene la desiderata dispersione statistica con una risoluzione spettrale pari a circa $B = 20/NT$; invece, segmentando in sottosequenze, per avere la stessa dispersione statistica, basta porre $L = 15$, la risoluzione spettrale del periodogramma di una sottosequenza diventa $B' \simeq 15/NT$ e quindi un poco meglio rispetto al metodo di Blackman: abbiamo trattato i dati in modo diverso che sembra più efficiente e con risultati equivalenti a quelli che si avrebbero convolvendo lo spettro di un unico periodogramma con un rettangolo di larghezza B' .

12.7 MODIFICHE DEL PERIODOGRAMMA

In realtà succede di peggio; **mediando gli L periodogrammi di sequenze brevi ottenuti troncando i dati con una finestra corta, la polarizzazione della stima peggiora a causa dei troncamenti** e delle periodicizzazioni più fitte. Un disturbo monocromatico si ripercuote su tutto lo spettro con maggiore influenza. In presenza di disturbi fortemente colorati, è meglio smussare nel dominio delle frequenze il periodogramma della sequenza intera convolvendolo p.e. con un rettangolo di larghezza B , come visto alla fine del paragrafo precedente. In questo caso, gli effetti di un forte segnale monocromatico non si estendono al di là della banda B ed è come se avessimo rastremato a zero la funzione di autocorrelazione moltiplicandola per una finestra a *sinc*, come peraltro già osservato.

Per contro, mediando periodogrammi fatti su sottosequenze l'occupazione di memoria è minore ed anche i calcoli sono meno impegnativi. Nella figura 10 .4, si confrontino gli effetti dello smussamento del periodogramma (fig. 10 .4b) con quelli della media di periodogrammi fatti su sottointervalli (figg. 10 .4c - d - e - f - g); la qualità della stima in figura 10 .4b è ottenuta approssimativamente solo con la tecnica WOSA (fig. 10 .4g) di cui al prossimo paragrafo.

12.7.3 La tecnica WOSA

E' necessario evitare polarizzazioni dovute a eccessive infiltrazioni spettrali (spectral leaks): queste sono dovute, ad esempio, al troncamento nei tempi di una sinusoidale effettuato con una finestra rettangolare. Ciò corrisponde nelle frequenze alla convoluzione con un *sinc* e origina componenti spettrali a frequenze anche lontane dalla frequenza della sinusoidale di partenza. Quindi, per formare le sottosequenze, è necessario **finestrare i dati** con finestre abbastanza "dolci" come ad esempio coseno rialzato per evitare questo fenomeno, specialmente se il segnale è disturbato da forti segnali monocromatici. Così facendo, tuttavia, la risoluzione spettrale si riduce e soprattutto si pesano poco alcuni dati; quindi, la dispersione della stima spettrale è maggiore del necessario perché non si utilizzano in egual modo tutti i dati. Allora, non volendo o non potendo fare la cosa ottimale e cioè trattare tutti i dati con una finestra unica per poi smussare l'unico periodogramma ottenuto, la tecnica consigliata è quella di sovrapporre le finestre (Window Overlap) con le quali i dati vengono sezionati (vedi figura 10 .8) in modo da trascurare il minor nu-

CAPITOLO 12 ANALISI SPETTRALE NON PARAMETRICA

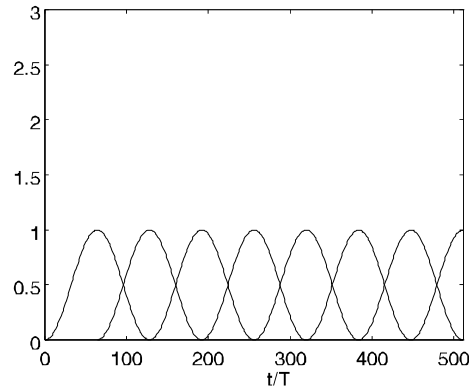


Figura 12.6: Disposizione delle finestre nella tecnica WOSA.

mero di dati possibile. Ancora, è possibile tenere conto, **in media**, della finestra con cui sono stati sezionati i dati, al momento della stima dello spettro medio. Infatti, per smussare quest'ultimo spettro si moltiplica la funzione di autocorrelazione per una seconda finestra, in generale più stretta della finestra utilizzata per limitare la lunghezza dei dati. Questa seconda finestra può essere sagomata in modo opportuno da compensare la polarizzazione che in media la prima finestra ha generato sui dati. Questa tecnica di analisi spettrale è nota con il nome di Window Overlap Spectral Analysis (WOSA; Nuttall; Proc. IEEE 1982) e assomiglia molto alla tecnica parametrica MA che vedremo al prossimo capitolo.

12.8 Sommario

Riassumiamo qui le principali proprietà del periodogramma e delle metodologie derivate.

- 1) Il segnale viene troncato con una finestra p.e. rettangolare su un intervallo $NT = T_{tot}$.
- 2) Ne viene fatto il modulo della trasformata di Fourier, e diviso per N . Questo corrisponde al calcolo dell'autocorrelazione campionaria, polarizzata.
- 3) La risoluzione spettrale è massima ($1/T_{tot}$), ma la dispersione è anche massima.

4) La polarizzazione della stima dovuta al troncamento brusco del segnale è molto elevata.

Cosa fare per migliorare?

Per una minore polarizzazione: troncatura con finestre più dolci come la Hanning o la Hann. Questo migliora la polarizzazione ma peggiora la risoluzione spettrale (lobo principale della trasformata della finestra più largo).

Per una minore dispersione: si smussa il periodogramma e cioè si finestra l'autocorrelazione campionaria trovata.

Ancora, si può suddividere la sequenza in L sottosequenze e mediare gli L periodogrammi: questo migliora la dispersione così come uno smussamento, ma peggiora la polarizzazione perchè il troncamento è fatto con una finestra più corta.

12.9 Un esempio

Si considerino 2 sequenze casuali reali di 8 campioni e se ne trovi lo spettro con il periodogramma, usando o due periodogrammi da 4 campioni o uno solo da 8.

sequenza 1 6.3 5.9 5.8 5.2 4.5 3.3 3.1 3.2

sequenza 2 4.3 -4.2 3.9 -3.5 3.5 -2.9 3.0 -2.7

I periodogrammi sono:

$$X_k = \sum_{n=0}^3 x_n W_4^{-nk}; W_4 = \exp\left(\frac{i2\pi}{4}\right) = i \quad (12.14)$$

$$\underline{\mathbf{X}}_k = \begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix} \begin{bmatrix} 6.3 \\ 5.9 \\ 5.8 \\ 5.2 \end{bmatrix} = \begin{bmatrix} 23.2 \\ 0.5 - 0.7i \\ 1.0 \\ 0.5 + 0.7i \end{bmatrix} \quad (12.15)$$

CAPITOLO 12 ANALISI SPETTRALE NON PARAMETRICA

$$\underline{\mathbf{X}}'_k = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix} \begin{bmatrix} 4.5 \\ 3.3 \\ 3.1 \\ 3.2 \end{bmatrix} = \begin{bmatrix} 14.1 \\ 1.4 - 0.1i \\ 1.1 \\ 1.4 + 0.1i \end{bmatrix} \quad (12.16)$$

I due periodogrammi hanno i seguenti valori

$$P_k = \frac{|X_k|^2}{4} = 135; 0.15; 0.25; \quad (12.17)$$

$$P'_k = \frac{|X'_k|^2}{4} = 50; 0.49; 0.30 \quad (12.18)$$

Il periodogramma a quattro campioni mediato è:

$$\mathbf{P}_k^m = \frac{1}{2} \left(\frac{|X_k|^2}{4} + \frac{|X'_k|^2}{4} \right) = \mathbf{92; 0.33; 0.28} \quad (12.19)$$

Il periodogramma a 8 campioni è:

$$175; 5; 0.5; 0.3; 0.5; 0.3; 0.5; 5 \quad (12.20)$$

Per la sequenza 2:

$$\underline{\mathbf{X}}_k = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix} \begin{bmatrix} 4.3 \\ -4.2 \\ 3.9 \\ -3.5 \end{bmatrix} = \begin{bmatrix} 0.6 \\ 0.4 + 0.6i \\ 15.8 \\ 0.4 - 0.6i \end{bmatrix} \quad (12.21)$$

$$\underline{\mathbf{X}}'_k = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix} \begin{bmatrix} 3.5 \\ -2.9 \\ 3 \\ -2.7 \end{bmatrix} = \begin{bmatrix} 0.9 \\ 0.5 + 0.2i \\ 12.1 \\ 0.5 - 0.2i \end{bmatrix} \quad (12.22)$$

I due periodogrammi hanno i seguenti valori

$$P_k = \frac{|X_k|^2}{4} = 0.09; 0.13; 62; \quad (12.23)$$

$$P'_k = \frac{|X'_k|^2}{4} = 0.2; 0.07; 36 \quad (12.24)$$

Il periodogramma a quattro campioni mediato è

$$\mathbf{P}_k^m = \frac{1}{2} \left(\frac{|X_k|^2}{4} + \frac{|X'_k|^2}{4} \right) = \mathbf{0.14; 0.1; 49} \quad (12.25)$$

Il periodogramma a 8 campioni è:

$$0.24; 0.04; 0.16; 0.82; \mathbf{98.}; 0.82; 0.16; 0.04 \quad (12.26)$$

Capitolo 13

Analisi Spettrale Parametrica

13.1 Introduzione

Si consideri una sequenza di campioni indipendenti $\{x_n\}$ e con essa si alimenti un filtro $H(\omega)$; se la trasformata della sequenza $\{x_n\}$ è indicata con $X(\omega)$, quella dell'uscita sarà:

$$Y(\omega) = X(\omega)H(\omega)$$

$$E[|Y(\omega)|^2] = E[|X(\omega)|^2] \times |H(\omega)|^2 \quad (13.1)$$

$$H(\omega) = \frac{B(\omega)}{A(\omega)} \quad (13.2)$$

La densità spettrale di potenza bilatera del processo bianco $\{x_n\}$ è costante perché per sequenze di campioni indipendenti di lunghezza N , le componenti della trasformata di Fourier discreta sono tutte di eguale valore quadratico medio (assumendo $T = 1$):

$$\frac{E[|X_k^2|]}{N} = \sigma_x^2 \quad (13.3)$$

L'analisi spettrale parametrica consiste nel determinare i parametri del filtro $H(\omega)$ (fig. 14.1) in modo che lo spettro in uscita del filtro assomigli quanto possibile allo spettro del segnale $\{y_n\}$ da analizzare, o similmente, la sua autocorrelazione sia eguale all'autocorrelazione misurata. Avremo analisi spettrale Moving Average (MA) se il filtro ha una trasformata z caratterizzata da tutti zeri; avremo il caso Auto Regressivo (AR) quando

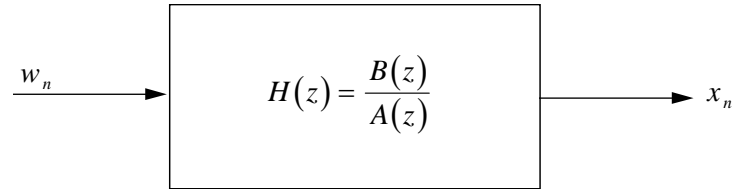


Figura 13.1: Filtro i cui parametri vengono calcolati nell'analisi spettrale AR e MA

il filtro é tutti poli e cioè autoregressivo: infine avremo il caso misto ARMA (Auto Regressive Moving average) nel caso piú generale di poli e zeri.

Vedremo che l'analisi spettrale parametrica tutti zeri praticamente coincide con le tecniche non parametriche modificate delle ultime sezioni del capitolo precedente. Le tecniche spettrali AR invece sono di natura ben diversa ed innovativa ed ad esse destineremo grande attenzione per tutta una serie di motivi che chiariremo in seguito.

L'analisi spettrale ARMA poi, é di uso meno frequente anche perché come noto, un qualsiasi filtro può essere rappresentato con soli zeri o soli poli e una rappresentazione mista serve solo per una descrizione della stessa (o di una simile) funzione di trasferimento con un numero inferiore di parametri.

La validità di una descrizione AR o MA o ARMA dipende comunque dalla qualità del modello così ottenuto.

13.2 Analisi tutti zeri (MA)

Si consideri un filtro FIR con risposta all'impulso $\{b_h\}$ la cui trasformata z é caratterizzata da tutti zeri (fig. 13.2):

$$B(z) = \sum_{h=0}^H b_h z^{-h} \quad (13.4)$$

Sia $\{x_n\}$ la sequenza bianca all'ingresso del filtro ed $\{y_n\}$ la sequenza,

13.2 ANALISI TUTTI ZERI (MA)

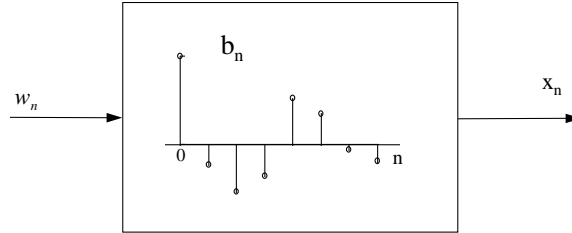


Figura 13.2: Filtro FIR con risposta impulsiva $\{b_n\}$

spetttralmente colorata, in uscita. La funzione di autocorrelazione della sequenza $\{y_n\}$ è lunga $2H + 1$ campioni ed è:

$$r_{y,m} = E \left[\sum_{h,k=0}^H b_h b_k x_{n-h} x_{n+m-k} \right] = \sigma_x^2 \sum_{h=0}^H b_h b_{m+h} \quad (13.5)$$

dove si è utilizzato il fatto che:

$$E[x_{n-h} x_{n+m-k}] = \sigma_x^2 \delta(h + m - k) \quad (13.6)$$

Ricordiamo che il nostro problema è quello di determinare un filtro $\{b_h\}$ o la sua trasformata $B(z)$ (la soluzione non è unica) a partire da una stima dell'autocorrelazione dei dati. Per ora supponiamo che la stima disponibile sia molto buona, per cui praticamente possiamo supporre di avere la funzione di autocorrelazione "vera" r_m .

Passando alla trasformata z si vede che $R(z)$, trasformata z della r_m è (fig. 13.4):

$$R(z) = \sigma_x^2 B(z) B^*(1/z) \quad (13.7)$$

Si osservi che $R(z)$ è un polinomio nella variabile z^{-1} che insieme ad ogni radice, ha anche la radice reciproca e coniugata. Se le radici sono sul cerchio unitario, saranno a coppie. Quindi, un modo per determinare un filtro $B(z)$, assegnata la funzione di autocorrelazione $R(z)$, è quello di trovare le radici di $R(z)$ e per esempio assegnare a $B(z)$ tutte le H radici all'esterno del cerchio unitario e a $B^*(1/z)$ tutte le altre (all'interno). In questo modo, $B^*(1/z)$ risulta a fase minima; altrimenti possiamo scegliere $B^*(1/z)$ a fase massima o fase mista, in 2^H modi diversi (fig. 13.4).

CAPITOLO 13 ANALISI SPETTRALE PARAMETRICA

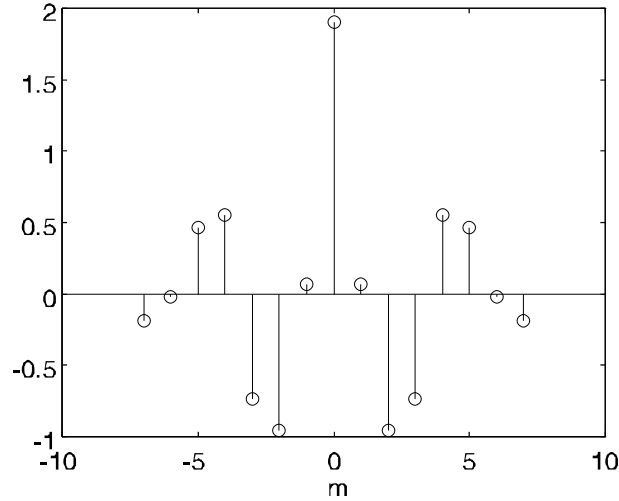


Figura 13.3: Funzione di autocorrelazione del filtro di fig.13.2

Si osservi che non basta che la $R(z)$ sia un qualsiasi polinomio simmetrico, perchè si possa individuare un filtro $B(z)$. Infatti si avrebbero coppie di zeri reciproci ma si potrebbero avere anche zeri semplici sul cerchio unitario (fig. 13.5); in questo caso non sarebbe possibile trovare la $B(z)$ perchè non si potrebbe associare un'altra radice in posizione reciproca. Ma in questo caso, il polinomio simmetrico $R(z)$ **non rappresenterebbe** una funzione di autocorrelazione in quanto la trasformata di Fourier della $r_{y,m}$ e cioè i valori assunti dalla $R(z)$ sul cerchio unitario, conterrebbero cambi di segno al passaggio per gli zeri e quindi anche valori negativi, mentre invece uno spettro di potenza, Trasformata di Fourier di una funzione di autocorrelazione vera, non sarà mai negativo.

In conclusione, se lo spettro di un polinomio simmetrico è tutto positivo,

- le radici sono sempre a coppie e non compaiono radici isolate,
- si può trovare un filtro $B(z)$ in 2^H modi diversi, se H è l'ordine del filtro.

Una sequenza simmetrica invece, può avere zeri semplici (e non doppi) sul cerchio unitario e quindi non corrispondere ad una funzione di autocorrelazione.

13.2 ANALISI TUTTI ZERI (MA)

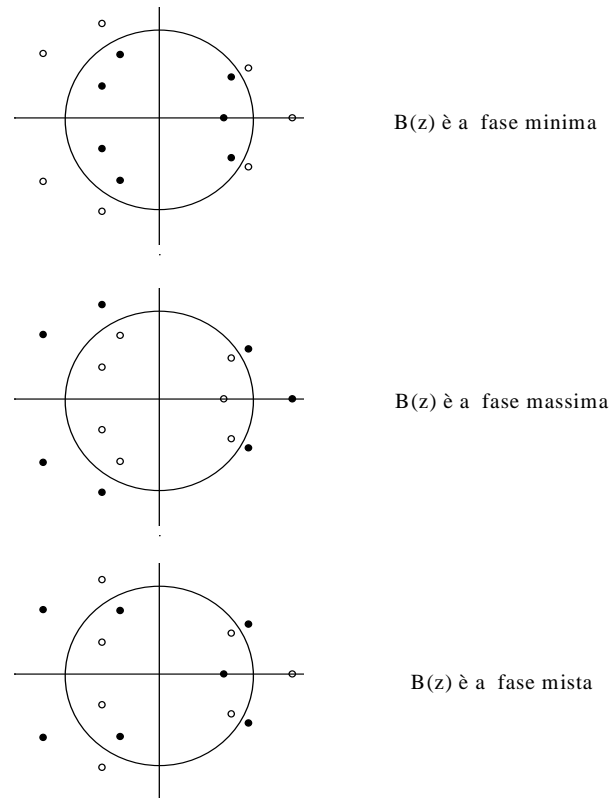


Figura 13.4: Esempi di assegnazione degli zeri della $R(z)$ alla $B(z)$ (zeri anneriti)

13.2.1 Troncamento della funzione di autocorrelazione

Si vede pertanto che se si ha:

- una funzione di autocorrelazione (spettro positivo)
- nulla per $m > H$.

allora è sempre possibile trovare 2^H filtri di lunghezza $H + 1$ che alimentati da sequenze bianche hanno in uscita quella autocorrelazione.

Tuttavia, sono disponibili solo stime delle funzioni di autocorrelazione: se la stima è fatta con la correlazione con se stessa della sequenza bordata di zeri, allora la sua trasformata di Fourier, il periodogramma, è sempre positivo. Tuttavia, in questo caso, la lunghezza del filtro risulta eccessiva perchè, a causa della dispersione della stima, i campioni

CAPITOLO 13 ANALISI SPETTRALE PARAMETRICA

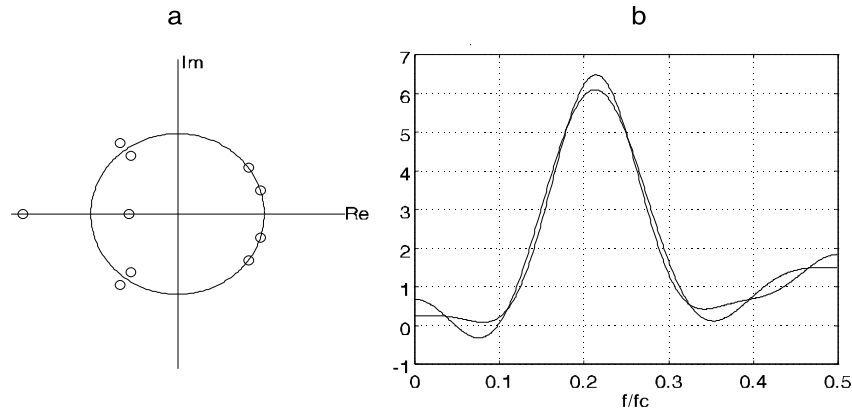


Figura 13.5: a) Zeri dell'autocorrelazione di fig 13.3 troncata con finestra rettangolare b) Spettro di potenza derivante dalla finestrazione rettangolare; è sovrapposto lo spettro vero. Lo spettro è vicino a quello vero, ma diventa negativo, cosa impossibile.

dell'autocorrelazione stimata non saranno mai nulli.

Se si desidera limitare la lunghezza H del filtro, si deve rastremare a 0 la funzione di autocorrelazione in H campioni, finestrandola in modo che lo spettro si mantenga positivo, e quindi moltiplicandola per una finestra che in pratica è sempre quella triangolare (fig. 13.6). Usando finestre più efficienti come il coseno rialzato, non si potrebbe garantire che la correlazione finestrata sia a spettro positivo e pertanto che la ricerca del filtro $B(z)$ possa poi andare a buon fine.

In conclusione, per fare una stima parametrica tutti zeri, è necessario finestrare la funzione di autocorrelazione con una finestra triangolare di lunghezza $2H$; quanto maggiore è H tanto maggiore sarà la risoluzione spettrale della stima parametrica. In sostanza, si vede che questa tecnica di stima spettrale coincide con quella dello smussamento del periodogramma presa in esame nel precedente capitolo: l'unica differenza sta nel fatto che ora è possibile dare un significato fisico alla semilunghezza della finestra sull'autocorrelazione (e cioè la lunghezza dell'antitrasformata del filtro usato per lo smussamento dello spettro).

13.3 ANALISI TUTTI POLI (AR)

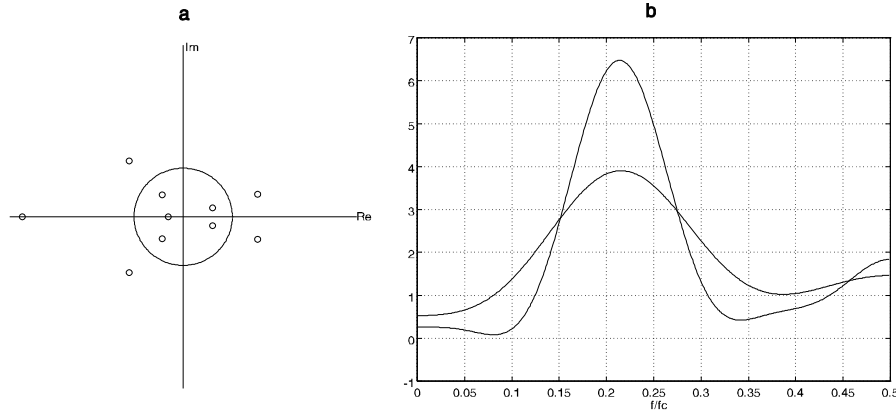


Figura 13.6: a) Zeri della autocorrelazione di fig. 13.3 troncata con finestra triangolare b) Spettro di potenza derivante dalla finestatura triangolare; è sovrapposto lo spettro vero. Lo spettro ora è tutto positivo, anche se è più lontano da quello vero.

13.3 Analisi tutti poli (AR)

Questa tecnica di stima spettrale è molto importante per tutta una serie di motivi che emergeranno via via che procederà lo studio: per ora si osservi che avendo i filtri IIR risposta all'impulso illimitata, possono produrre spettri di grande risoluzione spettrale con un numero limitato di parametri. Pertanto si prestano alla descrizione di fenomeni che hanno un lungo tempo di coerenza, in cui cioè il processo si decorrela assai lentamente.

Sia $\{x_n\}$ la sequenza in uscita di un filtro IIR di ordine N , $1/A_N(z)$ alimentato da una sequenza bianca $\{w_n\}$.

$$x_n = - \sum_{h=1}^N a_{h,N} x_{n-h} + w_n; \quad X(z) = \frac{W(z)}{A_N(z)}; \quad (13.8)$$

$$A_N(z) = 1 + \sum_{h=1}^N a_{h,N} z^{-h}$$

L'autocorrelazione della sequenza di uscita, che calcoliamo moltiplicando l'equazione (13.8) per il campione x_{n-m}^* anticipato di m campioni e

CAPITOLO 13 ANALISI SPETTRALE PARAMETRICA

coniugato e poi considerando il valore atteso, r_m è:

$$r_m = E[x_n x_{n-m}^*] = -E\left[\sum_{h=1}^N a_{h,N} x_{n-h} x_{n-m}^*\right] + E[w_n x_{n-m}^*] = \quad (13.9)$$

$$= -\sum_{h=1}^N a_{h,N} r_{m-h} + \sigma_w^2 \delta(m); \quad \forall m \quad (13.10)$$

Infatti la sequenza $\{w_n\}$ è incorrelata con i campioni precedenti x_{n-m} ; $m > 0$. Le equazioni (13.10) sono note come le equazioni di Yule - Walker e forniscono **tutta l'autocorrelazione dell'uscita**, dati i coefficienti del filtro $A_N(z)$ ed il v.q.m. del segnale d'ingresso σ_w^2 . Si osservi che nel caso di segnale e/o coefficienti $a_{h,N}$ complessi, le equazioni (13.10) vanno scritte tenendo conto che $r_{-m} = r_m^*$. Per motivi mnemonici, può essere utile osservare che la formula (13.10) assomiglia ad una convoluzione tra le sequenze r_m ed $a_{h,N}$; non è una convoluzione per la limitazione degli indici. L'equazione corrispondente a $m = 0$ contiene anche la varianza del segnale di ingresso σ_w^2 ; peraltro, il filtro non dipende da questo fattore di scala e si può sempre riscalarlo la funzione di autocorrelazione ponendo $r_0 = 1$ risolvendo poi le equazioni (Y-W) per $m = 1, \dots, N$. Il fatto che l'equazione (13.10) sia definita per tutti i valori di m , implica che assegnata qualsiasi $N - pla$ consecutiva di valori di r_m , poi possiamo determinare i valori di r_m per qualsiasi altro valore.

Ancora, le equazioni di Yule Walker si possono ottenere in forma matriciale in modo assai semplice indicando i seguenti vettori con i corrispondenti simboli:

$$\underline{\mathbf{a}}_N = \begin{bmatrix} a_{1,N} \\ a_{2,N} \\ a_{3,N} \\ \dots \\ a_{N,N} \end{bmatrix}; \quad \underline{\boldsymbol{\xi}}_N = \begin{bmatrix} x_{n-1}^* \\ x_{n-2}^* \\ x_{n-3}^* \\ \dots \\ x_{n-N}^* \end{bmatrix}; \quad \underline{\mathbf{r}}_N = \begin{bmatrix} r_x(1) \\ r_x(2) \\ r_x(3) \\ \dots \\ r_x(N) \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ \dots \\ r_N \end{bmatrix}$$

Allora si ha:

$$x_n = -\underline{\boldsymbol{\xi}}_N^* \underline{\mathbf{a}}_N + w_n \quad (13.11)$$

13.3 ANALISI TUTTI POLI (AR)

E moltiplicando a sinistra per $\underline{\xi}_N$ e considerando il valore atteso, visto che $E[\underline{\mathbf{x}}_N w_n] = \underline{\mathbf{0}}$:

$$E[\underline{\xi}_N x_n] = \underline{\mathbf{r}}_N = E[\underline{\xi}_N (-\underline{\xi}_N^* \underline{\mathbf{a}}_N + w_n)] = - E[\underline{\xi}_N \underline{\xi}_N^*] \underline{\mathbf{a}}_N = -\underline{\mathbf{R}} \underline{\mathbf{a}}_N$$

La matrice dei coefficienti delle equazioni (Y-W) è una matrice di Toeplitz; infatti è simmetrica (o Hermitiana, per sequenze complesse) e tutti gli elementi appartenenti alla stessa diagonale o subdiagonale sono eguali tra loro. La matrice è pertanto caratterizzata da N numeri. Il vettore dei termini noti è anche esso formato da campioni della funzione di autocorrelazione, r_m ; $m = 1, \dots, N$ col segno cambiato. Riscrivendo in forma matriciale le equazioni di Yule Walker infatti si ottiene:

$$\begin{bmatrix} r_0 & r_{-1} & r_{-2} & \dots & r_{-N+1} \\ r_1 & r_0 & r_{-1} & \dots & r_{-N+2} \\ r_2 & r_1 & r_0 & \dots & r_{-N+3} \\ \dots & \dots & \dots & \dots & \dots \\ r_{N-1} & r_{N-2} & r_{N-3} & \dots & r_0 \end{bmatrix} \begin{bmatrix} a_{1,N} \\ a_{2,N} \\ a_{3,N} \\ \dots \\ a_{N,N} \end{bmatrix} = \underline{\mathbf{R}} \underline{\mathbf{a}}_N = -\underline{\mathbf{r}}_N = - \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ \dots \\ r_N \end{bmatrix} \quad (13.12)$$

ed in termini vettoriali:

$$\underline{\mathbf{R}} \underline{\mathbf{a}}_N = -\underline{\mathbf{r}}_N \quad (13.13)$$

Per completezza, si aggiunge la formulazione completa, facilmente verificabile, in cui compare anche la prima equazione che contiene la varianza della sequenza bianca di ingresso σ_w^2 . Si stia attenti al fatto che in questa equazione non sempre le incognite sono contenute nel vettore $\underline{\mathbf{a}}_N$; a volte è la funzione di autocorrelazione a non essere nota, oppure il valore di σ_w^2 , ecc.

$$\begin{bmatrix} r_0 & r_{-1} & r_{-2} & \dots & r_{-N} \\ r_1 & r_0 & r_{-1} & \dots & r_{-N+1} \\ r_2 & r_1 & r_0 & \dots & r_{-N+2} \\ \dots & \dots & \dots & \dots & \dots \\ r_N & r_{N-1} & r_{N-2} & \dots & r_0 \end{bmatrix} \begin{bmatrix} 1 \\ a_{1,N} \\ a_{2,N} \\ \dots \\ a_{N,N} \end{bmatrix} = \begin{bmatrix} \sigma_w^2 \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix} \quad (13.14)$$

13.4 La stima spettrale AR

Una volta trovati gli $a_{h,N}$ a partire per esempio dai valori assunti dalla funzione di autocorrelazione per $m = 1, \dots, N$ è immediato determinare le componenti dello spettro continuo del segnale $\{x_i\}$; infatti la funzione di autocorrelazione, che ne è la trasformata, è determinata per tutti i valori dell'indice temporale m ; pari a quella assegnata per $m = 1 : N$ e poi desunta dalle equazioni di Yule Walker per tutti gli altri valori.

$$E[|X(\omega)|^2] = \frac{E[|W(\omega)|^2]}{|A(e^{j\omega T})|^2} = \frac{\sigma_w^2 T}{|1 + \sum_{h=1}^N a_{h,N} e^{-j\omega h T}|^2}; \quad (13.15)$$

Per dare un significato fisico a questa stima è necessario garantire che il filtro $1/A_N(z)$ sia stabile e pertanto che il polinomio $A_N(z)$ sia a minima fase; si dimostrerà nei prossimi capitoli che ciò accade **se e solo se** i campioni r_m che si utilizzano nelle equazioni di Yule Walker sono campioni di una funzione di autocorrelazione (a spettro positivo).

Di nuovo, per avere la stima spettrale AR di ordine N , non serve **tutta** la funzione di autocorrelazione ma ne bastano i primi N campioni; gli infiniti altri sono calcolabili ancora in base alle equazioni di Y-W a partire dal polinomio $A_N(z)$. In altre parole, il troncamento della funzione di autocorrelazione non comporta problemi; semplicemente, i valori dell'autocorrelazione predetti in base alle equazioni Y-W, non coincideranno con quelli effettivi, quando la stima spettrale è fatta con un ordine (il valore di N) troppo basso.

13.5 Un esempio

Si considerino due 4 sequenze casuali reali di 8 campioni

sequenza 1 6.3078 5.8842 5.8076 5.2822 4.4676 3.3751 3.1054 3.1517

sequenza 2 4.2887 -4.1611 3.9408 -3.5029 3.5127 -2.9277 3.0263 -2.7812.

Si determina ora la densità spettrale di potenza delle sequenze con la tecnica a media mobile e la tecnica autoregressiva di ordine 2.

13.5.1 La stima della funzione di autocorrelazione

I sequenze:

La stima polarizzata dei primi tre campioni della funzione di autocorrelazione è:

$$\hat{r}_0 = 23.4$$

$$\hat{r}_1 = 20.1$$

$$\hat{r}_2 = 17.$$

13.5.2 Stima MA

La stima MA di ordine 2 è, applicando la finestra triangolare per troncatura la funzione di autocorrelazione stimata:

$$(b_0 + b_1 z^{-1})(b_0 + b_1 z) = b_0^2 + b_1^2 + (z + z^{-1})b_0 b_1 \quad (13.16)$$

$$b_0^2 + b_1^2 = 23.4 \quad (13.17)$$

$$b_0 b_1 = 10 \quad (13.18)$$

$$(b_0^2 + b_1^2) - 2b_0 b_1 = (b_0 - b_1)^2 = 3.3 \rightarrow b_0 - b_1 = \pm 1.81 \quad (13.19)$$

$$(b_0^2 + b_1^2) + 2b_0 b_1 = (b_0 + b_1)^2 = 33.3 \rightarrow b_0 + b_1 = \pm 5.77 \quad (13.20)$$

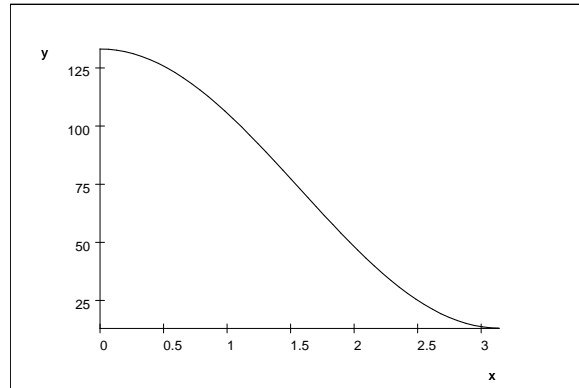
$$b_0 = \pm 7.58; b_1 = \pm 3.96 \quad (13.21)$$

$$b_0 = \pm 3.96; b_1 = \pm 7.58 \quad (13.22)$$

Lo spettro MA è quindi:

$$S_{MA} = |7.58 + 3.96e^{i\phi}|^2 \quad (13.23)$$

CAPITOLO 13 ANALISI SPETTRALE PARAMETRICA



13.5.3 Stima AR

La stima AR di ordine 2 si trova con la tecnica di Yule Walker:

$$\begin{bmatrix} 23.4 & 20.1 & 17 \\ 20.1 & 23.4 & 20.1 \\ 17 & 20.1 & 23.4 \end{bmatrix} \begin{bmatrix} 1 \\ a_{1,2} \\ a_{2,2} \end{bmatrix} = \begin{bmatrix} \sigma_w^2 \\ 0 \\ 0 \end{bmatrix} \quad (13.24)$$

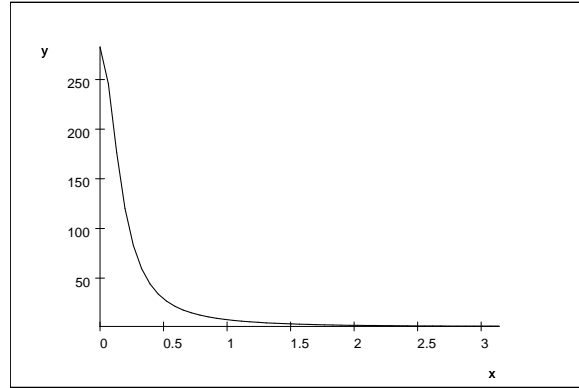
$$\begin{bmatrix} a_{1,2} \\ a_{2,2} \end{bmatrix} = - \begin{bmatrix} 23.4 & 20.1 \\ 20.1 & 23.4 \end{bmatrix}^{-1} \begin{bmatrix} 20.1 \\ 17 \end{bmatrix} = \begin{bmatrix} -0.89613 \\ 0.04326 \end{bmatrix} \quad (13.25)$$

$$\begin{bmatrix} 23.4 & 20.1 & 17 \end{bmatrix} \begin{bmatrix} 1 \\ -0.89613 \\ 0.04326 \end{bmatrix} = 6.12 = \sigma_w^2 \quad (13.26)$$

Lo spettro AR è quindi:

$$S_{AR} = \frac{6.12}{|1 - 0.89613e^{i\phi} + 0.04326e^{2i\phi}|^2} \quad (13.27)$$

13.5 UN ESEMPIO



Per la sequenza 2:

La stima polarizzata della funzione di autocorrelazione è:

$$\hat{r}_0 = 12.7$$

$$\hat{r}_1 = -11$$

$$\hat{r}_2 = 9.3$$

La stima MA di ordine 2 è, applicando la finestra triangolare per troncare la funzione di autocorrelazione stimata:

$$(b_0 + b_1 z^{-1})(b_0 + b_1 z) = b_0^2 + b_1^2 + (z + z^{-1})b_0 b_1 \quad (13.28)$$

$$b_0^2 + b_1^2 = 12.7 \quad (13.29)$$

$$b_0 b_1 = -5.5 \quad (13.30)$$

$$(b_0^2 + b_1^2) - 2b_0 b_1 = (b_0 - b_1)^2 = 23.7 \rightarrow b_0 - b_1 = \pm 4.86 \quad (13.31)$$

$$(b_0^2 + b_1^2) + 2b_0 b_1 = (b_0 + b_1)^2 = 1.7 \rightarrow b_0 + b_1 = \pm 1.3 \quad (13.32)$$

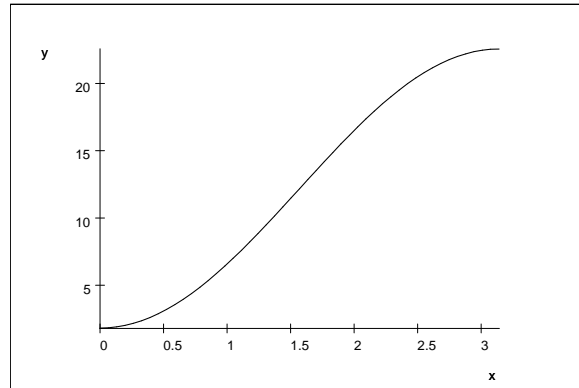
$$b_0 = \pm 3.05; b_1 = -\pm 1.7 \quad (13.33)$$

$$b_0 = \pm 1.7; b_1 = -\pm 3.05 \quad (13.34)$$

Lo spettro MA è quindi:

$$S_{MA} = |3.05 - 1.7e^{i\phi}|^2 \quad (13.35)$$

CAPITOLO 13 ANALISI SPETTRALE PARAMETRICA



La stima AR di ordine 2 si trova con Yule Walker:

$$\begin{bmatrix} 12.7 & -11 & 9.3 \\ -11 & 12.7 & -11 \\ 9.3 & -11 & 12.7 \end{bmatrix} \begin{bmatrix} 1 \\ a_{1,2} \\ a_{2,2} \end{bmatrix} = \begin{bmatrix} \sigma_w^2 \\ 0 \\ 0 \end{bmatrix} \quad (13.36)$$

$$\begin{bmatrix} a_{1,2} \\ a_{2,2} \end{bmatrix} = - \begin{bmatrix} 12.7 & -11 \\ -11 & 12.7 \end{bmatrix}^{-1} \begin{bmatrix} -11 \\ 9.3 \end{bmatrix} = \begin{bmatrix} 0.92827 \\ 7.1730 \times 10^{-2} \end{bmatrix} \quad (13.37)$$

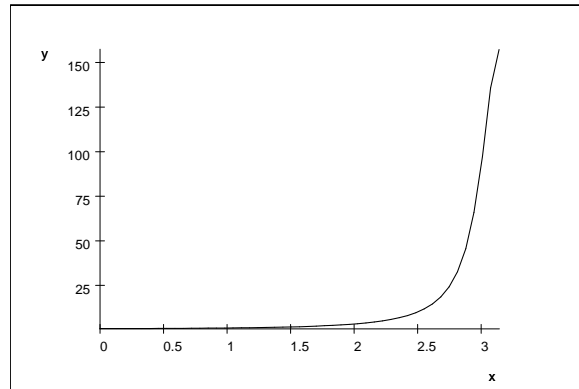
:

$$\begin{bmatrix} 12.7 & -11 & 9.3 \end{bmatrix} \begin{bmatrix} 1 \\ 0.92827 \\ 7.1730 \times 10^{-2} \end{bmatrix} = 3.16 = \sigma_w^2 \quad (13.38)$$

Lo spettro AR è quindi:

$$S_{AR} = \frac{3.16}{|1 + 0.93e^{i\phi} + 0.0717e^{2i\phi}|^2} \quad (13.39)$$

13.5 UN ESEMPIO



Si confrontino i risultati con quelli ottenuti con il periodogramma nel capitolo precedente.

Capitolo 14

La predizione lineare

Si consideri una sequenza $\{x_i\}$ a spettro non uniforme; per fissare le idee, si pensi a $\{x_i\}$ come ad una sequenza uscente da un filtro tutti poli con funzione di trasferimento $1/A_{N_0}(z)$, a sua volta alimentato da una sequenza a spettro bianco $\{w_i\}$ (fig14.1). Si voglia stimare l' i -esimo campione

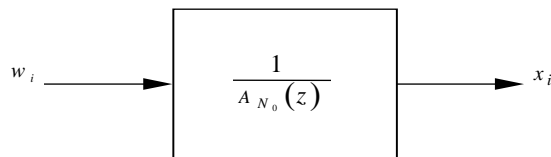


Figura 14.1: Filtro tutti poli

della sequenza, una volta che ne siano noti tutti i campioni precedenti, fino a quello di posto $i - 1$. In particolare, si desidera che la predizione sia combinazione lineare degli N campioni precedenti della sequenza con coefficienti che indicheremo con $a_{n,N}$; $n = 1, \dots, N$; indicando con \hat{x}_i il valore predetto del campione i -esimo:

$$\hat{x}_i = - \sum_{n=1}^N a_{n,N} x_{i-n} \quad (14.1)$$

CAPITOLO 14 LA PREDIZIONE LINEARE

Indicheremo con $\varepsilon_{i,N}$ l'errore di predizione corrispondente ad un predittore di ordine N e cioè realizzato con N coefficienti:

$$\varepsilon_{i,N} = x_i - \hat{x}_i = x_i + \sum_{n=1}^N a_{n,N} x_{i-n} \quad (14.2)$$

La determinazione dei coefficienti $a_{n,N}$ viene fatta in modo ottimo, minimizzando il valore atteso del quadrato dell'errore:

$$E[|\varepsilon_{i,N}|^2] = E[(x_i + \sum_{n=1}^N a_{n,N} x_{i-n})(x_i + \sum_{n=1}^N a_{n,N} x_{i-n})^*] \quad (14.3)$$

Per trovare gli $a_{n,N}$ basta minimizzare rispetto ad essi il valore dell'errore q.m. e quindi imporre che:

$$\frac{\partial E[|\varepsilon_{i,N}|^2]}{\partial a_{k,N}} = 2E[(x_i + \sum_{n=1}^N a_{n,N} x_{i-n})x_{i-k}^*] = 2E[\varepsilon_{i,N}x_{i-k}^*] = 0; \quad k = 1, \dots, N. \quad (14.4)$$

Queste equazioni corrispondono alla **condizione di ortogonalità**. Se l'errore di predizione è minimo, deve essere incorrelato con i dati utilizzati per calcolarlo; infatti, se correlazione vi fosse, si potrebbero utilizzare i dati ancora meglio per ridurre l'errore ancora di piú, contraddicendo l'ipotesi che l'errore fosse minimo. Inoltre, le equazioni (14.4) non sono altro che le equazioni di Yule Walker.

$$E[\varepsilon_{i,N}x_{i-k}^*] = r_k + \sum_{n=1}^N a_{n,N} r_{k-n} = 0; \quad k = 1, \dots, N. \quad (14.5)$$

Indichiamo ora con $E_N(z)$ la trasformata z dell'errore di predizione $\{\varepsilon_{i,N}\}$, con $X(z)$ la trasformata z della sequenza $\{x_i\}$, e con $A_N(z)$ il filtro causale tutti zeri lungo $N + 1$ campioni (e che ha quindi una trasformata z costituita da un polinomio di N grado nella variabile z^{-1}) corrispondente alla funzione di trasferimento tra la sequenza $\{x_i\}$ e l'errore di predizione $\{\varepsilon_{i,N}\}$ abbiamo:

$$E_N(z) = A_N(z)X(z) \quad (14.6)$$

$$A_N(z) = 1 + \sum_{n=1}^N a_{n,N} z^{-n} \quad (14.7)$$

14.1 L'ERRORE DI PREDIZIONE

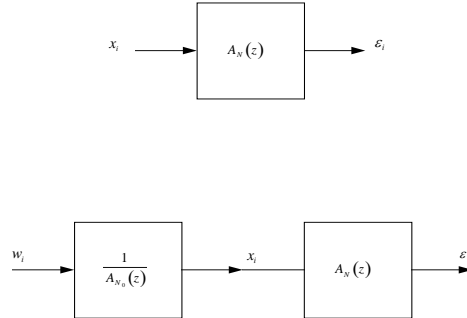


Figura 14.2: Progressivo sbiancamento di uno segnale con spettro tutti poli

Si osservi che $a_{0,N}$ è unitario; $A_N(z)$ è quindi un polinomio **monico**. Per la coincidenza delle equazioni di Yule - Walker, il polinomio $A_N(z)$ è quello stesso che abbiamo trovato a denominatore della funzione di trasferimento tutti poli che sagoma lo spettro bianco di un rumore nella stima spettrale autoregressiva rendendolo al massimo simile allo spettro della sequenza $\{x_i\}$, $E[X(\omega)X^*(\omega)]$. In quel caso, si colorava uno spettro bianco; in questo caso invece si sbianca lo spettro della sequenza $\{x_i\}$. Quindi, al crescere dell'ordine N dello stimatore, lo spettro di $E_N(z)$ tende ad essere tanto più uniforme in frequenza (bianco) perchè la stima spettrale autoregressiva diventa sempre più accurata.

14.1 L'errore di predizione

Indichiamo con:

$$\sigma_{\varepsilon,N}^2 = E[(\varepsilon_{i,N})^2] \quad (14.8)$$

Usando ancora le equazioni di Yule Walker e l'incorrelazione tra errore e dati vediamo che

$$E[|\varepsilon_{i,N}|^2] = E[\varepsilon_{i,N}x_i^*] = r_0 + \sum_{n=1}^N a_{n,N}r_n$$

CAPITOLO 14 LA PREDIZIONE LINEARE

Vista la stazionarietà delle statistiche, si tralascerà l'indice i nelle medie. Se il predittore è di ordine 0, l'errore di predizione coincide con il segnale da predire, e il suo valore quadratico medio coincide con il valore $\sigma_x^2 = E[x_i^2]$. Al crescere dell'ordine del predittore N , il valore quadratico medio dell'errore di predizione deve decrescere, o quantomeno non crescere. Infatti, si usa sempre più informazione per fare la predizione, che deve quindi migliorare progressivamente.

$$\sigma_x^2 \geq \sigma_{\varepsilon, N}^2 \geq \sigma_{\varepsilon, N+1}^2 \quad (14.9)$$

Questa proprietà sarà poi confermata dall'equazione (14.18). Peraltro, il valore quadratico medio dell'errore di predizione non potrà diventare nullo a meno che non sia possibile, per le caratteristiche statistiche della sequenza $\{x_i\}$ osservata, prevederne i valori futuri in modo esatto. Comunque, il miglioramento deve progressivamente rallentare e quindi i dati del passato sempre più remoto saranno sempre meno utili per migliorare la predizione. Se sono poco utili, sono incorrelati con il valore attuale dell'errore di predizione. Infatti, si è già osservato che la sequenza degli errori di predizione tende ad essere sempre più bianca, al crescere di N .

Questo concetto, affermato in modo intuitivo ora, verrà poi dimostrato analiticamente poi.

14.2 Predizione lineare e analisi AR

L'analisi spettrale autoregressiva ci ha fornito un metodo per calcolare un filtro tutti poli, $1/A_N(z)$ che, applicato ad un rumore bianco, ne sagoma lo spettro in modo da riprodurre lo spettro della sequenza osservata. La predizione lineare ci fornisce un filtro tutti zeri $A_N(z)$ che tende a sbiancare la sequenza $\{x_n\}$ osservata, sempre meglio al crescere di N . Questi filtri sono entrambi derivati con le stesse equazioni e dagli stessi parametri e cioè i primi N valori della funzione di autocorrelazione della sequenza $\{x_n\}$, r_0, \dots, r_N (r_0 , in sostanza, funziona da fattore di scala); quindi essi tendono ad essere l'uno il reciproco dell'altro, e lo saranno se dello stesso ordine. Supponiamo infatti di avere un processo bianco, reale, $\{w_i\}$ all'ingresso di un filtro tutti poli $1/A_{N_0}(z)$ e ne sia $\{x_i\}$ l'uscita. Se ora applichiamo alla sequenza $\{x_i\}$ un predittore lineare di ordine N , $N \leq N_0$ l'errore di predizione $\varepsilon_{i, N}$ sarà reale ma non perfettamente bianco; lo sarà perfettamente se $N \geq N_0$. In quelle condizioni, $\varepsilon_{i, N_0} = w_i$.

Una sequenza $\{x_n\}$ prodotta *colorando* una sequenza bianca con un filtro tutti poli di ordine N_0 , sarà perfettamente sbiancabile con un predittore di ordine $N \geq N_0$.

Se l'ordine del predittore cresce ancora, non succede altro perchè gli ulteriori campioni utilizzati per la predizione sono incorrelati con il segnale da predire e quindi il predittore non cambia ulteriormente. Se invece l'ordine N del predittore è $N < N_0$ lo sbiancamento non è perfetto e la sequenza $\varepsilon_{i,N}$ non avrà uno spettro uniforme. Troveremo che il predittore di ordine N corrisponde alla struttura riverberante costituita dai primi N tronchi di linea della struttura effettiva (di ordine N_0). Quindi, alla crescita di N , corrisponde la progressiva identificazione dei successivi coefficienti di riflessione della struttura riverberante, da c_1 a c_{N_0} .

14.2.1 Allungamento del predittore

Vogliamo trovare un criterio per determinare i coefficienti del predittore di ordine $N+1$ a partire da quelli del predittore di ordine N . In sostanza, non dobbiamo fare altro che risolvere le equazioni di Yule Walker, per l'ordine $N+1$, data la soluzione di ordine N .

Accrescere N di 1 implica utilizzare anche x_{i-N-1} per stimare meglio x_i ; tuttavia, parte dell'informazione contenuta in x_{i-N-1} è già nota attraverso i campioni utilizzati per la predizione x_{i-k} ; $1 \leq k \leq N$. *In termini intuitivi, se devo prevedere un evento di domani x_n , uso l'informazione reperibile dal giornale di oggi x_{n-1} e dalla copia arretrata di ieri x_{n-1} . Devo usare anche l'informazione reperibile dalla copia ancora più arretrata, dell'altro ieri x_{n-2} ? A cosa mi servirebbe? L'ovvia risposta è che l'informazione contenuta in una copia ancora più arretrata è utile se a) non è deducibile dal materiale già in mio possesso (le copie più recenti) e b) se questo materiale nuovo poi mi serve per davvero per prevedere l'evento di domani.*

Quindi, la prima operazione da fare è separare quanto già noto da x_{i-N-1} da e utilizzare solo quello che di nuovo (rispetto a prima) è contenuto in questo campione. In questo modo, il processo di miglioramento della predizione risulterà più comprensibile. Stimiamo quindi x_{i-N-1} utilizzando proprio quei campioni **successivi** x_{i-n} ; $n = 1, N$ che abbiamo usato per la predizione di x_i . Questa stima lineare ottima, fatta *all'indietro*, per la stazionarietà del processo, utilizza gli stessi coefficienti di predizione, ma applicati alla sequenza ribaltati nel tempo. *In termini*

CAPITOLO 14 LA PREDIZIONE LINEARE

intuitivi, cerco nel giornale dell'altro ieri le notizie che non sono prevedibili dai giornali di oggi e di ieri. Ma mi serviranno per prevedere degli eventi di domani? La stima ottima **all'indietro** di x_{i-N-1} è \tilde{x}_{i-N-1} :

$$\tilde{x}_{i-N-1} = - \sum_{n=1}^N a_{n,N} x_{i-N-1+n} \quad (14.10)$$

L'errore di predizione all'indietro $\beta_{i,N}$ (*quello che non sapevo*) è:

$$\beta_{i,N} = x_{i-N-1} - \tilde{x}_{i-N-1} \quad (14.11)$$

Ora $\beta_{i,N}$ è incorrelato con $x_{i-n}; n = 1, N$ per costruzione ed è quindi l'innovazione contribuita dal campione x_{i-N-1} ; (*l'informazione contribuita dal giornale dell'altro ieri, non deducibile dai contenuti dei giornali di oggi e di ieri*). Il nuovo stimatore di $\{x_i\}$ che fa uso di questa innovazione è:

$$\hat{x}_{i,N+1} = \hat{x}_{i,N} - c_{N+1} \beta_{i,N} \quad (14.12)$$

$$\varepsilon_{i,N+1} = x_i - \hat{x}_{i,N} + c_{N+1} \beta_{i,N} = \varepsilon_{i,N} + c_{N+1} \beta_{i,N} \quad (14.13)$$

dove c_{N+1} è un numero da determinare minimizzando l'errore q.m.; questo numero è stato indicato con c_{N+1} proprio perchè coincide con l' $N + 1$ coefficiente di riflessione della struttura riverberante corrispondente al filtro che ha prodotto la sequenza $\{x_i\}$. Calcoliamo c_{N+1} applicando il principio di ortogonalità; il nuovo errore di predizione deve quindi essere incorrelato con la nuova informazione $\beta_{i,N}$:

$$E[\varepsilon_{i,N+1} \beta_{i,N}] = 0 \quad (14.14)$$

$$c_{N+1} = - \frac{E[\beta_{i,N} \varepsilon_{i,N}]}{E[(\beta_{i,N})^2]} \quad (14.15)$$

Infatti

$$E[(\beta_{i,N})^2] = E[(\varepsilon_{i,N})^2] \quad (14.16)$$

e per la diseguaglianza di Schwartz, segue:

$$|c_{N+1}| \leq 1 \quad (14.17)$$

Si osservi che c_{N+1} è proporzionale alla correlazione tra gli errori di predizione in avanti e all'indietro e cioè tra l'errore che avrei fatto senza

14.3 LA RICORSIONE DI LEVINSON

la copia del giornale dell'altro ieri, e l'informazione additiva che posso ricavare da quel numero del giornale. Se i due errori di predizione (l'uno noto perchè dipende da dati in mio possesso, l'altro ignoto) sono correlati, posso usare l'uno per ridurre l'altro, migliorando la predizione. Inoltre:

$$E[(\varepsilon_{i,N+1})^2] = E[(\varepsilon_{i,N})^2] \times (1 - c_{N+1}^2) = \sigma_x^2 \prod_{n=1}^{N+1} (1 - c_n^2) \leq E[(\varepsilon_{i,N})^2] \quad (14.18)$$

Abbiamo ora la possibilità di calcolare direttamente c_{N+1} dalla funzione di autocorrelazione; infatti, profittando dell'incorrelazione tra errore di predizione e stima:

$$\begin{aligned} E[(\varepsilon_{i,N})^2] &= E[(x_i - \hat{x}_{i,N})(x_i - \hat{x}_{i,N})] = \\ &= E[(x_i - \hat{x}_{i,N})x_i] = r_0 + \sum_{n=1}^N a_{n,N}r_n \end{aligned} \quad (14.19)$$

Similmente:

$$E[\varepsilon_{i,N}\beta_{i,N}] = r_{N+1} + \sum_{n=1}^N a_{n,N}r_{N+1-n} \quad (14.20)$$

$$c_{N+1} = -\frac{r_{N+1} + \sum_{n=1}^N a_{n,N}r_{N+1-n}}{r_0 + \sum_{n=1}^N a_{n,N}^*r_n} = -\frac{r_{N+1} + \tilde{\mathbf{r}}_N^* \mathbf{a}_N}{r_0 + \tilde{\mathbf{r}}_N^* \tilde{\mathbf{a}}_N} \quad (14.21)$$

dove:

$$\tilde{\mathbf{r}}_N^* = \begin{bmatrix} r_{-N} \\ \dots \\ r_{-2} \\ r_{-1} \end{bmatrix}; \quad \mathbf{a}_N = \begin{bmatrix} a_{1,N} \\ a_{2,N} \\ \dots \\ a_{N,N} \end{bmatrix}; \quad \tilde{\mathbf{a}}_N = \begin{bmatrix} a_{N,N}^* \\ \dots \\ a_{2,N}^* \\ a_{1,N}^* \end{bmatrix};$$

14.3 La ricorsione di Levinson

Come abbiamo visto nel paragrafo precedente, il predittore di ordine $N + 1$ si deriva da quello di ordine N utilizzando l'errore di predizione all'indietro e cioè l'innovazione apportata dal campione di posto $i - N - 1$

CAPITOLO 14 LA PREDIZIONE LINEARE

rispetto agli altri campioni già utilizzati per la predizione. L'errore di predizione è quindi:

$$\varepsilon_{i,N+1} = x_i - \hat{x}_{i,N} + c_{N+1}\beta_{i,N} = \varepsilon_{i,N} + c_{N+1}\beta_{i,N} \quad (14.22)$$

In termini di trasformata z , indicando con

$$E_N(z) = A_N(z)X(z) \quad (14.23)$$

$$E_{N+1}(z) = E_N(z) + c_{N+1}B_N(z) \quad (14.24)$$

dove abbiamo indicato con $B_N(z)$ la trasformata z dell'errore di predizione all'indietro e cioè:

$$B_N(z) = z^{-N-1}A(z^{-1})X(z)$$

si ricava:

$$A_{N+1}(z) = A_N(z) + c_{N+1}z^{-N-1}A_N(z^{-1}) = A_N(z) + c_{N+1}z^{-1} \tilde{A}_N(z) \quad (14.25)$$

Abbiamo riottenuto la ricorsione di Levinson. **La progressiva identificazione di $A_{N_0}(z)$ avviene attraverso la progressiva identificazione dei coefficienti di riflessione della struttura riverberante corrispondente.**

Infatti, per $N > N_0$ le equazioni di Yule Walker ci garantiscono la corretta analisi spettrale ed un errore di predizione perfettamente bianco, se la sequenza di partenza era l'uscita di una struttura riverberante di ordine N_0 alimentata da una sequenza bianca. Accrescendo progressivamente l'ordine del predittore, a partire da $N = 1$, calcoliamo in successione i coefficienti di riflessione, che alla fine risultano essere quelli corretti, perchè lo sbiancatore (la funzione di trasferimento tra la sequenza $\{x_i\}$ e la sequenza errore di predizione $\{\varepsilon_i\}$) diventa perfetto. Quindi, i coefficienti di riflessione erano stati individuati correttamente sin dal momento in cui erano stati calcolati.

Pertanto, se si utilizzano gli stessi campioni della funzione di autocorrelazione, i coefficienti di riflessione che risultano dalla predizione lineare sono gli stessi della struttura riverberante che ha colorato la sequenza bianca $\{w_i\}$ iniziale. Ovviamente, se invece il processo nasce da meccanismo completamente diverso, è possibile che questi coefficienti non abbiano gran significato; tuttavia, l'esempio delle Direzioni di Arrivo che

14.3 LA RICORSIONE DI LEVINSON

verrà proposto in un capitolo successivo dimostra che questa tecnica di stima spettrale ha una validità che si estende a numerosissimi campi applicativi.

Per avere una motivazione intuitiva di questo importante risultato, si pensi ad una struttura riverberante disadattata da un lato, eccitata da una sequenza bianca dalla parte disadattata. Si supponga poi che il sistema sia eccitato da una successione di impulsi indipendenti, e che un coefficiente di riflessione c_1 corrisponda alla riflessione di una frazione c_1 di questi impulsi, mentre invece il restante segnale viene trasmesso. E' comprensibile che il primo campione della funzione di autocorrelazione dell'uscita dipenda solo dal primo coefficiente di riflessione c_1 ; una frazione c_1 dei campioni del segnale in ingresso si riflette e riemerge all'uscita dopo un ritardo pari a T sovrapponendosi al campione successivo (o alla sua parte che è stata trasmessa direttamente attraverso la prima interfaccia). Nasce così la correlazione a ritardo T , r_1 . Vediamo quale frazione del segnale viene ritardata di $2T$ e cioè contribuisce a formare r_2 : contribuiscono:

- i campioni dell'ingresso che hanno subito una riflessione alla seconda interfaccia (una frazione pari a c_2);
- quella frazione dei campioni che dopo avere fatto una prima volta il percorso ingresso - prima interfaccia, lo ripete. Questa frazione è pari a c_1^2 .

Questa frazione di segnale è però correlata anche con il segnale che abbiamo trovato dopo solo un intervallo T , e quindi aveva fatto solo una volta il percorso $c_1 \div c_0$; se pertanto cerchiamo di rimuovere questa componente cercando quanto è decorrelato con il campione precedente, la componente di correlazione residua (il coefficiente di correlazione parziale o PARCOR come viene chiamato in alcune trattazioni) è legato solo al coefficiente di riflessione c_2 e così via.

In altre parole, la correlazione tra gli errori di predizione in avanti ed all'indietro è legata solo alla frazione di segnale che ha percorso il tratto fino alla N -esima interfaccia senza subire riflessioni intermedie ma venendo solo trasmesso; quindi la sua ampiezza è proporzionale al solo coefficiente di riflessione c_N . Un altro risultato importante è che, per la ricorsione di Levinson, ripetendo la dimostrazione fatta, i polinomi $A_N(z)$ sono a fase minima, se i coefficienti di riflessione sono minori di 1, cosa peraltro dimostrata in precedenza.

14.3.1 Vantaggi computazionali

La ricorsione di Levinson comporta notevoli vantaggi computazionali nella soluzione delle equazioni di Yule - Walker: infatti la soluzione di un sistema di N equazioni in N incognite comporta un numero di moltiplicazioni proporzionale a N^3 ; invece, se si utilizza la ricorsione il costo diventa proporzionale a N^2 . Infatti, allungando il predittore dobbiamo calcolare l' $(N + 1)$ -esimo coefficiente di riflessione utilizzando la formula (14.15); questo comporta $2N$ moltiplicazioni ed una divisione; inoltre, dobbiamo calcolare i nuovi $\{a_{h,N+1}\}$ e questo comporta altre N moltiplicazioni; quindi, circa $3N$ moltiplicazioni per ogni passo e pertanto circa un totale di $1.5 \times N^2$ moltiplicazioni invece di una proporzionalità ad N^3 , con indubbi vantaggi.

14.4 Il filtro a traliccio non ricorsivo

Possiamo ora costruire un sistema numerico per il calcolo degli errori di predizione in avanti ed all'indietro, basato appunto sulla ricorsione di Levinson. Questo sistema sarà poi utilizzabile come uno sbiancatore, a minima fase, di sequenze colorate.

Supponiamo di disporre di $\varepsilon_{i,N}$ e $\beta_{i,N}$; possiamo ottenere, tenendo conto del ritardo crescente del campione stimato all'indietro, senza contraddizione con la (14.22):

$$\varepsilon_{i,N+1} = \varepsilon_{i,N} + c_{N+1}\beta_{i,N} \quad (14.26)$$

$$\beta_{i,N+1} = \beta_{i-1,N} + c_{N+1}\varepsilon_{i-1,N} \quad (14.27)$$

Ne nasce la struttura numerica di figura 14.3 nota come **filtro a traliccio non ricorsivo**; è evidente la somiglianza con il filtro a traliccio ricorsivo corrispondente ad una struttura riverberante: le differenze sono nei segni dei collegamenti incrociati e nei sensi della propagazione del segnale nei due rami. Il filtro a traliccio non ricorsivo serve non solo da sbiancatore della sequenza, fornendo in uscita $\varepsilon_{i,N}$, ma anche da predittore $\{\hat{x}_{i,N}\}$; la predizione è la differenza tra ingresso ed uscita del filtro ed è funzione solo dei dati passati; essa è disponibile **prima** dell'arrivo del prossimo campione della sequenza e quindi prima che sia disponibile $\varepsilon_{i,N}$.

I coefficienti di riflessione possono essere derivati dai valori della funzione di autocorrelazione, come già visto; però possono anche essere ottenuti dalla correlazione mutua tra gli errori di predizione in avanti e

14.5 DESCRIZIONI ALTERNATIVE DELL'AUTOCORRELAZIONE

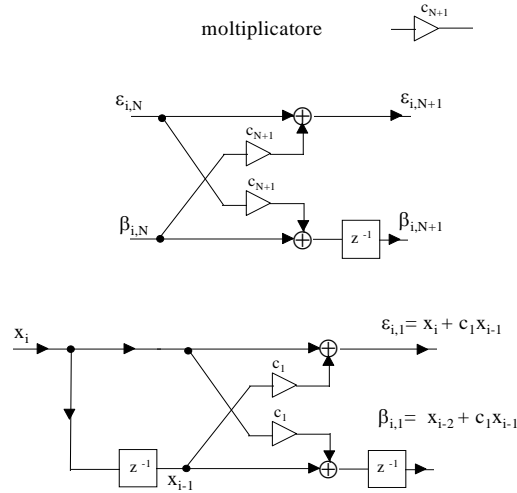


Figura 14.3: Filtro a traliccio non ricorsivo: cella intermedia e cella iniziale

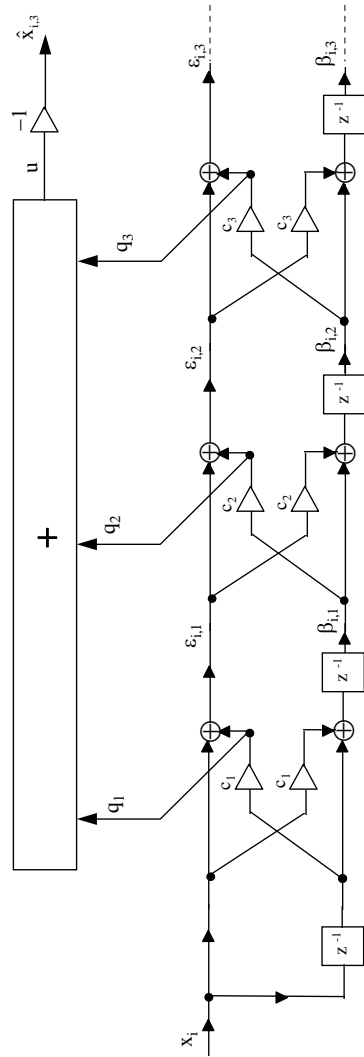
all'indietro, normalizzati. In questo modo il filtro a traliccio ricorsivo assume la forma di figura 14.5: l'elemento interessante di questa configurazione è che il sistema può essere reso adattativo, stimando la correlazione tra $\varepsilon_{i,N}, \beta_{i,N}$ con stimatori che p.e. utilizzano solo gli ultimi M campioni, **obliando** quelli passati, con dei criteri di progressività che vedremo nel capitolo sui sistemi adattativi.

Il filtro a traliccio nella configurazione di figura 14.5 è perciò un sistema autonomo, che può calcolare, al variare del tempo e delle statistiche dell'ingresso, i coefficienti di riflessione e pertanto lo spettro dell'ingresso, la predizione ottima ed infine l'errore di predizione e cioè la sequenza sbiancata.

14.5 Descrizioni alternative dell'autocorrelazione

Nei paragrafi precedenti abbiamo visto come i primi N valori di una funzione di autocorrelazione $r_0 \dots r_N$ di una qualsiasi sequenza colorata $\{x_n\}$ siano esprimibili in funzione dei coefficienti del polinomio $A_N(z)$

CAPITOLO 14 LA PREDIZIONE LINEARE



$$q_1 = \epsilon_{i,1} - x_i$$

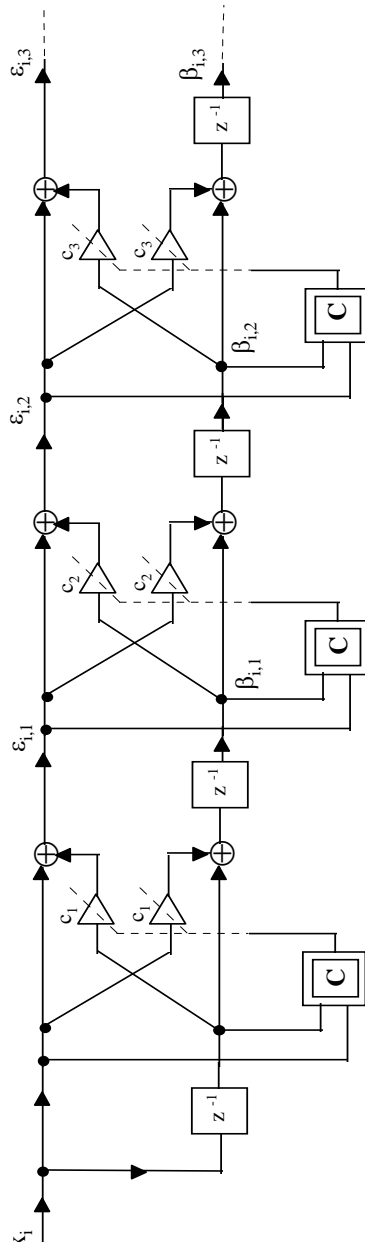
$$q_2 = \epsilon_{i,2} - \epsilon_{i,1}$$

$$q_3 = \epsilon_{i,3} - \epsilon_{i,2}$$

$$\text{quindi: } u = q_1 + q_2 + q_3 = \epsilon_{i,3} - x_i = -\hat{x}_{i,3}$$

Figura 14.4: Filtro a traliccio non ricorsivo: uso come predittore

14.5 DESCRIZIONI ALTERNATIVE DELL'AUTOCORRELAZIONE



Dove:

$$\begin{array}{c}
 \beta_{i,N} \quad \beta_{i,N+1} \\
 \epsilon_{i,N} \quad \epsilon_{i,N} \\
 \boxed{C} \\
 E[\beta_{i,N} \cdot \epsilon_{i,N}] \\
 E[(\beta_{i,N})^2]
 \end{array}$$

Figura 14.5: Filtro a traliccio adattativo

CAPITOLO 14 LA PREDIZIONE LINEARE

per la predizione $a_{n,N}$, $n = 1, \dots, N$:

$$A_N(z) = 1 + \sum_{n=1}^N a_{n,N} z^{-n}$$

nonchè al valore quadratico medio dell'errore di predizione residuo $E[(\varepsilon_{n,N}^2)]$; ancora, sono esprimibili in funzione degli N coefficienti di riflessione c_i , $i = 1, \dots, N$, ancora insieme a $E[(\varepsilon_{n,N}^2)]$. Un terzo modo di descrivere lo spettro è dato dalle posizioni degli zeri di $A_N(z)$ ed ancora altri modi saranno visti nel capitolo 15, con le LSP (Line Spectral Pairs). Di tutte queste descrizioni alternative di una funzione di autocorrelazione (o di uno spettro) quella dei coefficienti di riflessione è la più vicina ad una descrizione fisica. Difatti, basta assicurarsi che $|c_i| < 1$, $i = 1, \dots, N$ perchè automaticamente risulti che la matrice di covarianza \mathbf{R} sia semi definita positiva e cioè che per tutti i suoi autovalori risulti $\lambda_i \geq 0$. Se poi è comunque sempre possibile passare da un modo di descrizione ad un altro, è anche vero che i coefficienti di riflessione sono meno sensibili ad effetti di quantizzazione rispetto agli $a_{n,N}$, come visto nel capitolo 7. Infine, il difetto della descrizione mediante gli r_i è la difficile verifica della positività della matrice di covarianza \mathbf{R} , e quindi le derivanti difficoltà nel troncamento di una serie di valori dell'autocorrelazione (la finestrazione vista nel capitolo 10) che può portare a valori negativi dello spettro e quindi a qualche $\lambda_i < 0$; abbiamo già osservato che la soluzione a questo problema, nel caso della stima spettrale AR, è l'estrapolazione della sequenza delle autocorrelazioni con le equazioni di Yule Walker, cosa che poi consente di avere uno spettro positivo. Tutte queste difficoltà sono risolte dall'uso dei coefficienti di riflessione; infatti, con il metodo di stima spettrale AR di Burg, molto usato e che sarà visto nel prossimo paragrafo, la stima spettrale avviene con la stima diretta dei c_i .

14.6 Ordine del predittore

14.6.1 Risoluzione spettrale

Se i dati a disposizione hanno lunghezza infinita e vi è stazionarietà statistica, l'analisi spettrale non pone problemi: c'è modo di determinare la funzione di autocorrelazione con la qualità desiderata. Il periodogramma

14.6 ORDINE DEL PREDITTORE

potrà dare risultati con la risoluzione spettrale e la dispersione statistica volute, senza dovere fare compromessi tra l'una e l'altra. Il problema dell'analisi spettrale è quindi **sempre e solo** collegato ad una limitata disponibilità di dati. E' interessante valutarne gli effetti nel caso dell'analisi spettrale parametrica. L'affidabilità del modello da estrarre dai dati diminuisce, se ne aumenta la complessità. Al crescere dell'ordine N del predittore (o dello stimatore) tende a peggiorare il condizionamento della matrice \mathbf{R} ; inoltre, si dovranno stimare un numero crescente di campioni dell'autocorrelazione dei dati e l'effetto degli errori di questa stima sarà anch'esso crescente. Similmente, nel caso del periodogramma, l'affidabilità dei risultati decresce con il crescere del numero di parametri da stimare; il numero di parametri qui coincide con il numero delle celle in cui è suddiviso l'asse delle frequenze e pertanto vi è una relazione diretta tra affidabilità della stima e risoluzione spettrale.

In questo, la stima AR ha una grande differenza con il periodogramma, come si vedrà nel caso delle direzioni di arrivo (DOA): **la risoluzione spettrale AR non dipende sostanzialmente dalla lunghezza del predittore**, ma piuttosto dalla vicinanza al cerchio unitario degli zeri del polinomio $A_N(z)$. Questa vicinanza dipende sia dall'entità del rumore che accompagna i dati sia dalla larghezza di banda del segnale stimato; inoltre la dispersione della stima dei campioni dell'autocorrelazione utilizzati per costruire \mathbf{R} porta alla dispersione delle posizioni degli zeri di $A_N(z)$ e quindi ad errori di stima delle posizioni dei poli del modello AR ma non necessariamente a picchi spettrali meno pronunciati.

Nel caso di un segnale sinusoidale di frequenza ignota (*il caso dell'orologiaio*), se il rapporto segnale - disturbo è buono, se il segnale corrisponde ad una sola senoide, e se inoltre la stima della funzione di correlazione è buona, basterebbero anche 2 soli campioni (ma ben stimati) della funzione di autocorrelazione, per avere una risoluzione spettrale al limite infinita. Qui la differenza con la stima spettrale fatta con il periodogramma è veramente molto grande. Infatti, c'è l'ipotesi essenziale che vi sia una sola senoide: la stima spettrale si riduce alla stima della frequenza di quella senoide. Se il numero di sinusoidi fosse $m > 1$, avremmo invece bisogno almeno di $m + 1$ campioni della funzione di autocorrelazione per stimare le m frequenze ed il valore q.m. del segnale. Infine, osserviamo ancora che questo è il tipo di analisi spettrale che potrebbe essere fatto per misurare la frequenza di un singolo oscillatore: l'unica variabile da stimare è la frequenza. Poi, si possono usare le

equazioni di Yule Walker per estrapolare la funzione di autocorrelazione su tutto l'asse dei tempi completando l'analisi.

14.6.2 Lunghezza dei dati e criterio di Akaike

Se sono disponibili M campioni dei dati, l'ordine N del predittore sarà necessariamente $N < M$. Volendo, però, N può essere fatto crescere fino ad $M - 1$, senza rischi di instabilità, utilizzando la via della stima diretta dei c_i piuttosto che degli r_i . Un metodo per accrescere senza rischi l'ordine del predittore è il metodo di Burg (o della massima entropia). Invece di stimare i campioni della funzione di autocorrelazione, si stimano appunto direttamente i coefficienti di riflessione e per garantire che siano sempre minori di 1, si usa la formula:

$$c_{h+1} = -2 \frac{E[\varepsilon_{i,h} \beta_{i,h}]}{E[\varepsilon_{i,h}^2 + \beta_{i,h}^2]} \quad (14.28)$$

dove, non potendo realizzare l'operatore di valore atteso, $E[.]$, lo si sostituisce con la media su tutti i dati:

$$c_{h+1} = - \frac{\sum_{m=0}^M \varepsilon_{m,h} \beta_{m,h}}{\sum_{m=0}^M \frac{(\varepsilon_{m,h}^2 + \beta_{m,h}^2)}{2}} \quad (14.29)$$

Poi si calcolano gli errori di predizione di ordine $h + 1$ e così via, eventualmente fino ad avere $h = M - 1$; allora, ovviamente l'errore di predizione diventa eguale a zero. Il metodo di Burg (ma sarebbe la stessa cosa per la corrispondente stima spettrale AR) è indicato come metodo a massima entropia perchè si verifica che:

- assegnati i primi N valori della funzione di autocorrelazione;
- determinando tutti gli altri valori della funzione di autocorrelazione mediante le equazioni di Yule Walker;
- trasformando con Fourier la funzione di autocorrelazione e trovando la densità spettrale;
- calcolando l'entropia di questa densità spettrale (il valore medio del logaritmo);
- **si verifica che tale entropia è massima** tra tutte le densità spettrali (tutte positive) caratterizzate dall'avere i primi N valori dell'autocorrelazione eguali a quelli assegnati.

In altre parole, nella stima AR lo spettro è il più frastagliato possibile, compatibilmente con i valori assegnati della funzione di autocorrelazione. Al contrario, nella stima MA, lo spettro è il più liscio possibile (la funzione di autocorrelazione viene troncata appena possibile, compatibilmente con i campioni dell'autocorrelazione misurati e con l'avere un spettro tutto positivo).

Tuttavia, pur disponendo così di un metodo che garantisce di poter avere predittori di ordine elevatissimo, è sbagliato accrescere troppo N rispetto ad M ; i coefficienti di riflessione trovati sono solo una pessima stima di quelli *veri*. Con un'altra realizzazione del processo, sarebbero tutt'altri. Esiste un criterio per limitare l'ordine del predittore, noto come criterio di Akaike: si studia l'andamento del cosiddetto Final Prediction Error:

$$\overline{\varepsilon_F^2} = \frac{M + N - 1}{M - N - 1} \sum_{m=0}^M \varepsilon_{m,N}^2 \quad (14.30)$$

e si cerca il valore di N per cui lo FPE è minimo; se $N \ll M$ lo FPE decrescerà con N , come consueto; poi, se il numero di parametri del modello diventa esagerato, e N si avvicina troppo ad M , lo FPE tende a crescere di nuovo. Spesso è prudente mantenere N alquanto inferiore a quanto proposto da Akaike.

14.7 Condizionamento di \mathbf{R}

Abbiamo visto come il predittore \mathbf{a}_N venga determinato risolvendo, eventualmente con la ricorsione di Levinson, il sistema:

$$\mathbf{R}\mathbf{a}_N = -\mathbf{r}_N \quad (14.31)$$

Abbiamo anche visto che, se gli elementi della matrice \mathbf{R} derivano dai campioni di una funzione di autocorrelazione, allora il polinomio

$$A_N(z) = 1 + \sum_{n=1}^N a_{n,N} z^{-n} \quad (14.32)$$

è a fase minima; infatti, è denominatore della funzione di trasferimento corrispondente alla struttura riverberante che ha sagomato spettralmente

CAPITOLO 14 LA PREDIZIONE LINEARE

il processo, troncata all'ennesimo coefficiente di riflessione. Tuttavia $\underline{\mathbf{R}}$ può essere singolare, anzi in generale dovrebbe proprio esserlo. Infatti, se la sequenza $\{x_n\}$ che osserviamo è proveniente da un campionamento senza equivocazione di un segnale passa basso, il contenuto spettrale alle frequenze vicine alla frequenza di Nyquist dovrà essere praticamente nullo; ora poichè l'errore di predizione

$$\varepsilon_{i,N} = x_i + \sum_{n=1}^N a_{n,N} x_{i-n} \quad (14.33)$$

ha spettro $E[|E(z)|^2]$ bianco, se N è abbastanza grande, poichè

$$E[|E(z)|^2] = E[|X(z)|^2] |A_N(z)|^2 \quad (14.34)$$

allora $|A_N(z)|^2$ è molto grande, almeno vicino alla frequenza di Nyquist, dove $E[|X(z)|^2]$ è molto piccolo. Ciò implica che la matrice $\underline{\mathbf{R}}$ è prossima alla singolarità perchè il vettore che è soluzione del sistema ha un modulo grande, malgrado che sia termini noti che coefficienti siano dello stesso ordine di grandezza.

Pertanto, **non possiamo risolvere le equazioni di Y-W senza precauzioni.**

Da un altro punto di vista, gli autovalori λ_i della matrice $\underline{\mathbf{R}}$, se circolante, corrispondono alle energie delle componenti della Trasformata Discreta di Fourier della sequenza. Anche se la matrice $\underline{\mathbf{R}}$ non è circolante, ci attendiamo che gli autovalori siano collegati alle ampiezze dello spettro di potenza della sequenza $\{x_n\}$. Se la sequenza è passa basso, alcuni dei valori spettrali saranno pressochè nulli, e così saranno alcuni tra gli autovalori. Ancora, ci attendiamo che la $\underline{\mathbf{R}}$ sia singolare o almeno mal condizionata.

Per risolvere il problema del cattivo condizionamento, ci comporteremo come se vi fosse un disturbo sovrapposto alla sequenza osservata $\{x_i\}$, come vedremo nel prossimo paragrafo.

14.8 Stima con rumore

Supponiamo che la sequenza osservata $\{x_i\}$ sia disturbata da un rumore bianco $\{z_i\}$; supponiamo inoltre che $\{x_i\}$ sia una sequenza colorata; ad

esempio potrebbe essere l'uscita di una struttura riverberante con funzione di trasferimento $1/A(z)$ alimentata da una sequenza bianca $\{w_i\}$; per un'idea intuitiva, si pensi a $\{w_i\}$ come al suono generato da un *coro di voci bianche* (tanto per ricordare che si tratta di una sequenza bianca e gaussiana) in una *cattedrale* che causa delle riverberazioni corrispondenti alla funzione di trasferimento $1/A(z)$; (peraltro, una cattedrale vera corrisponde non ad una sola struttura riverberante ma al parallelo di tante di queste strutture e pertanto la funzione di trasferimento che si potrebbe misurare avrebbe non solo poli ma anche zeri; tuttavia, per proseguire l'esempio, supponiamo che la funzione di trasferimento sia costituita solo da poli). *Un ascoltatore rileva $\{y_i\}$, e cioè ascolta il coro $\{w_i\}$ riverberato attraverso la cattedrale $1/A(z)$ ma è disturbato dal brusio dei vicini $\{z_i\}$* ; si pongono due problemi nuovi:

(i) ricostruire $\{x_i\}$ usando N campioni del segnale riverberato e disturbato (la sequenza $\{y_i\}$), attenuando l'effetto del disturbo $\{z_i\}$; si vuole cioè fare una stima ottima $\{\hat{x}_{i,N}\}$ della sequenza $\{x_i\}$ combinando linearmente N valori della sequenza osservata $\{y_i\}$; per mantenere l'omogeneità con la trattazione precedente in cui si considera la predizione, supponiamo per ora che i campioni utilizzati per la stima siano tutti precedenti il campione y_i e siano cioè y_{i-1}, \dots, y_{i-N} . Si potrebbe tuttavia generalizzare il problema, utilizzando per la stima anche il campione y_i o addirittura i campioni successivi trasformando il problema da quello di stima predittiva a quello di un'interpolazione.

(ii) stimare direttamente il segnale di ingresso $\{w_i\}$ (*il coro*) con uno stimatore lungo N campioni, e cioè determinare $\{\hat{w}_{i,N}\}$.

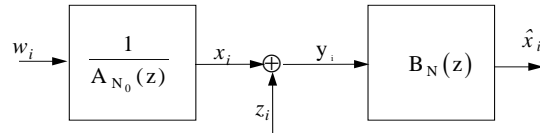


Figura 14.6: Somma di rumore bianco ad un processo colorato

Passando al calcolo, indichiamo con $\{y_i\}$ i campioni riverberati $\{x_i\}$ disturbati dal rumore additivo $\{z_i\}$:

$$y_i = x_i + z_i \quad (14.35)$$

CAPITOLO 14 LA PREDIZIONE LINEARE

Vogliamo stimare la sequenza $\{x_i\}$, combinando tra loro linearmente N campioni della sequenza $\{y_i\}$: questo problema è immediatamente riconducibile a quanto fatto finora, minimizzando l'errore di stima quadratico medio. Si applica quindi il principio di ortogonalità tra errore di stima (piuttosto che l'errore di predizione, ma in sostanza è la stessa cosa; cambia solo la posizione dei campioni assegnati rispetto a quella del campione da stimare) e dati; sia $\{\hat{x}_{i,N}\}$ il segnale stimato, funzione lineare dei vari campioni della sequenza $\{y_i\}$:

$$\hat{x}_{i,N} = - \sum_{n=1}^N b_{n,N} y_{i-n} \quad (14.36)$$

L'errore di stima è:

$$\gamma_{i,N} = x_i - \hat{x}_i \quad (14.37)$$

e se ne impone l'incorrelazione con i dati, per garantirne l'ottimalità:

$$E[\gamma_{i,N} y_{i-k}] = 0; \quad k = 1, \dots, N \quad (14.38)$$

È opportuno osservare che l'errore di stima $\{\gamma_{i,N}\}$ non è disponibile, poichè la sequenza $\{x_i\}$ non è nota; incidentalmente, si ricordi che l'errore di predizione $\{\varepsilon_{i,N}\}$ della sequenza $\{y_i\}$ è invece sempre disponibile, anche se con il ritardo di un campione. Utilizzando l'incorrelazione tra $\{x_i\}$ e $\{z_i\}$, lo stimatore

$$\underline{\mathbf{b}}_N = \{b_{i,N}\}; \quad i = 1, \dots, N \quad (14.39)$$

è dato dal sistema:

$$\begin{bmatrix} r_0 + \sigma_z^2 & r_{-1} & r_{-2} & \dots & r_{-N+1} \\ r_1 & r_0 + \sigma_z^2 & r_{-1} & \dots & r_{-N+2} \\ r_2 & r_1 & r_0 + \sigma_z^2 & \dots & r_{-N+3} \\ \dots & \dots & \dots & \dots & \dots \\ r_{N-1} & r_{N-2} & r_{N-3} & \dots & r_0 + \sigma_z^2 \end{bmatrix} \begin{bmatrix} b_{1,N} \\ b_{2,N} \\ b_{3,N} \\ \dots \\ b_{N,N} \end{bmatrix} = - \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ \dots \\ r_N \end{bmatrix} \quad (14.40)$$

e cioè

$$\underline{\mathbf{b}}_N = - (\underline{\mathbf{R}} + \sigma_z^2 \underline{\mathbf{I}})^{-1} \underline{\mathbf{r}}_N \quad (14.41)$$

Quindi, se al segnale da stimare è sommato del rumore, alla matrice \mathbf{R} si somma la matrice identica, moltiplicata per la varianza del rumore σ_z^2 . Gli autovalori della matrice $\mathbf{R} + \sigma_z^2 \mathbf{I}$ sono $\lambda_i + \sigma_z^2$; se la matrice \mathbf{R} è una matrice di covarianza, $\lambda_i \geq 0$; quindi $\lambda_i + \sigma_z^2 > 0$.

In conclusione, il problema (i) della stima di una sequenza $\{x_i\}$, disturbata da un rumore additivo $\{z_i\}$, effettuata con uno stimatore di lunghezza N , (la ricerca di $\{\hat{x}_{i,N}\}$) è sempre ben condizionato; infatti, non si devono effettuare amplificazioni illimitate perchè porterebbero a segnali troppo influenzati dal disturbo. Si osservi che per risolvere il problema (i) basta conoscere la matrice di covarianza \mathbf{R} del segnale osservato $\{y_i\}$, il valore quadratico medio del rumore aggiunto, ed il vettore \mathbf{r}_N ; inoltre, non è necessario che la stima sia riferita al campione attuale; in altre parole, si può decidere di stimare un campione precedente della sequenza $\{x_i\}$; ciò altera il vettore dei termini noti \mathbf{r}_N senza alterare la matrice \mathbf{R} . In quel caso, la ricorsione di Levinson non sarebbe più applicabile per la determinazione della soluzione, ma l'errore di stima sarebbe più piccolo, perchè utilizzeremmo per fare la stima dati utili anche se successivi al tempo di generazione del segnale.

Calcoliamo ora l'errore di stima quadratico medio risultante, approfittando dell'incorrelazione tra errore di predizione e dati $y_{i-n}; n = 1 : N$:

$$\begin{aligned} \gamma_{i,N} &= x_i - \hat{x}_i = x_i + \sum_{n=1}^N b_{n,N} y_{i-n}; \\ E[\gamma_{i,N}^2] &= \sigma_\gamma^2 = E \left[\left(x_i + \sum_{n=1}^N b_{n,N} y_{i-n} \right) \left(x_i + \sum_{m=1}^N b_{m,N} y_{i-m} \right) \right] \\ &= E \left[\left(x_i + \sum_{n=1}^N b_{n,N} (x_{i-n} + z_{i-n}) \right) x_i \right] \\ \sigma_\gamma^2 &= \sigma_x^2 + E \left[\sum_{n=1}^N b_{n,N} (x_{i-n} + z_{i-n}) x_i \right] \\ &= \sigma_x^2 + \sum_{n=1}^N b_{n,N} r_n = \sigma_x^2 - \mathbf{r}_N^T (\mathbf{R} + \sigma_z^2 \mathbf{I})^{-1} \mathbf{r}_N \end{aligned}$$

Possiamo ora risolvere il problema (ii) e da quella via dare un ulteriore significato anche all'errore di predizione; stimiamo con uno stimatore

CAPITOLO 14 LA PREDIZIONE LINEARE

lungo N campioni, la sequenza $\{w_i\}$ che alimenta il filtro tutti poli che genera la sequenza $\{x_i\}$; (*ricostruiamo il coro di voci bianche, senza la riverberazione della cattedrale e attenuando quanto possibile il brusio dei vicini*). Supponiamo che la stima questa volta venga fatta utilizzando anche il campione corrente y_i , e quindi la sommatoria corrispondente allo stimatore inizi da $n = 0$ e non da $n = 1$. I coefficienti dello stimatore sono sempre ottenuti mediante il principio di ortogonalità tra errore di stima e dati, per minimizzare l'errore quadratico medio.

$$y_i = x_i + z_i; \quad x_i = - \sum_{n=1}^{N_0} a_{n,N_0} x_{i-n} + w_i \quad (14.42)$$

$$\hat{w}_{i,N} = - \sum_{n=0}^N c_{n,N} y_{i-n}; \quad (14.43)$$

$$E[(w_i + \sum_{h=0}^N c_{h,N} y_{i-h}) y_{i-k}] = 0; \quad 0 \leq k \leq N \quad (14.44)$$

Mentre la matrice del sistema è sempre la matrice di autocorrelazione della variabile $\{y_i\}$ e cioè $\mathbf{R} + \sigma_z^2 \mathbf{I}$, il vettore dei termini noti è \mathbf{r}' con componenti $r'_k = - E[w_i y_{i-k}]$ e pertanto ha tutte le componenti nulle per la bianchezza della sequenza $\{w_i\}$ tranne la prima eguale a σ_w^2 .

$$\begin{bmatrix} r_0 + \sigma_z^2 & r_{-1} & r_{-2} & \dots & r_{-N} \\ r_1 & r_0 + \sigma_z^2 & r_{-1} & \dots & r_{-N+1} \\ r_2 & r_1 & r_0 + \sigma_z^2 & \dots & r_{-N+2} \\ \dots & \dots & \dots & \dots & \dots \\ r_N & r_{N-1} & r_{N-2} & \dots & r_0 + \sigma_z^2 \end{bmatrix} \begin{bmatrix} c_{0,N} \\ c_{1,N} \\ c_{2,N} \\ \dots \\ c_{N,N} \end{bmatrix} = - \begin{bmatrix} \sigma_w^2 \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix} \quad (14.45)$$

Vi sono due casi possibili.

14.8.1 Il caso $\sigma_z^2 = 0$

Si ritrovano esattamente le equazioni di Yule - Walker che, nel caso in cui il termine noto sia eguale a σ_w^2 , o a meno di un fattore di scala, danno come soluzione la sequenza $1, a_{1,N}, \dots, a_{N,N}$ la cui trasformata z è il solito

polinomio $A_N(z)$. Pertanto, in questo caso, la sequenza stimata $\{\widehat{w}_{i,N}\}$ coincide con $\{\varepsilon_{i,N}\}$ errore di predizione della sequenza $\{y_i\}$; ne deduciamo che $\{\varepsilon_{i,N}\}$ (a meno di un possibile fattore di scala pari a $\sigma_{\varepsilon}^2/\sigma_w^2$) è una stima della sequenza di ingresso della struttura tutti poli che ha generato la sequenza $\{x_i\}$; si comprenderà ancora meglio perchè $\{\varepsilon_{i,N} = \widehat{w}_{i,N}\}$ tenda ad essere una sequenza bianca, non appena lo stimatore abbia una lunghezza sufficiente. È infatti la stima di una sequenza bianca. Inoltre, si può ancora applicare la ricorsione di Levinson.

14.8.2 Il caso $\sigma_z^2 \neq 0$

Se $\sigma_z^2 \neq 0$, allora, anche per $N \rightarrow \infty$ la sequenza $\{\widehat{w}_{i,N}\}$ non diventa bianca a differenza della sequenza $\{\varepsilon_{i,N}\}$. Quindi, per $\sigma_z^2 \neq 0$, $\{\varepsilon_{i,N} \neq \widehat{w}_{i,N}\}$; in particolare la sequenza $\{\widehat{w}_{i,N}\}$ tenderà ad avere valori spettrali bassi alle frequenze per le quali la sequenza $\{x_i\}$ ha spettro di potenza inferiore al livello del disturbo additivo $\{z_i\}$. Per capire ciò, pensiamo al caso in cui *il brusio dei vicini diventi assordante*: la sequenza $\{y_i\}$ è praticamente coincidente con la sequenza $\{z_i\}$; quindi è imprevedibile e viene a coincidere con la sequenza $\{\varepsilon_{i,N}\}$. Invece la sequenza $\{\widehat{w}_{i,N}\}$ tende a zero; infatti, *il coro è impercettibile ed è opportuno tapparsi le orecchie*.

Per comprendere meglio questo effetto, è opportuno ripetere l'intera analisi nel dominio delle frequenze, come faremo nel prossimo paragrafo, o nel dominio degli autovettori, come verrà fatto nel paragrafo ancora successivo.

14.8.3 Predizione, stima e sbiancamento: un riassunto

Appare opportuno riassumere i vari casi di predizione o stima di una serie temporale di autocorrelazione assegnata sia essa discendente da un processo AR, che da un processo AR + rumore, che qualsiasi. Le equazioni che seguono corrispondono alla predizione del processo $\{y_i\}$, consistente in un processo AR $\{x_i\}$, cui è sovrapposto del rumore $\{z_i\}$, a varianza σ_z^2 . Il valore q.m. dell'errore di predizione è indicato con $\sigma_{\varepsilon,N}^2$; le incognite del sistema sono appunto $\sigma_{\varepsilon,N}^2$ e gli N coefficienti $a_{n,N}$. Facendo la predizione del campione attuale (i -esimo) usando gli N precedenti si

CAPITOLO 14 LA PREDIZIONE LINEARE

ha un errore $\varepsilon_{i,N}$:

$$\begin{aligned} y_i &= x_i + z_i \\ r_n &= E [x_i x_{i-n}^*] \\ \hat{y}_i &= - \sum_{n=1}^N a_{n,N} y_{i-n} \\ \varepsilon_{i,N} &= y_i - \hat{y}_i = y_i + \sum_{n=1}^N a_{n,N} y_{i-n} \end{aligned}$$

Le N equazioni per il calcolo degli $a_{n,N}$ sono derivanti dall'ortogonalità:

$$E [(y_i - \hat{y}_{i,N}) y_{i-m}^*] = E \left[\left(y_i + \sum_{n=1}^N a_{n,N} y_{i-n} \right) y_{i-m}^* \right] = 0 \quad m = 1, \dots, N$$

Un'altra equazione (qui messa come la prima, come viene fatto spesso in letteratura, ma con effetti a volte di confusione) invece dà il valore q.m. dell'errore di stima $\sigma_{\varepsilon,N}^2$:

$$\begin{aligned} E [(y_i - \hat{y}_{i,N}) y_i^*] &= E \left[\left(y_i + \sum_{n=1}^N a_{n,N} y_{i-n} \right) y_i^* \right] = \sigma_{\varepsilon,N}^2 \\ \begin{bmatrix} r_0 + \sigma_z^2 & r_{-1} & r_{-2} & \dots & r_{-N} \\ r_1 & r_0 + \sigma_z^2 & r_{-1} & \dots & r_{-N+1} \\ r_2 & r_1 & r_0 + \sigma_z^2 & \dots & r_{-N+2} \\ \dots & \dots & \dots & \dots & \dots \\ r_N & r_{N-1} & r_{N-2} & \dots & r_0 + \sigma_z^2 \end{bmatrix} \begin{bmatrix} 1 \\ a_{1,N} \\ a_{2,N} \\ \dots \\ a_{N,N} \end{bmatrix} &= \begin{bmatrix} \sigma_{\varepsilon,N}^2 \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix} \end{aligned} \quad (14.46)$$

Al crescere di σ_z^2 , gli $a_{n,N}$ tendono a 0 e $\sigma_{\varepsilon,N}^2$ diventa sempre più prossimo a σ_z^2 ; la predizione è inutile, se il dato è coperto da rumore bianco. Facciamo ora la stima della sequenza $\{w_i\}$ di varianza supposta nota σ_w^2 , modello di un possibile ingresso di un sistema autoregressivo di ordine N , alla cui uscita $\{x_i\}$ sarebbe aggiunto rumore bianco di varianza supposta nota σ_z^2 . Per questa stima usiamo anche il valore attuale della sequenza osservata y_i ; si noti il cambio di segno nell'espressione dello stimatore:

$$\hat{w}_{i,N} = \sum_{n=0}^N c_{n,N} y_{i-n}$$

14.8 STIMA CON RUMORE

Questa volta le incognite sono i $c_{n,N}$, in numero di $N + 1$. L'ortogonalità ci dà:

$$E[(w_i - \widehat{w}_{i,N}) y_{i-m}^*] = E\left[\left(w_i - \sum_{n=0}^N c_{n,N} y_{i-n}\right) y_{i-m}^*\right] = 0 \quad m = 0, \dots, N$$

$$\begin{bmatrix} r_0 + \sigma_z^2 & r_{-1} & r_{-2} & \dots & r_{-N+1} \\ r_1 & r_0 + \sigma_z^2 & r_{-1} & \dots & r_{-N+2} \\ r_2 & r_1 & r_0 + \sigma_z^2 & \dots & r_{-N+3} \\ \dots & \dots & \dots & \dots & \dots \\ r_{N-1} & r_{N-2} & r_{N-3} & \dots & r_0 + \sigma_z^2 \end{bmatrix} \begin{bmatrix} c_{0,N} \\ c_{1,N} \\ c_{2,N} \\ \dots \\ c_{N,N} \end{bmatrix} = \begin{bmatrix} \sigma_w^2 \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

Il valore q.m. dell'errore di stima $\sigma_{\varepsilon,N}^2$ è:

$$E[(w_i - \widehat{w}_{i,N}) w_i^*] = E\left[\left(w_i + \sum_{n=0}^N c_{n,N} y_{i-n}\right) w_i^*\right] = \sigma_{\varepsilon,N}^2 = \sigma_w^2 (1 - c_{0,N})$$

Al crescere di σ_z , tutti i $c_{n,N}$ tendono a 0; la stima diventa impossibile, il processo stimato ha ampiezza tendente a 0 e la varianza dell'errore tende a σ_w^2 . Invece se $\sigma_z = 0$ i due sistemi di equazioni visti sono identici.

Facciamo ora la stima del campione attuale (*i-esimo*) della sequenza $\{x_i\}$ uscita del sistema autoregressivo usando gli N campioni precedenti della sequenza $\{y_i\}$. Attenzione al segno dello stimatore:

$$\widehat{x}_{i,N} = - \sum_{n=1}^N b_{n,N} y_{i-n}$$

Le N equazioni per il calcolo degli $b_{n,N}$ sono derivanti dall'ortogonalità:

$$E[(x_i - \widehat{x}_{i,N}) y_{i-m}^*] = E\left[\left(x_i + \sum_{n=1}^N b_{n,N} y_{i-n}\right) y_{i-m}^*\right] = 0 \quad m = 1, \dots, N$$

CAPITOLO 14 LA PREDIZIONE LINEARE

$$\begin{bmatrix} r_0 + \sigma_z^2 & r_{-1} & r_{-2} & \dots & r_{-N} \\ r_1 & r_0 + \sigma_z^2 & r_{-1} & \dots & r_{-N+1} \\ r_2 & r_1 & r_0 + \sigma_z^2 & \dots & r_{-N+2} \\ \dots & \dots & \dots & \dots & \dots \\ r_N & r_{N-1} & r_{N-2} & \dots & r_0 + \sigma_z^2 \end{bmatrix} \begin{bmatrix} b_{1,N} \\ b_{2,N} \\ b_{2,N} \\ \dots \\ b_{N,N} \end{bmatrix} = - \begin{bmatrix} r_1 \\ r_2 \\ \dots \\ r_N \end{bmatrix}$$

Un'altra equazione ci dà il valore q.m. dell'errore di stima $\sigma_{\varepsilon,N}^2$:

$$E [(x_i - \hat{x}_{i,N}) x_i^*] = E \left[\left(x_i + \sum_{n=1}^N b_{n,N} y_{i-n} \right) x_i^* \right] = \sigma_{\varepsilon,N}^2$$

$$\sigma_{\varepsilon,N}^2 = r_0 + \sum_{n=1}^N b_{n,N} r_{-n}$$

Facciamo infine la stima della sequenza $\{x_i\}$ uscita del sistema autoregressivo usando sia gli N campioni precedenti che il campione attuale della sequenza $\{y_i\}$. Al solito, si noti la convenzione di segno diversa se si passa da predizione a stima,

$$\hat{x}_i = \sum_{n=0}^N d_{n,N} y_{i-n}$$

Le $N + 1$ equazioni per il calcolo dei $d_{n,N}$ sono sempre derivanti dall'ortogonalità:

$$E [(x_i - \hat{x}_i) y_{i-m}^*] = E \left[\left(x_i - \sum_{n=0}^N d_{n,N} y_{i-n} \right) y_{i-m}^* \right] = 0 \quad m = 0, \dots, N$$

$$\begin{bmatrix} r_0 + \sigma_z^2 & r_{-1} & r_{-2} & \dots & r_{-N} \\ r_1 & r_0 + \sigma_z^2 & r_{-1} & \dots & r_{-N+1} \\ r_2 & r_1 & r_0 + \sigma_z^2 & \dots & r_{-N+2} \\ \dots & \dots & \dots & \dots & \dots \\ r_N & r_{N-1} & r_{N-2} & \dots & r_0 + \sigma_z^2 \end{bmatrix} \begin{bmatrix} d_{0,N} \\ d_{1,N} \\ d_{2,N} \\ \dots \\ d_{N,N} \end{bmatrix} = \begin{bmatrix} r_0 \\ r_1 \\ \dots \\ r_N \end{bmatrix}$$

14.9 LA DECONVOLUZIONE NEL DOMINIO DELLE FREQUENZE

e di nuovo, per σ_z^2 crescente. i $d_{n,N}$ tendono a 0.

In generale, è bene sottolineare che i problemi di stima (o predizione) sono sempre caratterizzati dalla determinazione di un vettore di coefficienti di predizione $\underline{\alpha}_N$ mediante il prodotto dell'inverso (se esiste) della matrice di covarianza dei dati \mathbf{R} (**sempre la stessa!**) per un opportuno vettore di correlazioni $\underline{\rho}_N$ tra il segnale desiderato e i dati. Si ha quindi sempre una relazione del tipo:

$$\mathbf{R} \underline{\alpha}_N = \underline{\rho}_N \rightarrow \underline{\alpha}_N = \mathbf{R}^{-1} \underline{\rho}_N$$

ed il valore q.m. dell'errore di predizione o di stima è:

$$E[(x_i - \hat{x}_i) x_i^*] = r_0 - \underline{\rho}_N^* \mathbf{R}^{-1} \underline{\rho}_N$$

14.9 La deconvoluzione nel dominio delle frequenze

Il problema della stima di un processo può essere risolto sia nel dominio dei tempi che in quello delle frequenze. Il vantaggio di operare nel dominio delle frequenze, veramente grande, sta nel fatto che la **matrice di covarianza dei dati è diagonale** qualora si tratta di dati risultanti dalla trasformazione di Fourier discreta di sequenze stazionarie colorate o bianche. Certamente, come visto in precedenza, il filtraggio deve essere fatto in modo periodicizzato nei tempi. Poi, operando nel dominio delle frequenze, l'inversione della matrice di covarianza dei dati \mathbf{R} di ordine $[N \times N]$ necessaria per ottenere il vettore dei coefficienti $\underline{\alpha}_N$, a partire dal vettore delle correlazioni $\underline{\rho}_N$

$$\mathbf{R} \underline{\alpha}_N = \underline{\rho}_N; \rightarrow \underline{\alpha}_N = \mathbf{R}^{-1} \underline{\rho}_N$$

diventa semplicissima, perchè corrispondente a N equazioni ad una sola incognita. In questo caso, tuttavia, non sarà in generale possibile limitare il dominio su cui è definito il filtro, come è possibile fare nel dominio dei tempi, in cui si conoscono quali campioni vengono utilizzati per fare la stima (ad esempio, gli N campioni che precedono il campione da stimare). Ora l'asse dei tempi si estende da $-\infty$ ad ∞ . Passando all'analisi del problema, sia data la sequenza:

$$y_i = x_i + z_i = \sum_{n=-N/2}^{N/2} w_{i-n} h_n + z_i \quad (14.47)$$

CAPITOLO 14 LA PREDIZIONE LINEARE

dove $\{x_i\}$ e $\{z_i\}$ sono due sequenze, l'una, $\{x_i\}$ corrispondente al segnale (*le voci bianche* w_i riverberate da una *cattedrale* con risposta all'impulso h_n che ora può essere qualsiasi e non necessariamente corrispondente ad un sistema AR) e l'altra $\{z_i\}$ corrispondente a rumore additivo che per semplicità considereremo a spettro bianco (*il brusio dei vicini*). In quanto segue, la frequenza può essere rappresentata sia da un indice continuo (la solita pulsazione ω) sia, pensando alle trasformazioni di Fourier discrete, da un indice discreto k . Per semplicità, useremo l'indice continuo. Supponiamo noti sia lo spettro di potenza della sequenza $\{w_i\}$, e cioè $E[|W(\omega)|^2] = \sigma_w^2 T$, sia **la funzione di trasferimento del filtro** $H(\omega)$ che ha operato sulla sequenza (*la riverberazione della cattedrale*); nei casi precedentemente discussi, bastava invece conoscere l'autocorrelazione del filtro e quindi $|H(\omega)|^2$. La sequenza $\{z_i\}$ (*il brusio dei vicini*) è anche ignota, e se ne conoscono solo le caratteristiche spettrali. Si vuole stimare

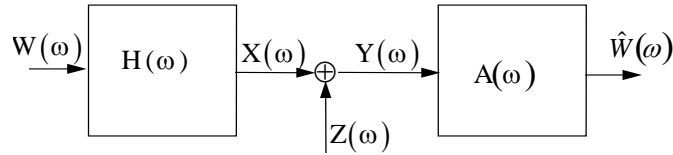


Figura 14.7: Deconvoluzione nel dominio delle frequenze

$\{w_i\}$ da $\{y_i\}$, nel dominio della frequenza. Pertanto, indicando con $A(\omega)$ la funzione di trasferimento del filtro che utilizzeremo per fare la stima della sequenza $\{w_i\}$ e indicando con lettere maiuscole le trasformate delle sequenze (fig. 14.7):

$$Y(\omega) = W(\omega)H(\omega) + Z(\omega); \quad E[|W(\omega)|^2] = \sigma_w^2(\omega)T \quad (14.48)$$

$$E[|Z(\omega)|^2] = \sigma_z^2(\omega)T \quad (14.49)$$

$$\widehat{W}(\omega) = A(\omega)Y(\omega) \quad (14.50)$$

Il principio di ortogonalità (derivante dall'ottimalità del risultato desiderato) comporta che sia nulla la correlazione tra l'errore di stima ed il dato (e quindi nelle frequenze sia nullo il valore atteso del prodotto tra la

14.9 LA DECONVOLUZIONE NEL DOMINIO DELLE FREQUENZE

trasformata dell'errore di predizione e la coniugata della trasformata del dato); la pulsazione ω è un parametro che non trascriveremo.

$$E[(W - AY)Y^*] = 0; \quad (14.51)$$

$$A = \frac{E[WY^*]}{E[YY^*]} = \frac{\sigma_w^2 H^*}{\sigma_w^2 |H|^2 + \sigma_z^2} = \frac{H^*}{HH^* + Q}; \quad (14.52)$$

$$Q(\omega) = \frac{\sigma_z^2(\omega)}{\sigma_w^2(\omega)} = \frac{\sigma_z^2}{\sigma_w^2}; \quad (14.53)$$

Si vede quindi che il risultato della stima nel dominio delle frequenze è un filtro $A(\omega) = 1/H(\omega)$, se $Q \ll |H(\omega)|^2$ e cioè se il disturbo è trascurabile. Quando invece il disturbo è più ampio del segnale ($Q \gg |H(\omega)|^2$), il filtro tende a dare uscita nulla (*alle frequenze in cui il segnale è coperto dal brusio, non ne tentiamo la ricostruzione*).

Quindi, possiamo ritornare sulla questione discussa nel precedente paragrafo e cioè chiarire bene la differenza tra $\{\varepsilon_{i,N}\}$ errore di predizione della sequenza $\{y_i\}$ e $\{\widehat{w}_{i,N}\}$ stima della sequenza bianca $\{w_i\}$ che alimenta il sistema riverberante che crea $\{x_i\}$. In assenza di rumore additivo $\{z_i\}$ non vi sono differenze: invece, se $\sigma_z^2 \neq 0$ la stima $\{\widehat{w}_{i,N}\}$ del segnale originario $\{w_i\}$ sarà diversa dall'errore di predizione $\{\varepsilon_{i,N}\}$ ed in particolare di ampiezza ridotta alle frequenze dove è più sensibile il rumore.

L'operazione così realizzata, e cioè la stima della sequenza $\{w_i\}$, si indica con il nome di **deconvoluzione**; infatti la sequenza $\{w_i\}$ viene osservata dopo la convoluzione con il filtro $\{h_n\}$ e l'aggiunta di rumore bianco; scopo della stima è quello di recuperare la sequenza $\{w_i\}$ rimuovendo l'effetto del filtro $H(\omega)$ con l'operazione appunto di deconvoluzione. Come si è visto, la cosa è fattibile applicando un filtro inverso, purchè il livello del segnale filtrato non scenda sotto il livello del rumore: a quelle frequenze, la stima va a zero.

Infine, calcoliamo il valore qm dell'errore di predizione, usando il fatto che l'errore di predizione è incorrelato con i dati o con ogni loro combinazione lineare come \widehat{W} :

$$\begin{aligned} E[(W - \widehat{W})(W - \widehat{W})^*] &= E[(W - \widehat{W})W^*] = \\ &= \frac{E[|W|^2]E[|Z|^2]}{E[|W|^2]|H|^2 + E[|Z|^2]} < E[|W|^2] \end{aligned}$$

CAPITOLO 14 LA PREDIZIONE LINEARE

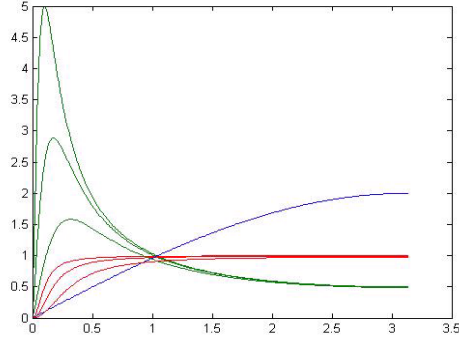


Figura 14.8: Funzione di trasferimento del filtro inverso di cui alla formula 14.54, per $Q=0,01; 0,03; 0,1$. La figura riporta anche la funzione di trasferimento complessiva, sempre più prossima ad 1.

Come esempio, si supponga che $H(z)$ sia un semplicissimo filtro passa basso, con un solo zero a frequenza zero ($z = 1$):

$$H(z) = 1 - z^{-1}$$

$$|H|_{z=e^{j\omega T}} = |1 - e^{-j\omega T}| = 2 \sin |\theta|; \theta = \frac{\omega T}{2}$$

Il filtro inverso ottimo $A(z)$ è:

$$A = \frac{H^*}{|H|^2 + Q} = \frac{1 - z}{(1 - z^{-1})(1 - z) + Q} = \frac{1 - z}{2 + Q + (z + z^{-1})} \quad (14.54)$$

$$|A|_{z=e^{j\omega T}} = \frac{\sin |\theta|}{2 \sin^2 \theta + Q/2}; \angle A = -\angle H$$

Si osservi che la sequenza $\{a_n\}$ antitrasformata z di $A(z)$ e corrispondente al filtro inverso, deve essere stabile e quindi non è nè causale, nè simmetrica nei tempi, in quanto i poli della trasformata nella formula 14.54 sono l'uno l'inverso dell'altro e quindi ve ne è uno fuori del cerchio unitario.

14.9.1 Deconvoluzione predittiva nelle frequenze

Operando nel dominio delle frequenze, non appare immediatamente possibile stabilire il supporto del filtro, cioè definire i campioni per i quali

14.9 LA DECONVOLUZIONE NEL DOMINIO DELLE FREQUENZE

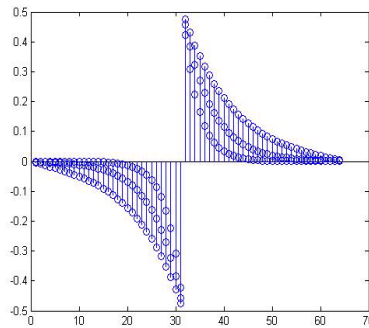


Figura 14.9: Risposta all'impulso del filtro inverso di cui alla formula 14.54, per $Q=0,01$; $0,03$; $0,1$. La soluzione, non causale, è trovata o nel dominio delle frequenze o in quello dei tempi, ma utilizzando dati sia prima che dopo l'istante a cui si vuole stimare il segnale sorgente.

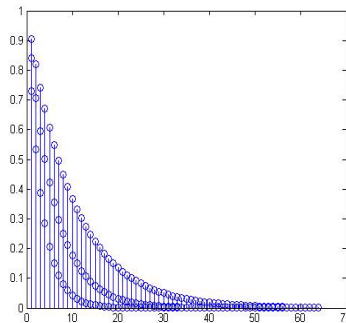


Figura 14.10: Risposta all'impulso del filtro inverso di cui alla formula 14.54, per $Q=0,01$; $0,03$; $0,1$. La soluzione è trovata nel dominio dei tempi utilizzando solo dati che precedono l'istante a cui si vuole stimare il segnale sorgente. Quindi il filtro è causale.

CAPITOLO 14 LA PREDIZIONE LINEARE

i coefficienti del filtro sono diversi da zero: quindi non sarebbe possibile trovare il filtro ottimo che determina l'errore di predizione, perchè questo implica che il filtro sia causale, cosa che appare difficile imporre. Esiste peraltro la tecnica di fattorizzazione spettrale (Wiener, Kolmogorov) che permette di risolvere il problema dell'errore di predizione anche nel dominio delle frequenze. Il principio è semplice. Con un filtro di lunghezza infinita, la sequenza degli errori di predizione deve:

- (a) essere bianca;
- (b) essere ottenuta dalla sequenza $\{y_i\}$ attraverso un filtro causale.

Ora, se lo spettro della sequenza $\{y_i\}$ è $E[|Y(\omega)|^2]$, il modulo della funzione di trasferimento del filtro sbiancatore sarà:

$$|A(\omega)|^2 = \frac{1}{E[|Y(\omega)|^2]} \quad (14.55)$$

Ne conosciamo quindi la caratteristica di ampiezza e sappiamo anche, dalle considerazioni fatte nel dominio dei tempi, che la sequenza di campioni che corrisponde al filtro predittore deve essere non solo causale (un predittore utilizza solo i dati "passati") ma anche a minima fase; in realtà avevamo dedotto la proprietà di minima fase solo per N finito qualsiasi; supponiamo ora che questa proprietà di minima fase si mantenga anche per N infinito. Pertanto, come visto nel capitolo 7, la caratteristica di fase della funzione di trasferimento sarà trasformata di Hilbert del logaritmo della caratteristica di ampiezza. In conclusione il predittore causale ottenuto dall'analisi nel dominio delle frequenze e di lunghezza infinita (in teoria) ha funzione di trasferimento, in fase ed ampiezza:

$$|A_\infty(\omega)| = \sqrt{\frac{1}{E[|Y(\omega)|^2]}} \quad (14.56)$$

$$\angle A_\infty(\omega) = -\frac{1}{2} \mathcal{H}\{\log E[|Y(\omega)|^2]\} \quad (14.57)$$

Questo risultato, scoperto negli anni della II guerra mondiale ed applicato alla predizione delle tracce radar per il pilotaggio delle batterie contraeree, era stato tenuto gelosamente segreto.

Sono da evitare zeri dello spettro $E[|Y(\omega)|^2]$, in particolare alla frequenza di Nyquist, per evitare singolarità del logaritmo e quindi problemi con la trasformazione di Hilbert. Se poi gli zeri spettrali sono ineludibili perchè troppo significativi, non è più possibile fare una stima spettrale

14.10 LE EQUAZIONI Y - W NEL DOMINIO DEGLI AUTOVETTORI

AR che diventa troppo lunga ma si deve passare ad una stima spettrale ARMA. In altre parole gli zeri vanno tenuti in conto separatamente, e eventualmente sbiancati in precedenza con i poli opportuni (piuttosto che con infiniti zeri). Per uno studio dell'analisi spettrale ARMA, si rimanda ai testi nella bibliografia proposta nella prefazione.

14.10 Le equazioni Y - W nel dominio degli autovettori

Ritornando ora nel dominio dei tempi, trasformeremo l'analisi fatta nel dominio delle frequenze in un'analisi nel dominio degli autovettori della matrice di covarianza $\underline{\mathbf{R}}$. Abbiamo visto che, se sulla diagonale principale della matrice $\underline{\mathbf{R}}$ è addizionata *luce bianca* (*white light, in inglese*) σ_z^2 , il sistema diventa ben condizionato, in quanto tutti gli autovalori di $\underline{\mathbf{R}}$ diventano positivi; si comprende l'importanza di garantire che la matrice $\underline{\mathbf{R}}$ sia semidefinita positiva; altrimenti, l'aggiunta di *luce bianca*, nel caso di autovalori negativi, poteva portare all'instabilità invece che alla stabilità. Quindi, una tecnica di garantire il condizionamento della matrice quindi la soluzione delle equazioni di Y - W è appunto quella di addizionare *luce bianca* anche se poi non vi è nessun rumore $\{z_i\}$ sovrapposto al segnale $\{x_i\}$. Sviluppando in autovalori ed autovettori di $\underline{\mathbf{R}}$ sia $\underline{\mathbf{R}}$ che $\underline{\boldsymbol{\alpha}}_N$ e $\underline{\boldsymbol{\rho}}_N$, si ha, partendo dall'equazione di stima generica:

$$\begin{aligned} \underline{\mathbf{R}} \underline{\boldsymbol{\alpha}}_N &= \underline{\boldsymbol{\rho}}_N \\ \underline{\mathbf{R}} \underline{\boldsymbol{\alpha}}_N &= \sum_{n=1}^N \lambda_n \underline{\mathbf{e}}_n \underline{\mathbf{e}}_n^* \underline{\boldsymbol{\alpha}}_N = \sum_{n=1}^N \underline{\mathbf{e}}_n \underline{\mathbf{e}}_n^* \underline{\boldsymbol{\rho}}_N = \underline{\boldsymbol{\rho}}_N \end{aligned} \quad (14.58)$$

Indicando con

$$\underline{\mathbf{e}}_n^* \underline{\boldsymbol{\alpha}}_N = a_{n,N} \quad \underline{\mathbf{e}}_n^* \underline{\boldsymbol{\rho}}_N = \rho_{n,N} \quad (14.59)$$

abbiamo che:

$$\begin{aligned} \lambda_n a_{n,N} &= \rho_{n,N} \\ a_{n,N} &= \frac{\rho_{n,N}}{\lambda_n} \end{aligned} \quad (14.60)$$

Addizionando $\sigma_z^2 \mathbf{I}$ ad $\underline{\mathbf{R}}$ abbiamo

$$a_{n,N} = \frac{\rho_{n,N}}{\lambda_n + \sigma_z^2} \quad (14.61)$$

CAPITOLO 14 LA PREDIZIONE LINEARE

Questa equazione corrisponde all'equazione (14.52) tranne che l'una è stata ottenuta nel dominio delle frequenze e l'altra nel dominio dei tempi (autovettori di $\underline{\mathbf{R}}$ e correlazione tra segnale desiderato e dati). In altre parole, di nuovo si trova che sono attenuate le componenti della soluzione $q_{n,N}$ lungo gli autovettori $\underline{\mathbf{e}}_n$ con autovalori λ_n piccoli rispetto al disturbo σ_z^2 . Se ora σ_z^2 è abbastanza piccola, ad esempio $r_0/100$ alteriamo, anche se percentualmente di poco, tutte le componenti $a_{n,N}$ comprese quelle corrispondenti agli autovalori più grandi di $\underline{\mathbf{R}}$, che non avevano motivo di essere alterate. In compenso, l'operazione è computazionalmente facile, perchè basta sommare una matrice identica alla matrice $\underline{\mathbf{R}}$. Una soluzione diversa, che prevede però di conoscere tutti gli autovalori di $\underline{\mathbf{R}}$, è quella di sostituire a $\lambda_i \rightarrow \max[\lambda_i, \sigma_z^2]$ (la tecnica è denominata *water filling* dall'effetto di riempire i buchi spettrali sostituendoli con una costante). In questo modo si alterano solo le componenti della soluzione relative agli autovalori minori di σ_z^2 , lasciando immutate le altre.

Se ora supponiamo di non conoscere perfettamente la matrice di covarianza e quindi nemmeno i λ_n , abbiamo una soluzione ancora diversa. Infatti, nell'equazione 14.60 deve comparire un termine di rumore ν_n per modellare questa incertezza. Omettendo i pedici in n , N per semplicità:

$$\lambda a + \nu = \rho \tag{14.62}$$

e la stima ai minimi quadrati della variabile a è:

$$\hat{a} = \mu \rho$$

Minimizzando l'errore q.m. di stima si ha:

$$\begin{aligned} \min_{\mu} E [(\mu \rho - a)^2] &\rightarrow E [(\mu \rho - a) \rho] = 0 \\ \mu &= \frac{E[\rho a]}{E[\rho^2]} = \frac{\lambda \sigma_q^2}{\lambda^2 \sigma_q^2 + \sigma_{\nu}^2} \\ \hat{a}_{n,N} &= \mu \rho_{n,N} = \frac{\lambda_n}{\lambda_n^2 + \frac{\sigma_{\nu}^2}{\sigma_q^2}} \rho_{n,N} \end{aligned} \tag{14.63}$$

dove con σ_{ν}^2 abbiamo indicato la varianza del disturbo ν_n . In termini matriciali, abbiamo:

$$\underline{\boldsymbol{\alpha}}_N = \{ \underline{\mathbf{R}}^2 + \frac{\sigma_{\nu}^2}{\sigma_q^2} \underline{\mathbf{I}} \}^{-1} \underline{\mathbf{R}} \underline{\boldsymbol{\rho}}_N \tag{14.64}$$

14.10 LE EQUAZIONI Y - W NEL DOMINIO DEGLI AUTOVETTORI

I vantaggi di questa ultima soluzione per le equazioni di Yule Walker sono che:

- il sistema non è mai singolare;
- non abbiamo bisogno di determinare autovalori ed autovettori di \mathbf{R} ;
- le componenti della soluzione peggio condizionate e cioè quelle corrispondenti agli autovalori più piccoli tendono ad essere poco amplificate invece di essere amplificate al massimo come nel caso dell'equazione (14.52).

Capitolo 15

Stima lineare ottima

15.1 Fattore di scala di un vettore noto

Consideriamo il seguente processo:

$$\underline{\mathbf{z}} = s\underline{\mathbf{c}} + \underline{\mathbf{m}} \quad (15.1)$$

dove le componenti del vettore $\underline{\mathbf{c}}$ formano una sequenza nota (il segnale), ed s , la sua ampiezza, è la grandezza da stimare. Questa corrisponde ad uno scalare (magari complesso) casuale a valor medio nullo e varianza σ_s^2 . Infine, le componenti di $\underline{\mathbf{m}}$ (il rumore) sono variabili casuali a valor medio nullo e correlate tra loro mediante la matrice di covarianza $\underline{\mathbf{M}}$:

$$E[\underline{\mathbf{m}}] = \underline{\mathbf{0}}; \quad E[\underline{\mathbf{m}}\underline{\mathbf{m}}^*] = \underline{\mathbf{M}}$$

Risulta:

$$E[\underline{\mathbf{z}}\underline{\mathbf{z}}^*] = \sigma_s^2 \underline{\mathbf{c}}\underline{\mathbf{c}}^* + \underline{\mathbf{M}}$$

Se si desidera stimare s , lo stimatore ottimo (ai minimi quadrati) è:

$$\hat{s} = \underline{\mathbf{b}}^* \underline{\mathbf{z}}$$

dove $\underline{\mathbf{b}}$ è un vettore da calcolare imponendo la minimizzazione dell'errore q. m. di stima e quindi l'ortogonalità tra il vettore dei dati $\underline{\mathbf{z}}$ e l'errore di stima $\varepsilon = s - \hat{s}$:

$$E[\varepsilon \underline{\mathbf{z}}^*] = E[(s - \hat{s}) \underline{\mathbf{z}}^*] = E[(s - \underline{\mathbf{b}}^* \underline{\mathbf{z}}) \underline{\mathbf{z}}^*] = \underline{\mathbf{0}} \rightarrow$$

CAPITOLO 15 STIMA LINEARE OTTIMA

$$\begin{aligned}\underline{\mathbf{b}}^* &= E[s\underline{\mathbf{z}}^*] E[\underline{\mathbf{z}}\underline{\mathbf{z}}^*]^{-1} \\ \underline{\mathbf{b}} &= \sigma_s^2 [\sigma_n^2 \underline{\mathbf{M}} + \sigma_s^2 \underline{\mathbf{c}}\underline{\mathbf{c}}^*]^{-1} \underline{\mathbf{c}}\end{aligned}$$

Volendo utilizzare la formula di inversione di una matrice costituita dalla combinazione lineare di una matrice identica ed una diade $\underline{\mathbf{p}}\underline{\mathbf{p}}^*$,

$$(\underline{\mathbf{I}} + \underline{\mathbf{p}}\underline{\mathbf{p}}^*)^{-1} = \underline{\mathbf{I}} - \frac{\underline{\mathbf{p}}\underline{\mathbf{p}}^*}{1 + \underline{\mathbf{p}}^* \underline{\mathbf{p}}} \quad (15.2)$$

si deve innanzitutto riportarsi a tale caso sbiancando il rumore e cioè premoltiplicando la 15.1 per la matrice

$$\underline{\mathbf{R}} = \underline{\mathbf{M}}^{-1/2}$$

che non è altro che la matrice $\underline{\mathbf{M}}$, con gli stessi autovettori, dove però gli autovalori originali $\lambda_{M,i}$ sono sostituiti con l'inverso della loro radice:

$$\lambda_{R,i} = \frac{1}{\sqrt{\lambda_{M,i}}}$$

Se la matrice $\underline{\mathbf{M}}$ è definita positiva, i $\lambda_{R,i}$ sono reali, positivi e finiti. L'equazione 15.1 diventa:

$$\underline{\mathbf{y}} = \underline{\mathbf{R}}\underline{\mathbf{z}} = \underline{\mathbf{M}}^{-1/2} (s\underline{\mathbf{c}} + \underline{\mathbf{m}}) = s\underline{\mathbf{p}} + \underline{\mathbf{n}}$$

dove ora però:

$$\begin{aligned}\underline{\mathbf{p}} &= \underline{\mathbf{M}}^{-1/2} \underline{\mathbf{c}} \\ E[\underline{\mathbf{m}}\underline{\mathbf{m}}^*] &= \underline{\mathbf{M}}^{-1/2} E[\underline{\mathbf{m}}\underline{\mathbf{m}}^*] \underline{\mathbf{M}}^{-1/2} = \underline{\mathbf{M}}^{-1/2} \underline{\mathbf{M}} \underline{\mathbf{M}}^{-1/2} = \underline{\mathbf{I}}\end{aligned}$$

Risulta allora:

$$\begin{aligned}\hat{s} &= \underline{\mathbf{a}}^* \underline{\mathbf{y}} \\ \underline{\mathbf{a}} &= \sigma_s^2 [\underline{\mathbf{I}} + \sigma_s^2 \underline{\mathbf{p}}\underline{\mathbf{p}}^*]^{-1} \underline{\mathbf{p}}\end{aligned}$$

Usando la 15.2:

$$\underline{\mathbf{a}} = \left[\sigma_s^2 - \frac{\sigma_s^4 \underline{\mathbf{p}}^* \underline{\mathbf{p}}}{(1 + \sigma_s^2 \underline{\mathbf{p}}^* \underline{\mathbf{p}})} \right] \underline{\mathbf{p}} = \left[\frac{\sigma_s^2}{1 + \sigma_s^2 \underline{\mathbf{c}}^* \underline{\mathbf{M}}^{-1} \underline{\mathbf{c}}} \right] \underline{\mathbf{M}}^{-1/2} \underline{\mathbf{c}}$$

$$\underline{\mathbf{b}} = \frac{\sigma_s^2 \underline{\mathbf{M}}^{-1} \underline{\mathbf{c}}}{1 + \sigma_s^2 \underline{\mathbf{c}}^* \underline{\mathbf{M}}^{-1} \underline{\mathbf{c}}}$$

15.1 FATTORE DI SCALA DI UN VETTORE NOTO

C'è polarizzazione poichè:

$$E[\hat{s}^2] = \left(\frac{\sigma_s^2 \underline{\mathbf{c}}^* \underline{\mathbf{M}}^{-1} \underline{\mathbf{c}}}{1 + \sigma_s^2 \underline{\mathbf{c}}^* \underline{\mathbf{M}}^{-1} \underline{\mathbf{c}}} \right)^2 \sigma_s^2 < \sigma_s^2$$

$$E[(s - \hat{s})^2] = E[\varepsilon^2] = \frac{\sigma_s^2}{1 + \sigma_s^2 \underline{\mathbf{c}}^* \underline{\mathbf{M}}^{-1} \underline{\mathbf{c}}}$$

Si osservi infine che per $\sigma_s^2 \rightarrow \infty$, la polarizzazione va a 0 e:

$$\lim_{\sigma_s^2 \rightarrow \infty} E[\varepsilon^2] = \frac{1}{\underline{\mathbf{c}}^* \underline{\mathbf{M}}^{-1} \underline{\mathbf{c}}}$$

Nel caso semplice che il rumore sia bianco

$$\underline{\mathbf{M}} = \sigma_n^2 \underline{\mathbf{I}} \quad (15.3)$$

si ha:

$$\underline{\mathbf{c}}^* \underline{\mathbf{M}}^{-1} \underline{\mathbf{c}} = \frac{\underline{\mathbf{c}}^* \underline{\mathbf{c}}}{\sigma_n^2}; \quad \underline{\mathbf{b}} = \frac{\sigma_s^2 \sigma_n^2 \underline{\mathbf{c}}}{\sigma_n^2 + \sigma_s^2 \underline{\mathbf{c}}^* \underline{\mathbf{c}}}; \quad E[\varepsilon^2] = \frac{\sigma_s^2 \sigma_n^2}{\sigma_n^2 + \sigma_s^2 \underline{\mathbf{c}}^* \underline{\mathbf{c}}}$$

Allo stesso risultato si poteva arrivare direttamente facendo uso del cosiddetto inversion lemma, che dà la matrice inversa della matrice somma di una matrice qualsiasi e di una diade:

$$(\underline{\mathbf{A}} + \underline{\mathbf{s}} \underline{\mathbf{s}}^*)^{-1} = \underline{\mathbf{A}}^{-1} - \frac{\underline{\mathbf{A}}^{-1} \underline{\mathbf{s}} \underline{\mathbf{s}}^* \underline{\mathbf{A}}^{-1}}{1 + \underline{\mathbf{s}} \underline{\mathbf{A}}^{-1} \underline{\mathbf{s}}^*}$$

15.1.1 Effetti del rumore colorato

Nel caso di rumore colorato, la stima deve avvalersi dei diversi valori della densità spettrale della sequenza di rumore m_n ; $n = 1, \dots, N$ in corrispondenza delle varie frequenze utilizzate dalla sequenza di segnale c_n , che non deve essere monocromatica. In sostanza, con la solita approssimazione di supporre le matrici circolanti, la matrice di covarianza del rumore ha come autovettori delle sinusoidi normalizzate $\underline{\mathbf{e}}_k$ alle pulsazioni $\omega_k = 2\pi k/NT$ e gli autovalori corrispondenti λ_k sono proporzionali alla densità spettrale del rumore a quelle pulsazioni, $S_m(\omega_k)$. Trasformando

CAPITOLO 15 STIMA LINEARE OTTIMA

in Fourier i vettori $\underline{\mathbf{b}}, \underline{\mathbf{c}}$, e cioè usando come base di decomposizione di questi vettori gli autovettori $\underline{\mathbf{e}}_k$ della matrice di covarianza $\underline{\mathbf{M}}$ risulta:

$$\underline{\mathbf{b}} = \sum_{k=1}^N B_k \underline{\mathbf{e}}_k; \underline{\mathbf{c}} = \sum_k C_k \underline{\mathbf{e}}_k$$

e utilizzando la ortonormalità degli autovettori:

$$\underline{\mathbf{c}}^* \underline{\mathbf{M}}^{-1} \underline{\mathbf{c}} = \left(\sum C_k \underline{\mathbf{e}}_k \right)^* \underline{\mathbf{M}}^{-1} \sum C_h \underline{\mathbf{e}}_h = \sum_k \frac{|C_k|^2}{\lambda_k} \quad (15.4)$$

$$\lim_{\sigma_s^2 \rightarrow \infty} \underline{\mathbf{b}} = \lim_{\sigma_s^2 \rightarrow \infty} \frac{\sigma_s^2 \underline{\mathbf{M}}^{-1} \underline{\mathbf{c}}}{1 + \sigma_s^2 \sum \frac{|C_k|^2}{\lambda_k}} = \frac{\underline{\mathbf{M}}^{-1} \underline{\mathbf{c}}}{\sum \frac{|C_k|^2}{\lambda_k}} = \frac{\sum \frac{C_k}{\lambda_k} \underline{\mathbf{e}}_k}{\sum \frac{|C_k|^2}{\lambda_k}}$$

$$B_k = \frac{1}{\sum \frac{|C_k|^2}{\lambda_k}} \frac{C_k}{\lambda_k} \quad (15.5)$$

$$\lim_{\sigma_s^2 \rightarrow \infty} E[\varepsilon^2] = \frac{1}{\underline{\mathbf{c}}^* \underline{\mathbf{M}}^{-1} \underline{\mathbf{c}}} = \frac{1}{\sum \frac{|C_k|^2}{S_m(\omega_k)}} \quad (15.6)$$

Se il segnale è monocromatico, se cioè $C_k \neq 0$, per un solo valore di $k = k_0$, allora

$$\lim_{\sigma_s^2 \rightarrow \infty} E[\varepsilon^2] = \sigma_{\varepsilon_{k_0}}^2 = \frac{S_m(\omega_{k_0})}{|C_{k_0}|^2} \rightarrow \frac{1}{\sigma_{\varepsilon_{k_0}}^2} = \frac{|C_{k_0}|^2}{S_m(\omega_{k_0})} \quad (15.7)$$

conta solo la densità spettrale del rumore a quella frequenza. Se invece la sequenza di segnale contiene molte componenti, a varie frequenze, e cioè $C_k \neq 0$, per molti valori di k , il sistema di stima ottimo del parametro ovviamente cerca di usare (pesa di più) le componenti alle frequenze con il minimo livello di rumore, pur utilizzando anche le altre componenti. Usando il concetto della media armonica (inverso = media della somma degli inversi), si vede che l'inverso dell'errore di stima ottimo è somma degli inversi degli errori di stima fatti sui vari sottocanali (le varie frequenze):

$$\frac{1}{\lim_{\sigma_s^2 \rightarrow \infty} E[\varepsilon^2]} = \frac{1}{\sigma_\varepsilon^2} = \sum \frac{|C_k|^2}{S_m(\omega_k)} = \sum \frac{1}{\sigma_{\varepsilon_k}^2} \quad (15.8)$$

15.1 FATTORE DI SCALA DI UN VETTORE NOTO

Conta quindi la più piccola delle $\sigma_{\varepsilon_k}^2$ (il cosiddetto *smoking gun*; il poliziotto si basa sull'indizio più significativo, mentre tutto il resto diventa quasi irrilevante; ciò è equivalente al greco $\text{o}\iota\delta\alpha$, ho visto e dunque so, o più semplicemente a sbirciare dal buco della serratura).

15.1.2 Stima di una costante in rumore bianco

Nel caso semplice in cui il rumore è bianco e il vettore $\underline{\mathbf{c}}$ è una costante di ampiezza A :

$$\underline{\mathbf{c}} = A \times \underline{\mathbf{1}} \rightarrow \underline{\mathbf{c}}^* \underline{\mathbf{c}} = N |A|^2; \quad \lim_{\sigma_s^2 \rightarrow \infty} E[\varepsilon^2] = \frac{1}{|A|^2 N}$$

Si osservi di nuovo che **lo stimatore è polarizzato** in quanto tende sistematicamente a sottostimare:

$$E[(\hat{s})^2] < \sigma_s^2$$

Inoltre, per $\sigma_n^2 \rightarrow \infty$, $\hat{s} \rightarrow 0$. A parole, se vi è troppo rumore, la stima tende a zero. Invece se $\sigma_s^2 \rightarrow \infty$, la stima diventa non polarizzata.

15.1.3 Stima di una costante in rumore colorato

Possiamo considerare due casi facili di rumore colorato, profittando di due matrici $[N \times N]$, $\underline{\mathbf{M}}_p$, $\underline{\mathbf{M}}_z$ l'una inversa dell'altra. La prima, $\underline{\mathbf{M}}_p$, corrisponde allo spettro di un polo ed è esattamente di Toeplitz. Una riga è

$$\underline{\mathbf{M}}_p \rightarrow \sigma_m^2 (\dots, \rho^2, \rho, 1, \rho, \rho^2 \dots)$$

$$\underline{\mathbf{M}}_p = \sigma_m^2 \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^N \\ \rho & 1 & \rho & \rho^2 & \dots \\ \rho^2 & \rho & 1 & \rho & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \rho^N & \dots & \rho^2 & \rho & 1 \end{bmatrix} \quad (15.9)$$

La seconda, $\underline{\mathbf{M}}_{p,inv}$, corrisponde allo spettro di uno zero, è con grande approssimazione ancora di Toeplitz, a parte la prima e l'ultima riga: una

CAPITOLO 15 STIMA LINEARE OTTIMA

riga centrale è

$$\underline{\mathbf{M}}_{p,inv} \rightarrow \frac{1}{\sigma_m^2 (1 - \rho^2)} (\dots 0, 0, 0, 0, -\rho, 1 + \rho^2, -\rho, 0, 0, 0..);$$

ma $m_{p,inv,11} = m_{p,inv,NN} = 1$.

$$\underline{\mathbf{M}}_{p,inv} = \frac{1}{\sigma_m^2 (1 - \rho^2)} \begin{bmatrix} 1 & -\rho & 0 & .. & 0 \\ -\rho & 1 + \rho^2 & -\rho & 0 & .. \\ 0 & -\rho & 1 + \rho^2 & -\rho & 0 \\ .. & .. & .. & .. & .. \\ 0 & .. & 0 & -\rho & 1 \end{bmatrix} \quad (15.10)$$

$$\underline{\mathbf{M}}_p \underline{\mathbf{M}}_{p,inv} = \underline{\mathbf{I}} \quad (15.11)$$

Questa accade perchè se:

$$H(z) = 1 - \rho z^{-1}; |H(z)|^2 = 1 + \rho^2 - \rho(z + z^{-1}); \quad (15.12)$$

anche

$$\frac{1}{|H(z)|^2} = \frac{1}{1 + \rho^2 - \rho(z + z^{-1})} = \frac{1}{1 - \rho^2} \left(\frac{1}{1 - z^{-1}\rho} + \frac{1}{1 - z\rho} - 1 \right) \quad (15.13)$$

da cui le righe centrali delle due matrici. Allora, se il rumore è colorato da un polo con la funzione di trasferimento $1/H(z)$, la sua matrice di covarianza è $\underline{\mathbf{M}}_p$ e l'errore di stima della costante nel rumore così colorato è

$$\lim_{\sigma_s^2 \rightarrow \infty} E[\varepsilon^2] = \frac{1}{\underline{\mathbf{c}}^* \underline{\mathbf{M}}_p^{-1} \underline{\mathbf{c}}} \sim \frac{\sigma_m^2 (1 - \rho^2)}{N (1 - \rho)^2} = \frac{\sigma_m^2 (1 + \rho)}{N (1 - \rho)} \quad (15.14)$$

In sostanza, se il rumore è colorato da un polo in ρ , tutto passa come se il numero di campioni utili per la stima fosse:

$$N' \sim N \frac{1 - \rho}{1 + \rho} \quad (15.15)$$

similmente con quanto visto in precedenza nel caso della stima delle potenze di segnali colorati. Solo che in quel caso, trattandosi della stima

15.1 FATTORE DI SCALA DI UN VETTORE NOTO

non lineare di una potenza, il risultato era leggermente diverso perchè compare ρ^2 invece di ρ . Qui il segno di ρ conta, invece, poichè se p.e. ρ è negativo, il rumore è passa alto, i campioni del disturbo sono anticorrelati, e mediandoli per stimare una costante (a frequenza zero) tendono a scomparire. Infatti la densità spettrale $S_{polo}(\omega)$ a frequenza zero (che è quella che conta nel caso della stima di un segnale costante, e cioè a frequenza 0), è pari a:

$$S_{polo}(\omega = 0) = \frac{\sigma_m^2}{(1 - \rho)^2} \quad (15.16)$$

Per fare un altro caso, si consideri la coloratura dovuta ad uno zero e cioè alla funzione di trasferimento $H(z)$. La matrice di covarianza del rumore ora è, per fare in modo che la varianza del disturbo sia sempre σ_m^2 :

$$\underline{\mathbf{M}}_z = \frac{\sigma_m^2}{1 + \rho^2} \begin{bmatrix} 1 & -\rho & 0 & .. & 0 \\ -\rho & 1 + \rho^2 & -\rho & 0 & .. \\ 0 & -\rho & 1 + \rho^2 & -\rho & 0 \\ .. & .. & .. & .. & .. \\ 0 & .. & 0 & -\rho & 1 \end{bmatrix} \quad (15.17)$$

La matrice inversa è:

$$\underline{\mathbf{M}}_{z,inv} = \frac{1 + \rho^2}{\sigma_m^2 (1 - \rho^2)} \begin{bmatrix} 1 & \rho & \rho^2 & .. & \rho^N \\ \rho & 1 & \rho & \rho^2 & .. \\ \rho^2 & \rho & 1 & \rho & .. \\ .. & .. & .. & .. & .. \\ \rho^N & .. & \rho^2 & \rho & 1 \end{bmatrix} \quad (15.18)$$

e quindi l'errore di stima di una costante cui è sovrapposto del rumore di potenza σ_m^2 , ma colorato da uno zero è:

$$\lim_{\sigma_s^2 \rightarrow \infty} E[\varepsilon^2] = \frac{1}{\underline{\mathbf{c}}^* \underline{\mathbf{M}}_{z,inv} \underline{\mathbf{c}}} \sim \frac{\sigma_m^2 (1 - \rho)^2}{N (1 + \rho^2)}$$

CAPITOLO 15 STIMA LINEARE OTTIMA

Infatti, se $\rho > 0$, i campioni del disturbo sono anticorrelati, e median-doli per trovare la costante, tendono a scomparire, come già visto nel caso della coloratura dovuta ad un polo. Infatti la densità spettrale alle basse frequenze (che è quella che conta nel caso della stima di un segnale costante, e cioè a frequenza 0, è pari a:

$$S_z(\omega = 0) = \frac{\sigma_m^2 (1 - \rho)^2}{1 + \rho^2} \quad (15.19)$$

Se invece p.e. $\rho = -0.9$, ; $N' \sim N/2$, come ragionevole, visto il mediocre filtraggio passa basso, che tende a correlare i campioni del rumore e quindi a ridurre circa del 50% l'efficienza della media campionaria. La dispersione della stima si riduce ed è come se avessimo solo la metà dei campioni (se fossero indipendenti).

15.1.4 Stima di una pendenza in rumore bianco

Indicando ora con il vettore $\underline{\mathbf{c}}$ il caso di una variazione lineare tra -1 e 1 , e quindi supponendo che il segnale osservato sia compreso tra $-L$, L , dove L è una variabile casuale

$$\underline{\mathbf{y}} = L\underline{\mathbf{d}} + \underline{\mathbf{n}}$$

Le componenti del vettore $\underline{\mathbf{d}}$, $\{d_i\}$ sono:

$$\{d_i\} = \frac{2i}{N}; \quad i = -N/2, \dots, N/2$$

$$\underline{\mathbf{d}}^* \underline{\mathbf{d}} = \sum_{i=-N/2}^{N/2} \left(\frac{2i}{N} \right)^2 \sim \frac{4}{N^2} \left[\frac{i^3}{3} \right]_{-N/2}^{N/2} = \frac{4}{N^2} \frac{N^3}{12} = \frac{N}{3}$$

La varianza della stima della variabile L è:

$$\sigma_L^2 = E[(L - \hat{L})L] \simeq \frac{\sigma_L^2 \sigma_n^2}{\sigma_n^2 + \sigma_L^2 N/3} \rightarrow \frac{3\sigma_n^2}{N}$$

e di nuovo si tende alla situazione di non polarizzazione se $\sigma_n^2 \rightarrow 0$. Si osservi poi che se si vuole stimare la pendenza l ponendo:

$$\underline{\mathbf{z}} = l\underline{\mathbf{r}} + \underline{\mathbf{n}}$$

15.2 STIME DI UNA FREQUENZA

$$\{r_i\} = i; \quad i = -N/2, \dots, N/2$$

$$\mathbf{r}^* \mathbf{r} = \sum_{i=-N/2}^{N/2} i^2 \sim \frac{N^3}{12}$$

$$\sigma_{\hat{m}}^2 \rightarrow \frac{12}{N^3} \sigma_n^2$$

e quindi l'effetto del braccio di leva si vede nel fattore N^3 invece del fattore N visto prima.

15.2 Stime di una frequenza

15.2.1 Stima lineare di una rampa di fase

Si consideri un segnale sinusoidale in presenza di rumore additivo:

$$\begin{aligned} x_i &= A \exp[j\theta_i] + w_i = A \exp[j(i\omega T + \phi)] + w_i; \\ \theta_i &= i\omega T + \phi; \quad i = -N/2 \dots N/2 \end{aligned}$$

in cui la fase ϕ e la pulsazione ω sono parametri ignoti e il rumore è gaussiano, bianco, complesso:

$$E[w_i w_l^*] = \sigma_w^2 \delta(i - l)$$

È possibile calcolare sia ω che ϕ facendo una stima lineare della pendenza con il tempo delle fasi dei campioni della sequenza $\{x_i\}$. E' semplice verificare che, per $\sigma \ll |A|$, le parti immaginarie dei campioni della sequenza $\{x_i\}$, normalizzate, sono rappresentate dal modello lineare con rumore additivo:

$$\text{Im} \left(\frac{x_i}{A} \right) \simeq \theta_i + \text{Im} \left(\frac{w_i}{A} \right) = i\omega T + \phi + n_i; \quad \sigma_n^2 = \frac{\sigma_w^2}{2|A|^2}$$

Diventa quindi possibile stimare la fase e la frequenza operando una stima lineare sulle parti immaginarie della sequenza complessa. L'approssimazione delle fasi con le parti immaginarie è accettabile se la rotazione in tutto l'intervallo, riferita al valore centrale $\phi + \omega NT/2$ e quindi pari a $\pm \omega NT/2$ si mantiene inferiore a $\pm \pi$. Si osservi però che se si converte il segnale a frequenza zero utilizzando la frequenza a cui il periodogramma

CAPITOLO 15 STIMA LINEARE OTTIMA

ha il massimo picco, si ottiene proprio un segnale la cui rotazione di fase in tutto l'intervallo di osservazione non supera $\pm\pi$. Ancora, **l'analisi è valida se il rumore è piccolo rispetto al segnale**. In queste ipotesi, la dispersione della stima della fase e della frequenza è quindi:

$$\sigma_{\hat{\phi}} = \frac{\sigma_w}{|A|} \frac{1}{\sqrt{2N}}$$
$$\sigma_{\hat{\omega T}} \sim \sqrt{\frac{12}{N^3} \frac{\sigma_w^2}{2|A|^2}} \approx \frac{\sigma_w \sqrt{6}}{|A| N^{3/2}}$$

trovando un fattore $N^{-3/2}$ invece del fattore N^{-1} trovato per il periodogramma e $N^{-1/2}$ per la tecnica AR quando si deve stimare l'autocorrelazione dai dati avendo a disposizione solo N campioni, come vedremo in un paragrafo successivo. Peraltro, se si volesse determinare una frequenza cercando la posizione del massimo locale in un periodogramma, anche cercando la frazione di cella corrispondente al massimo usando una tecnica di interpolazione, si troverebbe un risultato simile.

15.2.2 Discussione

Da quanto detto appare chiaro che per stimare una frequenza stabile nel tempo, sono necessarie due azioni complementari:

I: filtrare passa banda attorno alla frequenza di interesse, per ridurre il rumore irrilevante; il filtro sarà tanto più stretto quanto più lungo è l'intervallo di osservazione.

II: stimare la pendenza della rampa di fase residua, una volta che siamo già vicini al valore effettivo della frequenza; allora, se gli errori di fase sono al di sotto di π in tutto l'intervallo di stima, si effettua una stima lineare della pendenza residua di questi valori, come visto nella sezione precedente.

E' giusto domandarsi come mai quest'ultimo metodo sia tanto migliore della stima AR; con questa, si trova la frequenza come:

$$\hat{f} = \frac{1}{2\pi T} \angle \hat{r}_1$$

Per dare una risposta intuitiva, si può dire che nel caso AR **non** si sfrutta l'effetto di braccio di leva (un fattore N) dovuto alla progressiva rotazione

15.2 STIME DI UNA FREQUENZA

di fase. Essa è proporzionale al tempo di osservazione NT , vista la rigorosa costanza ipotizzata per la frequenza (coerenza del segnale). Con il metodo AR invece, si cerca di stimare al meglio, con una media su tutti i valori, l'ammontare della rotazione del segnale in un solo passo di campionamento, e pertanto si valuta la frequenza su un intervallo di tempo ristretto; poi si media. In sostanza, non si sfrutta appieno la coerenza a lungo termine del segnale, ma si sfrutta il buon rapporto segnale - rumore, quando c'è. Invece, si accetta che il segnale non sia una riga perfetta, ma un polo, magari vicinissimo al cerchio unitario, ma comunque con una banda di frequenza finita. Quindi, la rampa di fase corrispondente al processo potrebbe avere pendenza variabile e la analisi di questa pendenza non ci permetterebbe di migliorare la stima della frequenza. Essa potrebbe essere variabile nel tempo, magari casualmente. Per avvicinare i due tipi di stima, si può sottocampionare p.e. di un fattore M il processo filtrato passabanda, e quindi allungare l'intervallo di campionamento portandolo a MT ; questo avvicina le due tecniche, perchè da una parte affina la tecnica AR e dall'altra riduce il braccio di leva che diventa N/M e quindi peggiora la stima della pendenza di fase. Quando $M = N$, le due tecniche tendono a coincidere.

15.2.3 Stima non lineare di una frequenza in rumore colorato

Si farà ora un'analisi spettrale autoregressiva di uno segnale non perfettamente sinusoidale, ma costituito da uno spettro colorato, in presenza di rumore colorato. Si consideri un rumore bianco w_i che pilota un filtro passabanda la cui risposta all'impulso è una sinusoide complessa finestrata da una finestra rettangolare lunga N_0 campioni; a questo segnale fortemente colorato, (quasi monocromatico per $N_0 \rightarrow \infty$) si aggiunga un rumore lievemente colorato, p.e. passa - alto. Si abbia cioè:

$$y_i = A \sum_{h=1}^{N_0} w_{i-h} e^{jh\omega T} + n_i - n_{i-1} \quad (15.20)$$

$$E[|w_i|^2] = \sigma_w^2; \quad E[|n_i|^2] = \sigma_n^2; \quad (15.21)$$

Si voglia ora fare la stima spettrale AR della sequenza y_i ; supponiamo per ora di avere la funzione di autocorrelazione **esatta** $r_y(n)$; sia l'ordine

CAPITOLO 15 STIMA LINEARE OTTIMA

dello stimatore pari ad uno. Calcoliamo quindi:

$$r_y(0) = N_0|A|^2\sigma_w^2 + 2\sigma_n^2 \quad (15.22)$$

$$r_y(1) = (N_0 - 1)|A|^2\sigma_w^2 e^{j\varphi} - \sigma_n^2; \quad \varphi = \omega T; \quad (15.23)$$

Si vede anche che:

$$c_1 = -\frac{(N_0 - 1)|A|^2\sigma_w^2 e^{j\varphi} - \sigma_n^2}{N_0|A|^2\sigma_w^2 + 2\sigma_n^2} \quad (15.24)$$

Lo zero di $A_1(z)$ è in

$$\begin{aligned} z_0 &= \frac{(N_0 - 1)|A|^2\sigma_w^2 e^{j\varphi} - \sigma_n^2}{N_0|A|^2\sigma_w^2 + 2\sigma_n^2} = \\ &= \frac{(N_0 - 1)|A|^2\sigma_w^2}{N_0|A|^2\sigma_w^2 + 2\sigma_n^2} e^{j\varphi} - \frac{\sigma_n^2}{N_0|A|^2\sigma_w^2 + 2\sigma_n^2} \end{aligned} \quad (15.25)$$

La polarizzazione della stima spettrale è data dal vettore differenza d tra $e^{j\varphi}$ e z_0 ; in prima approssimazione:

$$d = e^{j\varphi} - z_0 \sim \frac{1}{N_0} \left[\left(1 + \frac{2\sigma_n^2}{|A|^2\sigma_w^2} \right) e^{j\varphi} + \frac{\sigma_n^2}{|A|^2\sigma_w^2} \right] \quad (15.26)$$

$$d \sim \frac{e^{j\varphi}}{N_0} + \frac{\sigma_n^2}{N_0|A|^2\sigma_w^2}$$

Quindi, invece di essere in $e^{j\varphi}$, lo zero dello stimatore (il polo dello spettro ricostruito) viene richiamato verso il centro del cerchio ed anche spostato. La lunghezza finita della risposta all'impulso sinusoidale porta all'allargamento della riga spettrale ricostruita; l'aggiunta del rumore colorato sposta il picco dello spettro stimato verso le frequenze più elevate.

Nel caso invece in cui $N_0 \rightarrow \infty$, e vi sia un rumore bianco additivo la stima di frequenza non è più polarizzata:

$$y_i = Ae^{ji\varphi} + n_i; \quad r_y(0) = \sigma_A^2 + \sigma_n^2; \quad r_y(1) = \sigma_A^2 e^{j\varphi}$$

$$z_0 = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_n^2} e^{j\varphi}$$

15.2.4 Effetti della lunghezza limitata dei dati

Nel paragrafo precedente abbiamo supposto che fosse disponibile l'autocorrelazione esatta; invece, se sono disponibili solo M dati, si ha solo una stima dell'autocorrelazione e quindi il valore trovato per z_0 (il valore atteso) sarà affetto da un disturbo casuale che vogliamo ora calcolare in modo approssimato. La dispersione della frequenza e quindi dell'argomento di \hat{z}_0 può essere stimata come segue, supponendo ora che il segnale sia sinusoidale e il rumore additivo sia bianco:

$$y_i = A e^{ji\omega T} + n_i ; i = 1, \dots, N \quad (15.27)$$

$$\hat{r}_y(1) = \sum \frac{y_i y_{i-1}^*}{N} \quad (15.28)$$

$$\begin{aligned} \arg(\hat{r}_1) &= \arg \sum (|A|^2 e^{j\phi} + |A| (n_i + n_{i-1}^*) + n_i n_{i-1}^*) = \\ &= \arg \left(e^{j\phi} + \frac{|A| (n_i + n_{i-1}^*) + n_i n_{i-1}^*}{N |A|^2} \right) \end{aligned}$$

e considerando solo l'effetto dei termini in quadratura con $e^{j\phi}$:

$$\hat{\phi} \sim \phi + \frac{|A|p}{\sqrt{2N}|A|^2} + \frac{q}{\sqrt{2N}|A|^2} \sim \phi + \frac{p}{\sqrt{2N}|A|} + \frac{q}{\sqrt{2N}|A|^2}$$

$$\sigma_p^2 = 2N\sigma_n^2; \quad \sigma_q^2 = N\sigma_n^4$$

$$\sigma_{\hat{\phi}}^2 \sim \frac{2N\sigma_n^2}{2N^2|A|^2} + \frac{N\sigma_n^4}{2N^2|A|^4} \sim \frac{\sigma_n^2}{N|A|^2}$$

$$\hat{f} = \frac{\hat{\phi}}{2\pi T} \rightarrow \sigma_{\hat{f}}^2 = \frac{\sigma_n^2}{4\pi^2 T^2 N |A|^2} = f_c^2 \frac{\sigma_n^2}{4\pi^2 N |A|^2}$$

In sostanza, con dei ragionevoli rapporti segnale rumore, la maggiore fonte di dispersione della posizione dello zero è proprio la dispersione della stima dell'autocorrelazione.

15.3 Stima di processi multidimensionali

Si supponga di volere stimare un processo multidimensionale e cioè un insieme di N sequenze temporalmente allineate: in particolare si voglia stimare la sequenza di vettori di N dimensioni p.e. incorrelati sia tra loro che nel tempo:

$$\underline{\mathbf{s}}_i = \begin{bmatrix} s_{1,i} \\ s_{2,i} \\ s_{j,i} \\ \dots \\ s_{N,i} \end{bmatrix}; \quad j = 1, \dots, N; \quad i = 1, 2, \dots;$$

$$E[\underline{\mathbf{s}}_i \underline{\mathbf{s}}_{i+k}^*] = \sigma_s^2 \delta(i-k) \mathbf{I}$$

Il segnale osservato è la sequenza $\underline{\mathbf{x}}_i$, e sussiste la seguente relazione matriciale:

$$\underline{\mathbf{x}}_i = \underline{\mathbf{A}} \underline{\mathbf{s}}_i + \underline{\mathbf{n}}_i; \quad E[\underline{\mathbf{n}}_i \underline{\mathbf{n}}_{i+k}^*] = \sigma_n^2 \delta(i-k) \mathbf{I}$$

Quindi, il rumore additivo è costituito da componenti incorrelate e tra loro e nel tempo. La matrice $\underline{\mathbf{A}}$, qualsiasi e magari a coefficienti complessi, ha dimensioni $[N \times N]$. Possiamo trovare subito la matrice $\underline{\mathbf{R}}$ di covarianza del vettore $\underline{\mathbf{x}}$:

$$\underline{\mathbf{R}} = E[(\underline{\mathbf{A}}\underline{\mathbf{s}} + \underline{\mathbf{n}})(\underline{\mathbf{A}}\underline{\mathbf{s}} + \underline{\mathbf{n}})^*] = \sigma_s^2 \underline{\mathbf{A}} \underline{\mathbf{A}}^* + \sigma_n^2 \mathbf{I}$$

Usando il principio di ortogonalità e sfruttando il formalismo matriciale troviamo subito la stima ottima del segnale $\underline{\mathbf{s}}$:

$$\hat{\underline{\mathbf{s}}} = \underline{\mathbf{B}} \underline{\mathbf{x}}; \quad E[(\underline{\mathbf{B}} \underline{\mathbf{x}} - \underline{\mathbf{s}}) \underline{\mathbf{x}}^*] = \mathbf{0} \rightarrow \underline{\mathbf{B}} \underline{\mathbf{R}} = E[\underline{\mathbf{s}} \underline{\mathbf{x}}^*] = \sigma_s^2 \underline{\mathbf{A}}^*$$

e la matrice $\underline{\mathbf{B}}$ è quindi:

$$\underline{\mathbf{B}} = \sigma_s^2 \underline{\mathbf{A}}^* \underline{\mathbf{R}}^{-1}$$

L'errore di stima è:

$$Q = \sum_{n=1}^N E[(\hat{s}_{i,n} - s_{i,n})^2]$$

15.3 STIMA DI PROCESSI MULTIDIMENSIONALI

$$Q = Tr \{E [(\underline{\mathbf{B}}\underline{\mathbf{x}} - \underline{\mathbf{s}}) (\underline{\mathbf{B}}\underline{\mathbf{x}} - \underline{\mathbf{s}})^*]\} = Tr \{E [(\underline{\mathbf{s}} - \underline{\mathbf{B}}\underline{\mathbf{x}}) \underline{\mathbf{s}}^*]\}$$

profittando dell'incorrelazione tra errore di stima e dati per il principio di ortogonalità. Ricordando poi che:

$$E[\underline{\mathbf{x}} \underline{\mathbf{s}}^*] = \sigma_s^2 \underline{\mathbf{A}}$$

risulta che:

$$Q = \sigma_s^2 Tr \{ \underline{\mathbf{I}} - \underline{\mathbf{A}}^* \underline{\mathbf{R}}^{-1} \underline{\mathbf{A}} \}$$

Si osservi che dalla sequenza di vettori $\underline{\mathbf{x}}_i$ è possibile ricavare la matrice $\underline{\mathbf{R}}$ e quindi, se σ_n^2 è nota, si può determinare la matrice $\underline{\mathbf{A}} \underline{\mathbf{A}}^*$; non è possibile però ricavare $\underline{\mathbf{A}}$. Infatti, si può osservare che

$$\underline{\mathbf{A}} \underline{\mathbf{T}} \underline{\mathbf{T}}^* \underline{\mathbf{A}}^* = \underline{\mathbf{A}} \underline{\mathbf{A}}^*$$

purchè:

$$\underline{\mathbf{T}} \underline{\mathbf{T}}^* = \underline{\mathbf{I}}$$

e quindi se la matrice $\underline{\mathbf{T}}$ è ortogonale, eguale all'inversa della sua trasposta. Allora, se è nota la matrice $\underline{\mathbf{R}}$ e quindi $\underline{\mathbf{A}} \underline{\mathbf{A}}^*$ non potremo ricavarne la matrice $\underline{\mathbf{A}}$, se non a meno di una matrice ortogonale $\underline{\mathbf{T}}$ che rimane incognita. Abbiamo visto un problema equivalente per la caratteristica di fase di una funzione di trasferimento, irrecuperabile dall'autocorrelazione dell'uscita da cui otteniamo lo spettro di potenza e quindi le ampiezze. Il problema è solubile se si possono utilizzare informazioni a priori sulla struttura della matrice $\underline{\mathbf{A}}$. Ad esempio, se $\underline{\mathbf{A}}$ è simmetrica (come dire fase zero), allora la matrice

$$\underline{\mathbf{T}} = \underline{\mathbf{I}}; \rightarrow \underline{\mathbf{A}} = \underline{\mathbf{A}}^*$$

gli autovettori di $\underline{\mathbf{R}}$ coincidono con quelli di $\underline{\mathbf{A}}$ (le matrici $\underline{\mathbf{A}}, \underline{\mathbf{A}}^*$ commutano) e gli autovalori di $\underline{\mathbf{R}}$ sono i quadrati di quelli di $\underline{\mathbf{A}}$ cui va sommato σ_n^2 . La formula che dà l'errore quadratico medio si semplifica:

$$Q = \sigma_s^2 Tr \{ \underline{\mathbf{I}} - \underline{\mathbf{A}}^* \underline{\mathbf{R}}^{-1} \underline{\mathbf{A}} \} = \sigma_s^2 \sum_{n=1}^N \left(1 - \frac{\lambda_{A,n}^2 \sigma_s^2}{\lambda_{A,n}^2 \sigma_s^2 + \sigma_n^2} \right)$$

$$Q = \sum_{n=1}^N \left(\frac{\sigma_s^2 \sigma_n^2}{\lambda_{A,n}^2 \sigma_s^2 + \sigma_n^2} \right)$$

CAPITOLO 15 STIMA LINEARE OTTIMA

profittando del fatto che la traccia di una matrice è la somma degli autovalori.

Allo stesso modo, nell'analisi fatta per la deconvoluzione di segnali nel dominio delle frequenze, si è visto che è impossibile determinare la fase della funzione di trasferimento a partire dall'autocorrelazione dell'uscita

15.3.1 Stima simultanea di una costante e di una pendenza

Se ora siamo nel caso di una stima simultanea sia di una costante A che di una pendenza L :

$$\underline{\mathbf{y}} = L\underline{\mathbf{d}} + A\underline{\mathbf{c}} + \underline{\mathbf{n}}$$

profittando dell'ortogonalità tra $\underline{\mathbf{c}}$ e $\underline{\mathbf{d}}$ e cioè del fatto che:

$$\sum_{i=0}^N c_i d_i^* = 0$$

si vede che le analisi fatte nei precedenti paragrafi sono ancora utilizzabili: basta stimare il vettore $[A \ L]$.

15.4 La separazione di sorgenti

Le considerazioni fatte nel paragrafo precedente hanno un interessante conseguenza nel caso della ricezione acustica binaurale (con le due orecchie) di un segnale costituito da due sorgenti che emettano contemporaneamente. In altre parole, si pensi ad una persona che ascolti due parlatori che parlino assieme. È ben noto che è possibile concentrare l'attenzione su uno dei due parlatori e praticamente cancellare la conversazione dell'altro (il cosiddetto *effetto cocktail*, che permette di seguire una conversazione in mezzo a tante). Tuttavia, la separazione è quasi impossibile se i due parlatori sono sovrapposti su un singolo canale fonico (p.e. parlano contemporaneamente alla televisione).

Questo problema di stima è analizzabile con le convenzioni stabilite nel paragrafo precedente e risulta:

$$\underline{\mathbf{x}}_i = \underline{\mathbf{A}} \underline{\mathbf{s}}_i + \underline{\mathbf{n}}_i ; \quad E [\underline{\mathbf{n}}_i \underline{\mathbf{n}}_{i+k}^*] = \sigma_n^2 \delta(i - k) \mathbf{I}$$

15.4 LA SEPARAZIONE DI SORGENTI

dove ora i vettori $\underline{\mathbf{x}}_i$, $\underline{\mathbf{s}}_i$, $\underline{\mathbf{n}}_i$ sono di rango $[2 \times 1]$ e la matrice $\underline{\mathbf{A}}$ di rango $[2 \times 2]$. Rappresentiamo le due conversazioni con $\underline{\mathbf{s}}$, i due segnali ricevuti (le due orecchie) con $\underline{\mathbf{x}}$ e la matrice di confusione anch'essa a priori ignota con $\underline{\mathbf{A}}$. Se si conoscesse $\underline{\mathbf{A}}$ sarebbe facile, a parte il disturbo, ritrovare le due sorgenti usando le tecniche viste prima; per trovare $\underline{\mathbf{A}}$ si potrebbe p.e. correlare il segnale misurato con quello della sorgente (se fosse nota, però), facendo:

$$\underline{\mathbf{A}} = E[\underline{\mathbf{x}} \underline{\mathbf{s}}^*] \quad (15.29)$$

Ma $\underline{\mathbf{s}}$ corrisponde alla conversazione che deve essere stimata è inaccessibile. Certamente si dispone di $\underline{\mathbf{x}}$ e quindi si dispone della sua matrice di covarianza $\underline{\mathbf{R}} = E[\underline{\mathbf{x}}\underline{\mathbf{x}}^*]$: il problema diventa più complesso. Si è visto infatti che dalla matrice di covarianza $\underline{\mathbf{R}}$ non è possibile ottenere la matrice $\underline{\mathbf{A}}$ se non a meno di una rotazione di coordinate $\underline{\mathbf{T}}$. In altre parole, anche in assenza di rumore non possiamo ottenere di nuovo $\hat{\underline{\mathbf{s}}}$ (la stima delle due sorgenti di segnale) da $\underline{\mathbf{x}}$ ma solo un vettore $\hat{\underline{\mathbf{s}}}'$ in cui le componenti di $\underline{\mathbf{s}}$, sono mescolate tra loro:

$$\hat{\underline{\mathbf{s}}}' = \underline{\mathbf{T}}\hat{\underline{\mathbf{s}}}; \quad E[\hat{\underline{\mathbf{s}}}' \hat{\underline{\mathbf{s}}}'^*] = E[\hat{\underline{\mathbf{s}}} \hat{\underline{\mathbf{s}}}^*] = \sigma_s^2 \underline{\mathbf{I}}$$

$$\underline{\mathbf{T}} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}; \quad \underline{\mathbf{T}}\underline{\mathbf{T}}^* = \underline{\mathbf{I}} \quad (15.30)$$

Peraltro, il fatto che si possano separare le due conversazioni dimostra che esistono dei meccanismi più efficienti della predizione lineare per fare l'operazioni di separazione delle sorgenti. Infatti, se invece di imporre l'incorrelazione tra gli elementi di $\hat{\underline{\mathbf{s}}}$, possibile qualsiasi sia la matrice $\underline{\mathbf{T}}$ e quindi inutile per effettuare una separazione delle sorgenti, si impone l'indipendenza tra quelle componenti, il risultato non potrebbe essere altro che quello corretto. In sostanza, si deve ricorrere a statistiche di ordine superiore al secondo (deconvoluzione cieca).

Per fare un esempio, consideriamo due processi identici s_1, s_2 ma indipendenti, non gaussiani, a curtosi $k \neq 3$; ricordiamo che la curtosi é:

$$k = \frac{E[s_1^4]}{(E[s_1^2])^2} = \frac{E[s_2^4]}{(E[s_2^2])^2} \quad (15.31)$$

CAPITOLO 15 STIMA LINEARE OTTIMA

Dopo la stima lineare ottima, noi abbiamo \hat{s}'_1, \hat{s}'_2

$$\underline{\hat{s}'} = \underline{\mathbf{T}}\underline{\hat{s}} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} \hat{s}_1 \\ \hat{s}_2 \end{bmatrix} \quad (15.32)$$

$$\hat{s}'_1 = \hat{s}_1 \cos \theta + \hat{s}_2 \sin \theta \quad (15.33)$$

$$\hat{s}'_2 = -\hat{s}_1 \sin \theta + \hat{s}_2 \cos \theta \quad (15.34)$$

che sono due combinazioni lineari di \hat{s}_1, \hat{s}_2 , **incorrelate ed a eguale potenza**, qualsiasi sia il valore di θ , peraltro ignoto. Quello che varia con θ , invece, é la curtosi di \hat{s}'_1, \hat{s}'_2 che é:

$$k_m = \frac{E[\hat{s}'_1]^4 (\cos^4 \theta + \sin^4 \theta) + 6(E[\hat{s}'_1]^2)^2 \sin^2 \theta \cos^2 \theta}{(E[\hat{s}'_1]^2)^2} \quad (15.35)$$

$$k_m(\theta) = \frac{3k+3}{4} + \frac{k-3}{4} \cos 4\theta \quad (15.36)$$

La curtosi misurata k_m oscilla tra due valori:

$$k < k_m < \frac{k+3}{2} \text{ oppure } k > k_m > \frac{k+3}{2} \quad (15.37)$$

a seconda che k sia maggiore o minore di 3. Si vede che, se $k \neq 3$, e quindi le due conversazioni sono rappresentate da segnali non gaussiani, la curtosi misurata di \hat{s}'_1, \hat{s}'_2 oscilla al variare di θ tra k e la media aritmetica tra k e 3. I valori di θ per cui k_m arriva al valore massimo corrispondono ad una separazione. Questi valori sono 4 e pari a:

$$\theta_m = m\pi/2; m = 0..3; \quad (15.38)$$

nel caso in cui il valore originario di θ è zero. Allora

$$k_m(\theta_m) = k \quad (15.39)$$

Peraltro, per $m = 1, 3$ scambiamo le sorgenti tra loro e per $m = 0, 2$ ne cambiamo i segni. In ogni modo, selezionando un qualsiasi massimo della curtosi, abbiamo separato le conversazioni e risolto il problema. Invece per

$$\theta = \pi/4 + m\pi/2; m = 0..3; \rightarrow k_m = \frac{k+3}{2} \quad (15.40)$$

15.4 LA SEPARAZIONE DI SORGENTI

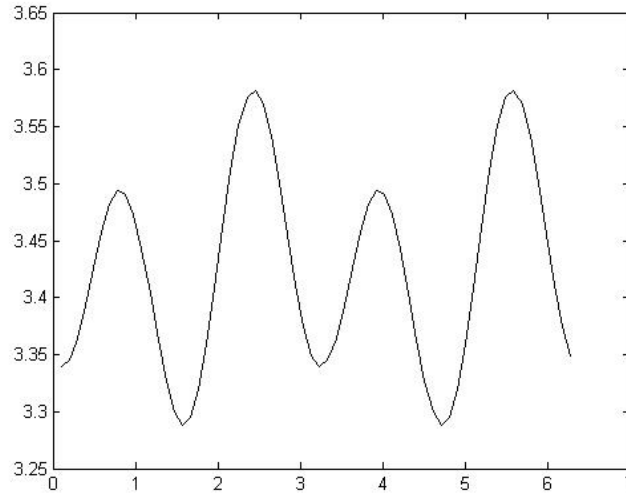


Figura 15.1: Curtosi del segnale combinato, al variare dell'angolo θ . Si notano i quattro massimi (circa 3.53) e i quattro minimi (circa 3.26).

non c'è separazione. Infine, se i processi s_1, s_2 sono gaussiani, anche le loro somme lo saranno e la separazione sarà impossibile.

Facciamo una semplice simulazione, utilizzando una sequenza di 10000 campioni indipendenti e gaussiani che, se azzerata con probabilità $1 - p$, ha curtosi $3/p$. Nella figura 15.1, si mostra l'andamento della curtosi del segnale \tilde{s}_1 al variare dell'angolo θ , nel caso in cui si abbia $p = 0.85$. Sono perfettamente visibili i 4 minimi e massimi della curtosi, proprio ai valori di θ corretti. Peraltro, a causa della lunghezza limitata della sequenza, i valori di curtosi stimati sono leggermente differenti da quelli attesi.

In conclusione, misurando la matrice di covarianza \mathbf{R} del vettore \mathbf{x} , dopo avere eventualmente sottratto la parte dovuta al rumore, otteniamo la matrice $\mathbf{A} \mathbf{T} \mathbf{T}^* \mathbf{A}^*$ e quindi, facendo la radice quadrata di questa matrice, otteniamo la matrice $\mathbf{A} \mathbf{T}$ che è una matrice ancora simmetrica, come la \mathbf{R} . Per trovare finalmente la matrice \mathbf{T} ed arrivare all'effettiva separazione delle sorgenti, dovremo sfruttare la non gaussianità delle

CAPITOLO 15 STIMA LINEARE OTTIMA

sorgenti. Per fare un esempio numerico sia:

$$\underline{\mathbf{A}} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}; \text{ quindi } \underline{\mathbf{A}}\underline{\mathbf{A}}^* = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 11 \\ 11 & 25 \end{bmatrix} \quad (15.41)$$

$$\underline{\mathbf{A}}\underline{\mathbf{T}} = \begin{bmatrix} 5 & 11 \\ 11 & 25 \end{bmatrix}^{1/2} = \begin{bmatrix} 1.2 & 1.89 \\ 1.89 & 4.63 \end{bmatrix}; \quad (15.42)$$

Se conoscessimo $\underline{\mathbf{A}}$, troveremmo la rotazione come:

$$\underline{\mathbf{T}} = \underline{\mathbf{A}}^{-1}\underline{\mathbf{A}}\underline{\mathbf{T}} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 1.2 & 1.89 \\ 1.89 & 4.63 \end{bmatrix} = \quad (15.43)$$

$$= \begin{bmatrix} -0.51 & 0.85 \\ 0.855 & 0.51 \end{bmatrix} = \begin{bmatrix} -\sin \theta & \cos \theta \\ \cos \theta & \sin \theta \end{bmatrix} \quad (15.44)$$

$$\theta = \arccos 0.855 = 0.545 \quad (15.45)$$

La rotazione di 0.54 radianti é appunto quella che ci ridarebbe la matrice $\underline{\mathbf{A}}$ originale invece di quella simmetrizzata. Questo valore di θ sará appunto quello ottenibile dalla massimizzazione (o minimizzazione) della curtosí.

Capitolo 16

Filtraggio adattativo

Uno dei numerosi motivi della diffusione delle tecniche di trattamento numerico dei segnali è la possibilità di modificare i filtri utilizzati in funzione delle caratteristiche spettrali del dato, variabili con il tempo e/o con lo spazio. Un tipico esempio di sistema tempo variante è il tratto vocale, che sagoma lo spettro del segnale generato dalle corde vocali. La stima spettrale di un segnale vocale dovrà quindi essere "adattativa" nel tempo.

Un esempio intuitivo di un sistema adattativo è quello del cancellatore di disturbo: si pensi ad un pilota di elicottero, che ascolta un segnale s_i ma è disturbato dal rumore del rotore; possiamo pensare di aggiungere negli auricolari del pilota un segnale che compensi ed annulli questo rumore esterno: per fare ciò è necessario ricostruire elettronicamente sia in fase che in ampiezza il rumore disturbante, per poi applicarlo con il segno cambiato. La struttura numerica che si può adottare è presentata in figura 13.1: anzitutto, viene prelevata una sequenza (indicata poi come il disturbo $\{d_i\}$) quanto possibile correlata con il rumore del rotore ascoltato dal pilota $\{n_i\}$, ma indipendente dalla sequenza di segnale $\{s_i\}$ che desideriamo arrivi indisturbata. Ad esempio, la sequenza $\{d_i\}$ potrebbe essere prelevata direttamente con un sensore posto vicino al rotore; si indichi poi con $\{y_i\}$ il segnale disturbato che arriva all'orecchio del pilota. Si desidera rimuovere da $\{y_i\}$ tutta l'energia correlata con il disturbo misurato $\{d_i\}$. Così facendo, si **decorrela** il segnale ascoltato dal pilota $\{y_i\}$ dal disturbo misurato $\{d_{i-n+1}\}$, $n = 1, \dots, N$. In altre parole, il residuo $\varepsilon_{i,N}$ che si ottiene sottraendo da $\{y_i\}$ il disturbo stimato $\{\hat{n}_i\}$ deve essere la sequenza $\{s_i\}$ per quanto possibile **incorrelata**

CAPITOLO 16 FILTRAGGIO ADATTATIVO

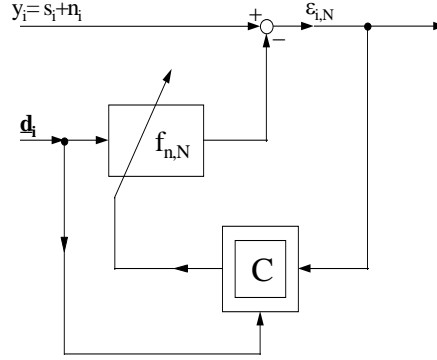


Figura 16.1: Filtro adattativo

con la sequenza di vettori \underline{d}_i aventi componenti $\{d_{i-n+1}\}$, $n = 1, \dots, N$. A sua volta, la sequenza $\{\hat{n}_i\}$ è ottenuta filtrando il disturbo misurato $\{d_{i-n+1}, n = 1, \dots, N\}$ con un filtro opportuno $\{f_{i,n}\}$, $n = 1, \dots, N$, che sarà tempo variante (ecco il perchè dell'indice i), ad esempio per compensare i ritardi tempo - varianti dovuti ai movimenti della testa del pilota.

$$y_i = s_i + n_i; \quad \varepsilon_i = s_i + n_i - \hat{n}_i = y_i - \sum_{n=1}^N f_{i,n} d_{i-n+1} \quad (16.1)$$

Quando si avrà che

$$E[\varepsilon_i d_{i-n+1}] = 0 \quad n = 1, \dots, N \quad (16.2)$$

avremo fatto il possibile per levare il rumore del rotore dall'auricolare. Difatti, nel segnale residuo $\varepsilon_i = s_i$ non ci dovrebbe più essere correlazione con il disturbo $\{n_i\}$, sempre che il filtro sia di durata sufficiente (N abbastanza grande) e che il disturbo misurato $\{d_i\}$ sia legato al disturbo da rimuovere $\{n_i\}$ attraverso una funzione di trasferimento che non ha zeri spettrali e senza che vi sia ulteriore disturbo sovrapposto; allora infatti il disturbo stimato $\{\hat{n}_i\}$ assomiglia molto al disturbo vero $\{n_i\}$.

Questo non è che un possibile esempio: un altro è quello di un elettrocardiogramma fetale; l'elettrocardiogramma di un nascituro è misurabile

16.1 LA TECNICA DI WIDROW E HOFF

solo se combinato con quello della madre; per isolare il primo, è necessario misurare un segnale correlato soprattutto al secondo e poi sottrarlo, sempre tenendo conto della variabilità delle condizioni dell'esperimento. Ancora, si pensi alla cancellazione di un'eco che arriva ad un auricolare telefonico; in questo caso, l'eco nasce soprattutto dall'accoppiamento fonico tra il telefono ed il microfono dell'ascoltatore lontano. Il segnale generato localmente viene trasmesso all'ascoltatore, rientra nel microfono lontano e ritorna come eco al parlatore. Per sopprimere quest'eco, basta decorrelare il segnale in arrivo dal segnale in partenza. La funzione di trasferimento è tempo variante; cambierà, ad esempio, se una barriera (una mano) riflette le radiazioni acustiche dal telefono al microfono dell'apparecchio lontano. Ma ancora, si può pensare che anche un predittore possa essere un sistema adattativo: in questo caso cercheremo di decorrelare il segnale attuale s_i da una sua versione ritardata di un campione e cioè porremo:

$$d_{i-n+1} = s_{i-n}; \quad n = 1, \dots, N$$

16.1 La tecnica di Widrow e Hoff

Per presentare in modo sintetico l'evoluzione del filtro utilizzeremo la notazione vettoriale (supponendo il segnale complesso, indichiamo con * le operazioni di trasposizione e coniugio, dove necessario); indichiamo poi con i vettori $\underline{\mathbf{f}}_i$, $\underline{\mathbf{d}}_i$ gli N coefficienti del filtro e gli ultimi N campioni della sequenza $\{d_i\}$ all'istante iT :

$$\underline{\mathbf{f}}_i = \{f_{i,n}\}; \quad n = 1, \dots, N \quad (16.3)$$

$$\underline{\mathbf{d}}_i = \{d_{i-n+1}\} \quad n = 1, \dots, N \quad (16.4)$$

La sottrazione del disturbo stimato viene quindi presentata così:

$$\varepsilon_i = s_i + n_i - \hat{n}_i = y_i - \hat{n}_i = y_i - \sum_{n=1}^N f_{i,n} d_{i-n+1} = y_i - \underline{\mathbf{f}}_i^* \underline{\mathbf{d}}_i \quad (16.5)$$

Inoltre aggiorniamo il filtro ad ogni istante con l'equazione (Widrow e Hoff):

$$\underline{\mathbf{f}}_{i+1} = \underline{\mathbf{f}}_i + \tilde{\gamma} \varepsilon_i^* \underline{\mathbf{d}}_i \quad (16.6)$$

Osserviamo anzitutto che $\underline{\mathbf{f}}_i$ è un vettore casuale, di cui cerchiamo di determinare l'evoluzione nel tempo. Dimosteremo ora che, purchè $\tilde{\gamma}$ sia un

CAPITOLO 16 FILTRAGGIO ADATTATIVO

fattore abbastanza piccolo, $\underline{\mathbf{f}}_i$ convergerà al filtro ottimo in qualche modo, ancora da illustrare. Anzitutto, è opportuno chiarire il perchè si utilizza l'equazione di Widrow e Hoff. Indichiamo con Q_i l'errore quadratico medio (media di insieme), all'istante i -esimo:

$$Q_i = E[|\varepsilon_i|^2] \quad (16.7)$$

ed il suo gradiente rispetto al filtro $\underline{\mathbf{f}}_i$;

$$\nabla_{\underline{\mathbf{f}}} Q_i = -2E[\varepsilon_i^* \underline{\mathbf{d}}_i] \quad (16.8)$$

pertanto l'equazione (16.6) corrisponde all'aggiornamento del vettore filtro $\underline{\mathbf{f}}_i$ nella direzione della minimizzazione dell'errore quadratico medio. Non essendo possibile ricorrere all'operatore di media d'insieme $E[.]$, bisogna accettarne una stima stocastica anche molto mediocre, e cioè $\varepsilon_i^* \underline{\mathbf{d}}_i$. Per evitare fenomeni di instabilità, come vedremo meglio al prossimo paragrafo, è necessario che il passo nella direzione del gradiente sia piccolo. Questo implica che l'aggiornamento del filtro $\underline{\mathbf{f}}_i$ è lento. Se osserviamo l'aggiornamento ogni K istanti, invece che ad ogni istante, trascurando la variazione del filtro in questo intervallo KT visto che $\tilde{\gamma}$ è comunque piccolo, abbiamo:

$$\underline{\mathbf{f}}_{i+K} = \underline{\mathbf{f}}_i + \gamma \frac{1}{K} \sum_{k=0}^{K-1} \varepsilon_{i+k}^* \underline{\mathbf{d}}_{i+k}; \quad \gamma = K \tilde{\gamma} \quad (16.9)$$

Quindi, se il sistema è stazionario, o almeno lo è durante l'aggiornamento del filtro, la stima stocastica del gradiente è ragionevole, purchè l'aggiornamento sia lento. Se quindi approssimiamo con:

$$\frac{1}{K} \sum_{k=0}^{K-1} \varepsilon_{i+k}^* \underline{\mathbf{d}}_{i+k} \simeq E[\varepsilon_i^* \underline{\mathbf{d}}_i] \quad (16.10)$$

possiamo ritenere valida l'equazione (16.9) che riscriviamo, decidendo anche di porre $K = 1$.

$$\underline{\mathbf{f}}_{i+1} = \underline{\mathbf{f}}_i + \gamma E[\varepsilon_i^* \underline{\mathbf{d}}_i] \quad (16.11)$$

Possiamo ora studiare l'evoluzione del filtro, che è divenuta deterministica per l'approssimazione fatta in precedenza (16.10):

$$\underline{\mathbf{f}}_{i+1} = \underline{\mathbf{f}}_i + \gamma E[\underline{\mathbf{d}}_i (y_i^* - \underline{\mathbf{d}}_i^* \underline{\mathbf{f}}_i)] \quad (16.12)$$

16.1 LA TECNICA DI WIDROW E HOFF

Indicando con $\underline{\mathbf{R}}$ la matrice di covarianza dei dati e con $\underline{\mathbf{r}}$ il vettore della correlazione tra segnale e dati:

$$\underline{\mathbf{R}} = E[\underline{\mathbf{d}}_i \underline{\mathbf{d}}_i^*]; \quad \underline{\mathbf{r}} = E[y_i^* \underline{\mathbf{d}}_i]$$

Sostituendo nell'equazione (16.12):

$$\begin{aligned} \underline{\mathbf{f}}_{i+1} &= \underline{\mathbf{f}}_i + \gamma E[\varepsilon_i \underline{\mathbf{d}}_i] \\ \underline{\mathbf{f}}_{i+1} &= (\underline{\mathbf{I}} - \gamma \underline{\mathbf{R}}) \underline{\mathbf{f}}_i + \gamma \underline{\mathbf{r}} \end{aligned}$$

Per $i \rightarrow \infty$, $\underline{\mathbf{f}}_{i+1} = \underline{\mathbf{f}}_i$ e quindi:

$$\underline{\mathbf{f}}_\infty = \underline{\mathbf{R}}^{-1} \underline{\mathbf{r}}$$

L'errore quadratico medio di predizione sarà, al solito sfruttando l'incorrelazione tra errore di stima e dati $\underline{\mathbf{d}}_i$:

$$\begin{aligned} E[\varepsilon^2] &= E[(y_i - \underline{\mathbf{f}}_i^* \underline{\mathbf{d}}_i)(y_i - \underline{\mathbf{f}}_i^* \underline{\mathbf{d}}_i)^*] \\ &= \sigma_y^2 - \underline{\mathbf{f}}_i^* E[\underline{\mathbf{d}}_i y_i^*] = \sigma_y^2 - \underline{\mathbf{r}}^* \underline{\mathbf{R}}^{-1} \underline{\mathbf{r}} \end{aligned} \quad (16.13)$$

La differenza tra la media di insieme del filtro corrente $\underline{\mathbf{f}}_i$ ed il filtro desiderato $\underline{\mathbf{f}}_\infty$ è il filtro errore $\underline{\mathbf{g}}_i$ (media di insieme) che evolve nel tempo con l'equazione:

$$\underline{\mathbf{f}}_{i+1} - \underline{\mathbf{f}}_\infty = \underline{\mathbf{g}}_{i+1} = (\underline{\mathbf{I}} - \gamma \underline{\mathbf{R}}) \underline{\mathbf{g}}_i = (\underline{\mathbf{I}} - \gamma \underline{\mathbf{R}})^{i+1} \underline{\mathbf{g}}_0 \quad (16.14)$$

Condizione necessaria perchè la media di insieme del filtro errore converga a zero è che il modulo degli autovalori della matrice $\underline{\mathbf{I}} - \gamma \underline{\mathbf{R}}$ sia inferiore ad uno, e cioè $|1 - \gamma \lambda_n| < 1$ Questo implica:

$$0 < \gamma < \frac{2}{\lambda_{max}}; \rightarrow \frac{2}{\lambda_{max}} > \frac{2}{\sum_{n=1}^N \lambda_n} = \frac{2}{Tr[\underline{\mathbf{R}}]} = \frac{2}{N\sigma_d^2} \quad (16.15)$$

pertanto imponendo:

$$\gamma \leq \frac{2}{N\sigma_d^2} \leq \frac{2}{\lambda_{max}} \quad (16.16)$$

siamo *convinti* ??? di ottenere il risultato di stabilità richiesto. In realtà, stiamo solo imponendo che il filtro errore medio converga a zero. Se vogliamo che anche la sua dispersione converga a zero (vedi ad un paragrafo successivo) il valore di γ deve essere almeno 10 volte più basso.

CAPITOLO 16 FILTRAGGIO ADATTATIVO

Con $Tr[\underline{\mathbf{R}}]$ abbiamo indicato la traccia della matrice $\underline{\mathbf{R}}$ e cioè la somma degli elementi della sua diagonale principale e quindi la somma degli autovalori; ciò corrisponde a N volte la varianza del disturbo stesso. Il risultato ottimale e cioè la convergenza più veloce si otterrà quando tutti gli autovalori sono uguali, come si vedrà nella prossima sezione, e poi ponendo $\gamma = 1/\lambda_{\max}$.

Si osservi che l'equazione di Widrow e Hoff, nella sua grande semplicità, ha un significato notevole: in sostanza, il sistema fa la stima $\underline{\mathbf{f}}_i$ della funzione di trasferimento e poi, per migliorarla, utilizza l'esperienza e quindi sia l'errore commesso ε_i che l'ingresso $\underline{\mathbf{d}}_i$ che ha condotto a quell'errore con quella stima. Questo è un meccanismo di autocorrezione valido per chiunque.

16.2 Dispersione degli autovalori

La convergenza a zero del filtro errore $\underline{\mathbf{g}}_i$ dipende dagli autovalori λ_n e quindi le componenti del filtro errore lungo le direzioni degli autovettori di $\underline{\mathbf{R}}$ con autovalori piccoli si aggiornano lentamente; peraltro, non è possibile accrescere troppo γ , limitato dall'autovalore più grande. Infatti, indicando con $g_{i,n}$ la componente di $\underline{\mathbf{g}}_i$ lungo $\underline{\mathbf{e}}_n$, autovettore corrispondente all' n -esimo autovalore λ_n della matrice $\underline{\mathbf{R}}$:

$$g_{i,n} = \underline{\mathbf{e}}_n^* \underline{\mathbf{g}}_i = (1 - \gamma \lambda_n)^{i+1} g_{0,n} \quad (16.17)$$

Sia:

$$\gamma = \alpha \frac{2}{\lambda_{\max}}; 0 < \alpha < 1$$

Ponendo $\alpha = 1/2$, la componente relativa all'autovalore più grande converge in un solo passo (in media). Quelle relative agli altri autovettori, più lentamente, e mai se $\lambda_n = 0$.

$$g_{i,n \max} = \underline{\mathbf{e}}_n^* \underline{\mathbf{g}}_i = (1 - 0.5 \frac{2\lambda_{\max}}{\lambda_{\max}})^{i+1} g_{0,n} = 0$$

$$g_{i,n} = \underline{\mathbf{e}}_n^* \underline{\mathbf{g}}_i = (1 - \frac{\lambda_n}{\lambda_{\max}})^{i+1} g_{0,n} = 0$$

Si ricordi che i λ_i sono le potenze del disturbo misurato $\underline{\mathbf{d}}_i$ p.e. alle varie frequenze: il sistema si aggiorna solo alle frequenze a cui i λ_i sono

16.2 DISPERSIONE DEGLI AUTOVALORI

elevati (*si pensi al proverbio toscano: chi non piange non puppa*). Il sistema risponde solo se *ci si lamenta*. Se ad esempio il disturbo è stato filtrato passa basso, come già visto in precedenza, la matrice \mathbf{R} sarà mal condizionata e qualcuno degli autovalori sarà piccolo. La componente del filtro errore verso quell'autovettore e in sostanza alle frequenze corrispondenti ai minimi dello spettro, si aggiorneranno lentamente, visto che sono poco eccitate dal disturbo \mathbf{d}_i , nullo o quasi in quella banda. Peggio ancora, può accadere che a causa di problemi numerici queste componenti del filtro errore possano crescere oltre misura senza controllo. Il problema non appare grave, se il sistema si mantiene lineare, visto che queste componenti del filtro errore agiscono su componenti di disturbo trascurabili, e quindi il loro effetto è trascurabile. Tuttavia, può accadere che i campioni del filtro (le prese, o *taps* in inglese), crescendo senza controllo nella banda poco controllata, si avvicinino ai valori di trabocco per i coefficienti, rendendo così il problema non lineare e distruggendo la capacità del filtro ad aggiornarsi. Il problema del cosiddetto *tap wandering* ossia "girovagare delle prese" deve essere considerato con cura nel progetto di qualsiasi filtro adattativo.

Per evitarlo, e quindi per fare in modo che tutte le componenti del filtro errore si annullino alla stessa velocità si dovrebbe fare in modo che la matrice \mathbf{R} sia proporzionale alla matrice identica, e cioè che la sequenza $\{d_i\}$ sia un processo a spettro bianco. Quindi, prima di utilizzare la sequenza $\{d_i\}$ per fare la cancellazione, è opportuno sbiancarla. Questo implica stimare il suo spettro; alternativamente, si può utilizzare un filtro a traliccio adattativo per effettuare lo sbiancamento. Un'altra tecnica efficace è quella di decomporre il segnale in bande con un banco di filtri, ad esempio polifase; per ogni banda, in cui si può pensare che lo spettro del segnale sia uniforme, si effettua un filtraggio adattativo, con un coefficiente di aggiornamento diverso.

Le operazioni di sbiancamento sono anch'esse tempo varianti come quelle di sottrazione ottima. E' naturale porsi il problema di accelerare al massimo tali procedure di stima per potere garantire in tutte le condizioni un rapido ritorno ad elevate reiezioni del disturbo. Peraltro, come vedremo nel prossimo paragrafo, accelerare la convergenza implica una maggiore sensibilità al rumore; *se si è troppo sensibili alla moda ...*

16.3 Rapidità della convergenza

In questo paragrafo analizzeremo le conseguenze dell'approssimazione fatta con l'equazione (16.10); in particolare, verrà delineata la necessità di un compromesso nella determinazione del passo γ di aggiornamento del filtro da adattare: infatti vi è sempre una sequenza di segnale $\{s_i\}$, sovrapposta ma indipendente dal disturbo che si vuole cancellare; questo segnale influenza in modo casuale il filtro $\underline{\mathbf{f}}_i$ (in questo paragrafo indichiamo con $\underline{\mathbf{f}}_i$ il vettore casuale e non la sua media) e impedisce che esso raggiunga stabilmente il valore desiderato $\underline{\mathbf{f}}_\infty$ anche per $i \rightarrow \infty$, anche se le statistiche non cambiano (in assenza di fenomeni transitori, cioè). Ancora, si deve tenere conto della casualità del disturbo misurato $\underline{\mathbf{d}}_i$ e quindi dell'aggiornamento del filtro stesso. Tuttavia, se γ fosse molto piccolo, l'aggiornamento sarebbe molto lento e queste due diverse variabili casuali agirebbero solo attraverso la loro media, fatta su un numero grande di campioni; gli effetti di dispersione avranno scarsa influenza, decrescente con γ . A regime quindi, il filtro avrà piccoli scostamenti dal valore asintotico corrispondente alla media di insieme. Invece, se γ non è piccolo, anzi si avvicina al limite teorico, avremo che il filtro $\underline{\mathbf{f}}_i$, **pur convergendo più velocemente in media di insieme** al valore di regime (16.14), **avrà a regime una dispersione maggiore**, per la grande sensibilità ai valori (casuali) sia della sequenza di segnale $\{s_i\}$ che del disturbo $\underline{\mathbf{d}}_i$; per un'idea intuitiva, si ricordino i detti popolari "ondeggiante come una banderuola" (γ grande) e l'altro "inerte come un sasso" (γ piccolo).

In conclusione, si comprende come all'inizio della convergenza il passo debba essere grande; il passo si deve poi ridurre progressivamente fino ad avere un valore modestissimo in condizioni di regime quasi - statico (lontananza dal transitorio iniziale) e convergenza praticamente raggiunta. Naturalmente, se c'è un nuovo cambiamento e si ripropone la necessità di un aggiornamento, il filtro deve riacquisire la sua capacità di adeguarsi e quindi γ deve crescere di nuovo.

16.3.1 Dispersione del filtro errore $\underline{\mathbf{g}}_i$

L'analisi della dispersione del filtro errore porterà per il passo di aggiornamento a risultati più restrittivi di quelli ottenuti con l'analisi dell'evoluzione della media di insieme del filtro errore. Un modo semplice

16.3 RAPIDITÀ DELLA CONVERGENZA

per analizzare il problema è il seguente. Si consideri un sistema monodimensionale in cui, all'insaputa del filtro, il rumore n è assente. Il filtro ottimo ha come media di insieme:

$$E[f] = 0 \quad (16.18)$$

ma l'equazione di aggiornamento è sempre la stessa:

$$y_i = s_i ; \quad f_{i+1} = f_i + \gamma \varepsilon_i d_i ; \quad \varepsilon_i = y_i - f_i d_i ;$$

in cui, al solito, si misura il disturbo d_i per poi valutare il filtro f . Si suppongano s, d indipendenti per cui il filtro asintotico è $f_\infty = 0$, situazione a cui si arriva a partire da una condizione iniziale $f_0 = g_0 \neq 0$. La dinamica effettiva del filtro errore è la seguente:

$$g_{i+1} = g_i + \gamma (s_i - g_i d_i) d_i = g_i (1 - \gamma d_i^2) + \gamma d_i s_i ; \quad (16.19)$$

Ora, a regime, quando il filtro errore è ormai ridotto a zero **in media**, levando gli indici temporali:

$$\sigma_g^2 = E \left[(g (1 - \gamma d^2) + \gamma ds)^2 \right] = E \left[g^2 (1 - \gamma d^2)^2 \right] + \gamma^2 E [d^2 s^2] \quad (16.20)$$

Qualora d sia gaussiana, $E[d^4] = 3\sigma_d^4$ e quindi:

$$\begin{aligned} \sigma_g^2 &= \sigma_g^2 (1 + 3\gamma^2 \sigma_d^4 + 2\gamma \sigma_d^2) + \gamma^2 \sigma_d^2 \sigma_s^2 \\ \sigma_g^2 (2 - 3\gamma \sigma_d^2) &= \gamma \sigma_s^2 \\ \sigma_g^2 &= \frac{\gamma \sigma_s^2}{2 - 3\gamma \sigma_d^2} \sim \frac{\gamma \sigma_s^2}{2} \end{aligned} \quad (16.21)$$

Ad esempio se

$$\gamma \lesssim \frac{1}{6\sigma_d^2} \ll \frac{2}{\sigma_d^2} \rightarrow \sigma_g^2 \lesssim \frac{\gamma \sigma_s^2}{1.5} \quad (16.22)$$

e l'effetto delle sbandate del guadagno (il denominatore che va verso zero) diventa piccolo rispetto al disturbo a regime. A regime infatti, il valore atteso del filtro errore è nullo, ma non il suo valore quadratico medio, che cresce con γ . E' possibile ora fare un'analisi di un modo sistematico con cui si deve ridurre il passo di aggiornamento, a partire dall'equazione (16.19) in cui però approssimiamo d_i^2 con il suo valore atteso:

$$g_{i+1} = g_i (1 - \gamma \sigma_d^2) + \gamma d_i s_i = \rho g_i + \mu_i$$

CAPITOLO 16 FILTRAGGIO ADATTATIVO

Questa è una semplice equazione alle differenze finite che, sovrapponendo gli effetti dell'ingresso μ_i e della condizione iniziale g_0 , variabili casuali indipendenti, ci dà che, se γ è fisso, la dispersione σ_g^2 varierà nel tempo come:

$$\begin{aligned}\sigma_{g,i}^2 &\sim \rho^{2i} \sigma_{g,0}^2 + \frac{\sigma_\mu^2}{1-\rho^2} \sim \sigma_{g,0}^2 e^{-2i\gamma\sigma_d^2} + \frac{\gamma\sigma_s^2}{2}; \\ \rho^{2i} &= e^{2i \log(1-\gamma\sigma_d^2)} \sim e^{-2i\gamma\sigma_d^2} \\ \frac{\sigma_\mu^2}{1-\rho^2} &= \frac{\gamma^2 \sigma_s^2 \sigma_d^2}{(1-\rho)(1+\rho)} \sim \frac{\gamma^2 \sigma_s^2 \sigma_d^2}{\gamma\sigma_d^2(1+\rho)} \sim \frac{\gamma\sigma_s^2}{2}\end{aligned}\quad (16.23)$$

Quindi, per valori crescenti di γ , cresce sia la rapidità di convergenza che l'asintoto. L'involuppo inferiore di queste curve parametrizzate da γ si trova annullando la derivata delle curve rispetto al parametro (l'intersezione tra due curve caratterizzate da valori molto vicini del parametro) e quindi eliminandolo.

$$\frac{\partial \sigma_{g,i}^2}{\partial \gamma} = -2i\sigma_d^2 \sigma_{g,0}^2 e^{-2i\gamma\sigma_d^2} + \frac{\sigma_s^2}{2} = 0$$

Ora γ è funzione dell'indice i :

$$\gamma(i) = \frac{1}{2\sigma_d^2 i} \ln \left(\frac{\sigma_d^2}{\sigma_s^2} 4\sigma_{g,0}^2 i \right)$$

e il risultato ottimizzato è quindi:

$$\sigma_{g,i}^2 \sim \frac{1}{i} \frac{\sigma_s^2}{4\sigma_d^2} \left[1 + \ln \left(4i\sigma_{g,0}^2 \frac{\sigma_d^2}{\sigma_s^2} \right) \right];$$

Si comprende come, all'inizio dell'adattamento, γ debba essere vicino al limite massimo ma poi debba scendere a valori molto più bassi non appena il sistema si avvicina all'equilibrio, proprio per ridurre σ_g^2 . Inoltre, per avere stabilità del filtro errore oltre che in media (vedi 16.14) anche in dispersione, il passo γ deve essere ben più piccolo del valore trovato in base all'equazione di evoluzione delle medie. Infine, se il sistema parte da condizioni iniziali ignote, ma poi è stazionario, il valore quadratico medio dell'errore con cui viene determinato il filtro deve potersi ridurre con il tempo in modo approssimativamente iperbolico.

Aumentando il livello di complessità e passando ad un'analisi multidimensionale ($N > 1$), possiamo quindi analizzare il filtro errore $\underline{\mathbf{g}}_i$

16.3 RAPIDITÀ DELLA CONVERGENZA

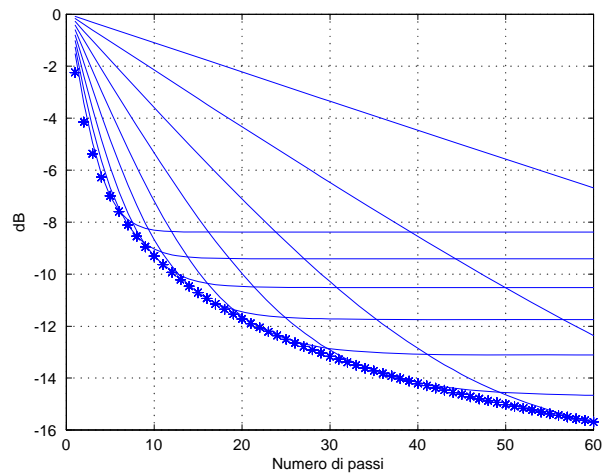


Figura 16.2: Dispersione del filtro errore in funzione del tempo, al variare della lunghezza del passo γ . Al crescere di γ , la convergenza è più rapida, ma l'asintoto è più elevato. La curva involucro, ottenuta variando il passo nel tempo, è rappresentata con asterischi.

CAPITOLO 16 FILTRAGGIO ADATTATIVO

(e cioè la differenza tra il filtro effettivo $\underline{\mathbf{f}}_i$ (vettore casuale) ed il filtro desiderato a regime, $\underline{\mathbf{f}}_\infty$, vettore deterministico) come un vettore casuale di cui si cerca media e varianza, al variare del tempo. La situazione non è stazionaria, in quanto la media (di insieme) del filtro errore convergerà a 0, sempre se $\gamma < \gamma_{max}$. Non così la sua dispersione statistica, come vedremo nel seguito. Anzitutto, riscriviamo l'equazione di aggiornamento del filtro $\underline{\mathbf{f}}_i$, visto come un vettore casuale:

$$\underline{\mathbf{f}}_{i+1} = \underline{\mathbf{f}}_i + \gamma \varepsilon_i^* \underline{\mathbf{d}}_i = \underline{\mathbf{f}}_i + \gamma (y_i^* - \underline{\mathbf{d}}_i^* \underline{\mathbf{f}}_i) \underline{\mathbf{d}}_i \quad (16.24)$$

L'errore di stima del disturbo ε_i è somma di due termini, l'uno dovuto al segnale sovrapposto s_i e l'altro al disturbo residuo dovuto all'imperfetto filtraggio, perchè il filtro errore $\underline{\mathbf{g}}_i$, anche esso casuale, non è nullo:

$$\begin{aligned} \varepsilon_i &= s_i + n_i - \hat{n}_i = y_i - \underline{\mathbf{d}}_i^* \underline{\mathbf{f}}_i = y_i - \underline{\mathbf{d}}_i^* (\underline{\mathbf{g}}_i + \underline{\mathbf{f}}_\infty) = \\ &= y_i - \underline{\mathbf{d}}_i^* \underline{\mathbf{f}}_\infty - \underline{\mathbf{d}}_i^* \underline{\mathbf{g}}_i \simeq s_i - \underline{\mathbf{d}}_i^* \underline{\mathbf{g}}_i \end{aligned} \quad (16.25)$$

sempre che si supponga $n_i \simeq \underline{\mathbf{d}}_i^* \underline{\mathbf{f}}_\infty$ e cioè che il filtro a regime $\underline{\mathbf{f}}_\infty$ sia efficace. Per ottenere dei risultati semplici, dobbiamo fare delle ipotesi semplificative, che tuttavia non oscurano la fisica del problema. In particolare supponiamo che, da un istante al successivo, i vettori $\underline{\mathbf{d}}_i$ siano indipendenti tra loro e dal segnale s_i .

$$E[\underline{\mathbf{d}}_i \underline{\mathbf{d}}_j^*] = \underline{\mathbf{0}}; \quad i \neq j$$

Inoltre, per semplificare le formule, supporremo che le componenti di $\underline{\mathbf{d}}_i$ siano indipendenti tra loro e con eguale varianza e pertanto che la matrice $\underline{\mathbf{R}}$ sia proporzionale alla matrice identica:

$$\underline{\mathbf{R}} = E[\underline{\mathbf{d}}_i \underline{\mathbf{d}}_i^*] = \sigma_d^2 \underline{\mathbf{I}}$$

Calcoleremo la media di insieme del vettore $\underline{\mathbf{g}}_{i+1}$,

$$\underline{\mathbf{g}}_{i+1} = (\underline{\mathbf{I}} - \gamma \underline{\mathbf{d}}_i \underline{\mathbf{d}}_i^*) \underline{\mathbf{g}}_i + \gamma \underline{\mathbf{d}}_i s_i \quad (16.26)$$

tenendo conto dell'indipendenza tra il segnale s_i ed il disturbo $\underline{\mathbf{d}}_i$.

Prima però di passare a qualsiasi operazione di media, osserviamo che la matrice casuale $\underline{\mathbf{d}}_i \underline{\mathbf{d}}_i^*$ o una sua media mobile a pesi positivi $\sum_i p_i \underline{\mathbf{d}}_i \underline{\mathbf{d}}_i^*$ è sempre semidefinita positiva in quanto per qualsiasi vettore $\underline{\mathbf{a}}$, $\underline{\mathbf{d}}_i$, si ha

$$\sum p_i (\underline{\mathbf{a}}^* \underline{\mathbf{d}}_i \underline{\mathbf{d}}_i^* \underline{\mathbf{a}}) = \sum p_i (\underline{\mathbf{d}}_i^* \underline{\mathbf{a}})^2 \geq 0 \quad \text{se } p_i \geq 0$$

16.3 RAPIDITÀ DELLA CONVERGENZA

Questo è un fatto importante perchè non permette il cambio di segno del γ nella proiezione su qualsiasi autovettore della matrice $\underline{\mathbf{d}}_i \underline{\mathbf{d}}_i^*$ e quindi possibili instabilità. Passando alle medie di insieme:

$$\begin{aligned} E[\underline{\mathbf{g}}_{i+1}] &= E[(\underline{\mathbf{I}} - \gamma \underline{\mathbf{d}}_i \underline{\mathbf{d}}_i^*) \underline{\mathbf{g}}_i] + E[\gamma \underline{\mathbf{d}}_i s_i] = \\ &= (\underline{\mathbf{I}} - \gamma \underline{\mathbf{R}}) E[\underline{\mathbf{g}}_i] = (\underline{\mathbf{I}} - \gamma \underline{\mathbf{R}})^{i+1} \underline{\mathbf{g}}_0 \end{aligned} \quad (16.27)$$

In quest'ultima equazione abbiamo utilizzato l'ipotesi che ogni nuovo valore di $\underline{\mathbf{d}}_i$ è indipendente dai precedenti e quindi anche indipendente da $\underline{\mathbf{g}}_i$. Quindi la media del loro prodotto è il prodotto delle medie. Ancora, si è supposto che il valore iniziale del filtro errore $\underline{\mathbf{g}}_0$ non sia una variabile casuale; esso è peraltro sconosciuto.

La **media di insieme** del filtro errore pertanto si evolve e converge a zero secondo l'equazione determinata precedentemente (16.14) utilizzando l'ipotesi di lenta variazione del filtro (γ piccolo) e stazionarietà delle statistiche del disturbo. Calcoliamo ora il valore quadratico medio del filtro errore; di nuovo, per evitare complicazioni ulteriori, calcoliamo questo v.q.m. a regime ($i = \infty$), e cioè quando la media di insieme del filtro errore ha raggiunto il valore nullo, desiderato. La varianza a regime pertanto corrisponde al valore quadratico medio delle componenti, a regime. Utilizziamo quindi la formula (16.26) di cui calcoliamo il valore quadratico medio di entrambi i membri, utilizzando l'indipendenza tra $\underline{\mathbf{d}}_i$, $\underline{\mathbf{g}}_i$, s_i e tralasciando l'indice ($i = \infty$), per snellire. Inoltre, riscriviamo l'equazione di aggiornamento **non sostituendo** immediatamente all'espressione $\underline{\mathbf{d}}_i \underline{\mathbf{d}}_i^*$ la sua media di insieme e cioè la matrice $\underline{\mathbf{R}}$, proprio per tenere conto della casualità dell'aggiornamento:

$$\begin{aligned} E[\underline{\mathbf{g}}^* \underline{\mathbf{g}}] &= E[\underline{\mathbf{g}}^* (\underline{\mathbf{I}} - \gamma \underline{\mathbf{d}} \underline{\mathbf{d}}^*) (\underline{\mathbf{I}} - \gamma \underline{\mathbf{d}} \underline{\mathbf{d}}^*) \underline{\mathbf{g}}] + E[\gamma^2 s^2 \underline{\mathbf{d}}^* \underline{\mathbf{d}}] = \\ &= E[\underline{\mathbf{g}}^* (\underline{\mathbf{I}} - 2\gamma \underline{\mathbf{d}} \underline{\mathbf{d}}^* + \gamma^2 \underline{\mathbf{d}} \underline{\mathbf{d}}^* \underline{\mathbf{d}} \underline{\mathbf{d}}^*) \underline{\mathbf{g}}] + E[\gamma^2 s^2 \underline{\mathbf{d}}^* \underline{\mathbf{d}}] \end{aligned} \quad (16.28)$$

e quindi:

$$N\sigma_g^2 = N\sigma_g^2(1 - 2\gamma\sigma_d^2 + \gamma^2(N + K - 1)\sigma_d^4) + \gamma^2 N\sigma_d^2\sigma_s^2$$

Infatti,

$$E[\underline{\mathbf{d}} \underline{\mathbf{d}}^*] = \sigma_d^2 \underline{\mathbf{I}}$$

CAPITOLO 16 FILTRAGGIO ADATTATIVO

e, per l'indipendenza tra $\underline{\mathbf{d}}_i, \underline{\mathbf{g}}_i$:

$$E[\underline{\mathbf{g}}^* \underline{\mathbf{d}} \underline{\mathbf{d}}^* \underline{\mathbf{g}}] = E[\underline{\mathbf{g}}^* E[\underline{\mathbf{d}} \underline{\mathbf{d}}^*] \underline{\mathbf{g}}] = N \sigma_g^2 \sigma_d^2$$

$$E[\underline{\mathbf{d}} \underline{\mathbf{d}}^* \underline{\mathbf{d}} \underline{\mathbf{d}}^*] = E[\underline{\mathbf{d}}^* \underline{\mathbf{d}} \times \underline{\mathbf{d}} \underline{\mathbf{d}}^*] = (N + K - 1) \sigma_d^4 \mathbf{I}$$

il k -esimo elemento della diagonale principale dell'ultima matrice è eguale a:

$$E\left[\sum_{i=1}^N d_i^2 \times d_k^2\right] = (N - 1) \sigma_d^4 + K \sigma_d^4$$

Abbiamo indicato con K il rapporto tra il valore medio della quarta potenza e il quadrato della varianza:

$$K = \frac{E[d_k^4]}{\sigma_d^4} \quad (16.29)$$

K è la **curtosi** di una variabile casuale a valor medio nullo e per variabili gaussiane $K = 3$. Quindi, in condizioni di regime, lontano dai transitori, quando la varianza σ_g^2 del filtro errore è dovuta **solo** all'effetto della varianza σ_s^2 del segnale $\{s_i\}$ che provoca aggiornamenti casuali del filtro, si ha:

$$\sigma_g^2 = \frac{\gamma \sigma_s^2}{2 - \gamma(N + K - 1) \sigma_d^2} \quad (16.30)$$

Se γ è piccolo,

$$\sigma_g^2 \simeq \frac{\gamma \sigma_s^2}{2} \quad (16.31)$$

In sostanza, a regime, la varianza del filtro errore cresce con la varianza del segnale σ_s e con il passo di aggiornamento. Quindi, una volta raggiunta in pratica la convergenza, γ va ridotto. Inoltre, la dispersione del filtro errore può essere illimitata ($\sigma_g^2 \rightarrow \infty$), se non si verifica che:

$$\gamma \leq \frac{2}{(N + K - 1) \sigma_d^2} < \frac{2}{N \sigma_d^2}$$

In sostanza, ci sono valori del passo di aggiornamento γ per cui il filtro errore convergerebbe (in media) a zero, ma poi ha una dispersione enorme.

16.4 Ottimizzazione del passo di aggiornamento

Riassumendo, se il ritmo di aggiornamento del filtro è veloce ($\gamma \leq \gamma_{max}$), i parametri del filtro si aggiornano rapidamente, ma poi, a regime, i parametri sono affetti da disturbo. Se invece il filtro si aggiorna *lentamente*, a regime sarà poco disturbato. Ne risulta inoltre un altro limite per il passo: infatti, come si è visto:

$$\gamma < \gamma_{max} = \frac{2}{(N + K - 1)\sigma_d^2} \quad (16.32)$$

Se $N \gg 1$, ci si potrebbe avvicinare alla condizione di massima velocità di aggiornamento del filtro media di insieme ($\gamma = \gamma_{max}$); però, la dispersione a regime del filtro errore tenderebbe ad infinito a causa dell'inevitabile presenza del segnale ($\sigma_s^2 \neq 0$) che impedisce al filtro di inseguire solo il disturbo. Se $\gamma \ll \gamma_{max}$, il filtro si aggiorna molto lentamente ma poi a regime la dispersione è minore perché si è meno sensibili all'influenza del segnale casuale mediata e rimossa a causa del limitatissimo passo di aggiornamento.

Pertanto, all'inizio dell'aggiornamento, quando si è lontani dalla situazione di regime (il filtro errore ha valori grandi), sarà opportuno porre $\gamma \simeq \gamma_{max}/2$ per accelerare la convergenza anche a costo di un'elevata dispersione a regime; appena il filtro errore si è ridotto, bisogna ridurre γ per evitare inutili dispersioni dei coefficienti del filtro attorno ad un valore medio vicino a valori già buoni. Gergalmente, in inglese, questo modo di operare è chiamato *gear shifting: scalare le marce*; l'aggiornamento deve essere veloce all'inizio del transitorio e lento alla fine, quando si è prossimi alla situazione asintotica. Si osservi che nei periodi iniziali del transitorio (dopo una variazione delle statistiche), ci si attende un cambiamento celere (*rivoluzione*). La media del filtro errore è grande e c'è grande differenza tra il comportamento atteso e quello effettivo del filtro. Sono possibili variazioni repentine del filtro, apparentemente casuali ($\gamma \simeq \gamma_{max} \rightarrow \sigma_g^2 \simeq \infty$). Nelle situazioni lontane dai transitori (*conservazione*) il sistema si adegua ordinatamente ma forse troppo lentamente ($\gamma \ll \gamma_{max} \rightarrow \sigma_g^2 \simeq 0$).

Se poi le statistiche del sistema cambiano continuamente, ci sarà un valore ottimale di γ dipendente sia dalla velocità di variazione delle statistiche che dall'ampiezza del rumore. In generale, il valore ottimo per γ

CAPITOLO 16 FILTRAGGIO ADATTATIVO

sarà quello che minimizza il valore atteso (quadratico medio) del filtro errore: *se il traffico è irregolare, è bene rimanere in seconda ...*

Si osservi che per $N = 1$ la situazione è peggiore, e si deve mantenere $\gamma \ll \gamma_{\max}/3 < \gamma_{\max}$, nel caso di variabili gaussiane. Infatti, sono possibili valori istantanei elevati del disturbo che mandano il sistema in instabilità. Non appena N cresce e l'aggiornamento è basato su stime fatte su un numero maggiore di variabili casuali indipendenti, c'è minore dispersione della stima, la situazione si stabilizza immediatamente, e γ può essere accresciuto. Questo effetto di instabilità dipende dalla possibilità di avere forti picchi istantanei in qualche componente $d_{i,n}$ di $\underline{\mathbf{d}}_i$; si ricordi che il fattore $(N + K - 1)/N$ per le variabili gaussiane vale $(N + 2)/N$ e vale 3 per $N = 1$; la situazione peggiora ancora con distribuzioni con curtosi più elevata che non la gaussiana (e cioè se vi è probabilità significativa di avere picchi elevati); infatti, sono possibili densità di probabilità con varianza finita e curtosi illimitata. La situazione invece migliora per variabili casuali *più tranquille* e cioè caratterizzate da distribuzioni con code limitate o addirittura assenti come accade per la distribuzione rettangolare per la quale $K = 9/5 < 3$; il valore minimo possibile per la curtosi, $K = 1$, è raggiunto solo nel caso di una variabile casuale che assume solo i valori $\pm q$. In questo caso e solo in questo caso si può mantenere $\gamma \leq 1$ anche per $N = 1$; in altre parole, il segnale di pilotaggio è così tranquillo che si può rimanere sempre *con le marce alte*; quest'ultimo non è un caso irrilevante in quanto corrisponde ad un segnale di comunicazioni del tipo BPSK.

16.4.1 *Genesi e collasso delle società idrauliche*

Un interessante caso di evoluzione sociale, congrua con quanto detto per i sistemi adattativi, è rappresentato dalle cosiddette società idrauliche (K. Wittfogel, Oriental Despotism: A Comparative Study of Total Power). Queste società si sviluppano attorno a gigantesche opere di canalizzazione per irrigazione, ed a volte collassano. Ve ne sono numerosi esempi in Eurasia (Cina del Nord, Egitto antico, Mesopotamia) ed in America (Maya).

Queste società si adattano sempre meglio al loro ambiente, costruendo dei sistemi di irrigazione indispensabili per nutrire l'accresciuta popolazione, ma fragilissimi. Per garantire che ognuno svolga il lavoro assegnatogli, si rende necessaria una burocrazia sempre più pesante e potente.

16.5 UN ESEMPIO DI FILTRO ADATTATIVO

Infatti, anche una modesta mancanza di qualcuno (p.e. l'ostruzione anche parziale di un canale irriguo) potrebbe poi causare danni gravissimi. Gli onnipotenti burocrati, per i quali la principale virtù è l'obbedienza, diventano nemici di tutte le innovazioni, che sono viste come un rischio per lo status quo e per l'equilibrio (γ piccolissimo).

Enormi ed inutili lavori pubblici, come le piramidi, i tipici monumenti di queste civiltà, hanno anche lo scopo di mantenere coesa la società. Imprevedibili siccità possono però provocare dei collassi globali (J. Diamond: Collapse: How Societies Choose to Fail or Survive) mentre invece altre società più primitive, ma più resilienti, possono passare indenni attraverso queste crisi climatiche.

16.5 Un esempio di filtro adattativo

Una sequenza disturbante x_n è ottenuta filtrando un processo bianco u_n con un filtro FIR

$$x_n = u_n + 2u_{n-1} + u_{n-2} = \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} u_n \\ u_{n-1} \\ u_{n-2} \end{bmatrix} = \underline{\mathbf{a}}^* \underline{\mathbf{u}}$$

Il segnale desiderato s_n è osservato dopo la somma del disturbo x_n :

$$y_n = s_n + x_n$$

Le sequenze u_n, s_n sono bianche, indipendenti tra loro e gaussiane con valore medio nullo e varianza unitaria. Ricostruiamo con il filtro adattativo $\underline{\mathbf{f}}_n$ i valori della sequenza s_n a partire dai valori del vettore $\underline{\mathbf{d}}_n$:

$$\underline{\mathbf{d}}_n = \begin{bmatrix} x_{n+1} \\ x_n \\ x_{n-1} \end{bmatrix} + \begin{bmatrix} w_{n+1} \\ w_n \\ w_{n-1} \end{bmatrix}$$

$$\hat{x}_n = \underline{\mathbf{f}}_n^* \underline{\mathbf{d}}_n$$

mediante la relazione iterativa:

$$\underline{\mathbf{f}}_{n+1} = \underline{\mathbf{f}}_n + \gamma [y_n - \hat{x}_n] \underline{\mathbf{z}}_n$$

CAPITOLO 16 FILTRAGGIO ADATTATIVO

La matrice di covarianza del vettore dei dati $\underline{\mathbf{z}}$ e la correlazione tra i dati ed lo scalare desiderato y_n .

$$\underline{\mathbf{R}} = E[\underline{\mathbf{z}}\underline{\mathbf{z}}^*] = \begin{bmatrix} 6 & 4 & 1 \\ 4 & 6 & 4 \\ 1 & 4 & 6 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 7 & 4 & 1 \\ 4 & 7 & 4 \\ 1 & 4 & 7 \end{bmatrix} \quad (16.33)$$

$$\underline{\mathbf{r}} = E[y_n \underline{\mathbf{z}}^*] = \begin{bmatrix} 4 \\ 6 \\ 4 \end{bmatrix}; \quad \underline{\mathbf{f}}_\infty = \underline{\mathbf{R}}^{-1} \underline{\mathbf{r}} = \begin{bmatrix} 0.16667 \\ 0.66667 \\ 0.16667 \end{bmatrix} \quad (16.34)$$

:Gli autovalori e gli autovettori di $\underline{\mathbf{R}}$ sono;

$$\begin{bmatrix} 1.82 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 13.17 \end{bmatrix}; \begin{bmatrix} -0.4775 & -0.7 & 0.5215 \\ 0.7376 & 0 & 0.6753 \\ -0.4775 & 0.7 & 0.5215 \end{bmatrix} \quad (16.35)$$

Si osservi come l'autovalore massimo corrisponde all'autovettore passa basso (la terza colonna della matrice di destra). Si ha che assolutamente:

$$\gamma < \frac{2}{21} = 0.09 \quad (16.36)$$

Supponendo di partire dal filtro $\underline{\mathbf{f}}_0 = \underline{\mathbf{0}}$, le componenti lungo i 3 autovettori del filtro errore $\underline{\mathbf{g}}_0$ sono

$$- \begin{bmatrix} -0.4775 & -0.7 & 0.5215 \\ 0.7376 & 0 & 0.6753 \\ -0.4775 & 0.7 & 0.5215 \end{bmatrix}^T \begin{bmatrix} 0.16667 \\ 0.66667 \\ 0.16667 \end{bmatrix} = - \begin{bmatrix} 0.33257 \\ 0.0 \\ 0.62404 \end{bmatrix} \quad (16.37)$$

Queste convergeranno a zero con la legge

$$\left(1 - \alpha \frac{2}{21} \lambda_i\right)^n \rightarrow \quad (16.38)$$

16.5 UN ESEMPIO DI FILTRO ADATTATIVO

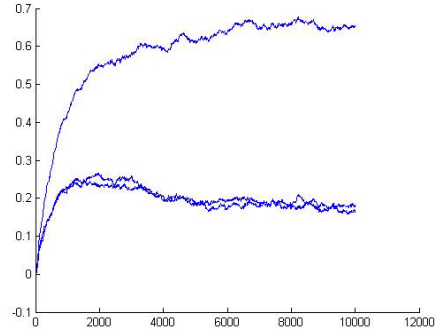


Figura 16.3: Convergenza al valore asintotico delle 3 componenti del filtro $\underline{\mathbf{f}}_n$. Il valore di γ è piccolo, e l'oscillazione a regime piuttosto ridotta. $\alpha = 0.002$

Prendendo $\alpha = 0.5$ sembrerebbe che

$$\left(1 - \frac{1}{21}13.82\right)^n \rightarrow 0.91^n; 0.71^n, 0.34^n \quad (16.39)$$

In realtà si deve prendere un α molto minore ($\alpha = 0.002; 0.2$) per garantire la stabilità (vedi figure). A transitorio esaurito la potenza dell'errore è, per l'incorrelazione tra errore di stima e dati :

$$E[(y_n - \hat{x}_n)^2] = E[(y_n - \hat{x}_n)y_n] = E[(s_n + x_n - \underline{\mathbf{f}}_\infty^* \underline{\mathbf{d}}_n)(s_n + x_n)] = \quad (16.40)$$

$$= \sigma_s^2 + \sigma_x^2 - \underline{\mathbf{r}}^* \underline{\mathbf{f}} = 1.66 \quad (16.41)$$

CAPITOLO 16 FILTRAGGIO ADATTATIVO

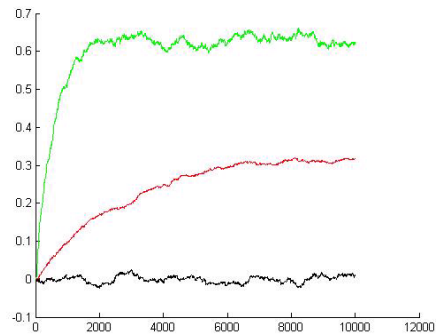


Figura 16.4: Convergenza al valore asintotico delle 3 componenti del filtro errore $\underline{\mathbf{g}}_n$ lungo i 3 autovettori della matrice \mathbf{R} . Il valore di γ è piccolo, e l'oscillazione a regime piuttosto ridotta. Si osservi la diversa velocità di convergenza. $\alpha = 0.002$

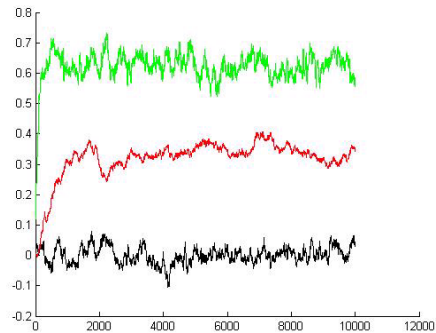


Figura 16.5: Convergenza al valore asintotico delle 3 componenti del filtro $\underline{\mathbf{f}}_n$. Il valore di γ è molto maggiore, e l'oscillazione a regime notevole. La densità di probabilità della sequenza $\{w_n\}$ è uniforme, a pari varianza. $\alpha = 0.008$

Capitolo 17

Le direzioni di arrivo

Fino ad ora abbiamo visto l'analisi spettrale di sequenze in cui l'indice rappresenta l'asse temporale; studieremo ora situazioni in cui l'indice si riferisce a posizioni nello spazio. In particolare vogliamo affrontare il problema della misura delle Direzioni di Arrivo su una schiera di sensori (Directions of Arrival: DOA) di segnali temporalmente monofrequenziali (o a spettro molto limitato).

Si disponga di un cavo lungo il quale sono regolarmente disposti dei sensori e siano $\{x_i\}$, $i = 1, \dots, N+1$, i segnali da essi misurati. La distanza tra il sensore i -esimo e l'inizio del cavo sia $d_i = (i - 1)\Delta$; supponiamo che esistano M sorgenti di radiazioni, monofrequenziali o a banda stretta, a grande distanza dal cavo; l'angolo tra la direzione della sorgente m e la normale al cavo (vedi fig. 17.1) è indicato con θ_m ; quindi la radiazione arriva al ricevitore $\{x_i\}$ con il ritardo rispetto al ricevitore $i = 1$ pari a $\tau_i = (i - 1) \times \Delta \sin \theta_m / c$. Inoltre, indichiamo con $A_m e^{j\omega n T}$ l'ampiezza sinusoidale della radiazione e la sua evoluzione nel tempo, e con c la velocità di propagazione delle onde. Il segnale misurato dal sensore i all'istante n è (per semplicità abbiamo posto $M = 1$):

$$x_{i,n} = A \exp(j\omega n T + j \frac{(i - 1)\omega \Delta \sin \theta_m}{c}) = A e^{j\omega n T} \times e^{jk_m(i-1)\Delta}; \quad (17.1)$$

$$k_m = \frac{\omega \sin \theta_m}{c}$$

Il segnale misurato nella stessa posizione al variare del tempo è una sinusoidale complessa di pulsazione ω ; il segnale misurato allo stesso istante

CAPITOLO 17 LE DIREZIONI DI ARRIVO

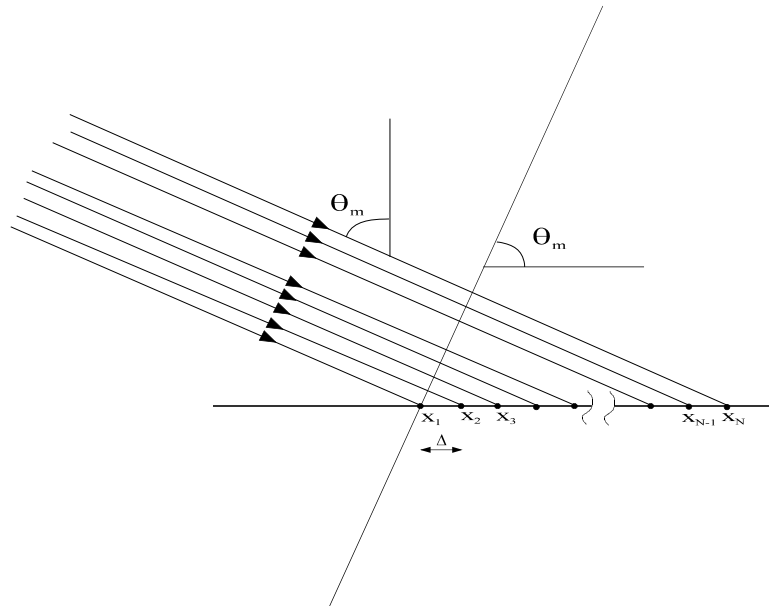


Figura 17.1: Schiera di sensori

nT da tutta la schiera di sensori (al variare della posizione nello spazio) è una **sinusoide complessa**, nel dominio degli spazi, avente numero d'onda k_m (per le sinusoidi spaziali, è la grandezza corrispondente alla pulsazione delle sinusoidi temporali). Misurare la direzione di arrivo equivale a misurare il numero d'onda della sinusoide complessa registrata.

Fare un'analisi spettrale della sequenza $\{x_i\}$, per un dato istante di tempo n , corrisponde quindi a determinare numero e direzione di arrivo delle radiazioni incidenti sul cavo. Si osservi che se la radiazioni in arrivo non sono monofrequenziali ma la loro banda è sufficientemente ampia, è possibile farne una decomposizione di Fourier lungo l'asse dei tempi, sensore per sensore, e poi analizzare ogni singola componente isofrequenziale nel dominio degli spazi. In questo modo si avranno più misure (alle varie frequenze temporali) della stessa direzione di arrivo, misure che potranno essere opportunamente combinate.

Si osservi ancora che in questo caso, l'ipotesi di avere segnali sinusoidali (nello spazio) è sostanzialmente corretta. Le tecniche che utilizzeremo per la misura della frequenza ed ampiezza di queste sinusoidi saranno quella del periodogramma e quella AR. Ve ne sono altre (Pisarenko, etc.)

17.1 DOA CON IL PERIODOGRAMMA

che non verranno discusse.

La distanza massima tra i sensori per evitare equivocazione spaziale è:

$$|k| = \frac{\omega \sin \theta}{c} \leq k_{Nyquist} = \frac{\pi}{\Delta}; \quad \Delta \leq \frac{c}{2f} = \frac{\lambda}{2} \quad (17.2)$$

Per esempio, con segnali a $20kHz$, se $c = 1500m/sec$, allora $\Delta \leq 0.0375m$. Scegliendo questo valore per Δ avremo che alla direzione $\theta = 0$ corrisponde $k = 0$; alle direzioni $\theta = \pm\pi/2$ corrisponde $k = \pm k_{Nyquist}$

17.1 DOA con il periodogramma

Determineremo ora l'accuratezza della misura delle direzioni di arrivo, utilizzando per l'analisi spettrale la tecnica del periodogramma. In questo caso, indicando con $L_{tot} = N\Delta$ la lunghezza del cavo la risoluzione spettrale è

$$\delta k = \frac{2\pi}{L_{tot}} \quad (17.3)$$

A questa incertezza di numero d'onda Δk corrisponde un'incertezza angolare $\delta\theta$:

$$\delta k = \frac{dk}{d\theta} \delta\theta = \frac{\omega \cos \theta}{c} \delta\theta = \frac{2\pi \cos \theta}{\lambda} \delta\theta \quad (17.4)$$

$$\delta\theta = \frac{\lambda}{2\pi \cos \theta} \frac{2\pi}{L_{tot}} = \frac{\lambda}{L_{tot} \cos \theta} \quad (17.5)$$

Questa è la direttività di un'antenna lunga $L_{tot} \cos \theta$ pari alla proiezione della lunghezza del cavo nella direzione ortogonale alla radiazione. In sostanza, utilizzare il periodogramma per fare l'analisi spettrale coincide con il combinare assieme i segnali in un'antenna sintetica lunga quanto il cavo. Si osservi che l'uso di questa tecnica corrisponde all'ipotesi che radiazioni provenienti da direzioni diverse (potenze in celle disgiunte del periodogramma) siano tra loro indipendenti. Se l'analisi fosse bidimensionale, questa tecnica produrrebbe quindi un'immagine dell'energia illuminante suddivisa per direzioni angolari di arrivo: come un telescopio, pertanto. Infatti, una lente sottile (o uno specchio ricurvo opportunamente) funziona come un trasformatore di Fourier (a 2 dimensioni): l'immagine nel piano focale è la trasformata di Fourier del segnale in arrivo

CAPITOLO 17 LE DIREZIONI DI ARRIVO

sulla lente. Quindi, per leggere un titolo di giornale (5cm di risoluzione) da 300 km di distanza serve un telescopio di diametro almeno eguale a:

$$\frac{0.05}{300000} = \frac{\lambda}{D}; \text{ Per il colore verde: } \lambda = 0.5\mu\text{m} \rightarrow D = 3\text{m}$$

Ma la risoluzione può crescere di molto se si possono imporre ipotesi stringenti sul numero di sorgenti, come vedremo nella prossima sezione.

17.2 Stima AR: la risoluzione angolare

L'analisi spettrale che si adatta bene al caso in cui si ha un numero limitato di sorgenti lontane è quella parametrica, tutti poli. Infatti il segnale in arrivo è caratterizzato da un numero discreto M di fronti d'onda, ognuno con la sua ben definita direzione di arrivo. La sequenza $\{x_i\}$ sarà somma di M sinusoidi complesse corrispondenti agli M fronti. Le loro pulsazioni spaziali (i numeri d'onda) sono:

$$k_m = \frac{\omega \sin \theta_m}{c}; \quad m = 1, \dots, M \quad (17.6)$$

Supporremo $M = 1$, che l'ampiezza della radiazione A sia una variabile casuale complessa di varianza $E[AA^*] = \sigma_A^2$ e che sui dati sia sovrapposto del rumore bianco w_i , complesso, di varianza σ_w^2 . Non riportiamo il fattore complesso $e^{j\omega n T}$ avente modulo unitario, comune a tutti i dati. Ponendo quindi:

$$x_i = A e^{jk_1(i-1)\Delta} + w_i$$

si vede che l'analisi spettrale AR si presta bene, anche se poi non esiste una vera e propria corrispondenza fisica diretta e cioè non siamo nel caso di un segnale generato da un rumore bianco che pilota un filtro IIR.

Per determinare la direzione della sorgente con l'analisi spettrale AR abbiamo bisogno dei campioni della funzione di autocorrelazione spaziale della sequenza $\{x_i\}$ che sono:

$$E[x_i x_{i-n}^*] = r_n = r_{-n}^* = \sigma_A^2 e^{jk_1 n \Delta} + \sigma_w^2 \delta(n); \quad 0 \leq n \leq N \quad (17.7)$$

Si osservi che possiamo stimare con precisione i campioni della funzione di autocorrelazione facendo delle operazioni di media **nel tempo**; non abbiamo necessità di fare delle medie nello spazio.

17.2.1 2 sensori, una sorgente

Nel caso semplice di due sensori, e quindi una stima AR di ordine 1, il polo dello spettro stimato è in

$$z_1 = \frac{e^{jk_1\Delta}}{1 + \sigma_w^2/\sigma_A^2} = (1 - \varepsilon) e^{jk_1\Delta}; \quad \varepsilon \sim \frac{\sigma_w^2}{\sigma_A^2}; \quad \delta k = \frac{2\varepsilon}{\Delta} \quad (17.8)$$

La risoluzione angolare è ben migliore che nel caso del periodogramma fatto con la stessa schiera:

$$\delta k = \frac{2\pi \cos \theta}{\lambda} \delta \theta = \frac{2\sigma_w^2}{\Delta \sigma_A^2} \rightarrow \delta \theta \sim \frac{\lambda}{\Delta \cos \theta} \times \frac{1}{\pi} \times \frac{\sigma_w^2}{\sigma_A^2} \quad (17.9)$$

Per $N = 1$, si migliora il risultato del periodogramma di un fattore $\sigma_w^2/(\pi\sigma_A^2)$.

17.2.2 N+1 sensori, una sorgente

Nel caso di $N + 1$ sensori ed una sorgente, utilizzando le equazioni di Yule Walker, si osserva che la matrice di covarianza $\underline{\mathbf{R}}$ è Hermitiana (la trasposta coincide con la coniugata) ed è in questo caso rappresentabile come combinazione lineare di una matrice identica e di una diade, prodotto esterno di due vettori complessi. Si verifica che la matrice inversa è ancora combinazione lineare della matrice identica e della stessa diade. Nel caso più generale $M > 1$, e cioè con M sorgenti indipendenti, la matrice $\underline{\mathbf{R}}$ sarebbe somma di M diadi e la soluzione analitica del problema richiederebbe la soluzione di una matrice $M \times M$. Ovviamente, il problema delle direzioni di arrivo è solubile solo se $M \leq N$, ed in generale $M \ll N$. Lo spettro AR è determinato dalle relazioni di Yule Walker. Per $M = 1$ la soluzione analitica è semplice; si osservi che con poco rumore, l'ordine N del sistema necessario per misurarne la frequenza potrebbe essere limitato ad 1, (bastano due sensori, quindi) in quanto una sinusoidale complessa corrisponde ad un sistema ad un solo polo. Tuttavia, la presenza del disturbo $\{w_i\}$ fa sì che l'errore di predizione decresca progressivamente con N , come vedremo, rendendo utili schiere di più sensori. Sia quindi al solito:

$$\underline{\mathbf{R}}\mathbf{a}_N = -\mathbf{r}_N$$

CAPITOLO 17 LE DIREZIONI DI ARRIVO

$$\underline{\mathbf{R}} = \begin{bmatrix} r_0 & r_{-1} & r_{-2} & \dots & r_{-N+1} \\ r_1 & r_0 & r_{-1} & \dots & r_{-N+2} \\ r_2 & r_1 & r_0 & \dots & r_{-N+3} \\ \dots & \dots & \dots & \dots & \dots \\ r_{N-1} & r_{N-2} & r_{N-3} & \dots & r_0 \end{bmatrix} = \sigma_A^2 \underline{\mathbf{c}} \underline{\mathbf{c}}^* + \sigma_w^2 \underline{\mathbf{I}}; \quad (17.10)$$

$$\underline{\mathbf{c}} = \begin{bmatrix} e^{jk_1\Delta} \\ e^{j2k_1\Delta} \\ e^{j3k_1\Delta} \\ \dots \\ e^{jNk_1\Delta} \end{bmatrix}; \quad \underline{\mathbf{r}}_N = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ \dots \\ r_N \end{bmatrix} = \sigma_A^2 \underline{\mathbf{c}} \quad (17.11)$$

$$\underline{\mathbf{R}}^{-1} = \frac{\underline{\mathbf{I}}}{\sigma_w^2} - \frac{\sigma_A^2 \underline{\mathbf{c}} \underline{\mathbf{c}}^*}{\sigma_w^2 (\sigma_w^2 + \sigma_A^2 \underline{\mathbf{c}}^* \underline{\mathbf{c}})} \quad (17.12)$$

In un caso piú generale:

$$(\underline{\mathbf{A}} + \underline{\mathbf{s}} \underline{\mathbf{s}}^*)^{-1} = \underline{\mathbf{A}}^{-1} - \frac{\underline{\mathbf{A}}^{-1} \underline{\mathbf{s}} \underline{\mathbf{s}}^* \underline{\mathbf{A}}^{-1}}{1 + \underline{\mathbf{s}} \underline{\mathbf{A}}^{-1} \underline{\mathbf{s}}^*}$$

Inoltre, indicando con $\underline{\mathbf{a}}_N$ il vettore incognito delle equazioni di Yule Walker (vedi al Capitolo sull'Analisi spettrale parametrica) e cioè lo stimatore ottimo del segnale:

$$\underline{\mathbf{a}}_N = - \underline{\mathbf{R}}^{-1} \underline{\mathbf{r}}_N = - \left(\frac{\sigma_A}{\sigma_w^2 + N\sigma_A^2} \right) \underline{\mathbf{c}} \quad (17.13)$$

poichè

$$\underline{\mathbf{c}}^* \underline{\mathbf{c}} = N \quad (17.14)$$

Risulta infine che la soluzione è:

$$a_{n,N} = - \frac{\sigma_A^2}{\sigma_w^2 + N\sigma_A^2} e^{jk_1 n \Delta} \quad (17.15)$$

Il denominatore della funzione di trasferimento che viene identificata con il metodo autoregressivo è:

$$A(z) = 1 + \sum_{n=1}^N a_{n,N} z^{-n}; \quad z = e^{jk\Delta}$$

17.2 STIMA AR: LA RISOLUZIONE ANGOLARE

La stima dello spettro di potenza unilatero del segnale visto dalla schiera dei sensori è ottenuta dallo spettro dell'errore di predizione, di cui è nota la potenza ma non il colore e pertanto è supposto bianco, spettro che poi deve essere diviso per la funzione di trasferimento del predittore. La potenza dell'errore di predizione è quella del segnale di ingresso, ridotta del fattore dipendente dagli N coefficienti di riflessione c_n ; $n = 1 \div N$ e, per $\sigma_A^2 \gg \sigma_w^2$:

$$c_N = a_{N,N} = -\frac{\sigma_A^2}{\sigma_w^2 + N\sigma_A^2} e^{jk_1 N \Delta}; \quad N = 1, \dots \quad (17.16)$$

$$1 - |c_1|^2 = 1 - \left(\frac{\sigma_A^2}{\sigma_w^2 + 1 \cdot \sigma_A^2} \right)^2 = \frac{2\sigma_w^2 \sigma_A^2 + \sigma_w^4}{(\sigma_w^2 + \sigma_A^2)^2} \sim \frac{2\sigma_w^2}{\sigma_A^2}$$

$$\text{Per } N > 1; \quad 1 - |c_N|^2 = 1 - \left(\frac{\sigma_A^2}{\sigma_w^2 + N\sigma_A^2} \right)^2 \sim 1 - \frac{1}{N^2}$$

Poichè:

$$\prod_{i=2}^{\infty} \left(1 - \frac{1}{i^2} \right) = \frac{1}{2}$$

si ha:

$$(1 - |c_1|^2)(1 - |c_2|^2)(1 - |c_3|^2) \dots \sim \quad (17.17)$$

$$\sim \frac{2\sigma_w^2}{\sigma_A^2} \left(\frac{3}{4} \right) \left(\frac{8}{9} \right) \dots \quad (17.18)$$

L'errore di predizione q.m. è:

$$E[|\varepsilon_{i,N}|^2] = (\sigma_A^2 + \sigma_w^2) \prod_{n=1}^N (1 - |c_n|^2) \quad (17.19)$$

è pari a circa $2\sigma_w^2$ per $N = 1$ e si riduce a σ_w^2 per N crescenti. Lo spettro stimato è:

$$\begin{aligned} \widehat{S}(k) &= \frac{2\Delta(\sigma_A^2 + \sigma_w^2) \prod_{n=1}^N (1 - |c_n|^2)}{\left| 1 + \sum_{n=1}^N a_{n,N} e^{-jkn\Delta} \right|^2} \\ &= \frac{2\Delta(\sigma_A^2 + \sigma_w^2) \prod_{n=1}^N (1 - |c_n|^2)}{\left| 1 - \frac{\sigma_A^2}{\sigma_w^2 + N\sigma_A^2} \sum_{n=1}^N e^{-j(k-k_1)n\Delta} \right|^2} \end{aligned} \quad (17.20)$$

CAPITOLO 17 LE DIREZIONI DI ARRIVO

$\widehat{S}(k)$ ha un massimo al numero d'onda k_1 dove vale:

$$\begin{aligned}\widehat{S}(k_1) &= \frac{2\Delta(\sigma_A^2 + \sigma_w^2) \prod_{n=1}^N (1 - |c_n|^2)}{\left|1 - \frac{N\sigma_A^2}{\sigma_w^2 + N\sigma_A^2}\right|^2} = \\ &= \frac{2\Delta(\sigma_A^2 + \sigma_w^2) \prod_{n=1}^N (1 - |c_n|^2)}{\left(\frac{\sigma_w^2}{\sigma_w^2 + N\sigma_A^2}\right)^2} \sim \alpha\Delta \frac{N^2\sigma_A^4}{\sigma_w^2}\end{aligned}$$

Il coefficiente α passa da 4 per $N = 1$ a 2 per $N \rightarrow \infty$. Si osservi che questo massimo cresce e con N^2 e con il rapporto segnale rumore; peraltro, nello spettro vero troveremmo un impulso ideale: quindi l'ampiezza del massimo non corrisponde comunque all'ampiezza del segnale. La qualità della stima deriva però dall'elevata risoluzione spettrale che ora calcoliamo. Infatti, l'intervallo δk di numeri d'onda intorno a k_1 che corrisponde ad una variazione di 3 dB dell'ampiezza $\widehat{S}(k)$ rispetto al massimo della stima spettrale è, per alto rapporto SNR :

$$\delta k = \frac{4\sigma_w^2}{N\sigma_A^2} \frac{1}{(N+1)\alpha\Delta} \quad (17.21)$$

Ciò si dimostra facilmente sviluppando attorno al suo minimo lo spettro del predittore:

$$\begin{aligned}1 + \sum_{n=1}^N a_{n,N} e^{-j(k-k_1)n\Delta} &= 1 - \sum_{n=1}^N \frac{\sigma_A^2}{\sigma_w^2 + N\sigma_A^2} e^{jn\phi} \sim \\ &= \frac{\sigma_w^2}{\sigma_w^2 + N\sigma_A^2} - j\phi \frac{\sigma_A^2 N(N+1)}{2(\sigma_w^2 + N\sigma_A^2)}; \quad \text{se } |\phi| = |(k - k_1)\Delta| \ll 1 \\ \phi_{3dB} &= \frac{\delta k \alpha \Delta}{2} \sim \frac{2\sigma_w^2}{N^2\sigma_A^2}\end{aligned}$$

Peraltro la potenza totale nella banda a 3 dB è proprio σ_A^2 : ricordando che

$$\int S(\omega) \frac{d\omega}{2\pi} = r_0$$

si ha:

$$P_{est} = \frac{4\sigma_w^2}{N^2\sigma_A^2} \frac{1}{\Delta} \Delta \frac{N^2\sigma_A^4}{\sigma_w^2} \frac{1}{2\pi} = \frac{2}{\pi} \sigma_A^2$$

a meno del fattore $2/\pi$ già incontrato, dovuto all'approssimazione di spettro uniforme (e non 1 polo) nella banda 3dB. La risoluzione angolare (spettrale) è quindi:

$$\begin{aligned}\Delta\theta &= \delta k \frac{c}{\omega \cos\theta} = \frac{4\sigma_w^2}{N\Delta(N+1)\sigma_A^2} \frac{\lambda}{2\pi \cos\theta} \simeq \\ &= \frac{\lambda}{L_{tot} \cos\theta} \times \frac{2}{\pi} \times \frac{\sigma_w^2}{(N+1)\sigma_A^2}; \quad N > 1\end{aligned}\quad (17.22)$$

tanto migliore rispetto a quanto ottenuto con il periodogramma per quanto più elevato è il rapporto Segnale - Rumore; inoltre cresce con il numero N di sensori utilizzati. Praticamente, è inutile fare crescere N oltre 2 o 3. In questo caso, la sinusoidale è ricostruita quasi perfettamente e l'errore residuo è solo il rumore additivo, imprevedibile. Invece, per $N = 1$, la sinusoidale è stimata in modo imperfetto e l'errore ha potenza doppia.

17.2.3 Numero dei sensori della schiera

Dall'occhio del falco alla punta del cane da caccia

Da quanto abbiamo visto finora, si vede che, se disponiamo di una schiera di $N + 1$ sensori e se utilizziamo il periodogramma, possiamo determinare contemporaneamente direzioni ed ampiezze di N sorgenti da tutte le direzioni accessibili (i contenuti delle N celle del periodogramma). Come vedremo nella sezione sulle lenti sottili nel capitolo sui filtri 2D di natura fisica, questo è il modo di funzionamento dell'occhio. Una lente agisce come un filtro *chirp* (vedi il paragrafo in proposito) bidimensionale e opera una trasformazione di Fourier (a 2D nel caso dell'occhio) che focalizza su varie posizioni della retina le radiazioni che provengono da varie direzioni. L'innervamento della retina porta poi alla valutazione delle ampiezze (e dei colori) delle sorgenti. Così, per fissare le idee in modo intuitivo, *l'occhio del falco* permette di osservare dall'alto un intero territorio e trovare la posizione delle prede. Siamo quindi nella condizione di N molto grande, e scopo dell'osservazione è quello di determinare, tra tante, posizioni e ampiezze delle sorgenti di interesse. La conoscenza a priori delle posizioni è nulla, perchè siamo nella condizione di avvistamento. Se riduciamo N , ci mettiamo nella condizione di potere osservare solo un numero minore di sorgenti simultanee, e quindi dobbiamo utilizzare delle informazioni a priori sulla loro natura e posizione.

CAPITOLO 17 LE DIREZIONI DI ARRIVO

L'altro caso estremo si ha per $N + 1 = 2$; si pensi per esempio alla *punta di un cane da caccia*. Questo drizza le orecchie per identificare con la massima precisione la posizione in un cespuglio di una preda che non vede. La schiera è solo di due sensori, orientata in modo ortogonale all'unica direzione di arrivo di interesse ($\cos \theta = 1$). La precisione della localizzazione, come abbiamo visto, dipende dal rapporto segnale rumore (*orecchie grandi e dispiegate*). Visto che si tratta di un'unica sorgente, bastano due sensori, che devono essere massimamente sensibili ed alla massima distanza possibile tra loro. Le risorse sono così impiegate al meglio.

In altre parole, il numero degli elementi di una schiera ha due possibili motivazioni: il numero delle sorgenti da stimare contemporaneamente, e l'agilità con cui la direzione della schiera può essere modificata. Se la schiera è agile e la sorgente è una sola, la massima efficienza è raggiunta con solo due sensori massimamente sensibili, da disporre alla massima distanza reciproca possibile per evitare l'equivocazione, che dipende dalla frequenza usata. Per fare un esempio, i pipistrelli usano ultrasuoni, ma hanno la testa piccola (vedi figura 17.2). La distanza tra i due elementi della schiera è limitata solo dall'incertezza a priori della posizione della sorgente. Disporre in questo caso altri sensori posti tra gli estremi sarebbe uno spreco: meglio usare le risorse per migliorare la sensibilità dei due sensori estremali (*solo due orecchie, ma grandi. Tre orecchi proporrebbero due posizioni possibili di due prede, e lascerebbero il cane nell'incertezza. Un terzo orecchio centrale è quindi inutile: i cani cercano una preda per volta!*). Se invece l'agilità della schiera è limitata, si devono disporre i sensori più vicini tra loro, per poterla orientare elettronicamente in un arco di direzioni più vasto. La distanza tra i ricevitori si deve ridurre per evitare l'equivocazione dovuta ad arrivi laterali, e quindi servono più ricevitori per ottenere la risoluzione angolare voluta. In sostanza, il numero degli elementi della schiera riflette l'incertezza a priori del numero dei bersagli e la dispersione delle loro posizioni (*l'occhio del falco*). Come vedremo in seguito, una schiera può anche essere utile per separare (e sopprimere) sorgenti disturbanti o indesiderate. Il caso anomalo di voler inseguire sorgenti contemporanee con direzioni molto diverse porta ad un effetto come di strabismo, ingegneristicamente costoso. In questo caso, possono essere utilizzate delle schiere con pochi sensori ma a distanze molto non uniformi tra loro (schiere logaritmiche).

Infine, si deve prolungare al massimo la durata dell'osservazione per

17.3 REIEZIONE DIREZIONALE DEL DISTURBO

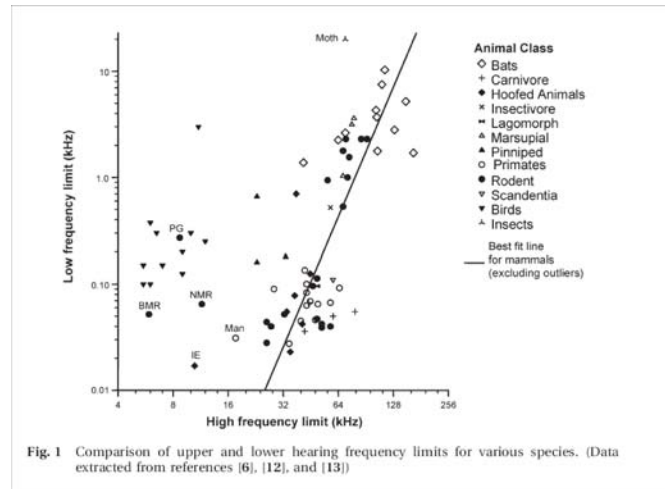


Figura 17.2: Massime e minime frequenze dell'apparato uditivo di vari animali. Da P Counter, Implantable hearing aids, Newcastle, UK

migliorare il rapporto segnale rumore e quindi la precisione della localizzazione. Ovviamente, è necessario che in questo intervallo di tempo il bersaglio non si muova o possa essere inseguito dalla schiera.

Per fare un ultimo esempio dal mondo animale, osserviamo che i delfini ascoltano ultrasuoni fino a circa 150KHz, con una lunghezza d'onda di circa 1cm ($c = 1500m/s$). Possono sentire una sardina (20cm) a 100m di distanza. La schiera potrebbe essere costituita dalla mascella ed i denti sarebbero i sensori: in questo caso è probabile che il valore di N sia elevato, riportandosi ad una situazione più simile a quella del periodogramma e dell'occhio del falco.

17.3 Reiezione direzionale del disturbo

La tecnica della misura delle DOA può avere varie applicazioni: p.e. nella telefonia cellulare, le stazioni fisse possono localizzare le direzioni dei chiamanti, per concentrare nella loro direzione la direttività delle antenne di ricezione e trasmissione; questo meccanismo potrebbe portare ad un decremento delle interferenze quindi all'accrescimento del numero degli utenti. Ancora, sono facilmente immaginabili applicazioni alla difesa: le

CAPITOLO 17 LE DIREZIONI DI ARRIVO

sorgenti possono essere sommergibili o aerei di cui si deve determinare l'azimuth.

In ogni caso, vi possono però essere forti sorgenti vicine che *disturbano* il sensore. La presenza di queste sinusoidi spaziali molto ampie *polarizza* la stima spettrale e la rende imprecisa, come già visto quando abbiamo studiato l'effetto delle finestre su dei segnali sinusoidali. Può essere utile rendere il sistema meno sensibile a queste sorgenti disturbanti; si adottano quindi tecniche per individuare la direzione da cui proviene il segnale disturbante per cercare poi di annullare in quella direzione la direttività dell'antenna sintetizzata con il sistema di sensori. (*Le visiere dei cappelli fanno da schermo alla luce solare*). Nel caso specifico, si può mettere un ulteriore sensore x_0 (vedi fig. 17.1) il cui scopo è quello di registrare la radiazione disturbante. Quindi, si cercherà di decorrelare tutti i sensori rispetto a x_0 ; in altre parole, usando x_0 cercheremo di stimare l'uscita di ogni sensore; la nuova uscita del sensore sarà costituita dall' *errore di predizione*. Così facendo, si eliminerà approssimativamente la radiazione disturbante, raggiungendo lo scopo. Indichiamo il segnale misurato con:

$$x_i = A_1 e^{jk_1 i \Delta} + A_2 e^{jk_2 i \Delta}; \quad i = 0, \dots, N; \quad (17.23)$$

$$E[|A_1|^2] = \sigma_1^2 \gg E[|A_2|^2] = \sigma_2^2; \quad (17.24)$$

$$\mu = \frac{\sigma_2^2}{\sigma_1^2} \ll 1 \quad (17.25)$$

Ampiezza e numero d'onda della radiazione disturbante sono indicati con A_1, k_1 ; quelle della radiazione disturbata con A_2, k_2 . Indichiamo con:

$$y_i = x_i - \alpha_i x_0; \quad i = 1, \dots, N; \quad (17.26)$$

l'errore di predizione dell'uscita del sensore i -esimo utilizzando come predittore il sensore x_0 ; perciò, imponiamo l'ortogonalità tra y_i e x_0 :

$$\alpha_i = \frac{E[x_i x_0^*]}{E[|x_0|^2]} \quad (17.27)$$

$$\begin{aligned} \alpha_i &= \frac{\sigma_1^2 e^{jk_1 i \Delta} + \sigma_2^2 e^{jk_2 i \Delta}}{\sigma_1^2 + \sigma_2^2} = \\ &= \frac{e^{jk_1 i \Delta} + \mu e^{jk_2 i \Delta}}{1 + \mu} \quad i = 1, \dots, N; \end{aligned} \quad (17.28)$$

17.4 STIMA SPETTRALE DI CAPON (MINIMAMENTE POLARIZZATA)

L'errore di predizione è:

$$y_i = B_1 (e^{jk_1 i \Delta} - e^{jk_2 i \Delta}) \quad (17.29)$$

$$B_1 = \frac{\mu A_1 - A_2}{1 + \mu}$$

Pertanto, se prima dell'operazione di predizione le ampiezze delle due radiazioni erano fortemente differenti, dopo la predizione, le ampiezze dei due residui sono eguali, e quindi la polarizzazione spettrale sarà ridotta. Sostanzialmente abbiamo realizzato un sistema che ha un guadagno direzionale con un minimo fortemente pronunciato nella direzione della radiazione disturbante. Se il sistema è adattativo, e quindi le operazioni di media di insieme sono invece realizzate con delle medie temporali con un opportuno fattore di oblio, anche se la direzione dell'onda disturbante varia, la direzione del minimo del guadagno dell'antenna equivalente al sistema si sposterà opportunamente.

17.4 Stima spettrale di Capon (minimamente polarizzata)

Una generalizzazione del problema precedente viene ottenuta dalla stima spettrale di Capon; con essa, si cerca il filtro che pur avendo funzione di trasferimento unitaria al numero d'onda k_0 , abbia in uscita la potenza minima. Quindi, si minimizza il disturbo e perciò la polarizzazione dovuta a tutte le altre sorgenti. La stima spettrale che ne risulta è minimamente polarizzata, ed ha andamento:

$$\hat{S}_C(k_0) = \frac{1}{\underline{\mathbf{a}}(k_0)^* \underline{\mathbf{R}}^{-1} \underline{\mathbf{a}}(k_0)}$$

dove con $\underline{\mathbf{a}}(k_0)$ abbiamo indicato lo steering vector o vettore di direzione:

$$\underline{\mathbf{a}}(k_0) = \begin{bmatrix} 1 \\ e^{jk_0 \Delta} \\ \dots \\ e^{j(N-1)k_0 \Delta} \end{bmatrix}$$

CAPITOLO 17 LE DIREZIONI DI ARRIVO

La dimostrazione é semplice: si consideri un filtro, dipendente da k_0

$$\underline{\mathbf{h}}(k_0) = \begin{bmatrix} h_0 \\ h_1 \\ \dots \\ h_N \end{bmatrix}$$

e il risultato $\{y_n\}$ della convoluzione di una sequenza $\{x_n\}$ con questo filtro. Indicando con $\underline{\mathbf{x}}_n$ un vettore che contiene N elementi consecutivi della sequenza $\{x_n\}$

$$\underline{\mathbf{x}}_n = \begin{bmatrix} x_n \\ x_{n-1} \\ \dots \\ x_{n-N+1} \end{bmatrix}$$

abbiamo

$$y_n = \underline{\mathbf{h}}(k_0)^* \underline{\mathbf{x}}_n$$

Indicando con $\underline{\mathbf{R}}$ la matrice di covarianza di $\underline{\mathbf{x}}_n$ il valore q.m. di y_n é:

$$E[y_n^2] = E[\underline{\mathbf{h}}^* \underline{\mathbf{x}}_n \underline{\mathbf{x}}_n^* \underline{\mathbf{h}}] = \underline{\mathbf{h}}^* \underline{\mathbf{R}} \underline{\mathbf{h}} \quad (17.30)$$

La risposta in frequenza del filtro $\underline{\mathbf{h}}$ é:

$$H(k, k_0) = \sum_{n=1}^N h_n(k_0) e^{-jkn\Delta} = \underline{\mathbf{a}}(k)^* \underline{\mathbf{h}}(k_0)$$

La stima spettrale di Capon individua il filtro passabanda $\underline{\mathbf{h}}(k_0)$ che minimizza la potenza in uscita ($E[y_n^2]$) pur rispettando il vincolo

$$\underline{\mathbf{h}}(k_0)^* \underline{\mathbf{a}}(k_0) = 1$$

Si ha cioè:

$$\underline{\mathbf{h}}(k_0) = \arg \min_{\underline{\mathbf{h}}} \underline{\mathbf{h}}^* \underline{\mathbf{R}} \underline{\mathbf{h}} ; \underline{\mathbf{h}}(k_0)^* \underline{\mathbf{a}}(k_0) = 1 \quad (17.31)$$

17.4 STIMA SPETTRALE DI CAPON (MINIMAMENTE POLARIZZATA)

La funzione di trasferimento del filtro $\underline{\mathbf{h}}$ é funzione della pulsazione k , del numero d'onda k_0 al quale si vuole effettuare la stima della densità spettrale del processo $\{x_n\}$ ed anche del contenuto spettrale della sequenza alle altre frequenze. La soluzione della (17.31) é, come visto nel capitolo sulla stima lineare ottima nel caso di stima non polarizzata di una segnale $\underline{\mathbf{a}}$ immerso in un rumore colorato che ha matrice di covarianza $\underline{\mathbf{R}}$:

$$\underline{\mathbf{h}} = \frac{\underline{\mathbf{R}}^{-1} \underline{\mathbf{a}}}{\underline{\mathbf{a}}^* \underline{\mathbf{R}}^{-1} \underline{\mathbf{a}}}$$

La risoluzione angolare di questo filtro si può approssimare con l'inverso della sua lunghezza ($N\Delta$) e quindi sostituendo nella (17.30) si ha:

$$\widehat{S}_C(k) = \frac{E[y_n^2]}{1/N\Delta} = N\Delta \left(\frac{\underline{\mathbf{R}}^{-1} \underline{\mathbf{a}}}{\underline{\mathbf{a}}^* \underline{\mathbf{R}}^{-1} \underline{\mathbf{a}}} \right)^* \underline{\mathbf{R}} \left(\frac{\underline{\mathbf{R}}^{-1} \underline{\mathbf{a}}}{\underline{\mathbf{a}}^* \underline{\mathbf{R}}^{-1} \underline{\mathbf{a}}} \right) = \frac{N\Delta}{\underline{\mathbf{a}}^* \underline{\mathbf{R}}^{-1} \underline{\mathbf{a}}}$$

Se il rumore fosse bianco saremmo in pratica di nuovo al caso della stima fatta con il periodogramma (però k_0 può essere qualsiasi e non solo corrispondente ad un numero intero di cicli su N campioni).

Nel caso di rumore colorato, supponiamo che la matrice $\underline{\mathbf{R}}$ sia circolante, e i suoi autovettori $\underline{\mathbf{e}}_m$ e autovalori λ_m siano al solito le sinusoidi campionate che fanno un numero intero di cicli su N campioni e i campioni del periodogramma. Siano poi a_{m,k_0} , h_{m,k_0} le componenti lungo gli autovettori $\underline{\mathbf{e}}_m$ dello steering vector $\underline{\mathbf{a}}(k_0)$ alla pulsazione k_0 e del filtro ottimo $\underline{\mathbf{h}}(k_0)$ per stimare lo spettro alla pulsazione k_0 :

$$a_{m,k_0} = \underline{\mathbf{a}}(k_0)^* \underline{\mathbf{e}}_m; \quad h_{m,k_0} = \underline{\mathbf{h}}(k_0)^* \underline{\mathbf{e}}_m \quad (17.32)$$

Se k_0 corrisponde ad un numero intero di cicli su N campioni, $\underline{\mathbf{a}}(k_0)$ coincide con $\underline{\mathbf{e}}_{k_0}$ ed allora $a_{m,k_0} = A\delta(m - k_0)$. Altrimenti, gli a_{m,k_0} sono diversi da zero, e corrispondono ai lobi laterali ed al lobo principale della sinusoidi dello steering vector $\underline{\mathbf{a}}(k_0)$, finestrata con una finestra rettangolare, se appunto la sua frequenza non corrisponde ad un numero intero di cicli nell'intervallo $N\Delta$. Il filtro $\underline{\mathbf{h}}(k_0)$ pesa tutti questi lobi con lo spettro inverso del segnale:

$$h_{m,k_0} = \frac{\sum_m \frac{a_{m,k_0}}{\lambda_m}}{\sum_m \frac{|a_{m,k_0}|^2}{\lambda_m}} \quad (17.33)$$

CAPITOLO 17 LE DIREZIONI DI ARRIVO

In sostanza, la stima spettrale è fatta utilizzando un banco di filtri diversi per ogni k_0 , ottimizzati tenendo conto dello spettro stimato, in modo da minimizzare gli effetti di polarizzazione. La risoluzione spettrale è intermedia tra quella, minima, del periodogramma e quella, massima, della stima spettrale AR. Potrebbe accadere che un lobo laterale, se capita in una zona dove lo spettro del segnale totale è piccolo, sia pesato anche di più del lobo principale. Sarebbe proprio il caso in cui si osserva una sinusoide finestrata dai lobi laterali e quindi *dal buco della serratura!*

La stima spettrale di Capon si applica anche al caso in cui la schiera non è regolare (campionamento non uniforme). In questo caso, la tecnica del periodogramma incontra delle difficoltà perchè, nel caso di campionamento irregolare, la matrice di covarianza della trasformata di Fourier non è più diagonale.

Capitolo 18

Il segnale vocale

Una delle principali applicazioni del trattamento numerico dei segnali è quella al segnale vocale e cioè il segnale generato dalla voce umana. Dopo una brevissima presentazione dei meccanismi di fonazione e dell'udito, tratteremo sia la codifica che la sintesi del segnale vocale; infine si farà anche qualche cenno alla codifica del segnale musicale.

Il segnale vocale viene formato nel tratto vocale (figura 18.1) che va dalle corde vocali alle labbra. E' questo un sistema riverberante, disadattato all'uscita (le labbra) ed abbastanza ben adattato all'ingresso (le corde vocali). La sagoma del sistema riverberante dipende dalla posizione della bocca e della lingua; il segnale di ingresso è generato dalle corde vocali ed è di due tipi: vocalizzato (quando le corde vocali generano una sequenza di impulsi di energia acustica) e non vocalizzato (quando le corde vocali generano un fruscio praticamente gaussiano a spettro abbastanza uniforme). La sequenza di impulsi generata durante la formazione di segnale vocalizzato ha una frequenza di ripetizione che va da circa $100Hz$ (voce roca, maschi) fino a $250 - 300Hz$ (voce melodiosa, bambini). Ogni impulso ha una durata di qualche millisecondo e quindi uno spettro che si estende fino a $8 - 10kHz$. Il timbro, e cioè la frequenza di emissione di questi impulsi (in inglese, pitch) può avere valori di significato, ad esempio interrogativo, (se crescente col tempo), o dichiarativo (se decrescente, *p.e. quando si desidera concludere una telefonata*).

Gli impulsi traversano il tratto vocale e sono quindi sagomati spettralmente dalla funzione di trasferimento tutti poli corrispondente. Ovviamente, trattandosi di un sistema distribuito e non di una cascata di tronchi di linea uniformi, non è possibile determinare il numero dei poli

CAPITOLO 18 IL SEGNALE VOCALE

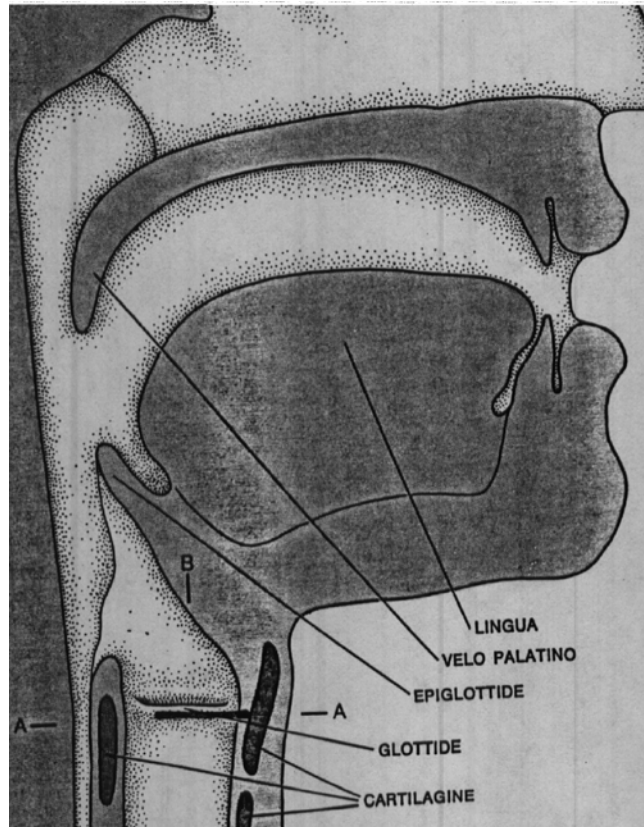


Figura 18.1: Apparato vocale umano

equivalente al sistema senza fare delle approssimazioni. Tuttavia, se si osserva che la lunghezza del tratto vocale è di circa 20cm e che la velocità del suono è circa 330m/s si deduce che il ritardo totale di propagazione è di circa $.6\text{msec}$. Se ora vogliamo approssimare il meccanismo del parlato fino a 4kHz la frequenza di campionamento deve essere almeno 8kHz e pertanto l'intervallo di campionamento inferiore a $125\ \mu\text{s}$; quindi il ritardo unitario di ogni cella deve essere inferiore a $62,5\ \mu\text{s}$ ed il numero di celle contenute nel tratto superiore a 10; infatti, in genere si sceglie di rappresentare il tratto vocale con 10 celle e quindi con 10 poli.

Al massimo possiamo avere 5 picchi di risonanza, pertanto; in realtà se ne contano, per ogni vocale, da 3 a 4; queste frequenze di risonanza sono chiamate formanti e la loro posizione caratterizza abbastanza le

vocali (vedi Tabella 1); si stia ben attenti a non pensare che distinguere le vocali porti alla comprensione del parlato; gli inflessi dialettali e numerosi giochi di società insegnano il contrario. Infatti, l'intelligibilità del parlato è soprattutto legata alle consonanti e cioè a transitori tra una vocale (che può essere mantenuta per lungo tempo) ed un'altra. In sostanza, le vocali corrispondono alla portante e le consonanti alla modulante del segnale acustico corrispondente al parlato.

Tabella I: Frequenze delle formanti dei suoni vocalici inglesi

Vocale	Frequenza	Frequenza	Frequenza
inglese	I formante	II formante	III formante
but	520	1190	2380
bat	660	1720	2410
feet	270	2290	3010
hot	730	1090	2440
food	440	1020	2240
fought	570	840	2410
foot	300	870	2240
bird	490	1350	1690
bit	390	1990	2550
bet	530	1840	2480

I suoni non vocalizzati sono generati dal fruscio emesso dalle corde vocali quando non generano impulsi (si pensi al parlato sottovoce). Infine, i suoni consonantici possono essere di vario tipo: le consonanti sibilanti sono generate dal fruscio emesso dalle corde vocali; ma vi sono consonanti di carattere plosivo (ad esempio b, p, t), corrispondenti a bruschi transitori del tratto vocale, ecc.

I toni nasali sono generati da interferenze tra emissioni acustiche attraverso il tratto vocale ed attraverso il setto nasale: quindi non cor-

rispondono a funzioni di trasferimento di soli poli, ma anche con zeri dovuti al parallelo di due reti tutti poli.

18.1 Tecniche di codifica del segnale vocale

Il segnale vocale può essere codificato con tecniche di complessità molto diversa; al limite, si può pensare ad un sistema di riconoscimento del segnale vocale, alla trasmissione alfanumerica del testo riconosciuto, e quindi alla sintesi vocale del testo effettuata in ricezione: questa tecnica, di massima riduzione di ridondanza, porterebbe ad un ritmo di trasmissione di circa $40 - 50 \text{ bit/sec}$. Il risparmio è massimo, senza tuttavia perdere la riconoscibilità del testo; si perderebbe invece la possibilità di individuare la personalità e lo stato d'animo del parlatore; inoltre, il ritardo dovuto al tempo necessario al riconoscimento è certamente maggiore dei circa 10 msec in genere consentiti. All'estremo opposto, possiamo pensare ad una codifica *trasparente* in cui la forma d'onda del segnale è trasmessa fedelmente al ricevitore: in questo caso, il ritmo di trasmissione può arrivare a $40 \text{ KHz} \times 16 \text{ bit/campione} = 640 \text{ kbit/sec}$; quindi, lo stesso segnale può essere trasmesso con 10.000 volte meno bit, ma naturalmente con delle perdite. Come vedremo, utilizzando sia i meccanismi di generazione del segnale, sia i meccanismi di percezione auditiva che rendono impercipienti alcune modifiche al segnale, è possibile avere un'ottima qualità del segnale vocale con $7 - 8 \text{ kbit/sec}$ ed un'ottima percepibilità con $3 - 4 \text{ kbit/sec}$. Un segnale musicale, molto più complesso e per cui è opportuna una maggiore fedeltà, ha bisogno di circa $150 - 200 \text{ kbit/sec}$. Naturalmente, queste cifre sono variabili con il tempo e con la tecnologia; tuttavia, il costo di trasmissione del timbro e delle sue variazioni è già di qualche centinaia di *bit/sec*.

18.2 Codifica delle forme d'onda

I campioni di un segnale spettralmente colorato quale è il segnale vocale sono correlati tra loro; quindi sono almeno parzialmente predicibili ed è conveniente trasmettere non il segnale ma l'errore di predizione. La convenienza non sta però nello sbiancamento, che non comporta alcuna riduzione della frequenza di campionamento. Essa nasce dal fatto

18.2 CODIFICA DELLE FORME D'ONDA

che nel caso piú frequente di segnale vocalizzato, l'errore di predizione (approssimativamente il segnale di ingresso generato dalle corde vocali) è una forma d'onda costituita da impulsi ripetuti, con molti campioni pressochè nulli, e che può essere descritta e trasmessa *parsimoniosamente* con un numero ridottissimo di parametri (ad esempio ampiezza, intervallo tra impulsi, e magari anche la forma d'onda dell'impulso, ripetitiva ma lentamente variabile con il tempo).

La possibilità di una codifica efficiente del segnale vocale nasce quindi dalla **non gaussianità** del processo stocastico corrispondente al segnale vocalizzato. Peraltro, i segnali non vocalizzati di durata abbastanza lunga corrispondono sostanzialmente a fruscii: essi sono abbastanza gaussiani, ma in compenso non serve una loro accurata descrizione, in quanto basta trasmetterne lo spettro di potenza piuttosto che la forma d'onda. In questo caso, la forma d'onda del segnale ricostruito non sarà affatto identica o simile a quella del segnale trasmesso, ma sarà incorrelata con esso avendo in comune solo densità di probabilità e spettro di potenza.

La tecnica di codifica del segnale piú diffusa (LPC, Linear Predictive Coding) si fonda appunto sulla trasmissione periodica dei parametri del tratto vocale (ad esempio, i 10 coefficienti di riflessione) ed inoltre dei parametri degli impulsi generati dalle corde vocali del parlatore. Tuttavia, a titolo introduttivo, si discuteranno le due tecniche di codifica della forma d'onda piú semplici, la Modulazione Delta e l'ADPCM (Adaptive Differential Pulse Code Modulation).

18.2.1 La Modulazione Delta

Il modo piú elementare per sfruttare la dipendenza statistica dei successivi campioni di un segnale è quello di codificarne le differenze; nella Modulazione Delta questa codifica viene fatta con un solo bit, che quindi comunica se il campione successivo è maggiore o minore di quello precedente. In conseguenza viene aggiornato un contatore per permettere il successivo confronto (vedi figura 18.2). Il segnale di ingresso potrebbe anche essere analogico: in questo caso il sistema funzionerebbe anche da convertitore analogico - numerico. Il disturbo creato dalla Modulazione Delta è ridotto se i campioni successivi sono fortemente correlati e cioè se lo spettro del segnale è passa - basso. Il disturbo è di due tipi:

- rumore di granularità (idle noise o rumore di inattività);
- rumore di inseguimento (slope overload).

CAPITOLO 18 IL SEGNALE VOCALE

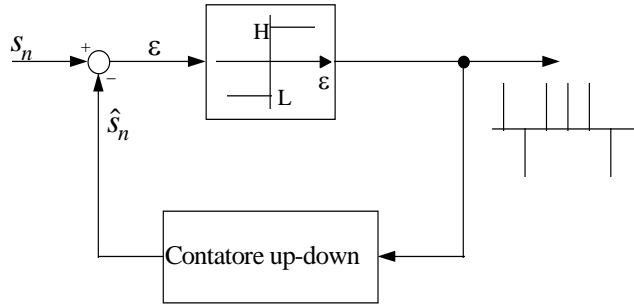


Figura 18.2: Modulatore delta

Supponiamo che il quanto a addizionato o sottratto al segnale per l'arrivo di un nuovo bit sia piccolo. Negli intervalli di tempo in cui il segnale varia poco, la trasmissione è fedele; tuttavia se c'è una variazione a gradino di ampiezza Ma , il sistema avrà bisogno di M passi per recuperare questa differenza (fig. 18.3). La pendenza massima descrivibile dal segnale ricostruito è $|a|/T$; se vi sono pendenze maggiori, siamo nella condizione di inseguimento, il sistema va in "sovraccarico di pendenza" (in inglese *slope overload*) e nascono delle differenze anche grandi. Se invece il quanto a è abbastanza grande, la pendenza massima sarà sufficiente, ma ne nascerà una sorta di errore di quantizzazione (errore di granularità). In funzione dello spettro del segnale e della frequenza di campionamento vi è un valore ottimale per il quanto; se il segnale è non stazionario, tuttavia, è opportuno rendere la Modulazione Delta adattativa variando il quanto p.e. con la regola di accrescerlo di un fattore P se due bit successivi hanno lo stesso segno (situazione di inseguimento) e di ridurlo di un fattore Q se invece hanno segni opposti (situazione di inattività). Si riesce ad ottenere un sistema efficiente se le permanenze di segno sono probabili come le variazioni: dopo n passi, dovendo il quanto avere circa lo stesso valore, ed essendo la variazione totale di ampiezza:

$$P^{n/2}Q^{n/2} = (PQ)^{n/2} = 1 \quad (18.1)$$

risulta:

$$PQ = 1$$

Agendo in questo modo, si possono avere al limite quanti di ampiezza troppo grande o troppo piccola; per limitare questa escursione si può

18.2 CODIFICA DELLE FORME D'ONDA

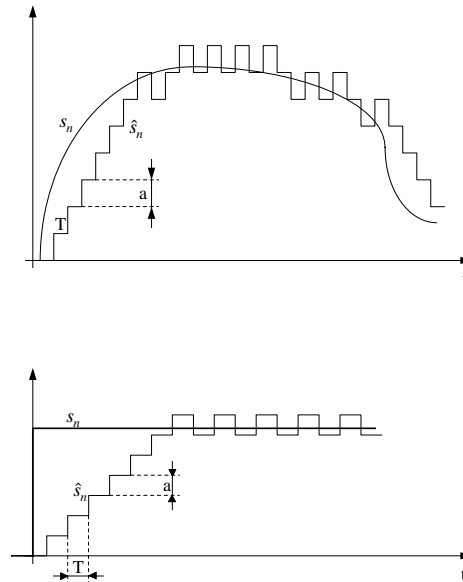


Figura 18.3: Funzionamento della modulazione delta e risposta al gradino

adottare la tecnica nota come CVSD (Continuously Variable Slope Delta) dove la regola di aggiornamento del quanto è la seguente:

$$a_n = \beta a_{n-1} + D_i ; \quad i = 1, 2 ; \quad D_1 \gg D_2 ; \quad 0 < \beta < 1 \quad (18.2)$$

In condizione di inseguimento (bit successivi eguali) l'indice $i = 1$, in condizione di inattività (bit successivi diversi) $i = 2$. Con un ingresso costante, dopo un numero grande di campioni, si arriva ad avere un quanto di ampiezza minima:

$$a_{min} = \frac{D_2}{1 - \beta}$$

Invece in situazioni di inseguimento prolungate, il quanto cresce fino al valore massimo:

$$a_{max} = \frac{D_1}{1 - \beta}$$

Se $\beta \simeq 1$ l'aggiornamento del quanto è lento; per progettare il sistema, si sceglie anzitutto il valore per β e quindi D_1 , D_2 in funzione dei valori

desiderati per i quanti minimo e massimo. Si vede ora un problema che pesa sui segnali codificati; essi sono molto vulnerabili agli errori di canale: infatti, se per qualche motivo un bit viene decodificato in modo errato, da quel momento in poi il ricevitore decodificherebbe in modo errato tutti i segnali ricevuti; questo mostra che è necessario prevedere dei periodici riassistentamenti, ad esempio trasmettendo alcuni campioni non in modo differenziale.

18.2.2 L'ADPCM

La modulazione Delta è il caso piú semplice di tutta una serie di codificatori differenziali: il sistema può essere migliorato in vari modi:

- predire meglio il prossimo campione, piuttosto che utilizzare un semplice mantentore come nella Modulazione Delta;

- usare un quantizzatore differenziale piú duttile, con un numero di livelli di restituzione pari a 2^m con $m = 3, 4$ ed inoltre con una caratteristica di quantizzazione non uniforme;

- rendere adattativo il sistema, con tecniche di predizione lineare adattativa, quali quelle viste al capitolo corrispondente. Lo schema a blocchi del sistema DPCM (Differential Pulse Code Modulator) è presentato in figura 18.4. La differenza tra la migliore predizione lineare del prossimo campione ed il campione effettivo viene quantizzata con un quantizzatore non uniforme; l'errore di predizione, quantizzato, aggiorna il segnale predetto e aggiorna la base dati per le future predizioni. Un punto importante è che il quantizzatore non è uniforme ma ha una caratteristica di quantizzazione come in figura 18.5. Quando l'ingresso del quantizzatore è grande (forte errore di predizione) l'errore di quantizzazione sarà anche grande; questo corrisponde però ad un momento di minore sensibilità percettiva e pertanto l'efficacia è buona. Il vantaggio di un quantizzatore non uniforme è soprattutto evidente per densità di probabilità non uniformi, quale è quella dell'errore di predizione del DPCM, molto vicina ad una Laplaciana, e cioè del tipo:

$$p(x) = \frac{x_0}{2} e^{-|x|/x_0} \quad (18.3)$$

Panter e Dite (1951) hanno dimostrato che l'errore quadratico medio dovuto alla quantizzazione si minimizza se i quanti sono approssimativamente proporzionali all'inverso della radice cubica della densità di probabilità: in questo caso risulta anche che i contributi di tutti i livelli al

18.2 CODIFICA DELLE FORME D'ONDA

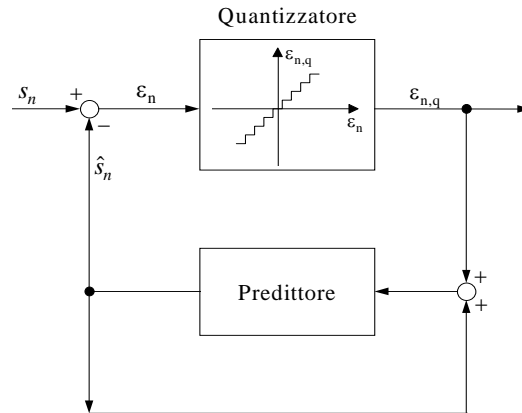


Figura 18.4: Schema a blocchi del DPCM (Differential Pulse Code Modulator)

valore q.m. dell'errore di quantizzazione totale sono circa eguali. In altre parole, i livelli esterni corrispondenti ai valori maggiori della variabile x , sono sì interessati più di rado dal quantizzatore, ma contribuiscono con errori maggiori all'errore totale, bilanciando la rarità. In conclusione, nel caso del DPCM, gli intervalli di quantizzazione crescono in modo approssimativamente esponenziale. Sia il predittore che il quantizzatore possono essere resi adattativi ed in questo caso il sistema viene indicato con il nome di ADPCM (Adaptive DPCM).

Codificatori ADPCM sono normalizzati ed utilizzati ad esempio con ritmo di informazione di 32 kbits/sec con un risparmio di un fattore 2 rispetto al consueto 64 kbit/sec del PCM (Pulse Code Modulation). In entrambi i casi la frequenza di campionamento è 8 KHz ; con il PCM, in questo caso, i campioni vengono trasmessi indipendentemente, usando 8 bit/campione. La caratteristica di quantizzazione usuale del PCM è di tipo logaritmico, ed assegna 5 bit alla mantissa e 3 alla caratteristica.

18.2.3 La codifica lineare predittiva

Una tecnica molto usata per al codifica del segnale vocale è quella della codifica lineare predittiva; il segnale campionato a 8 KHz viene seg-

CAPITOLO 18 IL SEGNALE VOCALE

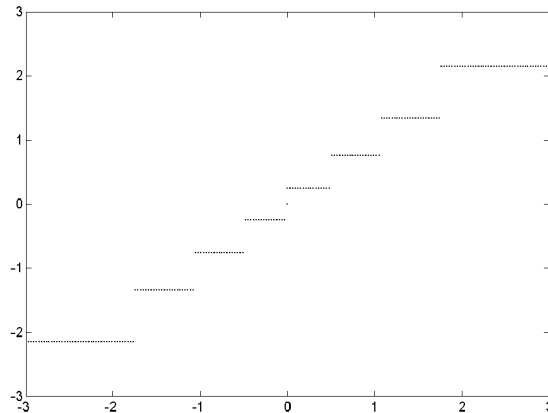


Figura 18.5: Quantizzatore non uniforme a passi crescenti. I livelli di restituzione sono posizionati in modo da minimizzare l'errore massimo

mentato in spezzoni della durata di circa 100 campioni ($12,5\text{msec}$, un intervallo di tempo pari circa a quello del timbro). Per ognuno di questi intervalli viene fatta un'analisi spettrale AR del segnale e vengono determinati 10 coefficienti di riflessione, che vengono trasmessi al ricevitore. L'errore di predizione corrispondente all'uscita del filtro tutti zeri caratterizzato da quei coefficienti di riflessione viene poi anch'esso codificato e trasmesso. Esistono numerose versioni del codificatore e le metodologie di codifica sia dei coefficienti di riflessione che dell'errore di predizione cambiano con il tipo di codificatore. Facciamo una breve valutazione del ritmo di informazione necessario: assegnando 4 bit/coefficiente di riflessione e 10 bit sia per l'ampiezza che per la frequenza del timbro, si arriva ad un totale di $60 \times 10^3 : 12.5 = 4800 \text{ bits/sec}$. Inoltre servono segnali per comunicare la situazione vocalizzata/non vocalizzata ecc. Si vede che anche per una versione semplificata del codificatore, sono richiesti comunque circa 5 kbit/sec . Esistono numerose possibilità per codificare i coefficienti di riflessione; abbiamo già fatto uso della loro resistenza al rumore di quantizzazione indicando la scarsa sensitività dei filtri a traliccio alla quantizzazione dei coefficienti. Si può fare meglio utilizzando le LSP (Line Spectral Pairs) e cioè le coppie di righe spettrali. Chiudendo il sistema riverberante corrispondente al tratto vocale su un undicesimo

18.3 LA CODIFICA DEL SEGNALE MUSICALE (MP3)

coefficiente di riflessione $c_{11} = \pm 1$ abbiamo diversi effetti:

a) i poli del sistema si portano tutti sul cerchio unitario, visto che il sistema è disadattato dalle due parti, e non c'è perdita di energia;

b) i poli che risultano con il coefficiente $c_{11} = -1$ si alternano con quelli corrispondenti a $c_{11} = 1$.

E' quindi possibile codificare le 10 distanze lungo il semicerchio unitario (e quindi differenze di frequenze) che separano i poli dell'uno e dell'altro tipo. In questa maniera si riesce ad ottenere il massimo dell'efficienza nella rappresentazione del tratto vocale. Inoltre, invece di trasmettere ad ogni intervallo di tempo di circa $12,5 \text{ msec}$ questa informazione, si possono trasmettere solo le variazioni di queste posizioni, quantizzate con tecniche *ADPCM*. Infine, l'errore di predizione relativo a 100 campioni può venire trasmesso scegliendo il più adatto tra numerosi vettori casuali, ancora di 100 campioni, memorizzati sia in trasmissione che in ricezione. Usando tecniche di questo tipo (CELP; Codebook Excited Linear Prediction) e ritmi di trasmissione di circa 8 kbit/sec , la qualità è sufficiente.

18.3 La codifica del segnale musicale (MP3)

Il segnale musicale ha banda di circa 20 kHz e natura molto più complessa del segnale vocale; se per quest'ultimo sono i meccanismi di generazione a dare lo spunto alle tecniche di codifica, nel caso del segnale musicale ciò non è possibile; si ricorre allora alle caratteristiche del processo uditivo ed in particolare al fenomeno del mascheramento. Se in un segnale è presente un forte tono ad una frequenza f_0 , il rumore che disturba le componenti alle frequenze $f > f_0$ è molto poco percepibile dall'orecchio umano. In sostanza, se appare un tono ad una certa frequenza f_0 , la banda di segnale che si estende per qualche centinaio di Hz prima e per circa 1KHz poi è percepita peggio. Le tecniche di codifica ne profitano (Jayant, N. et al.; Signal compression based on models of human perception, Proc. IEEE, Vol. 81, 10, Oct. 1993 1385 - 1422).

Descriveremo brevemente ora due famiglie di tecniche di codifica utilizzate per il segnale musicale. Con esse si riduce il ritmo di trasmissione di un fattore fino a 8, ma solo con quasi impercipienti perdite di qualità.

In un primo gruppo di metodologie, il segnale viene suddiviso in frequenza in sottobande con un banco ad esempio di 32 filtri polifase di

CAPITOLO 18 IL SEGNALE VOCALE

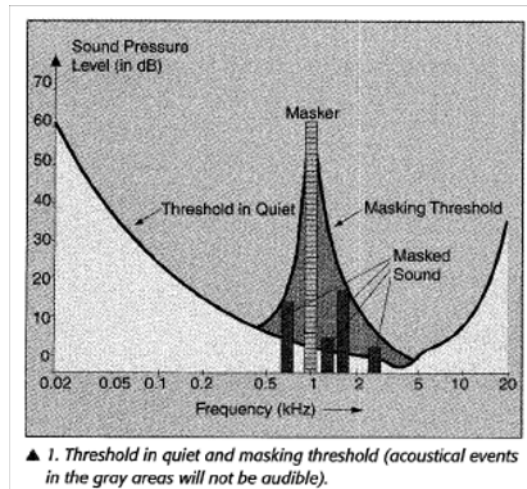


Figura 18.6: L'effetto di mascheramento (da Noll, P.; MPEG digital audio coding, Signal Processing Magazine, IEEE Volume 14, 5, Sept. 1997 pp:59 - 81

larghezza di banda anche non uniforme (le sottobande possono avere larghezza progressivamente crescente con la frequenza, da qualche centinaio di Hz a qualche KHz (fig. 18.6). I segnali in ogni sottobanda sono campionati alla frequenza minima e le componenti che sono equivocate a causa dell'imperfezione dei filtri polifase utilizzati sono poi cancellate dalla ricostruzione, come visto a suo tempo. I filtri usati sono in quadratura a specchio, lunghi 24 campioni, ed hanno un'attenuazione di 60 dB in banda oscura. Le sottobande possono poi essere ancora suddivise in 18 o 36 sotto sotto bande, per sfruttare al meglio il meccanismo di mascheramento. Se ci sono dei transitori (colpi di tamburo, uso del triangolo) si usano filtri meno selettivi per evitare eco anticipatorie.

Dopo il sottocampionamento, i segnali sono quantizzati con tecniche ADPCM. Il predittore usato per l'ADPCM può avere 2 poli e 6 zeri ed è adattativo nel tempo, anche all'indietro, visto che sono accettabili ritardi fino a $3msec$. Il numero di bit per campione (e quindi il ritmo di trasmissione necessario per ogni sottobanda) è determinato, per ogni intervallo di tempo di durata di circa $10 msec$, proprio sulla base della percettibilità del rumore di quantizzazione. Viene fatta cioè un'analisi spettrale (non parametrica) del segnale, viene trovata la frequenza a cui

18.4 CENNI AL RICONOSCIMENTO DEL PARLATO

il segnale ha potenza massima e viene ridotto il numero di bit assegnato a tutte le sottobande al di sopra di questa o per un certo intervallo di frequenza. Poichè l'analizzatore spettrale utilizzato ha una risoluzione in frequenza ben maggiore di 32 (usa una sequenza di circa 1024 campioni) può accadere che la riduzione della quantizzazione preceda, nel tempo, l'attacco del suono che causa il mascheramento. Questo causa delle pre-eco percepibili che vanno previste ed eliminate con metodi opportuni (mascheramento temporale). (H. G. Musmann: Genesis of the MP3 Audio Coding Standard, Consumer Electronics, IEEE Tr. Volume 52, 3, Aug. 2006 pp:1043 - 1049).

In questo modo si riesce a rendere molto meno percettibile il rumore di codifica, malgrado il notevole risparmio di bit; una buona codifica riesce ad avere dei risultati eccellenti con ritmi dell'ordine di $150 - 200 \text{ kbit/sec}$, molto al di sotto del ritmo corrispondente al Compact Disk ($44 \text{ KHz} \times 16 \text{ bit} \times 2(\text{Stereo}) = 1,408 \text{ Mbit/sec}$). Si raggiungono pertanto fattori di riduzione dell'ordine di 8, simili a quelli ottenuti per il segnale vocale.

Esistono poi delle tecniche alternative, perfettamente duali, basate sull'uso della MDCT (Modified Discrete Cosine Transform), vista brevemente nel capitolo dei filtri polifase. Il segnale viene segmentato nei tempi con finestre temporali parzialmente sovrapposte. Il contenuto di ogni finestra viene poi trasformato con una trasformazione simile alla trasformazione di Fourier discreta (una versione tutta reale nota come Modified Discrete Cosine Transform, MDCT). Le componenti della trasformata, (alle varie frequenze), vengono quantizzate con un numero di bit variabile da frequenza a frequenza in funzione del mascheramento e della sensibilità percettiva. Il campionamento in frequenza della MDCT crea un rumore di equivocazione temporale, che però viene cancellato nella fase di ricomposizione del segnale da un rumore eguale ed opposto proveniente dallo stesso ripiegamento che avviene nella finestra successiva. Come si vede, le tecniche sono duali e quindi anche equivalenti nei risultati.

18.4 Cenni al riconoscimento del parlato

Le tecniche di riconoscimento del parlato non sono ancora completamente mature, ma è comunque possibile fare un cenno.

Le prime tecniche utilizzate cercavano di analizzare l'evoluzione delle formanti, per riconoscere in qualche modo le vocali pronunziate dal parla-

CAPITOLO 18 IL SEGNALE VOCALE

tore. Tuttavia, si è visto come queste tecniche fossero abbastanza fragili; si è allora cercato di studiare l'evoluzione temporale dei coefficienti di riflessione ma per fare ciò è necessario:

individuare inizio e fine di ogni parola pronunciata;

infatti, il riconoscimento del parlato continuo è molto piú difficile del riconoscimento fatto parola per parola.

fare uno stiramento dinamico dell'asse temporale per compensare effetti di variazione della velocità di fonazione da parlatore a parlatore; ogni parola viene confrontata con ogni altra parola del vocabolario; le due parole a confronto sono *stirate* in modo da avere un'evoluzione temporale quanto possibile parallela. Infine, si cerca la parola piú simile alla parola pronunciata. Successivamente, la storia del vettore di 10 elementi corrispondente alle LSP è confrontata con quella di altre parole immagazzinate in memoria, fino al riconoscimento.

Capitolo 19

Segnali bidimensionali

19.1 Le singolarità a 2D

Il passaggio da una a due dimensioni non comporta gravi difficoltà per il trattamento dei segnali, a parte i costi computazionali e la memoria occupata, che cresce notevolmente. Dal punto di vista algoritmico, vi sono alcuni elementi innovativi, che cercheremo di riassumere; tratteremo inizialmente i sistemi continui. La prima differenza significativa tra i sistemi 2D e quelli 1D è nelle singolarità: mentre nei sistemi 1D l'unica singolarità era l'impulso, ora troviamo sia *spilli* e cioè impulsi bidimensionali veri e propri, sia *lame* e cioè segnali impulsivi in una dimensione, ma continui nell'altra. Evidentemente, le singolarità a lama non hanno corrispondenza a 1D. Consideriamo un generico segnale a 2D $f(x, y)$: esso può essere **separabile** e cioè

$$f(x, y) = a(x) \times b(y) \quad (19.1)$$

Si vedrà poi che si può sempre rappresentare un segnale qualsiasi come:

$$f(x, y) = \sum_{i=0}^{+\infty} a_i(x) \times b_i(y) \quad (19.2)$$

e quindi la separabilità è un caso particolare od un'approssimazione. Una semplice singolarità a *lama* è un caso particolare di segnale separabile:

$$f_1(x, y) = a_1(x) \times \delta(y) \quad (19.3)$$

CAPITOLO 19 SEGNALI BIDIMENSIONALI

Questa è una lama diretta lungo l'asse delle x e di altezza variabile $a_1(x)$; similmente possiamo pensare a singolarità a lama disposte in altro modo nel piano.

La grande differenza tra i sistemi 2D e quelli 1D nasce dal fatto che in 2D la rapidità del transitorio può dipendere dalla direzione del moto sulla superficie; *si trova sempre un sentiero per salire su una montagna anche se si sfiorano degli abissi*. Oppure, usando un'espressione proverbiale, solo se "si va sul filo di un rasoio" è possibile evitare violenti e indesiderati transitori. Una singolarità a *spillo* è il prodotto di due singolarità a lama:

$$\delta(x, y) = \delta(x) \times \delta(y) \quad (19.4)$$

Da queste definizioni si vede subito che moltiplicare per una lama e integrare a 2D equivale a fare un integrale di linea; invece moltiplicando per uno spillo e ancora integrando a 2D, si preleva il valore della funzione, come accade con le singolarità 1D:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) a_1(x) \delta(y) dx dy = \int_{-\infty}^{+\infty} f(x, 0) a_1(x) dx \quad (19.5)$$

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \delta(x, y) dx dy = f(0, 0) \quad (19.6)$$

Bisogna stare attenti ai fattori di scala delle variabili di cui le $\delta(\cdot)$ sono funzione. Infatti con semplici sostituzioni di variabili, si vede che:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \delta(\alpha x, \beta y) dx dy = \frac{f(0, 0)}{\alpha \beta} \quad (19.7)$$

In 2D ancora, risulta che lo *spillo* che nasce dall'incrocio di due singolarità a lama, dipende anche dall'angolo che esse formano tra loro:

$$I = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \delta[\alpha(x \cos \varphi_0 + y \sin \varphi_0), \beta(x \cos \varphi_1 + y \sin \varphi_1)] dx dy = \quad (19.8)$$

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi, \eta) \delta(\xi, \eta) d\xi d\eta = \frac{f(0, 0)}{\alpha \beta \sin |\varphi_1 - \varphi_0|}$$

La formula precedente è ottenibile con un cambiamento di coordinate 2D, in cui l'elemento d'area del nuovo sistema è ottenuto da quello vecchio moltiplicando per lo Jacobiano J (determinante della matrice di

19.1 LE SINGOLARITÀ A 2D

trasformazione di coordinate, locale) della trasformazione di coordinate. Ricordiamo qui che per un sistema di due funzioni in due variabili, come nel caso di una trasformazione di coordinate a 2D

$$\begin{aligned}\xi &= \xi(x, y) \\ \eta &= \eta(x, y)\end{aligned}$$

lo Jacobiano della trasformazione è:

$$J = \begin{vmatrix} \frac{\partial \xi}{\partial x} & \frac{\partial \xi}{\partial y} \\ \frac{\partial \eta}{\partial x} & \frac{\partial \eta}{\partial y} \end{vmatrix} = \frac{\partial(\xi, \eta)}{\partial(x, y)}$$

La formula può apparire poco comprensibile, ma basta pensare alla superficie (p.e. in m^2) che viene coperta se le due nuove coordinate crescono rispettivamente di $\delta\xi, \delta\eta$. Per esempio, nel caso di passaggio da trasformazione di coordinate da cartesiane a polari, la superficie coperta incrementando le nuove variabili è $\rho\delta\rho\delta\theta$ (*lontano dal centro le aree fabbricabili costano meno*). Infatti, il valore dello Jacobiano è dipendente dalle coordinate:

$$x = \rho \cos \theta; \quad y = \rho \sin \theta$$

$$dx = d\rho \cos \theta - \rho \sin \theta d\theta; \quad dy = d\rho \sin \theta + \rho \cos \theta d\theta$$

$$|J_{cp}| = \frac{\partial(x, y)}{\partial(\rho, \theta)} = \begin{vmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & -\rho \sin \theta \\ \sin \theta & \rho \cos \theta \end{vmatrix}$$

$$|J_{cp}| = \rho(\cos^2 \theta + \sin^2 \theta) = \rho = \sqrt{x^2 + y^2}$$

e pertanto gli elementi di area sono:

$$dxdy \Leftrightarrow \rho d\rho d\theta$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dxdy = \int_{-\pi}^{\pi} \int_0^{\infty} f(\rho, \theta) |\rho| d\rho d\theta$$

indicando in modo poco corretto ma intuibile:

$$f(\rho, \theta) = f(x = \rho \cos \theta, y = \rho \sin \theta)$$

CAPITOLO 19 SEGNALI BIDIMENSIONALI

Nel caso della trasformazione inversa da coordinate cartesiane a polari si ha:

$$\begin{aligned}\rho &= \sqrt{x^2 + y^2} \\ \theta &= \arctan \frac{y}{x}\end{aligned}$$

$$\begin{aligned}|J_{pc}| &= \begin{vmatrix} \frac{\partial \rho}{\partial x} & \frac{\partial \rho}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{vmatrix} \\ |J_{pc}| &= \begin{vmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ -\frac{y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{vmatrix} = \frac{1}{|\rho|} = \frac{1}{|J_{cp}|}\end{aligned}$$

Nel caso dell'incrocio di due lame, si ha invece che lo Jacobiano della trasformazione di coordinate che porta gli assi coordinati sulle due lame è una costante:

$$\xi = \alpha(x \cos \varphi_0 + y \sin \varphi_0); \quad \eta = \beta(x \cos \varphi_1 + y \sin \varphi_1);$$

$$d\xi d\eta = |J| dx dy; \quad |J| = \left| \frac{\partial(\xi, \eta)}{\partial(x, y)} \right| = |\alpha\beta \sin(\varphi_0 - \varphi_1)|$$

La motivazione di questo fattore è resa intuitiva pensando a due lame, di spessore finito Δ e di altezza h , che si intersecano; il volume del prisma di base rombica, loro intersezione e quindi costruito sull'area coperta con l'incremento delle due variabili, è pari a $\Delta^2 h \sin |\varphi_1 - \varphi_0|$ e cresce se le due lame tendono ad essere parallele (fig. 19.1). Un'importante applicazione delle singolarità impulsive 2D (a spillo), è la decomposizione di un segnale 2D in somma di impulsi:

$$f(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi, \eta) \delta(x - \xi, y - \eta) d\xi d\eta$$

19.2 Sistemi lineari 2D

Un sistema è caratterizzato da un ingresso ed un'uscita; se il sistema è lineare, vale la sovrapposizione degli effetti e visto che possiamo decomporre l'ingresso in somma di funzioni anche l'uscita sarà decomponibile nella

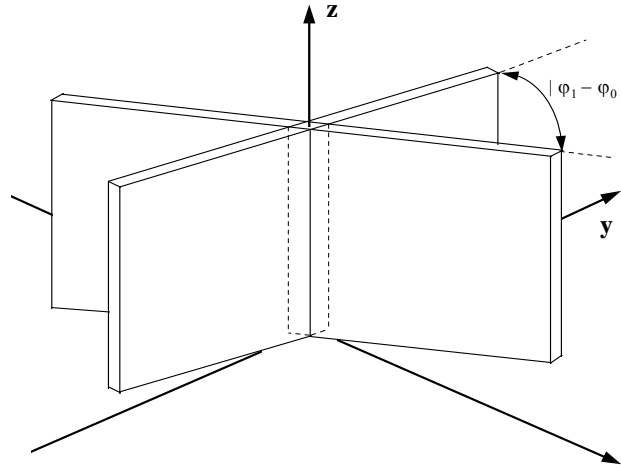


Figura 19.1: Intersezione di due lame di spessore finito

somma delle risposte a quelle funzioni. Consideriamo il caso dell'impulso, posizionato in x_0, y_0 ; in generale, la risposta del sistema è dipendente dalla posizione dell'impulso di ingresso $h(x, y; x_0, y_0)$. Risulta allora, che se all'ingresso di questo sistema vi è la funzione $f(x, y)$ l'uscita è:

$$g(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi, \eta) h(x, y; \xi, \eta) d\xi d\eta \quad (19.9)$$

In generale i sistemi saranno spazio - varianti e cioè la risposta all'impulso dipenderà dalla posizione dell'impulso stesso; per fare un esempio banale, il suono generato battendo una lastra metallica non uniforme dipende dalla posizione del punto di percussione. Se invece il sistema è **spazio invariante** allora la formula precedente si semplifica perchè:

$$h(x, y; x_0, y_0) = h(x - x_0, y - y_0) \quad (19.10)$$

$$g(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi, \eta) h(x - \xi, y - \eta) d\xi d\eta = f(x, y) ** h(x, y) \quad (19.11)$$

Questo è l'**integrale di convoluzione a 2D** che indicheremo con il segno **. La convoluzione di una funzione ed uno spillo centrato nel punto di coordinate x_0, y_0 è la stessa funzione traslata dall'origine nel punto di coordinate x_0, y_0 ; la convoluzione di due lame finite ortogonali è un

CAPITOLO 19 SEGNALI BIDIMENSIONALI

prisma rettangolare (fig. 19.5); la convoluzione di due *sinc* è un *sinc2D* (fig. 19.2).

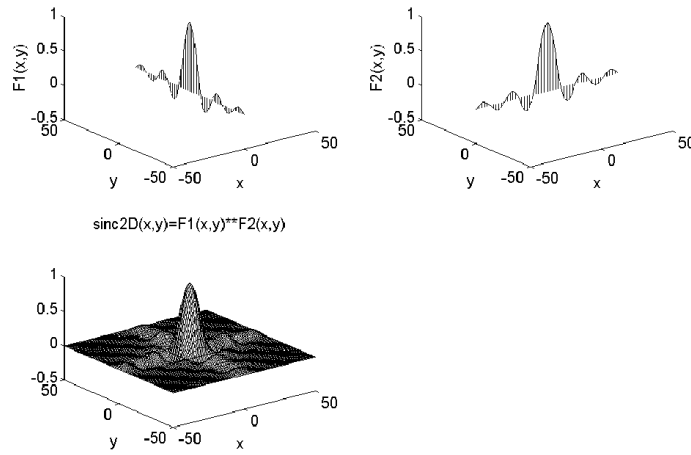


Figura 19.2: Risultato della convoluzione di due lame a forma di *sinc*.

19.2.1 Le funzioni cilindriche

Una funzione 2D che dipende solo da una variabile è detta una funzione cilindrica: ad esempio:

$$f_1(x, y) = a(x); \quad f_2(x, y) = b(x \cos \theta - y \sin \theta) \quad (19.12)$$

La funzione $f_1(x, y)$ è costante nella direzione y , che definiamo come direzione di cilindricità; similmente la funzione $f_2(x, y)$ ha come direzione di cilindricità la retta che forma l'angolo θ con l'asse delle y . Convolvere una funzione qualsiasi con una funzione cilindrica si ottiene ancora una funzione cilindrica, con la stessa direzione di cilindricità:

$$\begin{aligned} g(x) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi, \eta) h(x - \xi) d\xi d\eta = & (19.13) \\ &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(\xi, \eta) d\eta \right) h(x - \xi) d\xi = \int_{-\infty}^{+\infty} p(\xi) h(x - \xi) d\xi \end{aligned}$$

Indicheremo con:

$$p(\xi) = \int_{-\infty}^{+\infty} f(\xi, \eta) d\eta \quad (19.14)$$

la **proiezione** della $f(x, y)$ lungo l'asse y , direzione di cilindricità della $h(x)$. Ovviamente, può essere qualsiasi l'azimut di cilindricità (e cioè l'angolo con l'asse x della direzione lungo cui la funzione $f(x, y)$ è costante).

19.3 Le sinusoidi 2D

Nei sistemi 1D, convolvendo qualsiasi funzione per una sinusoida complessa, otteniamo ancora una sinusoida isofrequenziale, ma di diversa ampiezza complessa. In 2D, avviene la stessa cosa con le funzioni complesse:

$$s(x, y) = e^{j(ux+vy)} \quad (19.15)$$

Infatti:

$$\begin{aligned} g(x, y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi, \eta) e^{j[(u(x-\xi)+v(y-\eta))]} d\xi d\eta = \\ &= e^{j(ux+vy)} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi, \eta) e^{-j(u\xi+v\eta)} d\xi d\eta = e^{j(ux+vy)} F(u, v) \\ F(u, v) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi, \eta) e^{-j(u\xi+v\eta)} d\xi d\eta \end{aligned} \quad (19.16)$$

La funzione $F(u, v)$ è la trasformata di Fourier 2D della funzione $f(x, y)$ e dà la nuova fase ed ampiezza della sinusoida 2D all'uscita del sistema. E' interessante capire la forma di queste sinusoidi 2D: per fare questo passiamo da coordinate cartesiane a coordinate polari:

$$\begin{aligned} x &= \rho \cos \theta; & y &= \rho \sin \theta; & u &= \Omega \cos \varphi; & v &= \Omega \sin \varphi \\ \rho &= \sqrt{x^2 + y^2}; & \theta &= \tan^{-1} \frac{y}{x} & \Omega &= \sqrt{u^2 + v^2}; & \varphi &= \tan^{-1} \frac{v}{u} \\ s(x, y) &= e^{j(ux+vy)}; & s'(\rho, \theta) &= e^{j\Omega\rho \cos(\theta-\varphi)} \end{aligned} \quad (19.17)$$

La funzione $s'(\rho, \theta)$ è un sinusoida di **numero d'onda** (pulsazione spaziale) $\Omega \cos(\theta - \varphi)$ ed è quindi costante nella direzione $\theta = \varphi + \pi/2$ (le sinusoidi

CAPITOLO 19 SEGNALI BIDIMENSIONALI

2D sono quindi funzioni separabili cilindriche, con direzione di cilindricità $\varphi + \pi/2$) ed ha invece la pulsazione massima Ω nella direzione φ ; sostanzialmente è come un *cartone ondulato*; uniforme nella direzione $\varphi + \pi/2$, ondulata nella direzione φ (fig. 19.3).

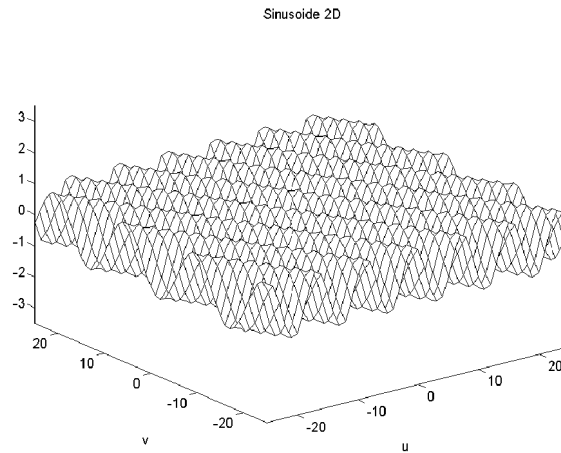


Figura 19.3: Sinusoide 2D (*tettoia ondulata*).

19.4 L'antitrasformazione di Fourier 2D

Come un impulso ideale 1D è somma delle sinusoidi a tutte le frequenze di ampiezza unitaria e fase zero, così un impulso 2D è somma di tutte le sinusoidi 2D a tutti i numeri d'onda, con fase nulla ed ampiezza costante. Infatti:

$$\begin{aligned} \delta(x, y) = \delta(x)\delta(y) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{jux} du \times \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{jvy} dv = \quad (19.18) \\ &= \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{j(ux+vy)} dudv \end{aligned}$$

Combinando la formula di composizione con la formula di trasformazione, se ne deduce la formula di antitrasformazione a 2D:

$$f(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi, \eta) \delta(x - \xi, y - \eta) d\xi d\eta = \quad (19.19)$$

19.4 L'ANTITRASFORMAZIONE DI FOURIER 2D

$$\begin{aligned} \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi, \eta) e^{j[u(x-\xi)+v(y-\eta)]} d\xi d\eta dudv = \\ \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(u, v) e^{j(ux+vy)} dudv \end{aligned}$$

Le formule di trasformazione sono utili anche in coordinate polari:

$$f(\rho, \theta) = \frac{1}{4\pi^2} \int_0^{+\infty} \int_{-\pi}^{+\pi} F(\Omega, \varphi) e^{j\Omega\rho \cos(\theta-\varphi)} \Omega d\Omega d\varphi \quad (19.20)$$

dove è stato inserito lo Jacobiano della trasformazione da coordinate cartesiane a polari, Ω . Nel caso dell'impulso ideale, esso è l'antitrasformata di una costante e pertanto si ha:

$$\delta(\rho) = \frac{1}{4\pi^2} \int_0^{+\infty} \int_{-\pi}^{+\pi} e^{j(\Omega\rho \cos(\theta-\varphi))} \Omega d\Omega d\varphi \quad (19.21)$$

Quest'ultima formula può anche essere interpretata osservando che per comporre un impulso ideale, per ogni numero d'onda Ω serve un *certo numero* di sinusoidi spaziali, componenti, che siano uniformemente distribuite su tutti gli azimut φ ; questo *numero di sinusoidi* cresce proporzionalmente ad Ω così da garantire una densità uniforme delle sinusoidi nel piano.

19.4.1 Funzioni a simmetria circolare

Se la funzione da trasformare è a simmetria circolare, lo sarà anche la trasformata, e viceversa. In questo caso la trasformazione si semplifica e diventa:

$$\begin{aligned} f(\rho) &= \frac{1}{4\pi^2} \int_0^{+\infty} \int_{-\pi}^{+\pi} F(\Omega) e^{j(\Omega\rho \cos(\theta-\varphi))} \Omega d\Omega d\varphi = \quad (19.22) \\ &= \frac{1}{2\pi} \int_0^{+\infty} \Omega F(\Omega) \left(\frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{j(\Omega\rho \cos(\theta-\varphi))} d\varphi \right) d\Omega = \\ &= \frac{1}{2\pi} \int_0^{+\infty} \Omega F(\Omega) J_0(\Omega\rho) d\Omega \end{aligned}$$

dove la parentesi è stata sostituita con il suo valore e cioè la funzione di Bessel di ordine 0, $J_0(\Omega\rho)$. La trasformazione 1D corrispondente all'equazione precedente viene indicata come trasformata di **Hankel**. Segue ora un'utile tabella di trasformate a simmetria circolare.

CAPITOLO 19 SEGNALI BIDIMENSIONALI

<i>Dominio spazi</i>	<i>Frequenze spaziali</i>
$e^{-\pi\rho^2}$	$e^{-\pi\Omega^2}$
$\frac{1}{\sqrt{a^2+\rho^2}}$	$\frac{1}{\Omega}e^{-2\pi a\Omega}$
$\frac{1}{(a^2+\rho^2)^{3/2}}$	$\frac{2\pi}{a}e^{-2\pi a\Omega}$
$\frac{1}{\rho}$	$\frac{1}{\Omega}$
$e^{-a\rho}$	$\frac{2\pi a}{(a^2+4\pi^2\Omega^2)^{3/2}}$
$\frac{e^{-a\rho}}{\rho}$	$\frac{2\pi}{\sqrt{a^2+4\pi^2\Omega^2}}$
$\frac{\delta(\rho)}{2\pi\rho}$	1
$\rho^2 e^{-\pi\rho^2}$	$\left(\frac{1}{\pi} - \Omega^2\right) e^{-\pi\Omega^2}$
$-\rho^2 f(\rho)$	$\frac{d^2 F}{d\Omega^2} + \frac{1}{\Omega} \frac{dF}{d\Omega} = \nabla^2 F$

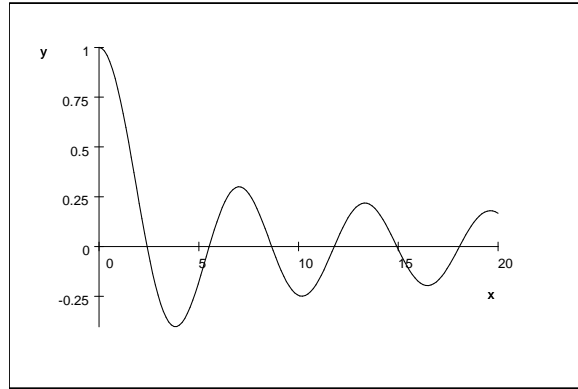
Da: Bracewell (1999) *The Fourier Transform and its applications*, McGraw Hill.

La circonferenza $\delta(\Omega - \Omega_0)$, ha per antitrasformata di Hankel la funzione $J_0(\Omega_0\rho)$ che può apparire simile ad un impulso ideale filtrato passa basso con un filtro ideale a simmetria circolare. L'impulso filtrato passa basso ha trasformata:

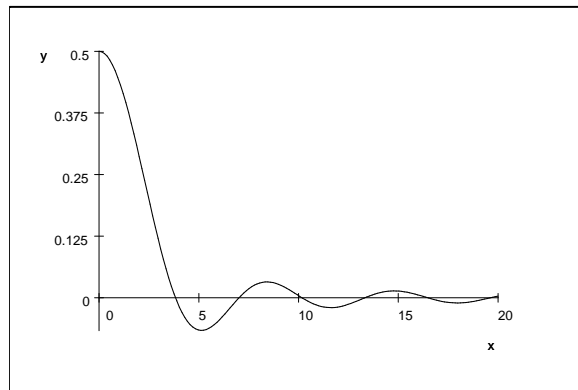
$$\frac{1}{2\pi} \int_0^{\Omega_0} \Omega J_0(\Omega\rho) d\Omega = \frac{\Omega_0 J_1(\Omega_0\rho)}{2\pi\rho} \quad (19.23)$$

Non ci si lasci ingannare: la convergenza a zero delle funzioni $J_m(x)$ (vedi le figure che seguono)

19.4 L'ANTITRASFORMAZIONE DI FOURIER 2D



La funzione di Bessel $J_0(x)$.



La funzione di Bessel $J_1(x)/x$

per $x \rightarrow \infty$, m intero, è solo come

$$J_m(x) \sim \sqrt{\frac{2}{\pi|x|}} \cos\left(x - \frac{m\pi}{2} - \frac{\pi}{4}\right)$$

e quindi un qualsiasi reticolo regolare di impulsi convoluti con $J_0(\Omega_0\rho)$ si *impasta* e sparisce o cambia drasticamente forma. Non così nel caso della 19.23. La cosa si può vedere anche nel dominio delle frequenze: convolvere con un reticolo regolare di impulsi equivale (come vedremo nel prossimo capitolo) a moltiplicare per un reticolo regolare di impulsi nelle frequenze spaziali. Questo reticolo, se moltiplicato per la circonferenza $\delta(\Omega - \Omega_0)$, trasformata della J_0 , si annulla ovunque, tranne al più in qualche posizione particolare.

19.5 Proprietà della trasformazione

Riepiloghiamo alcune delle proprietà della trasformata di Fourier 2D, estensioni di quella 1D; indichiamo rispettivamente con $f(x, y)$ la funzione nel dominio degli spazi e $F(u, v)$ la sua trasformata:

$$f(x, y) \leftrightarrow F(u, v) \quad (19.24)$$

$$f(\alpha x, \beta y) \leftrightarrow \frac{1}{\alpha\beta} F\left(\frac{u}{\alpha}, \frac{v}{\beta}\right); \quad (19.25)$$

$$\begin{aligned} \frac{\partial f(x, y)}{\partial x} &\leftrightarrow juF(u, v); & \frac{\partial f(x, y)}{\partial y} &\leftrightarrow jvF(u, v); \\ \frac{\partial^2 f(x, y)}{\partial x \partial y} &\leftrightarrow -uvF(u, v) \end{aligned} \quad (19.26)$$

Al solito, le relazioni di derivazione nascono da un'integrazione per parti. Se $f(x, y)$ è reale:

$$F(u, v) = F^*(-u, -v) \quad (19.27)$$

$$f(\rho, \theta + \psi) \leftrightarrow F(\Omega, \varphi + \psi) \quad (19.28)$$

La relazione di Parseval diventa:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |f(x, y)|^2 dx dy = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |F(u, v)|^2 du dv \quad (19.29)$$

Se abbiamo un'equazione differenziale alle derivate parziali, definita su tutto lo spazio (le condizioni al contorno sono all'infinito), essa corrisponde alla seguente equazione algebrica sulla trasformata:

$$\sum_{i,h=0}^{I,H} \frac{\partial^i \partial^h f(x, y)}{\partial x^i \partial y^h} \leftrightarrow \sum_{i,h=0}^{I,H} (ju)^i (jv)^h F(u, v) \quad (19.30)$$

19.5.1 Funzioni cilindriche e separabili

La trasformata di Fourier di un impulso 2D posizionato in x_0, y_0 è il piano di fase:

$$\delta(x - x_0, y - y_0) \leftrightarrow e^{-j(u x_0 + v y_0)} \quad (19.31)$$

Una cosinusoidale ha invece trasformata:

$$\cos(u_0 x + v_0 y + \varphi) \leftrightarrow \frac{1}{2} [e^{-j\varphi} \delta(u - u_0, v - v_0) + e^{+j\varphi} \delta(u + u_0, v + v_0)] \quad (19.32)$$

19.5 PROPRIETÀ DELLA TRASFORMAZIONE

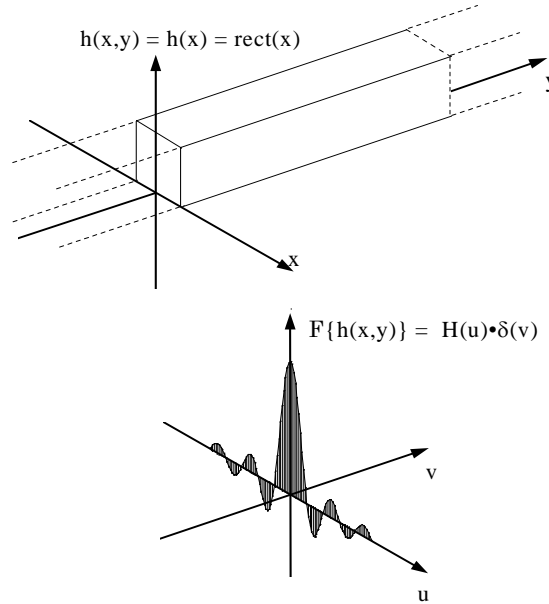


Figura 19.4: Trasformata di Fourier di una funzione cilindrica

La trasformata di Fourier di una funzione cilindrica è una *lama* nella variabile coniugata alla variabile lungo cui la funzione cilindrica è uniforme: ad esempio:

$$h(x, y) = h(x) \leftrightarrow \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x) e^{-j(ux+vy)} dx dy = H(u) \delta(v) \quad (19.33)$$

La superficie cilindrica $h(x)$ è costante nella direzione y e quindi nel dominio coniugato troviamo una funzione impulsiva in quella direzione (lama allungata lungo le u). Una lama uniforme, caso particolare di funzione cilindrica, ha per trasformata un'altra lama uniforme disposta lungo la coniugata della variabile ortogonale (una lama lungo le x ha come trasformata una lama lungo le v). La trasformata di una somma di due lame ortogonali è quindi ancora una somma di due lame ortogonali. Ancora, una stella di lame sarà eguale alla sua trasformata (a meno di una rotazione di $\pi/2$). La somma di infinite lame, come si vedrà al capitolo sulla tomografia, è una funzione del tipo $1/\rho$ ed ha come trasformata la stessa funzione $1/\Omega$.

Una funzione separabile, prodotto di due funzioni 1D, ha come trasformata 2D il prodotto delle due trasformate 1D.

19.6 Filtri ideali 2D

I filtri ideali a 2D hanno una funzione di trasferimento unitaria in alcune zone del dominio u, v e nulla altrove. A differenza che ad 1D, la forma del filtro potrà cambiare. Il caso semplice è quello del filtro a base rettangolare e cioè:

$$F(u, v) = 1; |u|, |v| < \frac{\pi}{\Delta} \quad (19.34)$$

$$F(u, v) = 0; |u|, |v| > \frac{\pi}{\Delta}$$

Questo è un filtro separabile: infatti è ottenibile come la cascata di due filtri passa basso ideali cilindrici in u, v successivamente (fig. 19.5). Un

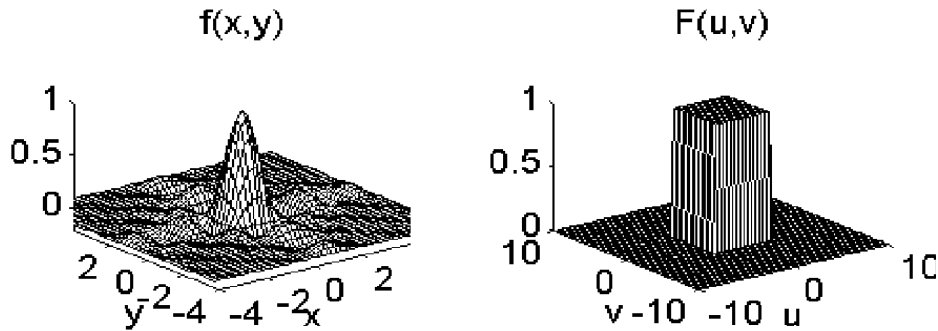


Figura 19.5: Trasformata 2D del *sinc*

filtro solo nelle u ha funzione di trasferimento a lama lungo le x e quindi ritroviamo il risultato già visto che:

$$F(u, v) \leftrightarrow f(x, y) = \frac{1}{\Delta^2} \frac{\sin(\pi x/\Delta)}{(\pi x/\Delta)} \frac{\sin(\pi y/\Delta)}{(\pi y/\Delta)} \quad (19.35)$$

Esiste però anche il filtro ideale circolare e cioè:

$$H(\Omega) = 1; \Omega < \Omega_0; \leftrightarrow h(\rho) = \Omega_0 \frac{J_1(\Omega_0 \rho)}{\rho} \quad (19.36)$$

Dove $J_1(\cdot)$ è la funzione di Bessel di ordine 1 (fig. 19.6).

19.7 FILTRI A VENTAGLIO

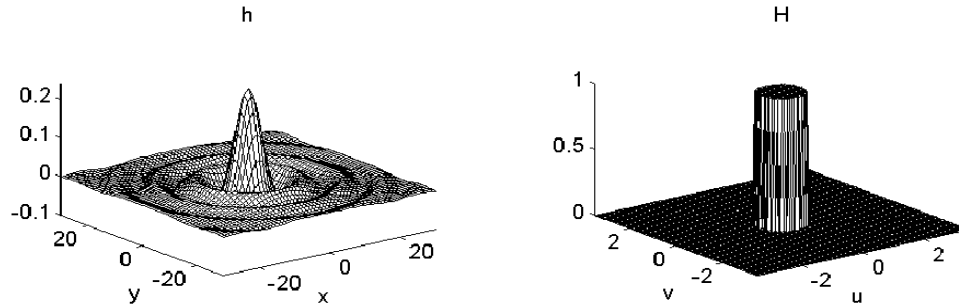


Figura 19.6: Il cilindro, trasformata 2D della funzione di Bessel di ordine 1

19.7 Filtri a ventaglio

Un altro tipo di filtri ideali inesistenti a 1D sono i filtri a ventaglio (fig. 19.7) e cioè filtri nulli in un settore angolare ed unitari altrove. Con

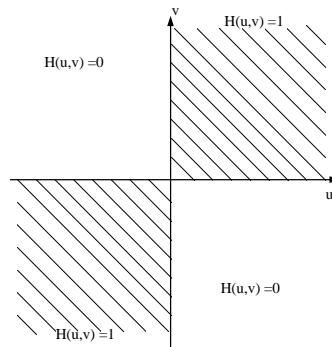


Figura 19.7: Funzione di trasferimento del filtro quadrantale

una cascata di rotazioni e scalamenti un qualsiasi filtro a ventaglio è trasformabile in un filtro quadrantale, nullo nel II e IV quadrante ed unitario altrove. Infatti, basta prima ruotare il filtro a ventaglio in modo che l'asse della zona "passa" sia ad esempio l'asse x. Poi, si scala l'asse y in modo che il settore angolare copra un angolo pari a $\pi/2$ ed infine lo si ruota di $\pi/4$ per portare il filtro a coincidere con il I e III quadrante.

CAPITOLO 19 SEGNALI BIDIMENSIONALI

Questo filtro ha per funzione di trasferimento la somma di due funzioni separabili (fig. 19.8:

$$H(u, v) = 1; \quad uv > 0; \quad H(u, v) = 0; \quad uv < 0 \quad (19.37)$$

$$H(u, v) = \frac{1}{2}[1 + \text{sign}(u) \times \text{sign}(v)] \leftrightarrow h(x, y) \quad (19.38)$$

$$h(x, y) = \frac{1}{2} \left[\delta(0, 0) - \frac{1}{\pi x} \times \frac{1}{\pi y} \right] \quad (19.39)$$

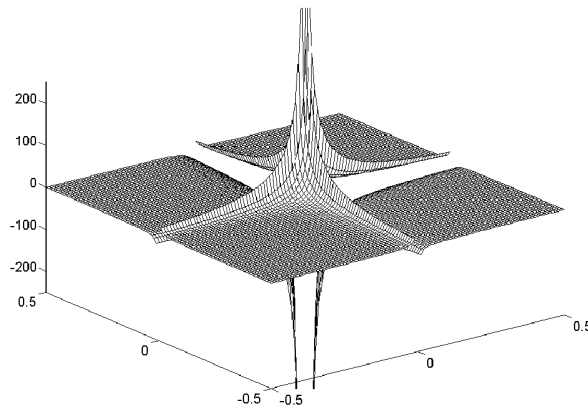


Figura 19.8: Risposta impulsiva del filtro quadrantale

Capitolo 20

Filtri 2D di natura fisica

Per formare un'intuizione dei meccanismi del filtraggio 2D, mostreremo ora come diversi fenomeni fisici possano essere interpretati come filtri lineari; in particolare ci riferiremo a delle equazioni differenziali alle derivate parziali ellittiche, iperboliche e paraboliche che origineranno funzioni di trasferimento a fase zero, di pura fase, e miste. Lo studente interessato è invitato a osservare le antitrasformate dei vari filtri qui derivati utilizzando ad esempio *MATLABTM*.

20.1 L'equazione dei potenziali

Si consideri il seguente problema fisico tridimensionale: nel sottosuolo (il semispazio $z < 0$) vi siano dei corpi di densità differente da quella del materiale che li circonda. Inoltre il semispazio $z > 0$ è vuoto (trascuriamo la densità dell'atmosfera; fig. 20.1). Per determinare la posizione e grandezza dei corpi disomogenei, si vuole determinare l'andamento del potenziale gravitazionale $f(x, y, z)$ a livello del terreno, e cioè per $z = 0$. Se il terreno è poco praticabile, perchè vi sono boschi, paludi ecc. è più agevole misurare il potenziale ad una certa quota z_0 , per esempio con misure eliportate, o aeroportate. Poi, si cercherà un algoritmo (un filtraggio lineare) per calcolare $f(x, y, z = 0)$ a partire da $f(x, y, z_0)$ (continuazione in basso di un campo gravitazionale). La funzione potenziale $f(x, y, z)$ in zone prive di sorgenti soddisfa l'equazione armonica e cioè:

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0 \quad (20.1)$$

CAPITOLO 20 FILTRI 2D DI NATURA FISICA

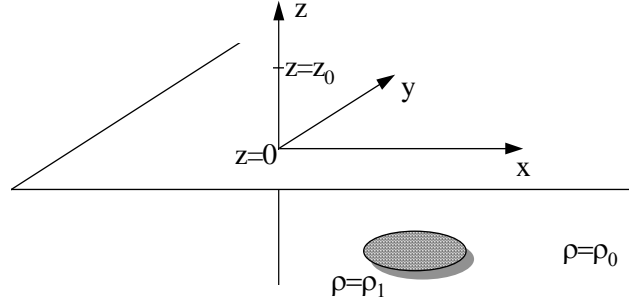


Figura 20.1: Modello di semispazio con materiali a densità diversi

Si desidera appunto, conosciuta $f(x, y, z_0); \forall x, y$ trovare $f(x, y, z = 0); \forall x, y$ e viceversa (continuazione in alto). Poichè la funzione $f(x, y, z_0)$ è conosciuta per tutti i valori della x e della y , possiamo trasformare gli elementi dell'equazione (20.1) secondo Fourier e ne risulta l'equazione differenziale semplice, in cui le variabili u, v coniugate ad x, y , sono dei parametri:

$$-(u^2 + v^2)F(u, v, z) + \frac{d^2 F(u, v, z)}{dz^2} = 0; \quad F(u, v, z) \leftrightarrow f(x, y, z) \quad (20.2)$$

L'equazione ha la soluzione:

$$F(u, v, z) = A(u, v)e^{-\Omega \Delta z} + B(u, v)e^{\Omega \Delta z} \quad (20.3)$$

$$\Delta z = z - z_0; \quad \Omega = \sqrt{u^2 + v^2}$$

Imponiamo $B(u, v) = 0$ per evitare che la soluzione diverga per $z \rightarrow \infty$, e ne risulta $A(u, v) = F(u, v, z_0)$; la soluzione è:

$$\begin{aligned} F(u, v, z) &= F(u, v, z_0)e^{-\Omega \Delta z} = \\ &= F(u, v, z_0) \times H(u, v, \Delta z); \end{aligned} \quad (20.4)$$

Decomponendo con la trasformazione di Fourier 2D il campo misurato a quota 0 in sinusoidi 2D, possiamo ottenere il campo misurabile a quota z_0 purchè l'ampiezza di ogni componente (appunto $F(u, v, z_0)$) sia scalfata del fattore $e^{-\Delta z \Omega}$. In sostanza, la continuazione in basso o in alto del campo è ottenuta filtrando linearmente la funzione potenziale con

20.1 L'EQUAZIONE DEI POTENZIALI

un filtro 2D a simmetria circolare $H(\Omega, \Delta z)$ funzione della differenza di quote Δz . Si osservi che la funzione di trasferimento $H(\Omega)$ è a fase zero, come dettato dalla natura del problema. Il filtro è passa basso se la continuazione è verso l'alto; se ci si allontana dalle sorgenti il potenziale è sempre più smussato fino ad essere uniforme a distanze molto grandi rispetto alle dimensioni dei corpi. Se invece ci si avvicina alle sorgenti, il filtro è passa alto e amplifica esponenzialmente le componenti ai numeri d'onda più elevati. Su questo punto torneremo nel prossimo paragrafo, perchè l'inevitabile presenza di rumore modifica la situazione. Il filtro passa basso ha la sua antitrasformata:

$$H(u, v, \Delta z) = e^{-\Delta z \sqrt{u^2 + v^2}} = e^{-\Omega \Delta z} \quad (20.5)$$

$$e^{-\Omega \Delta z} \leftrightarrow \frac{\Delta z}{(x^2 + y^2 + \Delta z^2)^{3/2}}; \quad \Delta z \geq 0$$

Non così il filtro inverso $\Delta z < 0$ che è divergente per numeri d'onda u, v abbastanza grandi. Utilizziamo questo esempio per mostrare come la tecnica di stima lineare viene in aiuto anche in questo caso per condurre a soluzioni significative.

20.1.1 Continuazione in basso e stima lineare

Supponiamo che la misura $f_m(x, y, z_0)$ del potenziale alla quota z_0 sia contaminata da rumore:

$$f_m(x, y, z_0) = f(x, y, z_0) + n(x, y) \quad (20.6)$$

Vogliamo stimare in modo lineare ottimo $f(x, y, z = 0)$ e sia $\hat{f}(x, y, z = 0)$ questa stima. Facciamo la stima nel dominio dei numeri d'onda, per semplicità. Trasformando l'equazione precedente si ha:

$$F_m(u, v, z_0) = F(u, v, 0)e^{-z_0\Omega} + N(u, v) \quad (20.7)$$

$$\hat{F}(u, v, 0) = H(u, v)F_m(u, v, z_0) \quad (20.8)$$

Vogliamo

$$E[(H(u, v)F_m(u, v, z_0) - F(u, v, 0))F_m^*(u, v, z_0)] = 0 \quad (20.9)$$

CAPITOLO 20 FILTRI 2D DI NATURA FISICA

e cioè incorrelazione tra errore di stima e dati.

$$H(u, v) = \frac{E[F_0 F_m^*]}{E[|F_m|^2] + \lambda_N^2}; \quad \lambda_N^2 = 4\Delta^2 \sigma_N^2 \quad (20.10)$$

dove Δ è l'intervallo di campionamento spaziale, e λ_N^2 , σ_N^2 sono rispettivamente la densità spettrale monolaterale del rumore ed il suo valore quadratico medio.

$$E[|F_m|^2] = E[|F_0|^2]e^{-2z_0\Omega} + \lambda_N^2; \quad E[F_0 F_m^*] = E[|F_0|^2]e^{-z_0\Omega} \quad (20.11)$$

$$H(u, v) = \frac{e^{-z_0\Omega}}{e^{-2z_0\Omega} + Q}; \quad Q = \frac{\lambda_N^2}{E[|F_0|^2]} \quad (20.12)$$

Ai numeri d'onda per cui $Qe^{2z_0\Omega} \ll 1$ il filtro opererà come filtro inverso e quindi passa alto; invece se $Qe^{2z_0\Omega} > 1$ il filtro si comporterà ancora come passa basso (figura 20.2).

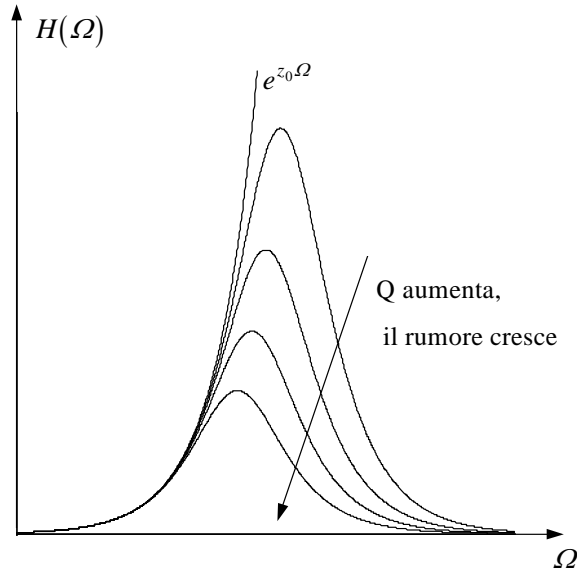


Figura 20.2: Filtri inversi per la continuazione in basso di potenziali

20.2 La propagazione di onde

In questo paragrafo discuteremo di un altro problema fisico, che condurrà allo studio di una funzione di trasferimento a fase pura. Il problema è quello della propagazione di onde in spazio libero: ad esempio le onde superficiali su un lago, oppure le onde acustiche nello spazio. L'equazione che governa la propagazione è

$$\frac{\partial^2 f(x, y, z)}{\partial x^2} + \frac{\partial^2 f(x, y, z)}{\partial y^2} + \frac{\partial^2 f(x, y, z)}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2 f(x, y, z)}{\partial t^2} \quad (20.13)$$

dove c è la velocità di propagazione delle onde. Anzitutto, eliminiamo la variabile z supponendo che vi sia uniformità in quella direzione e pertanto

$$\frac{\partial^2 f(x, y, z)}{\partial z^2} = 0 \quad (20.14)$$

Si desidera ora, noto ad esempio $f(x, y, t = 0); \forall x, y$ trovare $f(x, y, t); \forall x, y$ e quindi l'evoluzione temporale di un campo d'onde, in assenza di sorgenti, a partire da delle condizioni iniziali. Un altro problema che risolveremo (al prossimo paragrafo) è quello del telerilevamento e cioè:

determinare $f(x, y, t); \forall x, y, t$ dato: $f(x, y = 0, t); \forall x, t$

Come già fatto nel caso precedente, trasformiamo l'equazione nelle variabili note su tutto il dominio, da $-\infty \div \infty$. Poichè, nel primo caso dello studio dell'evoluzione temporale di un campo d'onde, la funzione è nota per tutti i valori della x e della y , possiamo trasformare in Fourier per quelle variabili e risulta l'equazione differenziale semplice:

$$-(u^2 + v^2)F(u, v, t) + \frac{1}{c^2} \frac{d^2 F(u, v, t)}{dt^2} = 0; \quad F(u, v, t) \leftrightarrow f(x, y, t) \quad (20.15)$$

$$F(u, v, t) = A(u, v)e^{jct\Omega} + B(u, v)e^{-jct\Omega} \quad (20.16)$$

La soluzione è somma di due termini che corrispondono rispettivamente alle onde che esplodono ed implodono. In particolare, la risposta all'impulso corrispondente alle onde che esplodono e cioè si allontanano dalla sorgente sarà:

$$F_0(u, v, t) = e^{j\Omega ct} \quad (20.17)$$

CAPITOLO 20 FILTRI 2D DI NATURA FISICA

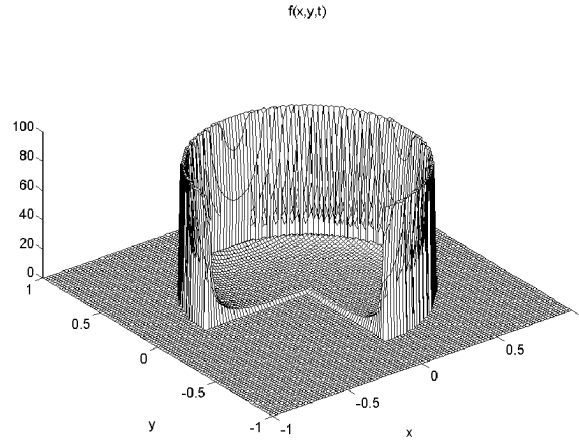


Figura 20.3: Sezione di un'onda circolare bidimensionale (p.e. un *cerchio sull'acqua*)

la cui antitrasformata è (figura 20.3):

$$f(x, y, t) = \frac{ct}{\sqrt{c^2t^2 - y^2 - x^2}^3}; \quad x^2 + y^2 \leq c^2t^2 \quad (20.18)$$

Questa risposta all'impulso corrisponde p.e. a dei *cerchi nell'acqua t secondi dopo avere lanciato un sasso in uno stagno*; piú precisamente, si può pensare ad un tuono, dopo un lampo (una sorgente rettilinea di perturbazione, di lunghezza praticamente illimitata in una direzione, che indichiamo con z). La perturbazione generata è contenuta in un cerchio di raggio ct nel caso del sasso nello stagno o in un cilindro di raggio ct diretto come z nel caso del lampo. Si osservi poi che vi e' differenza nella risposta ad un'eccitazione impulsiva nei casi 2D e 3D. Infatti, nel caso 3D la propagazione dell'onda generata da un impulso puntiforme avviene lungo un fronte sferico, e la perturbazione non permane, dopo che sia passato il primo fronte d'arrivo. Si pensi ora al caso 2D, nel caso della sorgente di perturbazione rettilinea; il piano di osservazione del fenomeno è ottenuto sezionando lo spazio 3D ortogonalmente all'asse z , direzione della linea delle sorgenti. Disponendo un ricevitore in questo piano, e applicando la sovrapposizione degli effetti, si comprende che questo ricevitore misurerà anzitutto l'arrivo delle perturbazione dovuta

20.2 LA PROPAGAZIONE DI ONDE

alla zona di sorgente situata nel piano. Successivamente, il ricevitore misurerà anche gli arrivi generati dalle zone della sorgente lineare situate al di fuori del piano che contiene il ricevitore (si pensi p.e. al *brontolio del tuono*). Quindi la forma del segnale che corrisponde al transitorio del passaggio dell'onda impulsiva in un sistema 2D **non è** una muraglia cilindrica, anche se il segnale si concentra intorno ad essa.

20.2.1 Il problema del telerilevamento

Il problema che ora abbiamo è il seguente: supponiamo p.e. di assistere alla rappresentazione di un'opera lirica; possiamo disporre dei microfoni lungo il bordo del palcoscenico (l'asse $x, y = 0$) e vogliamo combinare le loro uscite in modo da potere isolare la voce di qualsiasi cantante (le sorgenti s_1, s_2), come se disponessimo di un microfono in un punto qualsiasi della scena ($x, y \neq 0$) (fig. 20.4). Questo problema corrisponde al rilevamento e alla focalizzazione da lontano del segnale generato in un punto qualsiasi dello spazio. Supponendo che il campionamento spaziale (lungo il palcoscenico) e temporale sia fittissimo, misuriamo la funzione $f(x, y = 0, t); \forall x, t$; cerchiamo $f(x, y, t); \forall x, y, t$. L'equazione che utilizzeremo è ancora quella della propagazione:

$$-(u^2 - \frac{\omega^2}{c^2})F(u, y, \omega) + \frac{d^2 F(u, y, \omega)}{dy^2} = 0; \quad (20.19)$$

ma ora abbiamo trasformato le variabili x, t ; le variabili coniugate sono u, ω . Risolviamo l'equazione differenziale e troviamo il risultato desiderato e cioè che:

$$F(u, y, \omega) = A(u, \omega)e^{jy\sqrt{\frac{\omega^2}{c^2}-u^2}} + B(u, \omega)e^{-jy\sqrt{\frac{\omega^2}{c^2}-u^2}} \quad (20.20)$$

Di nuovo, le due soluzioni corrispondono ad onde che si avvicinano o si allontanano dall'asse $y = 0$ (il proscenio): per il problema che vogliamo risolvere, scegliamo le onde che si avvicinano e quindi anticipi per y crescenti.

$$F(u, y, \omega) = F(u, y = 0, \omega)e^{j\frac{\omega y}{c}\sqrt{1-\frac{u^2 c^2}{\omega^2}}} \quad (20.21)$$

Trovata $F(u, y, \omega)$, la funzione cercata $f(x, y, t)$ sarà ottenuta con una antitrasformazione di Fourier 2D.

$$f(x, y, t) = f(x, y = 0, t) * *FT^{-1} \left(e^{j\frac{\omega y}{c}\sqrt{1-\frac{u^2 c^2}{\omega^2}}} \right)$$

CAPITOLO 20 FILTRI 2D DI NATURA FISICA

Basta cioè convolvere la funzione $f(x, y, t)$ con un filtro opportuno, dipendente dalla variabile y .

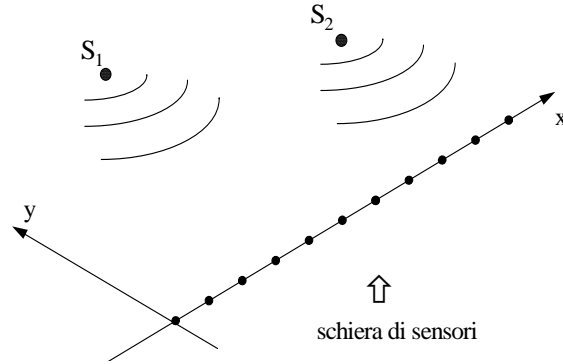


Figura 20.4: Disposizione di microfoni su un palcoscenico per isolare i suoni emessi dai singoli cantanti

20.2.2 Le onde piane

Il filtro dell'equazione precedente ha un'espressione complicata, che tuttavia può essere spiegata introducendo il concetto delle onde piane, già implicitamente utilizzato nel capitolo sulle Direzioni di Arrivo (DOA). Nel dominio x, t le sinusoidi bidimensionali sono segnali del tipo:

$$s(x, t) = Ae^{j(kx+\omega t)} \quad (20.22)$$

Come già visto nel capitolo 12 questi segnali corrispondono ad onde monocromatiche di ampiezza complessa A e di pulsazione ω , i cui fronti d'onda sono delle rette nel piano x, y che formano un angolo θ con l'asse delle ascisse x . Infatti, due punti a distanza $2\pi/k$ lungo l'asse delle x , oscillano in fase e pertanto sono separati da un tempo di propagazione $\tau = 2\pi/\omega$; si ha allora:

$$\begin{aligned} \frac{2\pi}{k} \sin \theta &= c \frac{2\pi}{\omega} = c\tau; \\ \sin \theta &= \frac{kc}{\omega}; \quad \cos \theta = \sqrt{1 - \frac{k^2 c^2}{\omega^2}} \end{aligned} \quad (20.23)$$

20.2 LA PROPAGAZIONE DI ONDE

Si osservi che se:

$$|k| > \left| \frac{\omega}{c} \right| = \frac{2\pi}{\lambda} \quad (20.24)$$

l'angolo θ diventa una variabile immaginaria; le onde non si propagano più, ma diventano evanescenti. Per capire questo fenomeno, basti pensare a più sorgenti puntiformi monocromatiche (a frequenza ω) disposte lungo l'asse x a distanza δ e irradianti insieme; inoltre, la fase dell' n -esima sorgente è ruotata di $e^{jkn\delta}$. Se la diseuguaglianza (20.24) è rispettata, il campo emesso non è rappresentabile come un'onda piana che fa un angolo θ con l'asse x ; le varie sorgenti interferiscono tra loro in modo distruttivo e il campo si riduce progressivamente (in modo esponenziale) via via che ci si allontana dall'asse x . Nel caso semplice in cui $\omega = 0$, siamo di nuovo al caso dell'elettrostatica o della gravità. Infatti nell'equazione 20.13 la derivata rispetto al tempo è nulla, non c'è propagazione, e siamo di nuovo all'equazione armonica.

Ritornando al problema del paragrafo precedente, per ricostruire il campo d'onda $f(x, y, t)$ in posizioni $y \neq 0$ a partire da $f(x, y = 0, t)$ si può farne anzitutto una decomposizione in onde piane. Questa è un'interpretazione della trasformazione di Fourier della $f(x, y = 0, t)$; le ampiezze delle onde piane a pulsazione k, ω sono appunto i valori $F(k, y = 0, \omega)$. Ogni onda piana si propaga in avanti o all'indietro nello spazio senza variazioni di ampiezza, ma subendo solo uno sfasamento. Essa quindi può venire retropropagata nello spazio (e/o nel tempo) fino alla posizione della sorgente applicando una rotazione di fase che dipende dalla pulsazione ω e dalla distanza y di cui deve essere retropropagata. La rotazione è :

$$\omega\tau = \frac{\omega y \cos \theta}{c} \quad (20.25)$$

Quindi per ricostruire il campo all'ascissa y basta retropropagare le onde sfasandole del fattore:

$$e^{j\omega\tau} = e^{j\frac{\omega y}{c} \cos \theta} = e^{j\frac{\omega y}{c} \sqrt{1 - \frac{k^2 c^2}{\omega^2}}} \quad (20.26)$$

come risulta anche dalla figura 20.5; la combinazione di tutte le onde piane dopo la loro retropropagazione ci dá la distribuzione delle sorgenti nello spazio e pertanto l'immagine cercata.

L'antitrasformata della funzione

$$e^{-jy\sqrt{\frac{\omega^2}{c^2} - k^2}} \quad (20.27)$$

CAPITOLO 20 FILTRI 2D DI NATURA FISICA

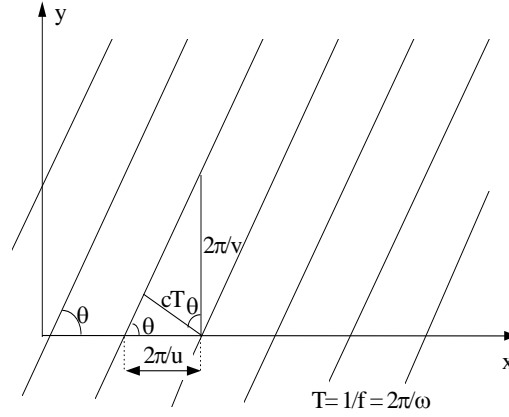


Figura 20.5: Retropropagazione di onde piane

ha la forma presentata in figura 20.6; essa corrisponde a quanto viene osservato lungo la retta $y = 0$ in corrispondenza di un impulso (una sorgente esplosiva) al tempo ed alla posizione $x = 0, t = 0, y$; il primo arrivo dell'onda all'ascissa x si ha al tempo:

$$t(x) = \frac{\sqrt{x^2 + y^2}}{c} \quad (20.28)$$

Quindi, nel piano x, t questi arrivi sono disposti lungo un'iperbole di equazione:

$$x^2 - c^2 t^2 = y^2 = c^2 t_0^2 \quad (20.29)$$

Disponendo ora della funzione $f(x, y = 0, t)$ possiamo ora determinare la funzione $f(x, y, t = 0)$ e cioè intuitivamente, *determinare dalle onde che arrivano alla spiaggia la distribuzione delle sorgenti in alto mare* oppure *potere appendere un microfono al collo di Pavarotti, malgrado che tutti i microfoni siano sul proscenio*. Per fare questo partiamo dall'equazione (20.21)

$$F(k, y = 0, \omega) = \int \int f(x, y = 0, t) e^{-j(kx + \omega t)} dt dx$$

$$f(x, y, t) = \int \int F(k, y = 0, \omega) e^{jy\sqrt{\frac{\omega^2}{c^2} - k^2} + jkx + j\omega t} \frac{dk d\omega}{(2\pi)^2}$$

20.2 LA PROPAGAZIONE DI ONDE

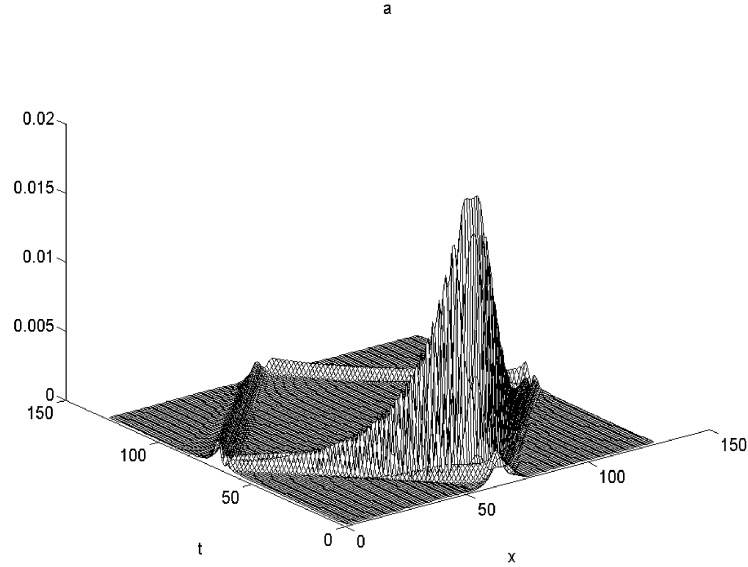


Figura 20.6: L'iperbole corrispondente ai tempi di arrivo di un segnale impulsivo ai vari microfoni sul palcoscenico; sono visibili effetti di periodicizzazione dell'asse delle ascisse dovute all'uso di DFT

(Il microfono al collo di Pavarotti). Ancora si ha:

$$f(x, y, t = 0) = \int \int F(k, y = 0, \omega) e^{jy\sqrt{\frac{\omega^2}{c^2} - k^2} + jkx} \frac{dkd\omega}{(2\pi)^2} \quad (20.30)$$

(Questo corrisponderebbe alla distribuzione delle posizioni, al largo, dove sono stati lanciati sassi al tempo 0, determinata sulla base delle onde che arrivano sulla spiaggia). Per ottenere, invece dell'espressione complessa (20.30), una semplice antitrasformata di Fourier 2D, basta effettuare la seguente trasformazione di coordinate:

$$\sqrt{\frac{\omega^2}{c^2} - k^2} = v \rightarrow \omega = c\sqrt{k^2 + v^2}; \quad d\omega = \frac{2cv dv}{\sqrt{k^2 + v^2}}$$

$$f(x, y, t = 0) = \int \int F(k, y = 0, \omega = c\sqrt{k^2 + v^2}) e^{j(vy+kx)} \frac{2cv}{\sqrt{k^2 + v^2}} \frac{dkdv}{(2\pi)^2}$$

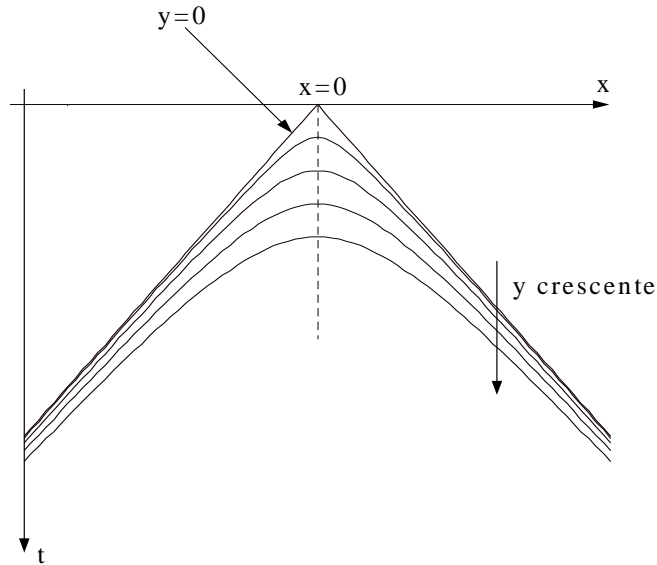


Figura 20.7: Iperboli dei tempi di arrivo ottenute con il principio della fase stazionaria

In altri termini, data la $F(k, y = 0, \omega)$, la si interpola nei punti k_i, ω_l corrispondenti ai valori di k_i, ω_l richiesti dal campionamento effettuato in queste due coordinate, e la si antitrasforma (Interpolazione di Stolt).

20.3 Il principio della fase stazionaria

Se si hanno delle complicate antitrasformazioni da compiere, è utile il principio della fase stazionaria. Consideriamo un integrale di Fourier del tipo:

$$Q(x) = \int_{-\infty}^{\infty} A(x, \xi) e^{jMG(x, \xi)} d\xi \quad (20.31)$$

dove M è un numero grande; se la fase dell'esponenziale scorre rapidamente l'integrale (una media) avrà valore assai basso; gli unici valori della variabile x per i quali l'integrale assume valori significativi si hanno quando **la fase è stazionaria** e cioè quando:

$$\frac{\partial G(x, \xi)}{\partial \xi} = 0 \quad (20.32)$$

20.4 IL *CHIRP*, LALENTE SOTTILE E L'OTTICA DI FOURIER

Ovviamente, ξ può anche essere una variabile multideimensionale. Ad esempio, nel caso dell'antitrasformata di Fourier:

$$\frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-j\frac{\omega y}{c} \sqrt{1 - \frac{u^2 c^2}{\omega^2}} + ux + \omega t} dud\omega \quad (20.33)$$

risulta:

$$G = -t_0 \omega \sqrt{1 - \frac{u^2 c^2}{\omega^2}} + ux + \omega t \quad (20.34)$$

$$\frac{\partial G(x, t; u, \omega)}{\partial \omega} = -\frac{\omega t_0}{\sqrt{\omega^2 - u^2 c^2}} + t = 0 \quad (20.35)$$

$$\frac{\partial G(x, t; u, \omega)}{\partial u} = \frac{t_0}{\sqrt{\omega^2 - u^2 c^2}} uc^2 + x = 0 \quad (20.36)$$

Da cui, eliminando ω , u :

$$\left(\frac{x}{c}\right)^2 - t^2 = t_0^2 \quad (20.37)$$

si trova appunto la (20.29). In sostanza, l'antitrasformata è nulla ovunque, a parte dove la 20.37 è verificata, e cioè lungo l'iperbole. Il principio della fase stazionaria si applica se $\omega t_0 \gg 1$ e quindi se i raggi percorrono un numero elevato di lunghezze d'onda e pertanto non siamo in condizione di "campo vicino".

20.4 Il *chirp*, la lente sottile e l'ottica di Fourier

Un altro caso di interesse è quello del *chirp* bidimensionale. Il nome *chirp* viene dal cinguettio di un uccello in cui la frequenza del suono emesso cresce progressivamente. Il chirp nel tempo è quindi corrispondente ad un segnale caratterizzato da una frequenza crescente linearmente con il tempo e quindi da una fase crescente quadraticamente con il tempo. La sua trasformata è ancora un chirp; si ricorda infatti la coppia di trasformate (vedi al capitolo 9):

$$c(t) = \exp\left(-j\frac{\pi t^2}{\tau^2}\right) \leftrightarrow C(f) = \tau \exp\left(-j\pi\tau^2 f^2 - j\frac{\pi}{4}\right)$$

CAPITOLO 20 FILTRI 2D DI NATURA FISICA

Il *chirp* spaziale bidimensionale corrisponde ad una frequenza spaziale crescente linearmente con lo spazio e quindi ad una fase crescente quadraticamente con la distanza dall'origine.

$$h(\rho) = \frac{\exp(j(x^2 + y^2)/4\sigma^2)}{2\pi\sigma^2} = \frac{\exp(jx^2/4\sigma^2)}{\sqrt{2\pi}\sigma} \cdot \frac{\exp(jy^2/4\sigma^2)}{\sqrt{2\pi}\sigma} \quad (20.38)$$

Si verifica subito che anche la funzione di trasferimento $H(\Omega)$ è un *chirp* bidimensionale. Infatti, la trasformata è immediatamente trovata, visto che le variabili sono in questo caso separabili, e si ritrova un *chirp* questa volta nello frequenza spaziale. Posto al solito $\Omega = \sqrt{u^2 + v^2}$, a parte una costante complessa:

$$H(\Omega) = \exp(-j\sigma^2 u^2) \cdot \exp(-j\sigma^2 v^2) = \exp(-j\sigma^2 \Omega^2) \quad (20.39)$$

L'interesse per il chirp nasce dal fatto che la convoluzione con un *chirp* approssima molto bene una trasformazione di Fourier. Infatti, operando ad $1D$ per semplicità, si ha che il risultato della convoluzione con il *chirp* di una funzione $f(t)$ che ha per trasformata $F(\omega)$ è la funzione $g(t)$ (che ha trasformata $G(\omega)$):

$$g(t) = \int F(\omega) e^{j(-\omega^2\sigma^2 + \omega t)} \frac{d\omega}{2\pi}$$

Calcoliamo $g(t)$ con la tecnica della fase stazionaria. Poichè si ha stazionarietà della fase quando:

$$-2\omega_0\sigma^2 + t_0 = 0 \rightarrow t_0 = 2\omega_0\sigma^2$$

si trova, a meno di un fattore di proporzionalità α ,

$$g(t_0) \simeq \alpha F\left(\omega_0 = \frac{t_0}{2\sigma^2}\right) \quad (20.40)$$

In altre parole, la funzione $g(t)$ è circa proporzionale alla trasformata di Fourier della $f(t)$.

Per una comprensione intuitiva di questo meccanismo, si pensi al chirp come alla successiva trascrizione in un'unica forma d'onda complessiva delle N sinusoidi finestrate di una trasformazione di Fourier di sequenze lunghe N campioni, cominciando da quella a frequenza di Nyquist

20.4 IL CHIRP, LALENTE SOTTILE E L'OTTICA DI FOURIER

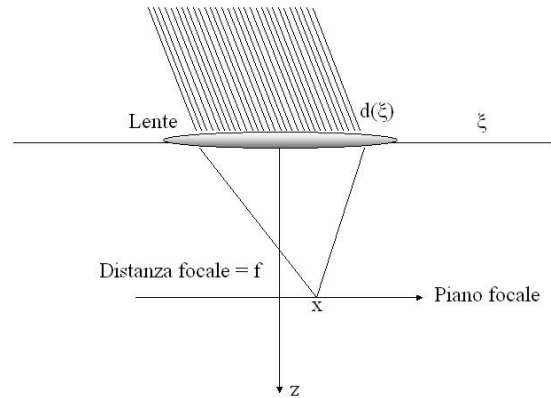


Figura 20.8: Lente sottile: nel piano focale si forma la trasformata di Fourier bidimensionale della funzione di illuminazione $d(\xi)$.

negativa, passando per quella a frequenza zero, per finire a quella a frequenza di Nyquist positiva. La lunghezza totale sarebbe di N^2 campioni. Potremo individuare la frequenza di una nuova sinusoide convolvendola con questo chirp lungo e poi trovando il tempo a cui la convoluzione è massima. Decimando il risultato della convoluzione di un fattore N , e prendendone il modulo, si otterrà circa il periodogramma, ma trascritto in una sequenza di valori. La risoluzione in frequenza crescerà con N . Ovviamente, nel caso del chirp, la frequenza cambia con continuità invece che a gradini, come nella trascrizione del periodogramma.

Vediamo ora come il caso del chirp si applica alla focalizzazione con delle lenti sottili ed alla formazione delle immagini nel loro piano focale (vedi figura 20.8). Indichiamo con $d(\xi)$ la funzione di illuminazione e cioè l'ampiezza complessa dell'onda elettromagnetica luminosa sull'apertura della lente. L'effetto della lente è quello di sfasare il segnale in arrivo di una fase pari a:

$$\exp j \frac{\omega}{c} \frac{\xi^2}{2f} ;$$

dove ω è al solito la pulsazione dell'onda, c la velocità della luce, f la distanza focale della lente. Per chi non ricordasse l'ottica geometrica,

CAPITOLO 20 FILTRI 2D DI NATURA FISICA

notiamo che:

$$\frac{1}{f} = \left(\frac{c - c_1}{c_1} \right) \left(\frac{1}{R_1} - \frac{1}{R_2} \right)$$

dove R_1 , R_2 , sono i raggi di curvatura (con i segni relativi) delle due facce della lente e c_1 la velocità della luce nella lente (piu' bassa). Il segnale irraggiato dalla lente, misurato in un piano parallelo alla lente e distante z da essa è, per sovrapposizione degli effetti:

$$m(x) = \int_{\text{apertura}} d(\xi) \exp \frac{jk\xi^2}{2f} \exp \left[-jk \left(z + \frac{(x - \xi)^2}{2z} \right) \right] d\xi \quad ; \quad k = \frac{\omega}{c} \quad (20.41)$$

Lo sfasamento additivo:

$$\exp \left[-jk \left(z + \frac{(x - \xi)^2}{2z} \right) \right]$$

tiene conto (con un'approssimazione del secondo ordine) dello sfasamento dovuto alla propagazione dal punto di ascissa ξ sulla lente a quello di ascissa x a distanza z dalla lente. Quindi:

$$m(x) = e^{-j\frac{kx^2}{z} - jkz} \int_{\text{apertura}} d(\xi) \exp jk \left[\frac{\xi^2}{2} \left(\frac{1}{f} - \frac{1}{z} \right) + \frac{x\xi}{z} \right] d\xi \quad (20.42)$$

e se $z = f$:

$$m(x) = e^{-j\frac{kx^2}{f} - jkf} \int_{\text{apertura}} d(\xi) \exp jk \left(\frac{x\xi}{f} \right) d\xi \quad (20.43)$$

Nel piano focale della lente quindi, troviamo la trasformata di Fourier della funzione di illuminazione a meno del fattore di fase $\exp \left(-j\frac{kx^2}{f} - jkf \right)$. Se poi l'oggetto (la funzione di illuminazione) è posizionato nell'altro fuoco della lente, questo fattore di fase si cancella. I due piani focali sono gli unici tra cui la trasformazione di Fourier avviene senza sfocamento. Ancora, ad ogni direzione di arrivo bidimensionale sulla lente (ogni pulsazione spaziale delle onde piane in cui si decompone la funzione di illuminazione sulla lente) corrisponde un preciso punto del piano focale (vedi figura 20.8) dove l'energia proveniente da questa direzione si concentra. L'immagine che si forma nel piano focale corrisponde quindi

20.4 IL *CHIRP*, LALENTE SOTTILE E L'OTTICA DI FOURIER

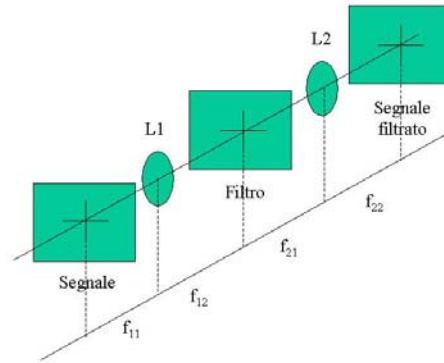


Figura 20.9: Tecniche di filtraggio ottico: il segnale è depositato sulla trasparenza nel primo fuoco della lente L1 ed il filtro su un'altra trasparenza nel secondo fuoco della lente L1 che coincide con il primo fuoco della lente L2. Il segnale filtrato appare nell'immagine nel secondo fuoco della lente L2.

appunto alla distribuzione dell'energia che proviene dalle varie direzioni. Se si osserva solo questa energia, come nel caso dell'occhio o di un telescopio, l'errore di fase di cui nell'equazione 20.43 è irrilevante. In ogni modo, la tecnica di analisi spettrale che viene implicitamente utilizzata nella visione è quindi quella del periodogramma, come peraltro già osservato in occasione della discussione delle DOA.

La disponibilità di una trasformata di Fourier 2D (o anche 1D, basta mettere delle lenti cilindriche invece che sferiche) in tempo reale ha portato a realizzare filtri spaziali con tecniche ottiche; basta mettere l'immagine da trasformare nel primo piano focale di una lente L1, la sagoma del filtro nell'altro piano focale della lente L1 che poi coincide con il primo piano focale di una seconda lente L2 (fig. 20.9). Il segnale filtrato è osservabile nel secondo piano focale della lente L2. I primi sistemi radar ad apertura sintetica utilizzavano proprio questo metodo per effettuare una focalizzazione sia pure approssimativa, ma in tempo reale, del segnale registrato. Questa tecnica appare oggi obsoleta, vista l'espansione delle metodologie numeriche di trattamento, ma potrebbe rinascere in qualsiasi momento, visto il continuo evolversi delle tecnologie e la grande potenza di calcolo che potrebbe garantire.

20.5 L'equazione della diffusione

Il quarto caso che prenderemo in considerazione è quello dell'equazione della diffusione, ad esempio termica; essa governa l'evoluzione temporale della temperatura $T(x, y, t)$ di una piastra indefinita che ha conducibilità termica uniforme c .

$$\frac{\partial^2 T(x, y, t)}{\partial x^2} + \frac{\partial^2 T(x, y, t)}{\partial y^2} = c \frac{\partial T(x, y, t)}{\partial t} \quad (20.44)$$

Supponiamo anzitutto di conoscere la temperatura della piastra al tempo $t = 0$; $\forall x, y$ e di volerne determinare l'evoluzione. Trasformiamo nelle variabili x, y ponendo, al solito $\Omega = \sqrt{u^2 + v^2}$:

$$-\Omega^2 T(\Omega, T) = c \frac{dT(\Omega, t)}{dt} \quad (20.45)$$

$$T(u, v, t) = T(u, v, 0) \times e^{-\frac{\Omega^2 t}{c}} \quad (20.46)$$

Si osservi che il valore della funzione di trasferimento $e^{-\frac{\Omega^2 t}{c}}$ per $\Omega = 0$, è unitario ed indipendente dal tempo. Passando alle variabili x, y :

$$T(x, y, t) = T(x, y, 0) ** \frac{c}{4\pi t} e^{-\frac{\rho^2 c}{4t}} \quad (20.47)$$

La temperatura iniziale è quindi convoluta per un filtro gaussiano di volume costante (la quantità di calore posseduta inizialmente dalla piastra rimane invariata) e invece *diametro* crescente con \sqrt{t} : è l'effetto della diffusione termica.

Un secondo problema è quello dell'osservazione della temperatura lungo la linea $y = 0, T(x, y = 0, t), \forall t$ e poi la determinazione della $T(x, y, t); y > 0$. Trasformando questa volta in x, t si ha:

$$(u^2 + j\omega c) \times T(u, y, \omega) = \frac{d^2 T(u, y, \omega)}{dy^2} \quad (20.48)$$

Risolvendo l'equazione differenziale ed eliminando la soluzione divergente per $y \rightarrow \infty$ abbiamo:

$$T(u, y, \omega) = T(u, y = 0, \omega) \times e^{-y\sqrt{u^2 + j\omega c}} \quad (20.49)$$

20.5 L'EQUAZIONE DELLA DIFFUSIONE

Nei tempi - spazi ciò corrisponde a:

$$T(x, y, t) = T(x, y = 0, t) * * \frac{y}{2\sqrt{\pi}(t/c)^{3/2}} \times e^{-\frac{\rho^2 c}{4t}} \quad (20.50)$$

Si osservi che il flusso di calore dalla superficie verso l'interno del corpo non è nullo. E' interessante anche osservare come l'ultima equazione affermi che la temperatura possa variare per valori anche elevati di x, y anche in tempi molto piccoli. Ciò è dovuto all'imperfezione della descrizione fisica del processo fatta con l'equazione (20.44) che trascura il tempo finito di propagazione delle variazioni termiche.

Capitolo 21

Sequenze 2D

Un segnale bidimensionale può essere campionato con le stesse regole del campionamento a 1D; in particolare, campionare implica replicarne lo spettro. Infatti la trasformata del reticolo bidimensionale di impulsi è un reticolo bidimensionale di impulsi nel dominio delle frequenze spaziali (fig.21.1).

$$\sum_{h,k=-\infty}^{+\infty} \delta(x - hx_0, y - ky_0) = \sum_{h=-\infty}^{+\infty} \delta(x - hx_0) \sum_{k=-\infty}^{+\infty} \delta(y - ky_0) \leftrightarrow \quad (21.1)$$

$$\sum_{l=-\infty}^{+\infty} \delta(u - lu_0) \sum_{m=-\infty}^{+\infty} \delta(v - mv_0) = \sum_{l,m=-\infty}^{+\infty} \delta(u - lu_0, v - mv_0) \quad (21.2)$$

$$u_0 = \frac{2\pi}{x_0}; \quad v_0 = \frac{2\pi}{y_0}; \quad (21.3)$$

Moltiplicare negli spazi per la funzione impulsiva ripetuta implica convolvere nelle frequenze per la sua trasformata e quindi lo spettro di una sequenza a due dimensioni è periodico sia lungo l'asse delle u che delle v . Per potere campionare è pertanto sufficiente che il segnale sia contenuto nel rettangolo:

$$|u| < \frac{|u_0|}{2}; \quad |v| < \frac{|v_0|}{2} \quad (21.4)$$

In effetti basta che le repliche non si sovrappongano; pertanto non è necessario che lo spettro stia in un rettangolo; basta che stia in una qualsiasi figura che replicata su un reticolo rettangolare non si sovrapponga. Ad esempio possiamo levare da un lato del rettangolo un'area per aggiungerla

CAPITOLO 21 SEQUENZE 2D

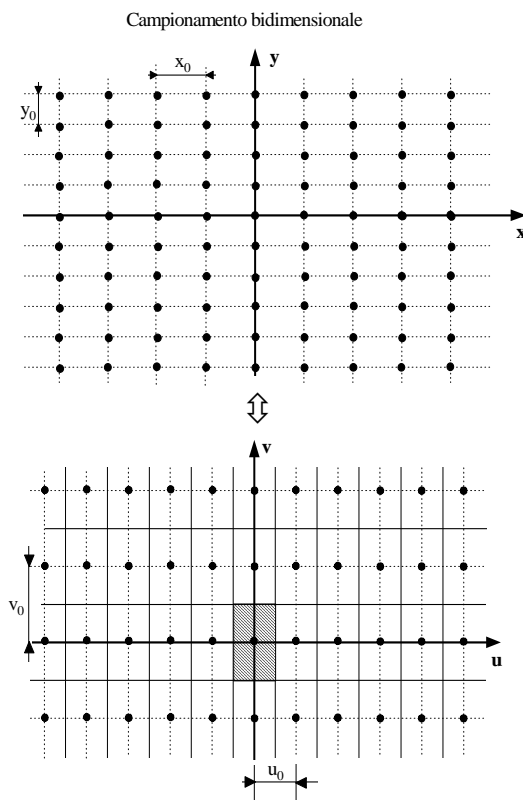


Figura 21.1: Campionamento bidimensionale

al lato opposto ecc (vedi figura 21.2) In sostanza, a 2D si hanno maggiori libertà.

21.1 L'equivocazione a 2D

Se lo spettro bidimensionale dell'immagine che viene campionata non è contenuto nel *dominio fondamentale*, le sue repliche rientrano e formano equivocazione. Nella figura 21.3 vediamo lo spettro relativo a una funzione cilindrica campionata (si pensi a dei filari di alberi); le repliche spettrali che escono ad esempio dal bordo destro della mattonella, rientrano dal bordo sinistro, formando una configurazione complessa. In sostanza, vediamo i filari degli alberi allinearsi sia secondo le direzioni originali

21.1 L'EQUIVOCAZIONE A 2D

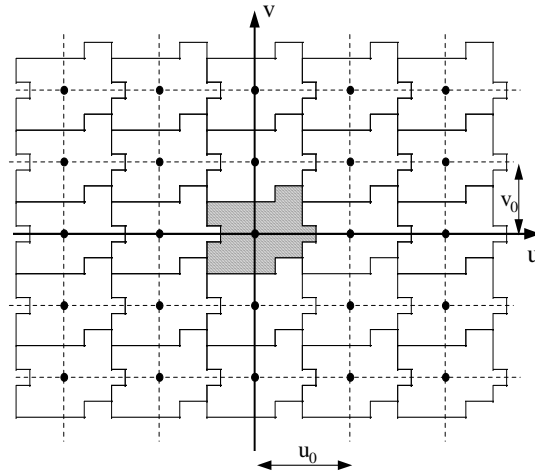


Figura 21.2: Un altro possibile dominio fondamentale per la griglia di campionamento

che secondo altre direzioni, equivocate, lungo cui appaiono allineamenti fittizi.

21.1.1 Effetti di moiré

Il campionamento 2D di segnali di tipo sinusoidale origina effetti di equivocazione spaziale abbastanza complessi, cui viene dato in generale il nome di *moiré*, dal tessuto omonimo che ha un aspetto *marezzato*. Consideriamo una sinusoide 2D, complessa:

$$s(x, y) = A \exp(ju_0x + v_0y)$$

La sinusoide campionata nei punti:

$$x_n = n\Delta; \quad y_m = m\Delta;$$

ha uno spettro replicato

$$S(u, v) = A \sum_{p, q = -\infty}^{\infty} \delta \left(u_0 + p \frac{2\pi}{\Delta}, v_0 + q \frac{2\pi}{\Delta} \right)$$

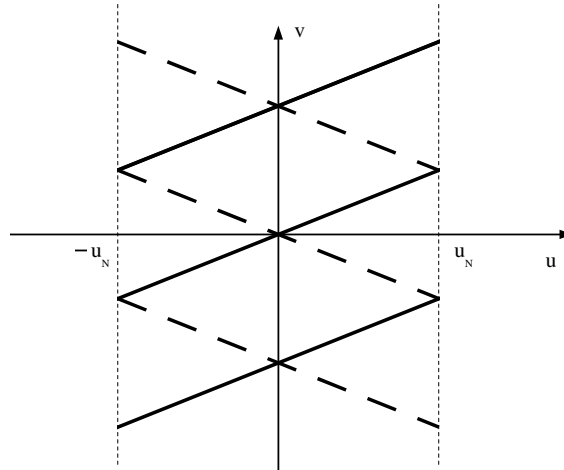


Figura 21.3: Lo spettro di una lama replicato per effetto del campionamento (orizzontale).

Il dominio spettrale fondamentale

$$|u, v| < \frac{\pi}{\Delta} \quad (21.5)$$

limitato dalla pulsazione di Nyquist, contiene solo una sinusoide 2D che coincide con la sinusoide non campionata quando il campionamento è corretto e cioè se:

$$|u_0, v_0| < \frac{\pi}{\Delta} \quad (21.6)$$

Altrimenti, la sinusoide equivocata è

$$u_{eq}, v_{eq} = u_0 + p \frac{2\pi}{\Delta}, v_0 + q \frac{2\pi}{\Delta}$$

dove p, q sono tali da garantire che la (21.6) sia soddisfatta. Si osservi in particolare che l'azimut della sinusoide 2D equivocata può essere totalmente differente da quello della sinusoide 2D originale.

Un esempio di *moiré* si ottiene dalla sovrapposizione di due reticoli leggermente diversi, che risulta p.e. se si guardano contro luce due veli eguali e sovrapposti, poco distanti tra loro. Per effetto della prospettiva, la trama del velo più lontano ci appare più fitta di quella del velo più vicino di un fattore pari a $D/(D+d)$, dove d è la distanza tra i veli e

21.1 L'EQUIVOCAZIONE A 2D

D la distanza che ci separa dal velo più vicino. L'immagine che vediamo in trasparenza è circa il prodotto dei due reticoli e pertanto lo spettro del segnale osservato è la convoluzione degli spettri. Quindi, se f_0 è la frequenza corrispondente alla trama di un velo (supponiamo il processo monodimensionale, per semplicità), e $f_0 D / (D + d)$ l'altra, il battimento (il prodotto) avrà una componente alla frequenza differenza:

$$\Delta f = f_0 - \frac{f_0 D}{D + d} \simeq \frac{d}{D} f_0 \quad \text{se } d \ll D$$

Se il reticolo è 2D, vedremo delle *frange interferenziali* da cui potremmo ricostruire dei dettagli della struttura microscopica dei reticoli di partenza; p.e. conoscendo $d, D, \Delta f$ potremmo determinare f_0 anche se ciò fosse impossibile ad "occhio nudo". Infine, si osservi che, allontanandosi dai veli, aumenta il periodo della frangia del *moiré*.

21.1.2 Il campionamento esagonale

Se il segnale da campionare ha spettro circolare, il reticolo di campionamento quadrato è inefficiente in quanto comporta un dominio fondamentale nelle frequenze spaziali anch'esso quadrato, sostanzialmente. Pertanto, sinusoidi 2D nelle direzioni delle diagonali del quadrato possono avere frequenza $\sqrt{2}$ volte maggiore che non nelle direzioni dei lati del quadrato stesso. Quindi, a seconda che il quadrato sia iscritto o circoscritto alla circonferenza che delimita lo spettro, o si deve tagliare lo spettro o si deve usare un campionamento troppo fitto. E' utile cercare tecniche di campionamento più isotrope, e cioè aventi un dominio fondamentale più assomigliante ad un cerchio. Ciò si ottiene con il campionamento esagonale che verrà ora presentato. Si consideri un reticolo la cui base sono dei triangoli equilateri (fig. 21.4); questo reticolo ha 3 direzioni di simmetria a 120° e quindi è più simmetrico di quello rettangolare che ne ha due a 90° . Si intuisce come questo reticolo di campionamento possa essere più efficiente di quello quadrato, per funzioni a spettro circolare. Indichiamo con a il lato del triangolo; il reticolo a base triangolare può essere visto come la combinazione di due reticoli a base rettangolare di lati $x_0 = a, y_0 = a\sqrt{3}$ rispettivamente; i due reticoli sono sfalsati di mezzo passo lungo le due direzioni. Nel dominio delle frequenze si ha che la

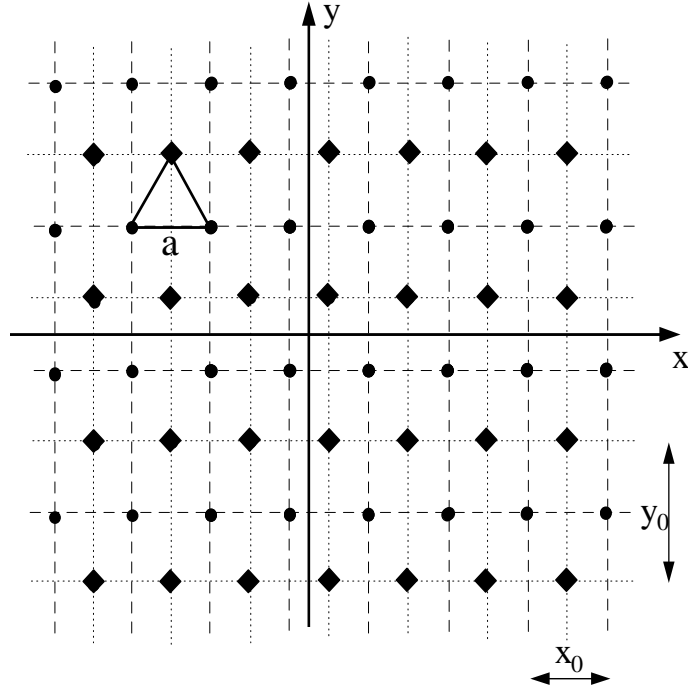


Figura 21.4: Reticolo corrispondente al campionamento esagonale (dominio degli spazi).

trasformata del doppio reticolo di impulsi è:

$$H(u, v) = [e^{j(\frac{ux_0}{4} + \frac{vy_0}{4})} + e^{-j(\frac{ux_0}{4} + \frac{vy_0}{4})}] \times \sum_{h,l=-\infty}^{+\infty} \delta(u - hu_0, v - lv_0); \quad (21.7)$$

$$u_0 = \frac{2\pi}{x_0} = \frac{2\pi}{a}; \quad v_0 = \frac{2\pi}{y_0} = \frac{2\pi}{a\sqrt{3}} = \frac{u_0}{\sqrt{3}} \quad (21.8)$$

$$H(u, v) = 2 \cos[(h + l)\frac{\pi}{2}] \times \sum_{h,l=-\infty}^{+\infty} \delta(u - hu_0, v - lv_0); \quad (21.9)$$

La combinazione dei due reticoli elimina le repliche per le quali $l + h$ è dispari, e raddoppia l'ampiezza di quelle per cui $l + h$ è pari. La sagoma del dominio base nelle frequenze è quindi un esagono regolare,

21.1 L'EQUIVOCAZIONE A 2D

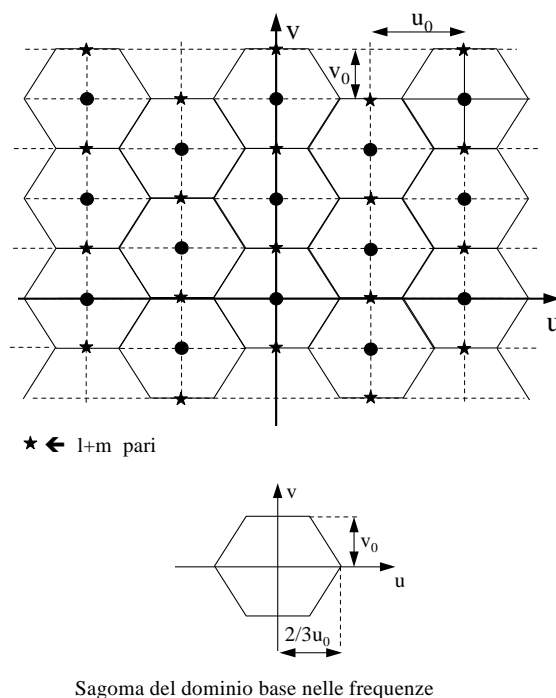


Figura 21.5: Reticolo corrispondente al campionamento esagonale (dominio delle frequenze).

come in figura 21.5. Il rapporto tra la massima frequenza spaziale che non viene equivocata (nelle direzioni $0^0, 60^0, 120^0$) e la minima (nelle direzioni $30^0, 90^0, 150^0$) è

$$\frac{2}{3}u_0 : v_0 = \frac{2}{\sqrt{3}} = 1,154 \quad (21.10)$$

mentre invece per il reticolo quadrato questo rapporto è $\sqrt{2}$. L'area per ogni campione è l^2 nel caso del reticolo quadrato di base l ; invece è di $a^2\sqrt{3}/2$ nel caso di reticolo triangolare di base a ; la massima frequenza non equivocabile è $1/2l$ in un caso e $1/(a\sqrt{3})$ nell'altro caso. Eguagliando questi due valori, e quindi imponendo di non equivocare questa frequenza qualsiasi sia il suo azimut, e ponendo quindi $a = 2l/\sqrt{3}$, si trova che, passando da reticolo quadrato a reticolo esagonale, il numero di campioni necessario decresce del fattore $2/\sqrt{3} = 1,154$.

21.2 Il campionamento del segnale televisivo

Nella trasmissione analogica di un segnale televisivo, il campionamento del volume (x, y, t) che corrisponde alla successione di infinite immagini a risoluzione spaziale e temporale illimitata avviene nel modo piuttosto complesso che verrà ora descritto.

Le immagini vengono scandite per righe su campi interallacciati. Il segnale viene cioè campionato verticalmente a distanza d ; nel sistema analogico PAL, un quadro è formato da 625 linee di cui solo 573 sono visibili. Le righe dispari, distanti perciò tra di loro $2d$, sono trasmesse in successione (semiquadro dispari o campo dispari); poi, con un ritardo di 20 msec, viene trasmesso il campo pari e quindi le linee di posizione pari. Dopo ancora 20 msec, si trasmette il campo dispari del quadro 2 e così via al ritmo di 25 quadri al secondo (fig.21.6) Il campionamento del

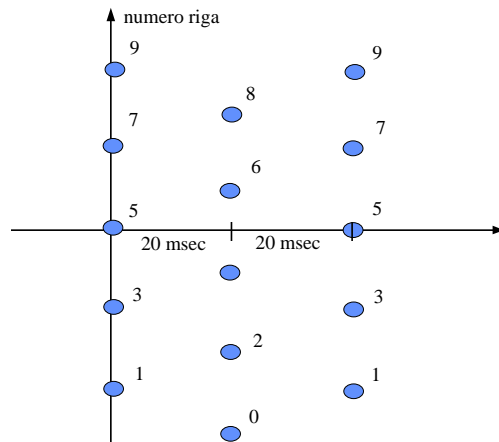


Figura 21.6: Campionamento interlacciato del segnale televisivo: dominio dei tempi

volume è perciò di tipo esagonale sulla coppia di assi y, t in quanto il reticolo della distribuzione delle righe (verticale) dei semiquadri dispari è sfalsato rispetto a quello dei semiquadri pari. Invece, l'asse x (le righe del segnale televisivo) non è campionato ma viene trasmesso in continuità nel tempo.

21.2.1 Equivocazione spazio - temporale

Vediamo ora cosa accade se il segnale non è opportunamente prefiltrato prima del campionamento spazio - temporale menzionato ed anche cosa accade se vogliamo una ricostruzione del volume x, y, t sopprimendo le armoniche del pettine di campionamento. L'asse x non è campionato ed è trasmesso con continuità; il filtraggio lungo questo asse è effettuato automaticamente se non altro dal filtraggio passa basso dovuto alla limitata banda disponibile a radiofrequenza e pertanto dai filtri che precedono la modulazione per la trasmissione.

Il segnale lungo le direzioni sia verticale che temporale è campionato e quindi dovrebbe subire un filtraggio anti equivocazione prima del campionamento. Invece, il prefiltraggio lungo la direzione verticale non avviene affatto se non per quanto è dovuto alla dimensione finita del pennello elettronico che esplora l'immagine e che convolve quindi con un filtro bidimensionale (la distribuzione delle ampiezze della sezione del pennello) l'immagine ripresa. Nel complesso questo filtraggio appare sufficiente, in quanto l'effetto di "moirè" e cioè' di conversione in basso di alte frequenze spaziali verticali è abbastanza limitato. Questo effetto è visibile quando un'immagine a righe orizzontali fitte viene campionata dalla scansione televisiva; per l'equivocazione nascono delle inesistenti striscie di bassa frequenza.

Problemi più gravi insorgono invece per il campionamento temporale, decisamente scarso come si vede dalle immagini riprese con la moviola televisiva. Ad esempio, un calciatore ripreso in due quadri televisivi successivi occupa posizioni distanti tra di loro di molto più dell'intervallo tra le righe d . Infatti, se si muove a 10 m/sec in 40 msec ha percorso 40 cm; se la sua immagine occupa ad esempio una cinquantina di righe lungo la verticale ed uno spazio proporzionale lungo l'orizzontale, d corrisponde a circa 4 cm di spazio (se è alto 2 m) e perciò si muove di circa 10 "pixels" (picture elements) per quadro. La rotazione di fase per quadro delle componenti alla frequenza corrispondente a quella di Nyquist π/d è circa 10π e quindi vi è equivocazione temporale di tutti i segnali che hanno una componente di frequenza spaziale maggiore di $1/10d$ nella direzione del moto.

Nel fermo immagine non si possono quindi proiettare assieme i due semiquadri perchè corrispondono a due immagini spostate relativamente di un intervallo ben maggiore della risoluzione spaziale. Si proietta un

CAPITOLO 21 SEQUENZE 2D

semiquadro solo, e da questo si interpola l'altro semiquadro generando un'immagine a risoluzione spaziale ridotta alla metà nella direzione verticale. Inoltre, la telecamera non può integrare la luminanza del segnale televisivo per tutta la durata di un quadro, come si faceva nel passato per migliorare il rapporto segnale - rumore, ma deve riprendere delle "istantanee" per evitare di avere poi delle immagini mosse alla moviola. Ovviamente, se le immagini fossero sempre proiettate in sequenza al loro ritmo, il prefiltraggio temporale effettuato con l'integrazione della telecamera sarebbe utile. La possibilità di dovere usare la moviola costringe ad introdurre del rumore di equivocazione. Peraltro l'osservatore umano, che interpreta le immagini in base a dei modelli a priori e non accetta l'ipotesi che il calciatore od il pallone cambino eccessivamente di forma tra un'immagine e la successiva, trova più accettabile il rumore di equivocazione al prefiltraggio, teoricamente corretto ma praticamente da respingere.

Da quanto osservato, risulta che il segnale televisivo è campionato temporalmente in modo da creare equivocazione e pertanto la ricostruzione dai campioni di un segnale continuo non può essere fatta senza commettere errori. Questo problema si pone quando volessimo interpolare immagini intermedie tra più quadri successivi per creare un segnale rallentato o un segnale a frequenza di quadro superiore (ad esempio 100 Hz) per limitare l'effetto di "sfarfallamento" (l'effetto fastidioso dovuto all'accensione periodica di una luce a frequenza al di sotto della frequenza di taglio del sistema visivo). Ad esempio, tecniche di interpolazione lineare interpolerebbero il moto di un pallone con delle dissolvenze tra un'immagine e la successiva che non potrebbero ricreare l'impressione del moto, che rimarrebbe a scatti; infatti nelle immagini interpolate il pallone scomparirebbe progressivamente da una posizione per riapparire nell'altra e non apparirebbe mai nelle posizioni intermedie. Ciò è possibile solo utilizzando tecniche di interpolazione complesse (a moto compensato) in cui:

a) si individuano i contorni dell'oggetto in moto confrontando le immagini successive;

b) si misura il vettore di moto, si prevede la futura posizione dell'oggetto, e si ricostruisce l'immagine.

Poiché l'interpolazione "a moto compensato" dipende dalle immagini da cui è rilevato il moto, si conclude che è una tecnica di interpolazione non lineare.

CAPITOLO 21 SEQUENZE 2D

Peraltro, si può fare di meglio, e lo dimostra la commercializzazione di sistemi in cui il segnale televisivo è ripresentato sullo schermo al ritmo di 100 Hz, interpolando le immagini intermedie. L'interpolazione viene fatta preferenzialmente con algoritmi "a moto compensato" proprio per evitare gli effetti di equivocazione menzionati in precedenza nel caso della moviola e dei calciatori.

Un dettaglio verticale mobile (B) si replica, equivocato, nel punto (B'), che è circa a frequenza f_c . Consideriamo infatti un segnale di alta risoluzione verticale, una riga bianca in campo nero; se la riga è in posizione fissa, non vi sono problemi. Ma se la riga scende di un intervallo di riga per ogni quadro, sono possibili due casi, (a) viene riprodotta sempre (in tutti i quadri, e quindi sarà vista con energia doppia); (b) non sarà mai visibile; una riga suborizzontale in moto discendente (o ascendente) sarà vista apparire e scomparire in modo casuale (equivocazione spazio-temporale). Questo fenomeno si osserva direttamente, p. e. in occasione di una partita di tennis, quando le righe del campo (suborizzontali) sono viste a gradinate ed i gradini tremolano quando la telecamera si muove. E' il solito problema dell'equivocazione, visto all'inizio del testo per i sistemi antincendio.

Dalla figura 21.7 si comprende anche come un filtro spazio temporale per la soppressione degli spettri replicati dovuti al pettine di campionamento debba avere il contorno tratteggiato e quindi il filtro non può essere realizzato p.e. con la cascata separabile di filtri spaziali e temporali, che porterebbe ad un filtraggio a base rettangolare.

21.3 Ricostruzione 2D

Per ricostruire un segnale continuo bidimensionale si deve convolvere la sequenza dei campioni per la risposta all'impulso di un filtro opportuno; ad esempio un mantentore bidimensionale $mant(x, y)$, la cui risposta all'impulso è il prisma a base quadrata o l'interpolatore bilineare; quest'ultimo corrisponde ad interpolare linearmente il segnale prima lungo le righe in corrispondenza dei campioni e poi lungo le colonne per tutti i valori (continui, ora) delle ascisse. Si può anche fare viceversa, il risultato sarà eguale. La risposta all'impulso dell'interpolatore bilineare è quindi costituita dalla convoluzione di due lame triangolari disposte lungo i due assi coordinati e larghe 2Δ ; si osservi che questa superficie $bilin(x, y)$ è

21.3 RICOSTRUZIONE 2D

separabile e **non** è una piramide a base quadrata; le superfici dei lati appartengono infatti ad iperboloidi iperbolic (fig.21.8).

$$mant(x, y) = 1; |x|, |y| < \Delta$$

$$bilin(x, y) = \left(1 - \frac{|x|}{\Delta}\right) \left(1 - \frac{|y|}{\Delta}\right); |x|, |y| < \Delta$$

Le trasformate di Fourier del mantentore e dell'interpolatore bilineare

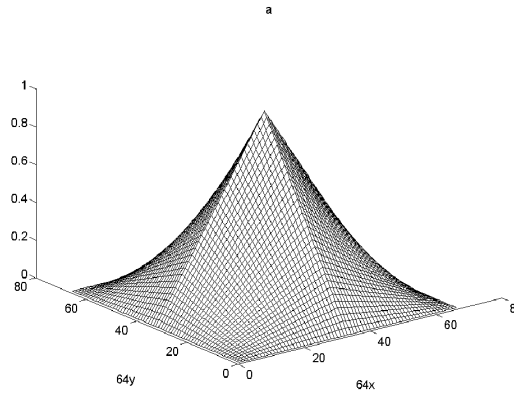


Figura 21.8: L'interpolatore bilineare $bilin(x, y)$: i lati della piramide *tenda* sono lembi di iperboloidi iperbolic.

sono:

$$H_m(u, v) = sinc(u\Delta)sinc(v\Delta); H_l(u, v) = [sinc(u\Delta)sinc(v\Delta)]^2; \quad (21.11)$$

La piramide *egizia* (fig 21.9) con la stessa base quadrata dell'interpolatore bilineare **non** è un buon interpolatore perchè interpola correttamente lungo le facce ma erratamente lungo gli spigoli; se la base fosse ancora quadrata ma con i vertici a terra sui quattro campioni adiacenti al campione assegnato lungo gli assi coordinati, la superficie interpolerebbe bene lungo gli spigoli e male lungo le facce. Per convincersene, basta pensare che il volume di queste piramidi è $4\Delta^2 H/3$ o $2\Delta^2 H/3$ invece di $\Delta^2 H$; nel caso di tutti i campioni eguali il volume totale eccederebbe o sarebbe inferiore a quello corretto. Un interpolatore con facce piane che rispetta i piani (e cioè se i campioni sono disposti su un piano, la superficie

CAPITOLO 21 SEQUENZE 2D

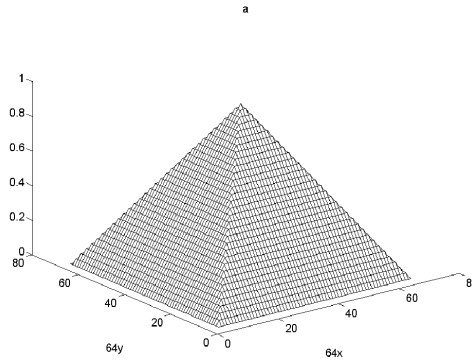


Figura 21.9: La piramide "egizia": un cattivo interpolatore a 2D.

interpolante coincide con il piano) è invece la piramide a base esagonale rappresentata in figura 21.10. Per convincersene basta tracciare una diag-

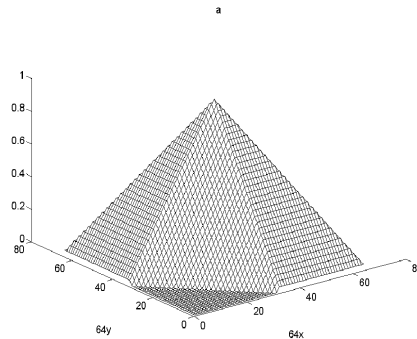


Figura 21.10: Piramide a base esagonale: un possibile interpolatore lineare a 2D anisotropo.

onale del quadrato base del reticolo per dividerlo in due triangoli. Una faccia della superficie interpolante è il piano che passa per i tre campioni posizionati nei vertici del triangolo cui appartiene il punto da interpolare. Ovviamente, vi sono due possibilità di tracciare la diagonale e quindi due superfici interpolanti che hanno diversa funzione di trasferimento lungo le due diagonali.(fig.21.11). Il loro volume è quello corretto (si veda per ulteriori dettagli il paragrafo sulla triangolazione di Delaunay, alla fine

21.3 RICOSTRUZIONE 2D

del capitolo).Facendone la media per avere una soluzione piú isotropa,

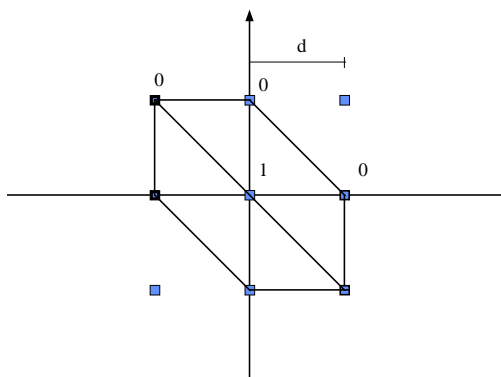


Figura 21.11: L'interpolatore a base esagonale visto in pianta; il comportamento differente nelle due direzioni a 90^0 è evidente.

si ottiene per la risposta all'impulso la figura 21.12, dove ancora la su-

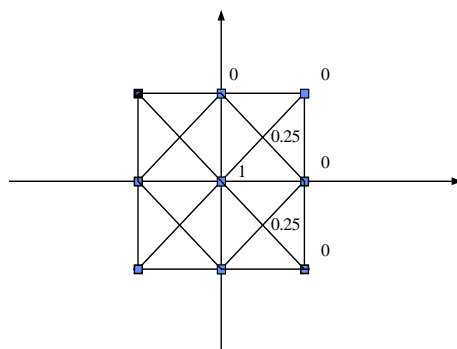


Figura 21.12: Media di due interpolatori a base esagonale orientati reciprocamente a 90^0 .

perficie interpolante ha facce piane, ma i tre punti che definiscono i vari piani sono i due campioni piú vicini al punto da interpolare ed il campione "supplementare" posto al centro del quadrato base e la cui ampiezza è la media delle ampiezze dei quattro campioni ai vertici del quadrato (0,25). Alla fine, abbiamo approssimato l'interpolatore bilineare con un poliedro a 16 facce piane di forma triangolare (fig. 21.13).Infine, una

CAPITOLO 21 SEQUENZE 2D

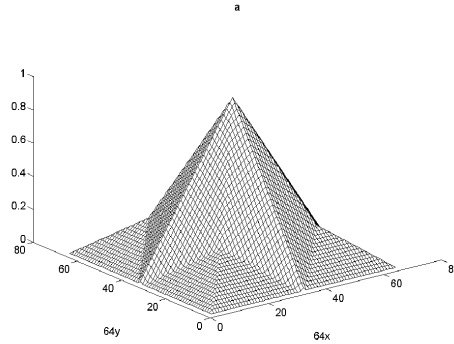


Figura 21.13: Approssimazione dell'interpolatore bilineare con un interpolatore a facce piane.

tecnica molto usata è quella degli splines bicubici: di nuovo si interpola ad esempio prima lungo la direzione orizzontale, questa volta con un interpolatore a bacchetta elastica; quindi, si interpolano lungo la direzione verticale, punto per punto, le funzioni già interpolate, ormai continue lungo la direzione orizzontale.

21.4 Trasformata di Fourier discreta a 2D

L'estensione a 2D della trasformazione di Fourier discreta è semplice: la periodicità tuttavia è ora su due assi invece che su uno e pertanto le sequenze antitrasformate sono definite su di una superficie torica piuttosto che su un circolo. Questo è congruente con quanto osservato per le trasformate di segnali campionati negli spazi; il campionamento è ora sia negli spazi che nelle frequenze spaziali (numeri d'onda). L'espressione della trasformazione discreta è la seguente:

$$X(h, k) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x_{n,m} W_N^{-nh} W_M^{-mk}; \quad W_N = e^{j2\pi/N} \quad (21.12)$$

L'antitrasformata è:

$$x(n, m) = \frac{1}{MN} \sum_{h=0}^{M-1} \sum_{k=0}^{N-1} X_{h,k} W_N^{nh} W_M^{mk}; \quad W_M = e^{j2\pi/M} \quad (21.13)$$

21.5 FILTRAGGI CON PRODOTTO DI TRASFORMATE

La trasformazione si può fare in cascata, prima per righe e poi per colonne:

$$x(n, m) = \frac{1}{MN} \sum_{h=0}^{M-1} W_N^{nh} \left(\sum_{k=0}^{N-1} X_{h,k} W_M^{mk} \right); \quad (21.14)$$

La quantità in parentesi è l'antitrasformata di Fourier effettuata per righe; la trasformata di Fourier 2D avviene quindi prima facendo la trasformazione per righe e poi per colonne (o viceversa). Il costo in termini di moltiplicazioni è:

$$C = \frac{MN}{2} \log_2 N + \frac{MN}{2} \log_2 M = \frac{MN}{2} \log_2 NM \quad (21.15)$$

21.5 Filtraggi con prodotto di trasformate

Con opportune aggiunte di zeri, in modo che le due funzioni abbiano eguale supporto e cioè siano definite sullo stesso dominio, è possibile farne la convoluzione facendo il prodotto delle trasformate. E' opportuno notare tuttavia che l'aggiunta di zeri è più costosa a 2D che a 1D; infatti se un filtro ha un supporto $M \times M$ ed il segnale $N \times N$, per eliminare gli effetti della periodicità, si devono aggiungere circa $2NM + M^2$ zeri, e cioè in proporzione circa il doppio che a 1D. In altre parole, il costo del *zero padding* è cresciuto con la dimensionalità dei dati. Il costo del filtraggio di una sequenza $N \times N$ sarà

$$2N^2 \left(2^{\frac{1 + \log_2 N}{2}} \right) \quad (21.16)$$

tenendo conto delle trasformate per righe e colonne, delle antitrasformate e del prodotto tra le trasformate per il filtraggio.

21.6 Filtri 2D

Il teorema fondamentale dell'algebra non si estende a 2 dimensioni; è vero che il prodotto di due polinomi in 2 variabili è la convoluzione bidimensionale dei loro coefficienti, ma non è vero che un polinomio in 2 variabili è sempre decomponibile nel prodotto di polinomi di grado minore. Quindi l'utilità della trasformata z a due dimensioni è minore perchè non si estende a 2D il concetto di zeri e quindi di poli e cioè sequenze semplici

CAPITOLO 21 SEQUENZE 2D

in cui possono venire decomposte le sequenze piú complesse. Pertanto il filtraggio 2D è prevalentemente realizzato con tecniche di prodotto delle trasformate di Fourier discrete oppure come convoluzione 2D con opportune sequenze, p.e. trovate finestrando le antitrasformate delle funzioni di trasferimento desiderate. E' invece utile ricorrere al concetto di separabilità per trasformare un problema 2D nella cascata di $2k$ problemi 1D, approssimando la sequenza 2D come:

$$h_{m,n} = \sum_{k=1}^K u_{k,m} v_{k,n} \quad (21.17)$$

Aggiungendo zeri, si può sempre far corrispondere la sequenza $h_{m,n}$ ad una matrice quadrata; questa matrice a sua volta può essere rappresentata in funzione dei suoi autovettori ed autovalori, come somma di diadi;

$$\underline{\mathbf{R}} = \sum_{n=1}^N \lambda_n \underline{\mathbf{e}}_n \underline{\mathbf{e}}_n^* \quad (21.18)$$

Pertanto $K \leq N$ se $|m, n| \leq N$. Una convoluzione bidimensionale per una funzione separabile può essere realizzata come cascata di due convoluzioni monodimensionali per le due funzioni che compaiono nel prodotto con un costo computazionale totale pari a $2NK$; questo modo di operare spesso è assai conveniente perchè si ottiene una buona approssimazione della $h_{m,n}$ con $K \ll N$. Un esempio molto utile è quello della trasformazione di Mc Clellan, che serve a generare filtri a simmetria circolare da filtri 1D.

21.6.1 La trasformazione di Mc Clellan

Si voglia realizzare un filtro con funzione di trasferimento a simmetria circolare ed a fase nulla.

$$H(u, v) = H(\Omega); \quad \Omega = \sqrt{u^2 + v^2} \quad (21.19)$$

Visto che la funzione di trasferimento 1D $H(\Omega)$ è anch'essa a fase 0, abbiamo:

$$H(\Omega) = a_0 + \sum_{n=1}^N a_n \cos(n\Omega\Delta) \quad (21.20)$$

La trasformazione è la seguente:

$$\cos \Omega\Delta \rightarrow -1 + \frac{1}{2}(1 + \cos u\Delta)(1 + \cos v\Delta) \quad (21.21)$$

somma di un impulso e di una funzione separabile. L'operazione di moltiplicazione (nelle frequenze) per $\cos u\Delta$ corrisponde alla convoluzione (negli spazi) per la sequenza di due campioni di ampiezza $1/2$ collocati nei campioni di coordinate $(-1, 0)$; $(1, 0)$; similmente per $\cos v\Delta$. Quindi il costo per campione di una convoluzione (la moltiplicazione della trasformata per $\cos \Omega\Delta$) è (a parte la moltiplicazione per il coefficiente) pari a 7 somme. Se ora si vuole realizzare l'intero filtro 2D e quindi convolvere per le antitrasformate di funzioni come $\cos n\Omega\Delta$ si può operare utilizzando la formula di prostaferesi:

$$\cos n\Omega\Delta = 2 \cos[(n-1)\Omega\Delta] \times \cos \Omega\Delta - \cos[(n-2)\Omega\Delta] \quad (21.22)$$

In questo modo, le convoluzioni corrispondenti al termine n -esimo dei polinomi trigonometrici non devono essere calcolate direttamente ma sono ottenute come combinazione dei risultati ottenuti per valori minori di n ripetendo poi la stessa operazione di convoluzione elementare (la trasformazione lineare indicata come $G(u, v)$ nella figura 21.15). Le isolinee della trasformazione di Mc Clellan per $n = 1$ sono rappresentate in fig.21.14; per n maggiori, la zona in cui l'approssimazione è accettabile si riduce in proporzione. e lo schema a blocchi è presentato in figura

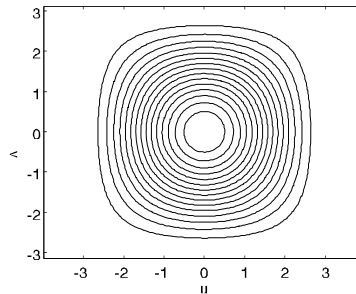


Figura 21.14: Isolinee della trasformazione di Mc Clellan

21.15.

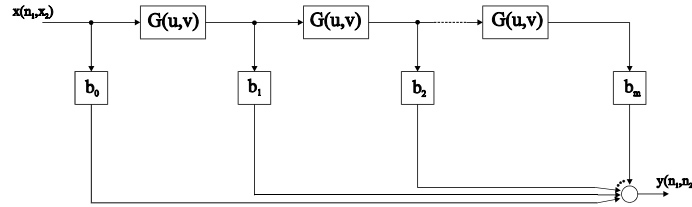


Figura 21.15: Schema a blocchi per il calcolo della trasformazione di McClellan

21.6.2 Filtri per la continuazione di campo

Come si è visto nel paragrafo sui filtri 2D derivanti dall'equazione del campo gravitazionale, il campo a quota $z + \Delta z$ è ottenuto da quello a quota z , applicando la funzione di trasferimento:

$$\exp(-|\Omega| \Delta z) \tag{21.23}$$

Nel caso del telerilevamento, si è visto che si può rilevare la morfologia del sottosuolo se p.e. si dispongono in superficie ($z = 0$) sorgenti impulsive in tutte le posizioni x, y e si registra ancora in tutte le posizioni il campo riflesso $f(x, y, z = 0, t)$. In questo tipo di esperimento radar, la posizione della sorgente coincide con la posizione del ricevitore. Come avevamo osservato, il campo riflesso misurato coincide con quello che sarebbe misurato in un mezzo a velocità di propagazione dimezzata se i riflettori emettessero un impulso di energia acustica al tempo 0, avente ampiezza proporzionale alla loro riflettività. L'ampiezza delle sorgenti fittizie (e quindi dei riflettori) a profondità z viene determinata ricostruendo il campo a profondità z e trovandone il valore al tempo 0. Il campo a profondità z viene ottenuto da quello in superficie, applicando la funzione di trasferimento (questa volta a 3D, piuttosto che a 2D)

$$\exp\left(jz\sqrt{\frac{\omega^2}{c^2} - \Omega^2}\right); \Omega = u^2 + v^2 \tag{21.24}$$

21.7 INTERPOLAZIONE MEDIANTE STIMA LINEARE

oppure convolvendo con l'antitrasformata di Fourier a 3D, indicata con F_{3D}^{-1} .

$$f(x, y, z, t) = f(x, y, z = 0, t) * * * F_{3D}^{-1} \left[\exp \left(jz \sqrt{\frac{\omega^2}{c^2} - \Omega^2} \right) \right] \quad (21.25)$$

e trovandone il valore al tempo 0. Nel dominio trasformato,

$$F(u, v, z = 0, \omega) = \int \int \int f(x, y, z = 0, t) \exp [-j (ux + vy + \omega t)] dx dy dt \quad (21.26)$$

$$F(u, v, z, \omega) = F(u, v, z = 0, \omega) \exp \left(jz \sqrt{\frac{\omega^2}{c^2} - \Omega^2} \right) \quad (21.27)$$

$$f(x, y, z, t = 0) = \int F(u, v, z, \omega) \frac{d\omega}{2\pi} \quad (21.28)$$

Invece di fare trasformazioni di Fourier a 3D, che implicherebbe la stazionarietà spaziale e quindi la costanza ovunque della velocità di propagazione, si può semplicemente convolvere a 2D (e cioè in x, y) la trasformata 1D (nel tempo t) del campo riflesso, con il vantaggio non trascurabile di avere uno strumento utilizzabile anche quando la velocità di propagazione è funzione di x, y , ed anche di z .

$$f(x, y, z, \omega) = f(x, y, z = 0, \omega) * * F_{2D}^{-1} \left[\exp \left(jz \sqrt{\frac{\omega^2}{c^2} - \Omega^2} \right) \right] \quad (21.29)$$

Filtri del tipo:

$$F_{2D}^{-1} [\exp (-|\Omega| \Delta z)], \quad F_{2D}^{-1} \left[\exp \left(jz \sqrt{\frac{\omega^2}{c^2} - \Omega^2} \right) \right] \quad (21.30)$$

sono appunto realizzabili con la trasformazione di McClellan, che può essere facilmente resa spazio variante.

21.7 Interpolazione mediante stima lineare

Una tecnica di interpolazione utile è quella che fa uso della stima lineare; i valori della funzione al di fuori dei punti del reticolo vengono stimati con uno stimatore lineare. Per discutere di questa tecnica è opportuno fare seguire un breve riassunto delle proprietà dei processi casuali bidimensionali e della loro stima.

21.7.1 Processi casuali a 2D

Supponiamo di avere un segnale 2D, $s(x, y)$ realizzazione di un processo stocastico, a valor medio nullo, campionato su un reticolo di lato Δ ; si può supporre inoltre di non avere a disposizione $s_{m,n}$ ma una versione disturbata da rumore bianco

$$s'_{m,n} = s(m\Delta, n\Delta) + z_{m,n}; \quad \forall m, n; \quad (21.31)$$

La sequenza 2D, disturbante, è costituita da campioni indipendenti $z_{m,n}$ e risulta:

$$\begin{aligned} E[z_{m,n}z_{h,k}] &= \sigma_z^2; \quad m = h; \quad n = k; \\ E[z_{m,n}z_{h,k}] &= 0; \quad m \neq h; \quad n \neq k \end{aligned} \quad (21.32)$$

La sequenza di segnale è ottenuta p.e. convolvendo un'altra sequenza "bianca" $w_{m,n}$ con un filtro $h_{m,n}$ e si ha:

$$s_{m,n} = w_{m,n} * h_{m,n} = \sum_{k,l=-N/2}^{N/2} w_{m-k,n-l} h_{k,l} \quad (21.33)$$

La funzione di autocorrelazione della sequenza 2D $s'_{m,n}$ è:

$$E[s'_{m,n} s'_{m+k,n+l}] = r'_{k,l} = \sigma_w^2 \sum_{i,j=-N/2}^{N/2} h_{k,l} h_{i+k,j+l} + \sigma_z^2 \delta(k, l) \quad (21.34)$$

21.7.2 Spettri di potenza a 2D

Similmente a quanto accade a 1D, lo spettro 2D di una sequenza si ottiene per trasformazione di Fourier discreta della funzione di autocorrelazione. Nel caso di rumore bianco a 2D si ha:

$$S_x(h, k) = F\{r_x(n, m)\} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} r_{n,m} W_N^{-nh} W_M^{-mk} = \sigma_x^2; \quad (21.35)$$

Similmente ancora a quanto accade a 1D, se la sequenza $x(i, j)$ con trasformata di Fourier discreta $X(h, k)$ è convoluta con un filtro a risposta all'impulso $h(i, j)$ e avente trasformata di Fourier discreta $H(h, k)$, lo spettro cambia come segue:

$$y(i, j) = x(i, j) * * h(i, j) \quad (21.36)$$

21.7 INTERPOLAZIONE MEDIANTE STIMA LINEARE

$$Y(h, k) = X(h, k) \times H(h, k)$$

$$E[|Y(h, k)|^2] = E[|X(h, k)|^2] \times |H(h, k)|^2 \quad (21.37)$$

In particolare, si possono considerare sequenze incorrelate in una direzione e correlate lungo la direzione ortogonale. In questo caso, la funzione di correlazione è a lama, e lo spettro è cilindrico.

21.7.3 Stima lineare a 2D

Noti i valori di $s'_{n,m}$ sul reticolo di campionamento si vuole ora stimare $s(x, y)$ in tutti i punti del piano. Costruendo lo stimatore come combinazione lineare degli $(H + 1)^2$ valori assunti dalla sequenza $s'_{n,m}$ in un intorno $H \times H$ del punto da stimare, risulta:

$$\hat{s}(x, y) = \sum_{n,m=-H/2}^{+H/2} a_{n,m}(x, y) s'_{n,m} \quad (21.38)$$

I coefficienti $a_{n,m}(x, y)$ sono funzione continua di x, y e sono determinabili in base al principio di ortogonalità.

$$E[\{s(x, y) - \hat{s}(x, y)\} s'_{n,m}] = 0; \quad |n, m| \leq H/2 \quad (21.39)$$

Questo sistema è detto il sistema di equazioni **normali** o di Wiener - Hopf ed è:

$$\sum_{p,q=-H/2}^{+H/2} a_{p,q}(x, y) r'[(p - n)\Delta, (q - m)\Delta] = \quad (21.40)$$

$$r(x - n\Delta, y - m\Delta); \quad |n, m| \leq H/2$$

$$r'(p, q) = r(x = p\Delta, y = q\Delta) + \sigma_w^2 \delta(p, q)$$

Generalizzando la stima al caso di reticolo irregolare, si possono indicare con un solo indice $i = 1, \dots, N$ gli N punti in cui la funzione $s(x, y)$ è assegnata. Definiamo quindi il vettore dei dati come:

$$\underline{\mathbf{s}} = \begin{bmatrix} s(x_1, y_1) \\ s(x_2, y_2) \\ \dots \\ s(x_N, y_N) \end{bmatrix}$$

CAPITOLO 21 SEQUENZE 2D

La stima della superficie in un punto qualsiasi (x, y) è combinazione lineare dei valori della funzione in tutti gli N punti dati con un vettore di coefficienti $\underline{\mathbf{a}}(x, y)$, dipendente dalla posizione del punto da stimare:

$$\widehat{s}(x, y) = \underline{\mathbf{s}}^* \underline{\mathbf{a}}(x, y)$$

Il vettore dei coefficienti $\underline{\mathbf{a}}(x, y)$ viene dal sistema di equazioni normali:

$$\underline{\mathbf{R}} \underline{\mathbf{a}}(x, y) = \underline{\mathbf{r}}(x, y)$$

Gli elementi della matrice di covarianza $\underline{\mathbf{R}}$ sono:

$$r_{m,n} = r(x_m - x_n, y_m - y_n) + \sigma_w^2 \delta(m, n); \quad m, n = 1, \dots, N$$

e il vettore $\underline{\mathbf{r}}(x, y)$ è la correlazione tra il valore della funzione nel punto da stimare e quello nei punti dove è assegnata:

$$\underline{\mathbf{r}} = \begin{bmatrix} r(x - x_1, y - y_1) \\ r(x - x_2, y - y_2) \\ \dots \\ r(x - x_N, y - y_N) \end{bmatrix}$$

Si verifica che l'interpolatore **non** onora il dato, se non nel caso particolare in cui $\sigma_w^2 = 0$. Quindi, l'interpolatore può funzionare anche come filtro passa basso, per smussare il segnale misurato qualora sia affetto da rumore a spettro bianco. Il vantaggio dell'interpolazione realizzata con la stima lineare è appunto nel fatto che le formule possono essere generalizzate a situazioni di campionamento irregolare ed a rumore con spettro qualsiasi, p.e. più uniforme in una direzione che in un'altra. Non tutte le funzioni $r(x, y)$ sono possibili funzioni di autocorrelazione; è necessario che la loro trasformata sia tutta positiva. Solo in questo caso è possibile garantire che, aggiungendo luce bianca σ_w^2 non compaiano zeri nello spettro degli autovalori di $\underline{\mathbf{R}}$ e quindi il sistema delle equazioni normali sia ben condizionato. Ad esempio sono utilizzabili come funzioni di autocorrelazione funzioni del tipo:

$$r_1(x, y) = \frac{1}{d_0^2} e^{-\frac{|x|}{|d_0|} - \frac{|y|}{|d_1|}}; \quad r_2(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}\right) \quad (21.41)$$

21.7 INTERPOLAZIONE MEDIANTE STIMA LINEARE

rispettivamente separabile (la prima) e separabile ed a simmetria circolare (la seconda, se $\sigma_x^2 = \sigma_y^2$). I rispettivi spettri di potenza saranno:

$$P_1(u, v) = \frac{1}{1 + u^2 d_0^2} \times \frac{1}{1 + v^2 d_1^2} \quad (21.42)$$

$$P_2(u, v) = \exp(-u^2 \sigma_x^2 - v^2 \sigma_y^2) \quad (21.43)$$

Non andrebbe bene un prisma; andrebbe bene la forma dell'interpolatore bilineare.

21.7.4 Stima lineare ed interpolazione

Si osservi che la tecnica della stima lineare, nel caso di $\sigma_w^2 = 0$, viene a coincidere con la tecnica di interpolazione discussa nel capitolo 2 in cui la funzione interpolata è combinazione lineare di tante funzioni lisce $g(t - t_j)$ centrate agli istanti t_j dove sono posizionati i campioni. I coefficienti della combinazione lineare sono tali da fare sì che la funzione interpolata rispetti i dati. La stessa cosa accade in questo caso; invece di un asse temporale, abbiamo due assi spaziali ed inoltre il ruolo della $g(t)$ è ora sostenuto dalla funzione di autocorrelazione $r(x, y)$. Infatti, si può verificare che la funzione interpolata è somma di tante funzioni di autocorrelazione centrate sui campioni e con coefficienti tali che la loro somma rispetti il dato; con scrittura vettoriale, abbiamo trovato:

$$\underline{\mathbf{R}} \underline{\mathbf{a}}(x, y) = \underline{\mathbf{r}}(x, y) \quad (21.44)$$

La matrice $\underline{\mathbf{R}}$ non dipende dalle variabili continue x, y ma solo dalla disposizione del reticolo di campionamento e ovviamente dalla forma della funzione di autocorrelazione. Inoltre, avendo indicato con $\underline{\mathbf{s}}$ il vettore dei campioni da interpolare:

$$\hat{s}(x, y) = \underline{\mathbf{s}}^* \underline{\mathbf{a}}(x, y) = \underline{\mathbf{s}}^* \underline{\mathbf{R}}^{-1} \underline{\mathbf{r}}(x, y) = \underline{\mathbf{q}}^* \underline{\mathbf{r}}(x, y) = \quad (21.45)$$

$$= \begin{bmatrix} q_1 & q_2 & \dots & q_N \end{bmatrix} \begin{bmatrix} r(x - x_1, y - y_1) \\ r(x - x_2, y - y_2) \\ \dots \\ r(x - x_N, y - y_N) \end{bmatrix}$$

Il vettore

$$\underline{\mathbf{q}} = \underline{\mathbf{R}}^{-1} \underline{\mathbf{s}}$$

non dipende dalle variabili continue x, y ma solo dai dati, dalla $\underline{\mathbf{R}}$, e dalla posizione reciproca dei punti di campionamento. Pertanto, come già detto, la stima $\widehat{s}(x, y)$ è una combinazione lineare di funzioni di autocorrelazione centrate sugli N campioni, e moltiplicate per coefficienti q_1, \dots, q_N tali da garantire il rispetto del dato, sempre che $\sigma_w^2 = 0$. In questo caso, infatti, per $x = x_i, y = y_i$ il vettore dei termini noti $\underline{\mathbf{r}}(x, y)$ è eguale ad una riga della matrice $\underline{\mathbf{R}}$ e il vettore $\underline{\mathbf{a}}(x = x_i, y = y_i)$ si riduce ad un vettore tutto di zeri ed un solo 1 nella posizione i -esima. La funzione stimata è quindi eguale al campione situato in quella posizione. In conclusione, abbiamo ora uno strumento per dare un'interpretazione fisica alla funzione $g(t)$ incontrata nel capitolo 2.

21.8 Poligoni di Voronoi e triangoli di Delaunay

Una distribuzione irregolare di campioni su un piano può essere interpolata con una tecnica molto semplice, che è poi una generalizzazione immediata del mantentore, nel caso di un reticolo regolare. Basta infatti associare ogni punto del piano al campione più vicino: la zona di influenza di ogni campione, poligonale, è chiamata poligono di Voronoi ed è delimitata dagli assi dei segmenti che congiungono ogni campione a quelli più vicini, cioè quelli con un lato di un poligono di Voronoi in comune. Se poi si connette ogni campione ai campioni vicini (e cioè quei campioni che condividono un lato di un poligono di Voronoi) si ottiene una suddivisione del piano in triangoli. Questa suddivisione è chiamata mosaico di Delaunay (Delaunay tessellation). La figura in basso illustra il procedimento. La procedura di decomposizione in triangoli di Delaunay identifica una tecnica di interpolazione semplicissima: basta imporre che ad ogni triangolo corrisponda un piano che passa per i tre punti: ne risulta una superficie a facce piane che onora il dato. Nel caso di un reticolo regolare ed una triangolazione di Delaunay uniforme ovunque (ne sono ammesse due differenti tra loro), la risposta all'impulso è una delle due piramidi a base esagonale viste in precedenza.

Nel caso di disposizioni casuali nel piano di N campioni, si può osser-

21.8 POLIGONI DI VORONOI E TRIANGOLI DI DELAUNAY

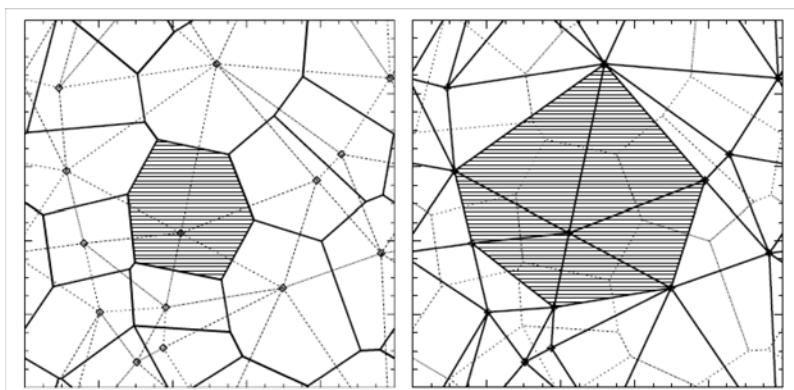


Figura 21.16: A sinistra i poligoni di Voronoi e a destra la triangolazione di Delaunay.

vare che il numero di poligoni di Voronoi é ovviamente eguale al numero N dei campioni. Quindi, la formula di Eulero

$$Facce + Vertici = Spigoli + 2 \quad (21.46)$$

applicata al poliedro costituito dai poligoni di Voronoi, ci dice che il numero medio di lati dei poligoni di Voronoi é 6. In altre parole, disponendo a caso dei campioni sul piano, ogni campione *confina* in media con altri 6. Infatti, i vertici dei poligoni di Voronoi sono sempre formati dall'incontro di 3 spigoli (a parte casi degeneri, come il reticolo regolare a base rettangolare, la probabilità che esistano punti equidistanti da 4 campioni o piú é trascurabile) e quindi:

$$N + V \sim \frac{3V}{2} \rightarrow V \sim 2N \rightarrow S \sim \frac{6N}{2} \quad (21.47)$$

ricordando di non contare 2 volte gli spigoli. I triangoli di Delaunay sono tanti quanti i vertici e quindi il doppio dei campioni. In una simulazione su $2 \cdot 10^8$ campioni distribuiti nel piano alla Poisson (e cioè con una probabilità λdA di avere un campione in un elemento di superficie dA ; piccolo a sufficienza di rendere trascurabile la probabilità che in quell'elemento cadano due campioni, Brakke), le frequenze relative del numero dei lati dei poligoni di Voronoi (da 3 a 11) sono state misurate così: 0.01, 0.1, 0.26, 0.29, 0.19, 0.09, 0.03, 0.007, 0.0015. La probabilità di avere poligoni con un numero ancora maggiore di lati é trascurabile anche se mai nulla.

21.9 Il campionamento a 3D

L'interpolazione lineare si generalizza semplicemente anche nel caso di funzioni campionate in modo casuale in spazi a 3 dimensioni. Vediamo anzitutto la generalizzazione del reticolo di campionamento ottimale. A 2D abbiamo visto che il reticolo a base esagonale è più funzionale di quello a base quadrata. Un modo per spiegare questo fatto sta nell'osservazione che, se dobbiamo impaccare sfere 2D (e quindi cerchi) di eguale diametro, la densità massima si ottiene quando i cerchi sono disposti su un reticolo a base esagonale. Se vediamo questo risultato nel dominio delle frequenze 2D, ne risulta che, a pari diametro del cerchio base e quindi a pari frequenza spaziale massima che si può riprodurre senza equivocazione, le repliche sono più vicine tra loro (più dense) se si adotta un reticolo esagonale. Ogni cerchio ne tocca altri sei, e peraltro anche con una distribuzione casuale dei campioni sul piano, avevamo verificato che il numero medio dei campioni vicini era sei. Reciprocamente, il reticolo dei campioni nello spazio sarà il più rado possibile, e quindi l'impaccamento fitto esagonale nelle frequenze porta al reticolo di campionamento ottimo, e cioè il più rado possibile per una data frequenza di taglio.

Similmente a 3D, si è di recente dimostrato quanto è intuitivo e cioè che il massimo impaccamento di sfere si ottiene con una disposizione quale quella di una *piramide di palle di cannone*; in sostanza, questa disposizione corrisponde alla sovrapposizione di strati di sfere con i centri su reticoli piani esagonali. Gli strati successivi hanno i centri sfalsati in maniera da fare penetrare le sfere dello strato superiore nelle cavità lasciate libere nello strato inferiore. Ogni sfera ne tocca altre 12.

Ancora, si possono estendere a 3D sia la triangolazione di Delaunay che i poligoni (poliedri) di Voronoi. Questi ultimi corrispondono a tutti i punti che sono vicini ad un campione più che a tutti gli altri. I poliedri sono separati da tanti piani in asse tra un campione ed uno vicino. Ogni vertice di un poliedro di Voronoi ha 4 spigoli. Questo vertice è equidistante da quattro campioni ed è il centro della sfera che li tocca. Quindi i piani che si incontrano in un vertice di un poliedro di Voronoi sono i quattro piani in asse tra i campioni equidistanti da quel vertice. Queste quaterne di campioni vicini, e cioè i vertici del tetraedro iscritto nella sfera identificata dai quattro campioni equidistanti dal vertice del poliedro in questione, formano un tetraedro di Delaunay, generalizzazione di un triangolo di Delaunay nella triangolazione piana. Così come si può

21.9 IL CAMPIONAMENTO A 3D

disporre un piano per 3 punti, si può disporre su quattro punti nello spazio una funzione che vari linearmente in funzione delle tre coordinate. In questo modo, è possibile definire una funzione continua in uno spazio 3D che assume i valori desiderati in corrispondenza dei campioni noti ed è localmente sempre variabile linearmente con le tre coordinate. In corrispondenza delle facce dei tetraedri di Delaunay, questa funzione ha delle discontinuità delle derivate prime.

Per curiosità, il numero medio di campioni adiacenti in uno spazio 3D, e con una distribuzione casuale alla Poisson dei campioni, è stato calcolato da Meijering (1953) in:

$$\frac{48\pi^2}{35} + 2 = 15,53.. \quad (21.48)$$

La relazione di Eulero, anche se generalizzata a 4D, non sarebbe sufficiente per trovare questo valore. È invece necessario fare un complesso studio di geometria aleatoria.

Capitolo 22

La Tomografia

22.1 Introduzione

Si pensi ad un bambino che osserva una torta con delle candeline; la torta è poggiata su di un tavolo ed il bimbo, che ha gli occhi all'altezza delle candeline, non può distinguere nè il loro numero nè la luminosità. Se però il bambino fa un giro attorno al tavolo su cui è poggiata la torta, vedrà le luci delle candeline allinearsi a volte lungo certe direzioni, a volte lungo altre. A volte ancora vedrà tutte le candeline distribuite lungo il diametro della torta. Ora, la tomografia è la tecnica con cui si riesce a trovare posizione e luminosità delle candeline pur senza mai vedere la torta dall'alto.

Passando dal caso particolare della torta con le candeline al caso generale della tomografia, abbiamo la seguente situazione. Si desidera rilevare una grandezza (ad esempio l'energia luminosa emessa dalle candeline) il cui andamento nel piano è descritto dalla funzione $f(x, y)$; questa grandezza tuttavia non può essere rilevata punto per punto; la cosa sarebbe possibile se disponessimo di un misuratore nel punto di coordinate x, y ma questo è invece inaccessibile. Stando all'esterno del corpo, invece, è possibile misurare solo **proiezioni** e cioè integrali della funzione $f(x, y)$ lungo le rette con direzione $\varphi + \pi/2$ (fig. 22.1); nel caso della torta possiamo misurare la somma di tutta l'energia luminosa generata dalle candeline lungo delle linee che indicheremo con ℓ .

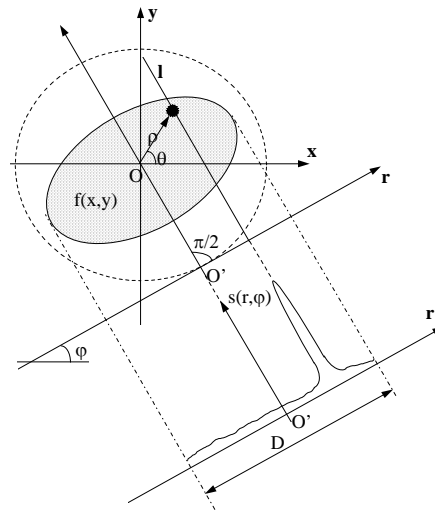
Indicheremo con \mathbf{r} l'asse lungo il quale sono situati i punti di rilevazione e con r le sue ascisse; l'asse è ortogonale alle linee ℓ ed è quindi

CAPITOLO 22 LA TOMOGRAFIA

inclinato di φ rispetto all'asse x . L'origine delle coordinate dell'asse \mathbf{r} è posta nel punto O' proiezione sulla retta \mathbf{r} dell'origine delle coordinate del piano x, y . Una proiezione è quindi data dalla funzione

$$s(r, \varphi) = \int_{-\infty}^{\infty} f(\xi, \eta) dl$$

Vedremo ora come fare il processo inverso e cioè passare dalle proiezioni $s(r, \varphi)$; $0 < \varphi < \pi$ all'immagine originale. Questa operazione di **inversione** è indicata con il nome di **tomografia** da $\tau\acute{\epsilon}\mu\nu\omega = \textit{taglio}$ in greco. Per comprendere il meccanismo dell'inversione tomografica, si rammenti che il problema è lineare e si può procedere per sovrapposizione degli effetti o utilizzare il dominio dei numeri d'onda.



D = dimensione massima dell'oggetto

Figura 22.1: Ripresa di proiezioni dall'esterno di un corpo

22.2 Trattamento nei numeri d'onda

Utilizzeremo il teorema detto delle sezioni di proiezione (*projection slices*) della trasformata di Fourier 2D e cioè utilizzeremo il fatto che la trasfor-

22.2 TRATTAMENTO NEI NUMERI D'ONDA

mata di Fourier di una proiezione (ad assegnato angolo φ con l'asse x), è la *sezione* della trasformata di Fourier bidimensionale $F(u, v)$ della $f(x, y)$ che ha eguale angolo φ con l'asse u . Determiniamo analiticamente l'espressione di una proiezione $s(r, \varphi)$ e della sua trasformata di Fourier:

$$s(r, \varphi) = \int_{-\infty}^{\infty} f(\xi, \eta) d\ell \quad (22.1)$$

dove ℓ è la retta di coordinate:

$$\ell \rightarrow \rho \cos(\theta - \varphi) = r \quad (22.2)$$

le variabili ρ, θ sono le coordinate polari nel piano x, y :

$$\rho = \sqrt{x^2 + y^2} \quad \theta = \tan^{-1} \frac{y}{x} \quad (22.3)$$

In coordinate cartesiane si avrebbe:

$$\ell \rightarrow x \cos \varphi + y \sin \varphi = r$$

Nel caso di $\varphi = 0$ l'espressione di una proiezione si semplifica in:

$$s(r, 0) = \int_{-\infty}^{\infty} f(r, y) dy \quad -\infty \leq r \leq \infty \quad (22.4)$$

e la variabile r coincide con la variabile x . Se la trasformata di Fourier dell'immagine è:

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j(ux+vy)} dx dy$$

la trasformata di Fourier della proiezione è, indicando con Ω la variabile coniugata alla variabile r :

$$\int_{-\infty}^{\infty} s(r, 0) e^{-j\Omega r} dr = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(r, y) e^{-j\Omega r} dr dy = F(\Omega, v = 0) \quad (22.5)$$

Ruotando gli assi del piano x, y di un angolo φ ruotano di altrettanto gli assi coordinati della trasformata $F(u, v)$. Quindi la trasformata di Fourier di una proiezione $s(r, \varphi)$ è $F(\Omega, \varphi)$ se introduciamo le coordinate polari nel dominio dei numeri d'onda u, v :

$$\Omega = \sqrt{u^2 + v^2}; \quad \varphi = \tan^{-1} \frac{v}{u} \quad (22.6)$$

CAPITOLO 22 LA TOMOGRAFIA

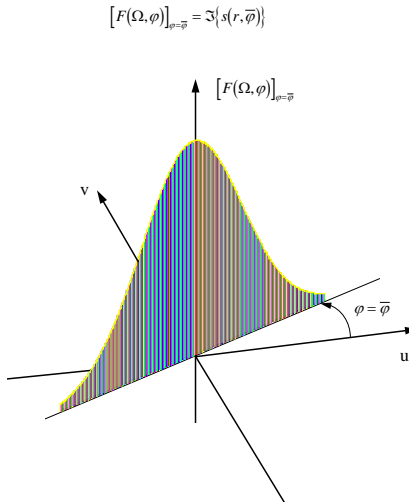


Figura 22.2: Una sezione di una trasformata di Fourier 2D, trasformata di una proiezione

Si osservi che ora $\Omega \geq 0$, e una sezione completa della trasformata di Fourier viene ottenuta considerando i valori della $F(\Omega, \varphi)$ sia per l'angolo $0 \leq \varphi \leq \pi$ che per l'angolo $\varphi + \pi$. In conclusione abbiamo mostrato come la trasformata di Fourier 1D della proiezione $s(r, \varphi)$ fatta lungo le rette ℓ che fanno un angolo $\varphi + \pi/2$ con l'asse x è eguale alla sezione della trasformata di Fourier 2D $F(u, v)$ dell'immagine $f(x, y)$ lungo la direzione che fa angolo φ con l'asse u e cioè $F(\Omega, \varphi)$, $F(\Omega, \varphi + \pi)$ (fig. 22.2).

$$s(r, \varphi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\Omega, \varphi) e^{j\Omega r} d\Omega \quad (22.7)$$

I valori di Ω negativi sono riferiti all'angolo $\varphi + \pi$.

La cosa diventa ovvia se si pensa che una proiezione è ottenuta convolvendo l'immagine originale per una lama di altezza costante, parallela all'asse della proiezione stessa; convolvere con una lama corrisponde a moltiplicare la trasformata di Fourier 2D per la lama ortogonale, trasformata della lama negli spazi. Pertanto, se sono note M proiezioni disposte regolarmente lungo un arco di 180° sono note M sezioni della trasformata

22.2 TRATTAMENTO NEI NUMERI D'ONDA

di Fourier 2D,

$$F(\Omega, \varphi_m); m = 1, \dots, M; \quad F(\Omega, \varphi_m + \pi); m = 1, \dots, M. \quad |\Omega| \geq 0;$$

Da queste sezioni che suddividono uniformemente l'angolo giro, con una interpolazione fatta nel dominio dei numeri d'onda, è possibile trovare la $F(u, v)$ su di un reticolo cartesiano regolare; con una antitrasformazione discreta si riottiene la $f(x, y)$ (fig. 22.3). Si può ora calcolare il numero di proiezioni necessario per fare una corretta interpolazione nel dominio dei numeri d'onda. I campioni della $F(\Omega, \varphi_m)$ sono massimamente distanti

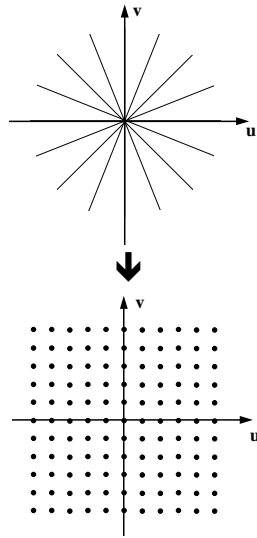


Figura 22.3: Passaggio da coordinate polari a cartesiane

tra loro per i valori massimi di $|\Omega|$; ora se indichiamo con L la larghezza del lobo principale di un filtro che delimita la risoluzione spaziale dell'immagine, si ha

$$|\Omega| \leq \frac{\pi}{L} \quad (22.8)$$

Se l'oggetto ha dimensione massima D , i campioni nel dominio spettrale devono essere distanti meno di $2\pi/D$ perchè le repliche dovute al campionamento nei numeri d'onda non si sovrappongano negli spazi; quindi

CAPITOLO 22 LA TOMOGRAFIA

(fig. 22.4):

$$\frac{\pi}{L} \times \frac{\pi}{M} \leq \frac{2\pi}{D} \quad \rightarrow \quad M \geq \frac{\pi D}{2L} \quad (22.9)$$

Dall'analisi fatta nel dominio dei numeri d'onda si conclude che è possibile

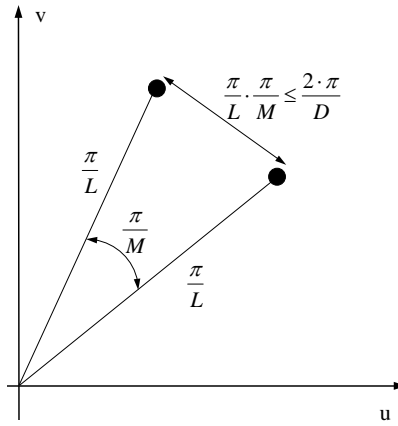


Figura 22.4: Condizione di corretto campionamento nel dominio dei numeri d'onda

effettuare una inversione tomografica:

- se la risoluzione spaziale dell'immagine è stata limitata e lo spettro dell'immagine è sostanzialmente nullo oltre il numero d'onda $|\Omega| = \pi/L$;
- se il numero delle proiezioni M è sufficiente e cioè $M \geq 1,57D/L$ dove D è il diametro dell'oggetto da tomografare.

Ripeteremo ora l'analisi rimanendo nel dominio degli spazi, per comprendere i meccanismi che dominano la tomografia e così avere gli strumenti concettuali per estendere la trattazione a situazioni di geometria irregolare. Per ora limitiamoci ad osservare che servono apparentemente più equazioni di quante siano le incognite in quanto con una risoluzione L ed un diametro D l'oggetto è descritto da $\simeq \pi(D/2L)^2$ valori diversi di ampiezza (la risoluzione spaziale corrispondente alle differenti luminosità di altrettante *candeline*). Le equazioni disponibili sono invece almeno $M \times D/L$ e cioè il doppio. Ciò implica che queste equazioni sono linearmente dipendenti tra loro; M relazioni sono trovate subito osservando che l'"area" di tutte le proiezioni (l'energia luminosa emessa da tutte le candeline) è costante e pari al "volume" racchiuso dalla funzione $f(x, y)$.

Questo equivale a osservare che "la componente a frequenza zero" di tutte le proiezioni è costante e pari a $F(0,0)$. Similmente, si intuisce che i campioni a bassa frequenza spaziale di tutte le proiezioni sono sì funzione di φ , ma dipendono dai pochi valori indipendenti assunti dalla $F(u, v)$ attorno all'origine delle coordinate, ecc.

22.3 Il sinogramma

Per comprendere il meccanismo della tomografia nel dominio degli spazi, faremo uso del principio di sovrapposizione degli effetti e quindi supporremo che la funzione $f(x, y)$ corrisponda ad un singolo impulso unitario (*la torta abbia una sola candelina*). Ad esempio, se l'impulso (la luce emessa dalla candelina) occupa l'origine degli assi O, le proiezioni misurate dall'osservatore avranno un picco all'ascissa O', proiezione sull'asse \mathbf{r} del punto O, qualsiasi sia il valore di φ . Si riportino tali proiezioni in un'immagine e cioè un grafico tridimensionale (detto **sinogramma** per i motivi che vedremo tra poco) che ha per ascissa r e per ordinata φ ; lungo la terza dimensione (il livello di grigio dell'immagine), non riportata nella figura 22.5, dovrebbe comparire il valore dell'integrale di linea della funzione $f(x, y)$ fatto integrando lungo la linea ℓ come si vede nella figura 19.1. Il risultato che si ottiene nel caso di una *candelina* nell'origine è un'immagine con una linea verticale in corrispondenza dell'ascissa 0 (la retta $r = 0$ nella fig. 22.5) il cui livello di grigio è proporzionale alla luminosità della *candelina*. Se invece sono presenti tre *candeline*, nel sinogramma compaiono, oltre la traccia verticale vista prima, altre due tracce di forma sinusoidale e di livello di grigio proporzionale alla luminosità delle altre due *candeline* (nel caso di fig. 22.6, una delle tre sorgenti (quella in B) ha segno negativo).

Supponiamo ora che le coordinate polari dell'unico impulso che forma la $f(x, y)$ siano ρ_0 e θ_0 . In ogni riga del sinogramma e cioè per ogni valore di φ l'unico valore di r per cui il sinogramma ha valore diverso da 0 oscilla tra $+\rho_0, -\rho_0$ con una traiettoria sinusoidale (ecco il perchè del nome sinogramma):

$$r = \rho_0 \cos(\theta_0 - \varphi) \quad (22.10)$$

Naturalmente, se invece di un impulso (*una candelina sulla torta*) ve ne sono molti, ad esempio in A, B, C, nella riga del sinogramma che corrisponde alla proiezione di angolo φ sarà registrato l'andamento relativo

CAPITOLO 22 LA TOMOGRAFIA

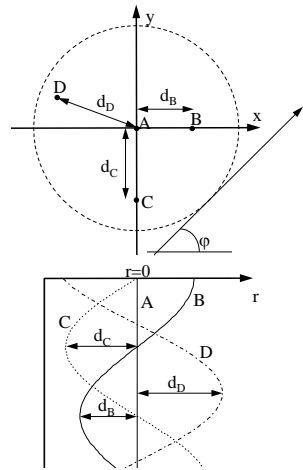


Figura 22.5: Sinogramma

sinogramma non filtrato

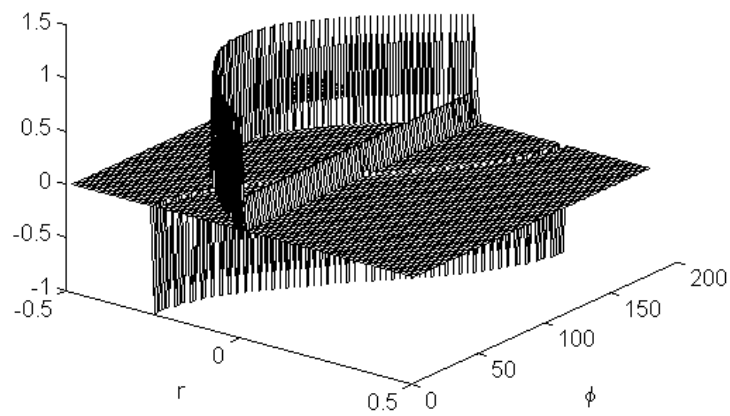


Figura 22.6: Sinogramma di 3 impulsi.

22.4 LA RETROPROIEZIONE

alla somma delle energie luminose di tutte le candeline posizionate ai punti di coordinate polari ρ, θ legate dalla relazione:

$$\rho \cos(\theta - \varphi) = r \quad (22.11)$$

Risulta quindi che l'ampiezza del sinogramma è:

$$s(r, \varphi) = \int_{-\infty}^{\infty} f(x, y) d\ell \quad (22.12)$$

e cioè l'integrale della $f(x, y)$ (rappresentata in coordinate polari come $f(\rho, \theta)$) lungo le linee ℓ .

22.4 La retroproiezione

Se si desidera ora passare dal sinogramma $s(r, \varphi)$ di nuovo all'immagine in coordinate polari $f(\rho, \theta)$, si deve effettuare l'operazione di **retroproiezione**. Un modo elementare di vedere il meccanismo di retroproiezione, direttamente corrispondente al nome, è il seguente. Si visualizzi l'immagine $f(x, y)$ come una distribuzione disuniforme di burro su di una fetta di pane; la proiezione lungo una direzione $\varphi + \pi/2$ dà l'informazione, ascissa per ascissa, di *quanto burro* sia disposto lungo quella direzione. La ricostruzione più ragionevole dell'immagine che si può fare, non disponendo di ulteriore informazione ma garantendo la congruenza con la proiezione, corrisponde all'operazione di *spalmare il burro sulla fetta di pane*. In altre parole, le ampiezze del segnale vengono prima sommate tra loro lungo la direzione ℓ (formazione della proiezione) e poi ridistribuite sull'immagine (*spalmate*) ancora lungo la stessa direzione (la retroproiezione); si passa così dall'immagine originale $f(x, y)$ ad una funzione cilindrica uniforme nella direzione $\varphi + \pi/2$, pur mantenendo eguale proiezione $s(r, \varphi)$. Un impulso, riproiettato, diventa una lama, passante per l'impulso originario, ortogonale alla direzione di proiezione. Completata l'operazione di retroproiezione relativa ad un direzione di proiezione (un azimut), si devono adesso utilizzare tutti gli altri azimut.

L'operazione di retroproiezione viene quindi completata mediando tra loro tutte le retroproiezioni effettuate lungo i vari azimut; ciò equivale a stimare la distribuzione originale del burro mediando tra loro *le quantità di burro rispalmate dalle varie direzioni*. Le stesse operazioni di

CAPITOLO 22 LA TOMOGRAFIA

retroproiezione dai vari azimut possono essere ottenute in modo equivalente operando come segue. Tutte le informazioni relative allo stesso punto dell'immagine posizionato nel punto P di coordinate polari (ρ_0, θ_0) , provenienti dalle proiezioni della $f(x, y)$ lungo i vari azimut, sono disposte nel sinogramma lungo la traiettoria sinusoidale descritta dall'equazione (22.10). Mediando i valori del sinogramma lungo questa traiettoria, si trova una stima del valore della $f(x, y)$, che indichiamo con $\hat{f}(x, y)$. Questa operazione di media coincide con la retroproiezione perchè si attribuisce alla $f(\rho, \theta)$ il valore medio delle proiezioni fatte sui punti di ascissa $r = \rho \cos(\theta - \varphi)$ al variare dell'angolo φ . Mediando su tutti gli azimut, la retroproiezione è quindi definibile come:

$$\hat{f}(\rho, \theta) = \frac{1}{2} \times \frac{1}{\pi} \int_0^{\pi} s[r = \rho \cos(\theta - \varphi), \varphi] d\varphi \quad (22.13)$$

formula che vedremo sarà utile nella derivazione formale della retroproiezione. La divisione per 2 è una conseguenza del sistema di riferimento scelto (ovvero $0 \leq \varphi < \pi$ mentre ρ può essere sia positivo che negativo e quindi la retroproiezione è equivalente, in pratica, ad effettuare l'operazione di media su 2π). Vediamo cosa accade con M proiezioni se la $f(x, y)$ è costituita da un singolo impulso ad esempio nell'origine O . Ogni retroproiezione è una lama di ampiezza costante che passa per O ; la $\hat{f}(x, y)$ è ottenuta dalla somma di M lame che passano tutte per O . Anche per $M \rightarrow \infty$ la $\hat{f}(x, y)$ non è l'impulso di partenza $\delta(x, y)$ ma la funzione:

$$q(x, y) = \frac{1}{\sqrt{x^2 + y^2}} = \frac{1}{\rho} \quad (22.14)$$

la cui trasformata di Fourier 2D è:

$$Q(\Omega) = \frac{1}{\sqrt{u^2 + v^2}} = \frac{1}{|\Omega|} \quad (22.15)$$

Infatti, basta pensare che le lame siano M e che ogni lama abbia uno spessore δ ; il volume totale della media delle lame in una corona circolare di raggio ρ e larghezza Δ è costante e pari a $M\delta\Delta/M$ (tale volume deve essere normalizzato rispetto al numero delle proiezioni da cui segue la divisione per M), quindi l'altezza equivalente in grado di fornire il medesimo volume delle M retroproiezioni è $M\delta/2\pi M\rho$ e decresce con ρ . Ne consegue che la retroproiezione ottenuta ridistribuendo il valore della

22.5 LA RETROPROIEZIONE FILTRATA

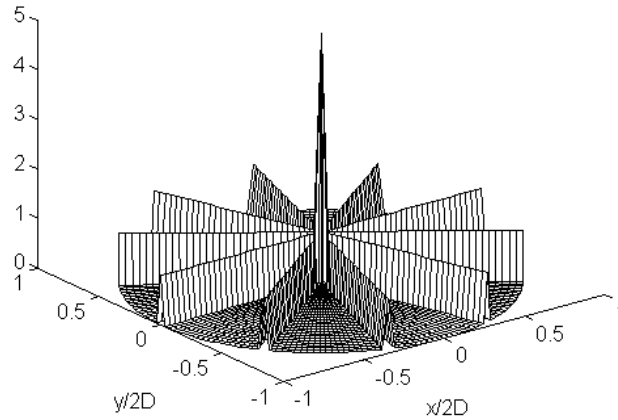


Figura 22.7: Retroproiezione di 5 proiezioni di un impulso nell'origine.

proiezione su tutti i campioni dell'immagine NON restituisce l'immagine di partenza bensì l'immagine convoluta con un filtro la cui risposta all'impulso è $q(x, y) = 1/\rho$. Per ottenere di nuovo un impulso "ideale" fino alla frequenza di Nyquist, si deve compensare (deconvolvere) la risposta all'impulso $q(x, y)$ e cioè moltiplicare (nelle frequenze) per $|\Omega|$, ovviamente ancora fino alla frequenza di Nyquist. Il motivo è che ad ogni lama nel dominio degli spazi (ad esempio $\delta(x)$) corrisponde una lama nel dominio dei numeri d'onda (in questo caso $\delta(v)$), e le due funzioni dell'equazione precedente sono entrambe somma di infinite lame che si incrociano nell'origine. Se il numero di proiezioni è limitato, si comprende come la risposta all'impulso del sistema di retroproiezione non sia un impulso ma piuttosto una stella di M raggi centrata là dove era situato l'impulso originario (fig.22.7, 22.8).

22.5 La retroproiezione filtrata

Abbiamo visto come in ogni punto del sinogramma si sommano i contributi relativi ai molti impulsi che compongono l'immagine; facendo la media del sinogramma lungo tutta la traiettoria relativa al punto di interesse P_0 di coordinate polari (ρ_0, θ_0) , i contributi degli altri impulsi **non**

CAPITOLO 22 LA TOMOGRAFIA

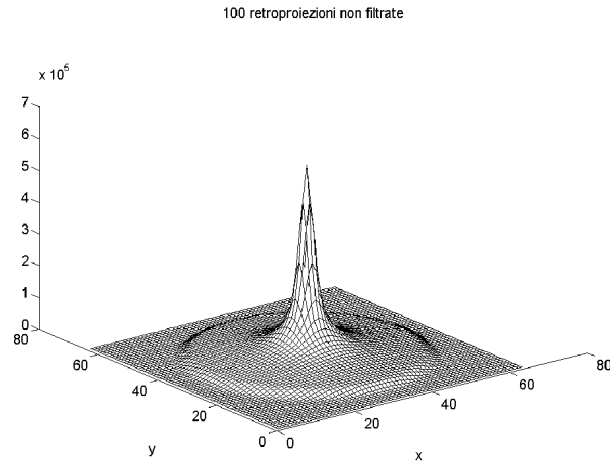


Figura 22.8: Retroproiezione di 100 proiezioni di un impulso.

si elidono. Se però deconvolviamo per la $q(x, y)$ (prima di sommare lungo le traiettorie sinusoidali nel sinogramma) e quindi convolviamo ogni proiezione con una opportuna funzione $\tilde{\rho}(r)$ la cui trasformata 1D vedremo essere $|\Omega|$, è possibile eliminare l'effetto dell'interferenza dei vari impulsi tra di loro (fig. 22.9). Supponiamo ad esempio che $\tilde{\rho}(0)$ abbia valore unitario; supponiamo inoltre che $\tilde{\rho}(r)$ abbia un lobo principale di larghezza circa L e lobi laterali negativi e positivi di larghezza L e di ampiezza decrescente in modo che il valor medio della $\tilde{\rho}(r)$ sia nullo (una specie di sinc). Se la funzione $f(x, y)$ è costituita da un solo impulso (ad

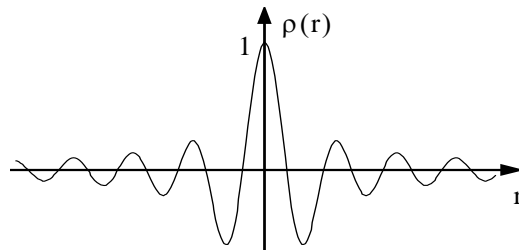


Figura 22.9: Il filtro per le retroproiezioni $\tilde{\rho}(r)$

22.5 LA RETROPROIEZIONE FILTRATA

esempio nell'origine), i valori del sinogramma lungo la traiettoria relativa alla posizione dell'impulso (la retta $r = 0$) non sono modificati dalla convoluzione con $\tilde{\rho}(r)$. Vediamo ora che accade se valutiamo l'ampiezza dell'immagine in un altro punto, distante ρ_1 dall'origine ($\rho_1 \gg L$). Per fare ciò dobbiamo sommare i valori assunti dal sinogramma lungo la traiettoria sinusoidale relativa a questo secondo impulso che traversa la retta $r = 0$ secondo una linea obliqua. Nella somma, si incontrano e si mediano successivamente tutti i lobi positivi e negativi della $\tilde{\rho}(r)$, che hanno per ipotesi somma nulla. Se ne conclude che, filtrando le proiezioni con la $\tilde{\rho}(r)$, l'interferenza tra impulsi abbastanza lontani viene soppressa (fig. 22.10, 22.11, 22.12, 22.13); ciò conferma che la moltiplicazione per $|\Omega|$ effettua la deconvoluzione. Non è così se i due impulsi sono più vicini di L . Allora, sommando lungo la seconda traiettoria, non si incontrano tutti i lobi sia positivi che negativi della funzione $\tilde{\rho}(r)$, ma ci si trattiene sempre nel lobo principale e perciò vi sarà interferenza. Da quanto visto fino ad ora, la risposta all'impulso del sistema tomografico con retroproiezione filtrata è prossima a quella ideale, a parte la limitazione a L della risoluzione spaziale, corrispondente alla limitazione della banda alla frequenza di Nyquist. Le lame che formano la stella nel caso della retroproiezione non filtrata hanno ora lobi positivi e negativi che tendono ad elidersi tra loro; se vi sono molte proiezioni, l'immagine ricostruita ha pertanto un lobo centrale isolato largo circa L . Vedremo ora l'effetto della limitazione del numero delle proiezioni e cioè l'equivocazione spaziale. Da quanto osservato finora si dedurrebbe che conviene fare L il più piccolo possibile; vedremo nella prossima sezione perchè questo non è vero.

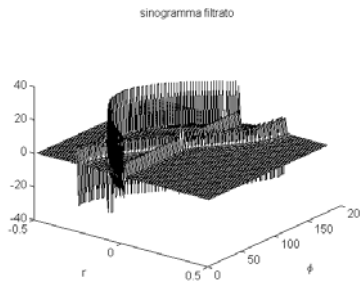


Figura 22.10: Sinogramma filtrato di 3 impulsi (Fig. 22.5).

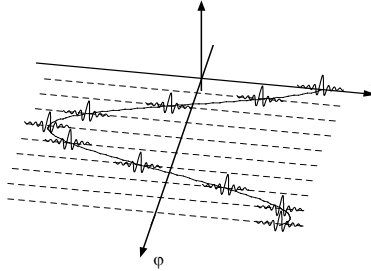


Figura 22.11: Il sinogramma filtrato

22.5.1 Equivocazione angolare

Se il campionamento angolare è insufficiente, vedremo ora che vi è interferenza tra gli impulsi se questi sono lontani tra loro: avremo cioè un effetto di equivocazione angolare se la differenza angolare tra due proiezioni successive $\Delta\varphi$ è

$$\Delta\varphi > 2L/D \quad (22.16)$$

dove D è la massima dimensione dell'oggetto osservato. Supponiamo infatti di avere due impulsi distanti tra loro D ; a partire dall'azimut φ_0 per cui i due impulsi sono sovrapposti nella proiezione, le loro due proiezioni si distanziano al crescere di φ di

$$D \sin(\varphi - \varphi_0) \simeq D(\varphi - \varphi_0)$$

Se ora $\varphi_n = n\Delta\varphi$, passando da una proiezione alla successiva, la $\tilde{\rho}(r)$ relativa al primo impulso viene campionata dalla traiettoria relativa al secondo impulso per valori dell'ascissa pari a

$$n\Delta x = nD\Delta\varphi$$

Ma se la (22.16) è vera, $\Delta x > 2L$, si saltano addirittura lobi della $\tilde{\rho}(r)$, la media non sarà più a valore nullo e quindi appare l'interferenza. In conclusione, si verifica che la risposta all'impulso dell'operazione di retroproiezione con campionamento angolare di intervallo $\Delta\varphi$ sarà caratterizzata da un lobo centrale di diametro L intorno a P_0 , da una zona oscura di diametro $2L/\Delta\varphi$ dove i lobi positivi e negativi delle varie lame si elidono, e infine a distanza ancora maggiore da P_0 , da una stella di tanti raggi quante sono le proiezioni (fig.22.12).

22.5 LA RETROPROIEZIONE FILTRATA

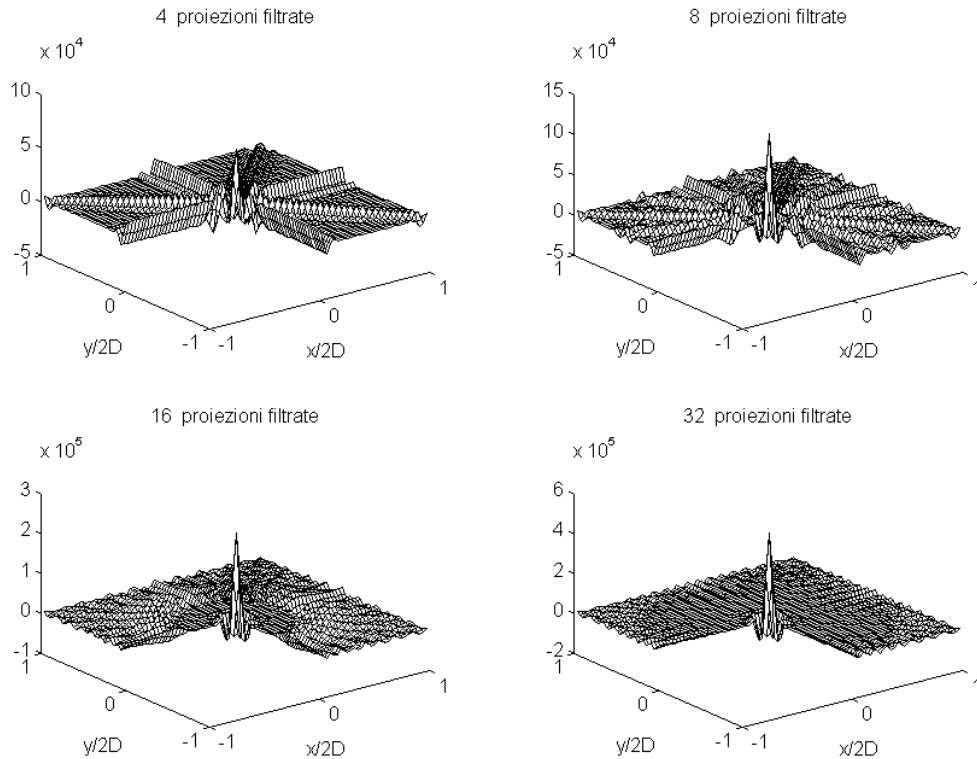


Figura 22.12: Ricostruzione dell'impulso da 4,8,16,32 proiezioni filtrate (particolare della ricostruzione).

22.5.2 Retroproiezioni e numeri d'onda

Possiamo ora interpretare l'operazione di retroproiezione avvalendoci del teorema delle *projection slices* e cioè delle sezioni della trasformata di Fourier 2D. Fare la retroproiezione, equivale a comporre l'immagine con tante funzioni cilindriche (le singole proiezioni retroproiettate) disposte lungo i vari azimut di proiezione. Lo spettro 2D dell'immagine così ricostruita è la media di tante lame, di ampiezza proporzionale alla corrispondente sezione della trasformata di Fourier 2D. Finestrando negli spazi, si smussa nel dominio dei numeri d'onda e quindi si ridistribuisce lo spettro da una distribuzione a lame ad un'altra più smussata che cor-

CAPITOLO 22 LA TOMOGRAFIA

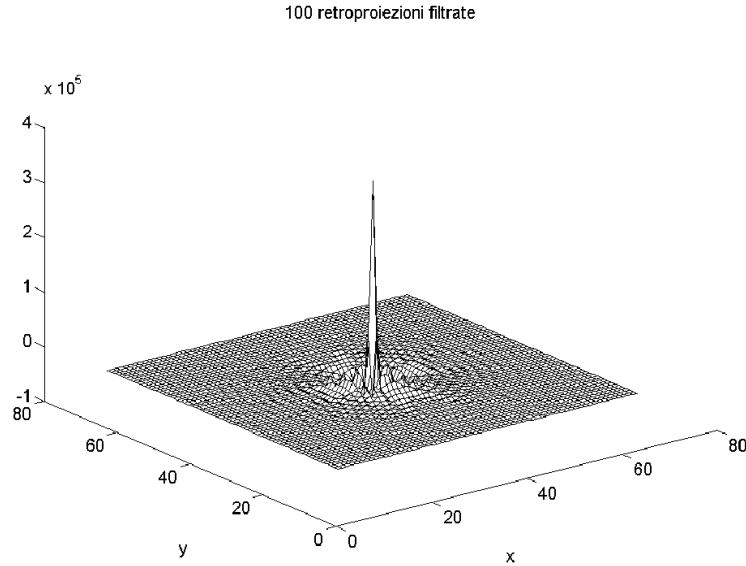


Figura 22.13: Ricostruzione da 100 retroproiezioni filtrate (confrontare con Fig.22.8).

risponde maggiormente al vero spettro 2D. Tuttavia, si deve evitare l'equivocazione e da ciò derivano le stesse condizioni già viste in precedenza sul numero delle proiezioni. Inoltre, le sezioni si affollano al centro delle coordinate del piano u, v e sono rare alla periferia; è necessario amplificare il risultato in proporzione alla distanza dal centro delle coordinate e quindi alla pulsazione spaziale $|\Omega|$. Una osservazione equivalente era già stata fatta prima, quando avevamo visto che la stella di lame corrispondenti alla retroproiezione non filtrata ha spettro $1/\Omega$.

Arriveremo ora a queste stesse conclusioni in modo più formale. Definiamo la retroproiezione come:

$$\hat{f}(\rho, \theta) = \frac{1}{2\pi} \int_0^{\pi} s[r = \rho \cos(\theta - \varphi), \varphi] d\varphi \quad (22.17)$$

Sostituendo l'espressione della proiezione in funzione della antitrasformata di Fourier dell'immagine si ha:

$$\hat{f}(\rho, \theta) = \frac{1}{2\pi} \int_0^{\pi} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} F(\Omega, \varphi) e^{j\Omega\rho \cos(\theta-\varphi)} d\Omega \right) d\varphi = \quad (22.18)$$

22.5 LA RETROPROIEZIONE FILTRATA

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{1}{2\pi} \int_0^{\infty} F(\Omega, \varphi) e^{j\Omega\rho \cos(\theta-\varphi)} d\Omega \right) d\varphi$$

Ma questa è esattamente l'espressione dell'antitrasformata di Fourier 2D **tranne il fattore** $|\Omega|$, corrispondente allo Jacobiano della trasformazione delle coordinate da cartesiane a polari:

$$\hat{f}(\rho, \theta) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_0^{\infty} \frac{F(\Omega, \varphi)}{|\Omega|} e^{j\Omega[\rho \cos(\theta-\varphi)]} \Omega d\Omega d\varphi \quad (22.19)$$

La retroproiezione non filtrata è l'antitrasformata della $F(\Omega, \varphi)/|\Omega|$; la moltiplicazione per la funzione $1/|\Omega|$ corrisponde alla convoluzione bidimensionale per la sua antitrasformata $1/\rho$ che ha **volume infinito**; infatti $1/|\Omega| = \infty$, nell'origine. In altre parole, ricostruendo l'immagine con le retroproiezioni senza filtraggio con i cosiddetti ρ -filtri, e cioè senza la convoluzione con la $\tilde{\rho}(r)$, otteniamo un'immagine con un'eccessiva riduzione di ampiezza alle frequenze spaziali piú elevate. Per ottenere un'immagine corretta si deve pertanto moltiplicare la trasformata della proiezione per la funzione $|\Omega|$. Inoltre, il numero d'onda massimo contenuto nelle proiezioni deve essere inferiore a π/L ; si osservi infine che invece di filtrare 2D l'immagine con il filtro $1/|\Omega|$, basta filtrare le singole proiezioni con un filtro passa alto la cui funzione di trasferimento è (asse delle pulsazioni bilatero):

$$\tilde{P}(\Omega) = |\Omega| = -j\Omega \times j \text{sign}(\Omega); \quad 0 \leq |\Omega| \leq \pi/L; \quad (22.20)$$

per ottenere il risultato corretto (fig. 22.14). Computazionalmente ciò è molto meno oneroso. Quindi, prima di fare la retroproiezione, nel caso continuo ed a risoluzione spaziale limitata, si dovrebbe fare la derivata delle proiezioni e poi fare seguire alla derivazione una trasformazione di Hilbert; il risultato della retroproiezione sarebbe quello desiderato. Nel dominio degli spazi, il ρ -filtro è descritto dalla seguente espressione:

$$\tilde{\rho}(r) = \frac{\pi}{L^2} \left[\text{sinc}\left(\frac{r}{L}\right) - \frac{1}{2} \text{sinc}^2\left(\frac{r}{2L}\right) \right]; \quad (22.21)$$

Nel caso discreto, si osservi che se il numero delle proiezioni è pari a M , il ρ -filtro $\tilde{P}(\Omega)$ deve avere un guadagno non nullo a frequenza zero e pari a $1/M$.

CAPITOLO 22 LA TOMOGRAFIA

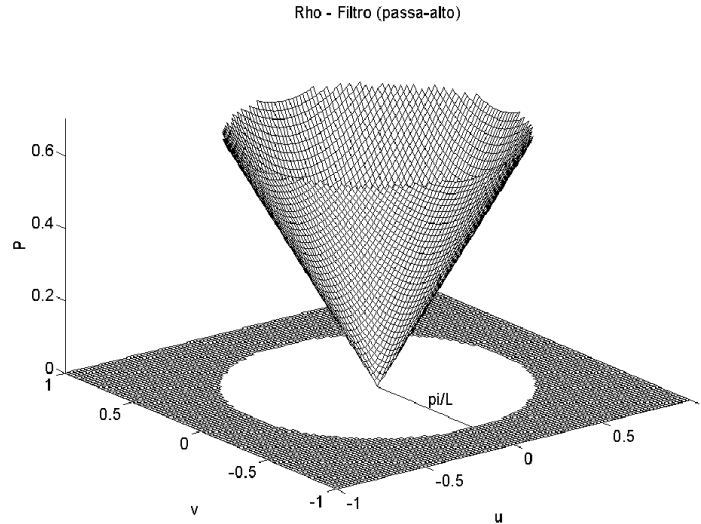


Figura 22.14: Filtro 2D passa alto equivalente al filtraggio delle retroproiezioni.

La tecnica delle retroproiezioni filtrate è computazionalmente più conveniente di quella dell'interpolazione su di un reticolo cartesiano della trasformata di Fourier bidimensionale nota su un reticolo polare, in quanto il mezzo di calcolo può essere utilizzato con maggiore uniformità temporale durante le acquisizioni. Nel caso delle retroproiezioni filtrate, è possibile fare la convoluzione con il filtro $\tilde{P}(\Omega)$ man mano che le proiezioni sono acquisite; si può anche fare immediatamente la retroproiezione della proiezione filtrata. Non appena viene acquisita e trattata l'ultima proiezione, il risultato è disponibile, senza che sia necessario un picco di utilizzo del mezzo di calcolo. Nel caso dell'interpolazione su reticolo cartesiano della trasformata di Fourier bidimensionale, è invece necessario ritardare l'antitrasformazione bidimensionale, e pertanto un forte uso di potenza di calcolo, ma solo al momento in cui sono disponibili tutte le proiezioni.

22.5 LA RETROPROIEZIONE FILTRATA

5 retroproiezioni filtrate nel dominio u,v

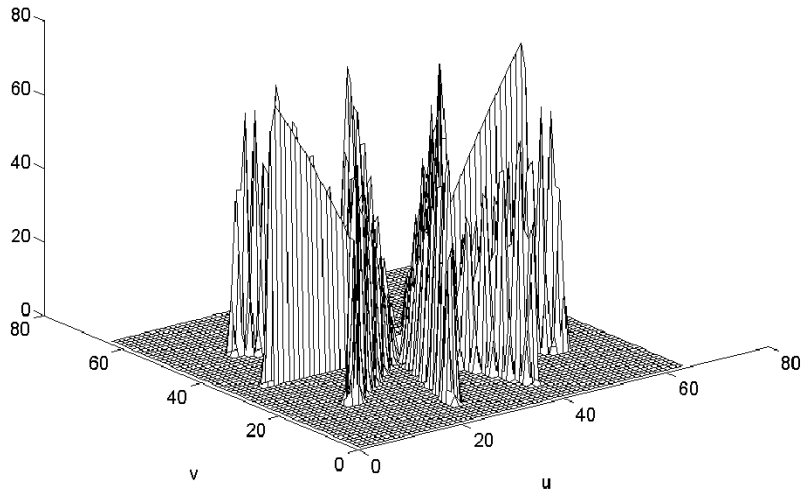


Figura 22.15: 5 retroproiezioni filtrate viste nel dominio dei numeri d'onda

22.5.3 Proiezioni incomplete o mosse

E' facile determinare cosa accade se non tutte le proiezioni vengono misurate, se cioè per qualche motivo mancano le *sezioni* dello spettro 2D del dato appartenenti ad un settore angolare $\Delta\varphi$. Il risultato è che il segnale viene filtrato da un filtro a ventaglio di larghezza $\Delta\varphi$ che sopprime le proiezioni appartenenti all'angolo non misurato. Ancora, può essere utile calcolare l'effetto di un movimento del soggetto ripreso durante la tomografia; supponiamo che durante la prima metà delle proiezioni il soggetto si trovi in O e durante la seconda metà delle proiezioni si trovi in O' . La risposta è di nuovo semplice; basta pensare di filtrare con il filtro a ventaglio corrispondente alle prime proiezioni l'impulso posizionato in O e poi con il filtro a ventaglio complementare l'impulso spostato nella seconda posizione O' . Il risultato sarà la combinazione delle due risposte all'impulso. Se p.e. le proiezioni relative ad una posizione sono corrispondenti ad un'apertura angolare di $\pi/2$ usando un sistema di coordinate

CAPITOLO 22 LA TOMOGRAFIA

conveniente, i due filtri a ventaglio sono rappresentati dall'espressione:

$$2 \times H_{1,2}(u, v) = [1 \pm \text{sign}(u)\text{sign}(v)] \leftrightarrow \delta(0, 0) \pm \frac{1}{\pi^2 xy} \quad (22.22)$$

Se le posizioni dei due impulsi sono corrispondenti a:

$$P_{1,2}(u, v) = e^{\pm j(ux_0 + vy_0)} \quad (22.23)$$

il risultato è:

$$H_1(u, v)P_1(u, v) + H_2(u, v)P_2(u, v) \leftrightarrow \quad (22.24)$$

$$\begin{aligned} & \delta(x - x_0, y - y_0) + \frac{1}{\pi^2(x - x_0)(y - y_0)} \\ & + \delta(x + x_0, y + y_0) - \frac{1}{\pi^2(x + x_0)(y + y_0)} \end{aligned}$$

Nel caso di segnale campionato e periodicizzato, le singolarità sono evitate; si utilizzi l'espressione della trasformazione di Hilbert riportata nel Capitolo 7.

22.6 Tomografia 3D

È interessante analizzare cosa accade estendendo a tre dimensioni il concetto di tomografia, discussa in 2 dimensioni; infatti, si è supposto che la propagazione avvenga nel piano (*della torta*) (x, y) detto assiale (infatti TAC: tomografia assiale computerizzata) e cioè nel piano ortogonale alla direzione z .

Anzitutto vediamo come nella trasformazione di Fourier a 3D, una retta è coniugata al piano che ha la giacitura ortogonale. In altre parole, una singolarità *a bacchetta* si trasforma in una singolarità *a lamina*; infatti il comportamento è uniforme nella direzione della bacchetta ed impulsivo nelle due direzioni ad essa ortogonali; correlativamente, nel dominio coniugato, abbiamo un comportamento impulsivo andando nella direzione della bacchetta, e troviamo una costante andando nelle due direzioni ortogonali. Indicando con u, v, w le variabili coniugate alle variabili x, y, z , la trasformazione di Fourier in 3D è:

$$F(u, v, w) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y, z) e^{-j(ux+vy+wz)} dx dy dz$$

$$f(x, y, z) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v, w) e^{j(ux+vy+wz)} du dv dw$$

Se:

$$f(x, y, z) = \delta(x)\delta(y) \rightarrow F(u, v, w) = \frac{1}{(2\pi)^3} \delta(w)$$

e se

$$f(x, y, z) = \delta(z) \rightarrow F(u, v, w) = \frac{1}{(2\pi)^3} \delta(u)\delta(v)$$

Quindi, convolvendo con *una bacchetta* si moltiplica lo spettro per una lamina; ancora, convolvendo con una lamina, si moltiplica lo spettro per una bacchetta. Tutte le osservazioni fatte per la tomografia 2D si estendono facilmente a 3D e si possono derivare delle relazioni che fissano il numero totale di proiezioni in funzione della risoluzione spaziale desiderata. Per esempio, nel caso di proiezioni lungo delle lamine (convolvendo la funzione $f(x, y, z)$ con delle lamine e cioè mediando i valori della funzione su interi piani) si ha che un volume di dati si proietta in una funzione 1D; inoltre, se D è il diametro dell'oggetto da tomografare (supposto sferico) e $\Delta\varphi$ l'angolo tra le normali a due lamine adiacenti,

$$|\Omega|_{\max} \Delta\varphi \leq \frac{2\pi}{D}; \quad |\Omega|_{\max} = \frac{\pi}{L} \rightarrow \Delta\varphi \leq \frac{2L}{D}$$

Infatti, per non avere equivocazione, il campionamento dello spettro, che avviene lungo una stella di rette, deve effettuarsi a intervalli di numeri d'onda inferiori a $2\pi/D$; il numero di proiezioni in una semisfera è pertanto almeno:

$$M_3 = \frac{2\pi |\Omega|_{\max}^2}{(\Delta\varphi)^2 |\Omega|_{\max}^2} \geq \frac{\pi D^2}{2 L^2} \quad (22.25)$$

indicando con L la risoluzione spaziale (la larghezza del lobo). Poichè ad ogni proiezione corrispondono D/L equazioni, si hanno in totale $\pi D^3/2L^3$ equazioni contro $\pi D^3/6L^3$ incognite con una ridondanza maggiore che non in 2D.

Un altro modo per intuire che la tomografia 3D può essere fatta proiettando il dato di volume lungo piani (convolvendo il volume con delle

CAPITOLO 22 LA TOMOGRAFIA

lamine) nasce dall'osservazione che nello spazio ci sono ∞^3 piani ed altrettanti punti. Quindi, il numero di proiezioni sui piani ha lo stesso ordine di infinità del numero dei punti dello spazio e conoscere le proiezioni ci porta a conoscere i valori puntuali, a parte gli effetti di indeterminazione considerati prima.

La situazione cambia facendo proiezioni su rette. Di queste, ce ne è ∞^4 e quindi è inutile misurare le proiezioni su "tutte" le rette: avremmo misure ridondanti rispetto ai dati. Facendo la convoluzione con tutte le rette ortogonali ad un dato piano otteniamo una proiezione bidimensionale (l'oggetto "schiacciato" lungo la direzione della normale al piano). La trasformata di Fourier 2D di questi dati "schiacciati" ci dà una sezione 2D dello spettro 3D del dato (un cerchio massimo della sfera). Di queste sezioni 2D ce ne sono ∞^2 , ma ce ne bastano ∞ . Per esempio, basta considerare solo le rette ortogonali ad un asse verticale z (che sono appunto ∞^3). Facendo proiezioni secondo queste linee, si può procedere a 3D operando in coordinate cilindriche e quindi ricadere nelle tecniche tomografiche 2D.

Peraltro, la proiezione lungo delle rette è più ricca di informazione e quindi ci permette di recuperare degli ulteriori parametri locali, come per esempio la variabilità angolare della riflettività locale alle onde o in generale parametri di variabilità angolare (azimut o elevazione) delle proprietà locali dell'oggetto sotto misura. Si pensi al caso di specchietti piani sospesi in un mezzo trasparente (come in un acquario). Illuminando tale volume con dei raggi luminosi provenienti da tutte le direzioni, riusciamo a trovare la distribuzione spaziale sia degli specchietti che dei loro puntamenti.

22.7 Tecniche iterative

22.7.1 Retroproiezione e gradiente

Si consideri un problema di tipo tomografico: sono assegnate delle misure, raccolte in un vettore $\underline{\mathbf{p}}$; si cerca un vettore di parametri $\underline{\mathbf{m}}$ tale che le misure fatte sul modello definito da questi parametri corrispondano quanto più possibile ai dati. In formula si ha:

$$\underline{\mathbf{A}}\underline{\mathbf{m}} = \underline{\mathbf{d}} \quad (22.26)$$

La matrice $\underline{\mathbf{A}}$ è in generale di dimensioni $[M \times N]$; i vettori $\underline{\mathbf{m}}$ e $\underline{\mathbf{d}}$ sono di dimensione $[N]$ ed $[M]$ rispettivamente. Ad esempio, il modello $\underline{\mathbf{m}}$ potrebbe corrispondere alla distribuzione delle lentezze (inverse delle velocità) di propagazione di un mezzo qualsiasi, e il vettore dei dati $\underline{\mathbf{d}}$ corrispondere ai tempi di propagazione di raggi che hanno traversato questo mezzo, nelle varie direzioni. Il tempo di propagazione è proporzionale all'integrale di linea delle lentezze lungo la direzione del raggio e quindi si ritrova il classico problema di tomografia. La limitata accessibilità del mezzo fa sì poi che non sono disponibili raggi in tutte le direzioni e pertanto la matrice $\underline{\mathbf{A}}$ può essere praticamente qualsiasi. La situazione è molto più semplice quando sono disponibili raggi in tutte le direzioni ed allora la struttura della matrice assume forme particolari. La tecnica consueta di risoluzione di questo sistema è quella della pseudoinversione; si cerca una soluzione ai minimi quadrati e cioè tale che l'energia del residuo ε^2 sia minima:

$$(\underline{\mathbf{A}}\underline{\mathbf{m}} - \underline{\mathbf{d}})^*(\underline{\mathbf{A}}\underline{\mathbf{m}} - \underline{\mathbf{d}}) = \varepsilon^2 \quad (22.27)$$

Come si vedrà più diffusamente nel prossimo capitolo, il minimo di questa forma quadratica si raggiunge per

$$\underline{\mathbf{m}} = (\underline{\mathbf{A}}^* \underline{\mathbf{A}})^{-1} \underline{\mathbf{A}}^* \underline{\mathbf{d}} \quad (22.28)$$

e cioè applicando al vettore $\underline{\mathbf{d}}$ la matrice $(\underline{\mathbf{A}}^* \underline{\mathbf{A}})^{-1} \underline{\mathbf{A}}^*$ detta appunto pseudo inversa della matrice $\underline{\mathbf{A}}$. Infatti, se la matrice $\underline{\mathbf{A}}$ è quadrata e non singolare si ritorna al caso classico. Se poi la matrice $\underline{\mathbf{A}}^* \underline{\mathbf{A}}$ è troppo vicina ad essere singolare, si può sempre aggiungere della luce bianca e cioè accrescere la diagonale principale aggiungendo $\lambda \underline{\mathbf{I}}$ dove $\underline{\mathbf{I}}$ è la matrice identità e λ un parametro "piccolo". Si verificherà che questo coincide con il minimizzare $\varepsilon^2 + \lambda(\underline{\mathbf{m}}^* \underline{\mathbf{m}})^2$. Cerchiamo di capire tutto ciò in termini di tomografia.

22.7.2 Retroproiezione e trasposizione

Anzitutto, determiniamo il significato fisico dell'operazione

$$\underline{\mathbf{m}}_0 = \underline{\mathbf{A}}^* \underline{\mathbf{d}} \quad (22.29)$$

e cioè di applicare al vettore dei dati la trasposta della matrice del modello. Questa operazione è la retroproiezione. Per vederlo, si consideri uno

CAPITOLO 22 LA TOMOGRAFIA

spazio suddiviso in cellette rettangolari. Ogni componente del vettore dei dati (le proiezioni $\underline{\mathbf{d}}$), $\{d_i\}$, è ottenuta sommando i contenuti $\{m_{j_i}\}$ delle cellette corrispondenti ad un'operazione di proiezione. La matrice $\underline{\mathbf{A}}$ che corrisponde all'operazione di proiezione, ha degli 1 negli elementi a_{i,j_i} che corrispondono alle celle j_i che si proiettano nell'elemento i -esimo del vettore delle proiezioni. La matrice $\underline{\mathbf{A}}^*$, che corrisponde all'operazione di retroproiezione, ha gli stessi 1 nelle celle $a_{j_i,i}$ relative agli elementi i_j del vettore delle proiezioni che sono alterati quando viene alterato il contenuto della cella m_{j_i} . Gli elementi della matrice corrispondenti a celle che non si proiettano in quell'elemento della proiezione conterranno degli zeri; se la proiezione è solo parziale si avranno numeri compresi tra 1 e 0. Il valore corretto del modello (il vettore $\underline{\mathbf{m}}$) si ottiene dall'equazione:

$$(\underline{\mathbf{A}}^* \underline{\mathbf{A}}) \underline{\mathbf{m}} = \underline{\mathbf{m}}_0 \quad (22.30)$$

Leggendo questa equazione si vede che, retroproiettando le proiezioni ottenute con il modello $\underline{\mathbf{m}}$, si deve ottenere il dato retroproiettato e cioè $\underline{\mathbf{m}}_0$. Si abbia un modello $\underline{\mathbf{m}}$ con una sola cella contenente un valore non nullo e cioè si voglia determinare la risposta all'impulso del sistema; questa risposta sarà spazio - variante tuttavia, e i risultati pertanto potranno cambiare al variare della posizione dell'impulso. Il risultato dell'applicazione dell'equazione (22.28) corrisponderà ad una approssimazione del modello $\underline{\mathbf{m}}_0$ in cui avremo sí un picco in corrispondenza dell'impulso di ingresso (la cella a contenuto non nullo) ma non solo in quel punto; molti altri elementi del modello saranno non nulli. Nel caso della tomografia queste due operazioni in cascata corrispondono alla proiezione seguita dalla retroproiezione non filtrata. La risposta all'impulso di questo sistema è una struttura a stella, centrata sull'impulso di ingresso e con tanti raggi quante sono le proiezioni.

Quindi, nel caso della tomografia, l'applicazione della matrice $(\underline{\mathbf{A}}^* \underline{\mathbf{A}})^{-1}$ corrisponde alla focalizzazione di questa stella di raggi. Questa operazione si realizza filtrando le proiezioni con il filtro $P(\Omega)$ avente una caratteristica di ampiezza proporzionale alla frequenza radiale Ω . Si osservi che la semplicità di quest'ultimo risultato è dovuta alla regolarità del rilevamento.

Il caso generale, quando cioè le geometrie non sono controllate e quindi la matrice $\underline{\mathbf{A}}$ è completamente libera, è ben piú complesso. Si osservi poi che nel caso piú generale, la matrice $\underline{\mathbf{A}}$ non è fatta solo di uni e

zeri, ma la retroproiezione su una cella dovrà tener conto della lunghezza (frazionaria) dell'intersezione tra la cella e il raggio di retroproiezione pesando la retroproiezione in funzione di questo dato. Inoltre, potrebbe darsi che la tomografia che stiamo analizzando è una tomografia di tempi di percorso, in cui cioè il modello ricercato è la distribuzione della lentezza di propagazione (inversa della velocità) dei raggi (acustici od elettromagnetici) all'interno del corpo e le misure sono i tempi di percorso, per l'appunto. In questo caso, i percorsi dei raggi dipendono, per la legge di Snell, dalla distribuzione delle velocità, e pertanto il sistema diventa non lineare, con complicazioni ancora maggiori.

22.7.3 Interpretazione stocastica della pseudo-inversione

Il problema della determinazione di $\underline{\mathbf{m}}$ a partire da $\underline{\mathbf{d}}$ è anche risolvibile con le ben note tecniche di stima. Infatti, si supponga che siano note le M misure degli N parametri, questa volta sommate a rumore:

$$\underline{\mathbf{d}} = \underline{\mathbf{A}}\underline{\mathbf{m}} + \underline{\mathbf{n}}$$

La stima di $\underline{\mathbf{m}}$ a partire da $\underline{\mathbf{d}}$ sarà ottenuta in base al principio di ortogonalità:

$$\begin{aligned} \hat{\underline{\mathbf{m}}} &= \underline{\mathbf{B}}\underline{\mathbf{d}} \\ E[(\underline{\mathbf{B}}\underline{\mathbf{d}} - \underline{\mathbf{m}})\underline{\mathbf{d}}^*] &= \underline{\mathbf{0}} \end{aligned}$$

Ma:

$$E[\underline{\mathbf{d}}\underline{\mathbf{d}}^*] = \underline{\mathbf{A}}E[\underline{\mathbf{m}}\underline{\mathbf{m}}^*]\underline{\mathbf{A}}^* + E[\underline{\mathbf{n}}\underline{\mathbf{n}}^*]$$

e supponendo per semplicità

$$E[\underline{\mathbf{m}}\underline{\mathbf{m}}^*] = \sigma_m^2 \underline{\mathbf{I}}_N; \quad E[\underline{\mathbf{n}}\underline{\mathbf{n}}^*] = \sigma_n^2 \underline{\mathbf{I}}_M$$

dove con $\underline{\mathbf{I}}_N$, $\underline{\mathbf{I}}_M$ è stata indicata la matrice identica negli spazi a N ed M dimensioni rispettivamente,

$$E[\underline{\mathbf{d}}\underline{\mathbf{d}}^*] = \sigma_m^2 \underline{\mathbf{A}}\underline{\mathbf{A}}^* + \sigma_n^2 \underline{\mathbf{I}}_M$$

$$E[\underline{\mathbf{m}}\underline{\mathbf{d}}^*] = \sigma_m^2 \underline{\mathbf{A}}^*$$

da cui

$$\begin{aligned}\underline{\mathbf{B}} (\sigma_m^2 \underline{\mathbf{A}} \underline{\mathbf{A}}^* + \sigma_n^2 \underline{\mathbf{I}}_M) &= \sigma_m^2 \underline{\mathbf{A}}^* \\ \underline{\mathbf{B}} &= \underline{\mathbf{A}}^* \left(\underline{\mathbf{A}} \underline{\mathbf{A}}^* + \frac{\sigma_n^2}{\sigma_m^2} \underline{\mathbf{I}}_M \right)^{-1}\end{aligned}\quad (22.31)$$

Ora, usando il cosiddetto lemma di inversione delle matrici, verificabile per sostituzione, per cui:

$$(\underline{\mathbf{B}} \underline{\mathbf{C}} \underline{\mathbf{D}} + \underline{\mathbf{A}})^{-1} = \underline{\mathbf{A}}^{-1} - \underline{\mathbf{A}}^{-1} \underline{\mathbf{B}} (\underline{\mathbf{C}}^{-1} + \underline{\mathbf{D}} \underline{\mathbf{A}}^{-1} \underline{\mathbf{B}})^{-1} \underline{\mathbf{D}} \underline{\mathbf{A}}^{-1}$$

si ha:

$$\underline{\mathbf{B}} = \left(\underline{\mathbf{A}}^* \underline{\mathbf{A}} + \frac{\sigma_n^2}{\sigma_m^2} \underline{\mathbf{I}}_N \right)^{-1} \underline{\mathbf{A}}^* \quad (22.32)$$

e, per un buon rapporto segnale rumore, si ottiene la (22.28). Si osservi che la (22.32) prevede l'inversione di una matrice $[N \times N]$ mentre invece la (22.31) richiede l'inversione di una matrice $[M \times M]$, di dimensioni possibilmente ben maggiori.

22.7.4 Interpretazione tomografica del gradiente

Possiamo ora tentare di dare un'interpretazione tomografica delle tecniche di ottimizzazione con il metodo del gradiente e del gradiente coniugato.

Cerchiamo di risolvere l'equazione (22.30) con il metodo del gradiente; a partire da un modello iniziale, ad esempio $\underline{\mathbf{m}}_0$, calcoliamo il gradiente della forma quadratica in (22.27) e procediamo lungo questa direzione fino ad arrivare ad un minimo.

$$\underline{\mathbf{m}}_1 = \underline{\mathbf{m}}_0 + \alpha_0 \underline{\mathbf{g}}_0; \quad \underline{\mathbf{g}}_0 = \underline{\mathbf{A}}^* (\underline{\mathbf{A}} \underline{\mathbf{m}}_0 - \underline{\mathbf{d}}) \quad (22.33)$$

In sostanza il nuovo modello $\underline{\mathbf{m}}_1$ è ottenuto dal precedente aggiungendo un vettore proporzionale al gradiente; questo non è altro che la **retro-proiezione del residuo** e cioè la retroproiezione della differenza tra la proiezione del modello iniziale ed i dati; il valore di α_0 è determinato imponendo che il nuovo gradiente (quindi il nuovo residuo) $\underline{\mathbf{A}}^* (\underline{\mathbf{A}} \underline{\mathbf{m}}_1 - \underline{\mathbf{d}})$ sia ortogonale al gradiente precedente (e cioè il residuo precedente). Poi si retroproietta il nuovo residuo e si calcola il nuovo gradiente, si determina il passo con la condizione di ortogonalità e così via. In altre

22.7 TECNICHE ITERATIVE

parole abbiamo mostrato l'equivalenza tra la retroproiezione del residuo e il metodo del gradiente.

Il metodo del gradiente coniugato che si studierà nel prossimo capitolo invece prevederà la retroproiezione di qualcosa di diverso dal residuo, e cioè la parte del residuo corrispondente ad un modello ortogonale coniugato con il modello utilizzato sino a questo punto. Così si decompone lo spazio dei modelli secondo vettori ortogonali coniugati sotto la norma indotta dalla matrice $\underline{\mathbf{A}}^* \underline{\mathbf{A}}$ e cioè tali che i modelli corrispondenti siano ortogonali **dopo il filtraggio con il ρ -filtro**. In questo modo, l'effetto di questo filtraggio può venire tenuto in conto e la soluzione viene determinata progressivamente.

Capitolo 23

Tecniche di ottimizzazione

23.1 Il gradiente

Supponiamo di volere risolvere il sistema di equazioni lineari:

$$\mathbf{R} \mathbf{a} = \mathbf{b} \quad (23.1)$$

dove \mathbf{R} e \mathbf{a} , \mathbf{b} sono rispettivamente una matrice $[N \times N]$ definita positiva e due vettori $[N \times M]$ corrispondenti alla matrice del sistema ed ai termini incogniti e noti rispettivamente. Nel seguito, come di consueto, supporremo $M = 1$. Invece di risolvere il sistema, possiamo cercare il minimo della forma quadratica:

$$\frac{1}{2} \mathbf{a}^* \mathbf{R} \mathbf{a} - \mathbf{b}^* \mathbf{a} = Q(\mathbf{a}) \quad (23.2)$$

Si può minimizzare $Q(\mathbf{a})$ con la tecnica del gradiente. Indicando con \mathbf{a}_0 il valore iniziale assegnato al vettore \mathbf{a} e \mathbf{f}_0 l'errore iniziale si ha:

$$\mathbf{a}_0 = \mathbf{a} + \mathbf{R} \mathbf{f}_0 \quad (23.3)$$

$$\mathbf{R} \mathbf{a}_0 = \mathbf{R}(\mathbf{a} + \mathbf{f}_0) = \mathbf{b} + \mathbf{f}_0 \quad (23.4)$$

L'evoluzione con le successive iterazioni della soluzione approssimata può essere rappresentata con la convergenza a 0 della sequenza dei vettori (delle matrici) degli errori $\mathbf{f}_0, \mathbf{f}_1, \mathbf{f}_N$. Il gradiente della forma quadratica nel punto di coordinate \mathbf{a}_0 è (supponiamo $M = 1$):

$$\underline{\varepsilon}_0 = \nabla_{\mathbf{a}} Q(\mathbf{a}_0) = \mathbf{R} \mathbf{a}_0 - \mathbf{b} = \mathbf{R} \mathbf{f}_0 \quad (23.5)$$

CAPITOLO 23 TECNICHE DI OTTIMIZZAZIONE

Con la tecnica della massima pendenza (steepest descent) aggiorniamo la soluzione all'iterazione k -esima ($\underline{\mathbf{a}}_k$) aggiungendo un vettore la cui direzione è quella del gradiente $\underline{\boldsymbol{\varepsilon}}_k$ e la cui ampiezza è determinata con un'ottimizzazione della variabile α_k :

$$\underline{\mathbf{a}}_{k+1} = \underline{\mathbf{a}}_k + \alpha_k \underline{\boldsymbol{\varepsilon}}_k; \quad \underline{\boldsymbol{\varepsilon}}_k = \underline{\mathbf{R}} \underline{\mathbf{f}}_k \quad (23.6)$$

Ricordando che

$$\underline{\mathbf{a}}_k = \underline{\mathbf{a}} + \underline{\mathbf{f}}_k \quad (23.7)$$

risulta:

$$\underline{\mathbf{f}}_{k+1} = \underline{\mathbf{f}}_k + \alpha_k \underline{\mathbf{R}} \underline{\mathbf{f}}_k = (\underline{\mathbf{I}} + \alpha_k \underline{\mathbf{R}}) \underline{\mathbf{f}}_k \quad (23.8)$$

Volendo mantenere α costante, risulta:

$$\underline{\mathbf{f}}_{k+1} = (\underline{\mathbf{I}} + \alpha \underline{\mathbf{R}})^{k+1} \underline{\mathbf{f}}_0 \quad (23.9)$$

Per avere stabilità nelle iterazioni, come già visto per il filtraggio adattativo, si ha:

$$\alpha < \frac{2}{Tr \underline{\mathbf{R}}} \quad (23.10)$$

In questo modo, gli autovalori della matrice in parentesi nell'equazione (23.9) sono minori di 1 e quindi l'errore converge a 0, seppure con infiniti passi. Di nuovo, si incontra il problema di una lenta convergenza nelle componenti di $\underline{\mathbf{f}}_0$ dirette verso gli autovettori di $\underline{\mathbf{R}}$ con autovalori piccoli; il problema è grave se gli autovalori di $\underline{\mathbf{R}}$ sono di ampiezze assai differenti tra loro, anche se sono tutti positivi visto che $\underline{\mathbf{R}}$ è positiva definita per ipotesi. Il valore ottimo di α_k è determinato minimizzando ad ogni passo la forma quadratica $Q(\underline{\mathbf{a}})$:

$$\frac{\partial}{\partial \alpha_k} (\underline{\mathbf{f}}_k + \alpha_k \underline{\mathbf{R}} \underline{\mathbf{f}}_k)^* \underline{\mathbf{R}} (\underline{\mathbf{f}}_k + \alpha_k \underline{\mathbf{R}} \underline{\mathbf{f}}_k) = 0 \quad (23.11)$$

e pertanto:

$$\alpha_k = -\frac{\underline{\mathbf{f}}_k^* \underline{\mathbf{R}}^2 \underline{\mathbf{f}}_k}{\underline{\mathbf{f}}_k^* \underline{\mathbf{R}}^3 \underline{\mathbf{f}}_k} = -\frac{\underline{\boldsymbol{\varepsilon}}_k^* \underline{\boldsymbol{\varepsilon}}_k}{\underline{\boldsymbol{\varepsilon}}_k^* \underline{\mathbf{R}} \underline{\boldsymbol{\varepsilon}}_k} \quad (23.12)$$

Si osservi che i vettori $\underline{\mathbf{R}} \underline{\mathbf{f}}_k = \underline{\boldsymbol{\varepsilon}}_k$ e $\underline{\mathbf{R}} \underline{\mathbf{f}}_{k+1} = \underline{\boldsymbol{\varepsilon}}_{k+1}$ sono ortogonali.

$$\underline{\boldsymbol{\varepsilon}}_k = \underline{\mathbf{R}} \underline{\mathbf{f}}_k \perp \underline{\mathbf{R}} \underline{\mathbf{f}}_{k+1} = \underline{\boldsymbol{\varepsilon}}_{k+1} \quad (23.13)$$

23.2 IL GRADIENTE CONIUGATO

In altre parole, con la tecnica della discesa lungo la linea di massima pendenza e con passi di lunghezza ottimizzata, si prosegue lungo la direzione di partenza fino a che non si sia su un minimo e perciò tangenti ad una curva di livello. Pertanto, le direzioni successive sono ortogonali tra loro. Questo limita di molto l'efficienza nel caso di dispersione degli autovalori, come verrà meglio chiarito nel paragrafo successivo sul gradiente coniugato.

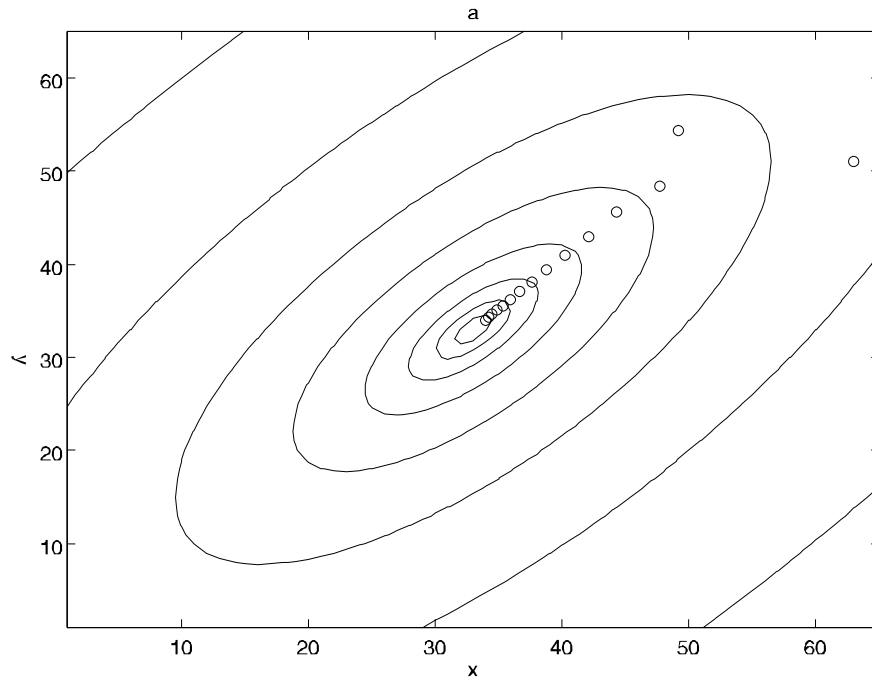


Figura 23.1: Ottimizzazione con la tecnica del gradiente

23.2 Il gradiente coniugato

23.2.1 Ortogonalità coniugata e incorrelazione

Prima di discutere la tecnica del gradiente coniugato, è opportuno introdurre una interpretazione di tipo stocastico per una maggiore comprensione. Si consideri \underline{x} , un vettore casuale di N elementi, caratterizzato da

CAPITOLO 23 TECNICHE DI OTTIMIZZAZIONE

una matrice di covarianza $\underline{\mathbf{R}}$. Si abbia cioè

$$E[\underline{\mathbf{x}}\underline{\mathbf{x}}^*] = \underline{\mathbf{R}}$$

La matrice $\underline{\mathbf{R}}$ è di dimensioni $[N \times N]$. Si considerino ora delle combinazioni lineari delle N componenti $\{x_n\}$ del vettore $\underline{\mathbf{x}}$ e cioè

$$y_i = \underline{\mathbf{a}}_i^* \underline{\mathbf{x}} \quad (23.14)$$

ottenute facendo un prodotto scalare tra $\underline{\mathbf{x}}$ ed un qualsiasi vettore $\underline{\mathbf{a}}_i$ di N componenti. Risulta che:

$$E[y_i y_j] = E[\underline{\mathbf{a}}_i^* \underline{\mathbf{x}} \underline{\mathbf{x}}^* \underline{\mathbf{a}}_j] = \underline{\mathbf{a}}_i^* \underline{\mathbf{R}} \underline{\mathbf{a}}_j \quad (23.15)$$

Le variabili casuali y_i, y_j sono incorrelate se $E[y_i y_j] = 0$ e cioè se

$$\underline{\mathbf{a}}_i^* \underline{\mathbf{R}} \underline{\mathbf{a}}_j = 0 \quad (23.16)$$

I due vettori $\underline{\mathbf{a}}_i, \underline{\mathbf{a}}_j$ sono allora detti essere **ortogonali coniugati** tra di loro sotto la norma della matrice $\underline{\mathbf{R}}$. Indicando con $\{a_{i,n}\}$ le componenti del vettore $\underline{\mathbf{a}}_i$, lungo gli autovettori $\underline{\mathbf{e}}_n$ della matrice $\underline{\mathbf{R}}$, risulta:

$$\sum_{n=1}^N \lambda_n a_{i,n} a_{j,n} = 0 \quad (23.17)$$

I due vettori sarebbero anche ortogonali tra loro se

$$\sum_{n=1}^N a_{i,n} a_{j,n} = 0; \quad \lambda_n = \text{cost} \quad (23.18)$$

e quindi se la matrice $\underline{\mathbf{R}}$ fosse una matrice identica e il processo fosse bianco. Si ricordi che $\underline{\mathbf{R}}$ è una matrice semi definita positiva visto che $E[y_i^2] \geq 0$ per qualsiasi vettore $\underline{\mathbf{a}}_i$ non nullo. L'onnipresenza del rumore ci consente poi di trascurare il caso in cui la matrice $\underline{\mathbf{R}}$ abbia autovalori nulli. Abbiamo perciò visto come all'incorrelazione delle due variabili casuali y_i, y_j corrisponda l'ortogonalità coniugata dei vettori deterministici $\underline{\mathbf{a}}_i, \underline{\mathbf{a}}_j$ sotto la norma della matrice $\underline{\mathbf{R}}$. Vediamo ora come anche le tecniche di stima lineare hanno una corrispondenza nel settore della minimizzazione di forme quadratiche. Utilizzando il vettore casuale $\underline{\mathbf{x}}$, si

23.2 IL GRADIENTE CONIUGATO

cerchi di stimare in modo lineare ottimo una variabile casuale ζ ; ciò si ottiene mediante:

$$\hat{\zeta} = \underline{\mathbf{a}}_{\zeta}^* \underline{\mathbf{x}} \quad (23.19)$$

una combinazione lineare delle componenti di $\underline{\mathbf{x}}$ ottenuta facendone il prodotto scalare con un vettore $\underline{\mathbf{a}}_z$ da determinare in modo ottimale. Senza ledere la generalità si può stimare anzichè ζ la variabile casuale *residuale*:

$$z = \zeta - \underline{\mathbf{a}}_0^* \underline{\mathbf{x}} \quad (23.20)$$

dove $\underline{\mathbf{a}}_0$ è un vettore qualsiasi, tipicamente un'ipotesi iniziale di stimatore, da raffinare poi con il calcolo successivo. Dobbiamo pertanto determinare in modo ottimale $\underline{\mathbf{a}}_z$ per calcolare $\underline{\mathbf{a}}_z^* \underline{\mathbf{x}}$ stimatore ottimo della variabile casuale z .

$$\hat{z} = \underline{\mathbf{a}}_z^* \underline{\mathbf{x}} \quad (23.21)$$

Il vettore $\underline{\mathbf{a}}_z$ si trova con la minimizzazione dell'errore quadratico medio, oppure con il principio di ortogonalità:

$$E[\varepsilon^2] = E[z - \underline{\mathbf{a}}_z^* \underline{\mathbf{x}}]^2 = E[z^2] + \underline{\mathbf{a}}_z^* \underline{\mathbf{R}} \underline{\mathbf{a}}_z - 2\underline{\mathbf{a}}_z^* E[z\underline{\mathbf{x}}] \quad (23.22)$$

Il minimo si raggiunge per

$$\underline{\mathbf{R}} \underline{\mathbf{a}}_z = \underline{\mathbf{b}} = E[z\underline{\mathbf{x}}] \quad (23.23)$$

Le componenti di $\underline{\mathbf{b}}$ e cioè $\{b_i\}$ sono date da

$$b_i = E[zx_i] \quad (23.24)$$

Lo scalare $E[\varepsilon^2]$, i vettori $\underline{\mathbf{a}}_z$, $\underline{\mathbf{b}}$ e la matrice $\underline{\mathbf{R}}$ hanno il ruolo che avevano rispettivamente nel paragrafo precedente $Q(\underline{\mathbf{a}})$, $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$ e la matrice $\underline{\mathbf{R}}$. Inoltre il vettore gradiente del paragrafo precedente corrisponde alla correlazione tra errore di stima e vettore dati:

$$\underline{\varepsilon} = \underline{\mathbf{R}} \underline{\mathbf{a}} - \underline{\mathbf{b}} = E[\varepsilon\underline{\mathbf{x}}] \quad (23.25)$$

Quando il gradiente è nullo, i dati sono incorrelati con l'errore di stima e il problema è risolto.

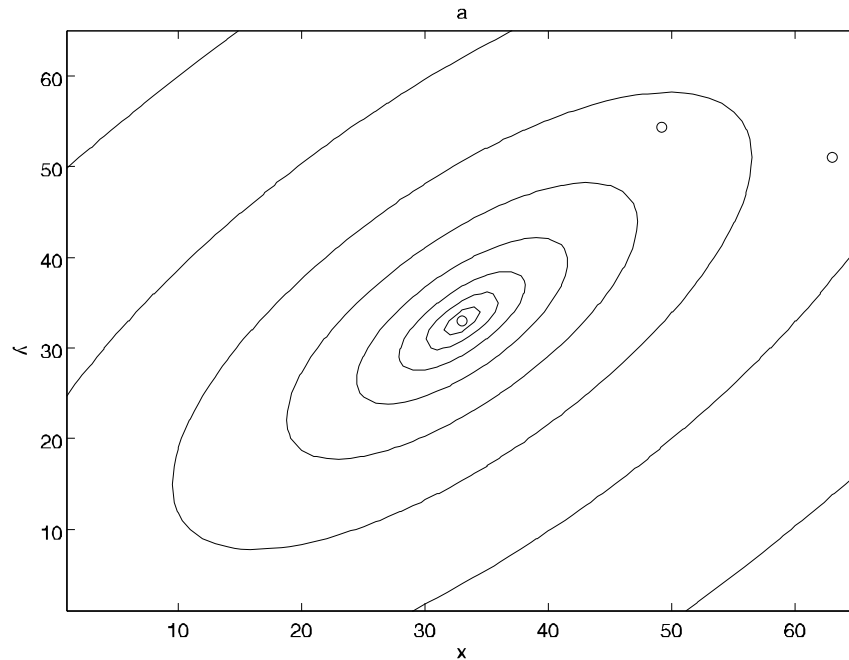


Figura 23.2: Ottimizzazione con la tecnica del gradiente coniugato

23.2.2 Connessione con la ricorsione di Levinson

Si voglia ora stimare la variabile casuale z con una tecnica corrispondente a quella ricorsiva di Levinson. In altre parole, si inizi a stimare z usando solo la prima componente di \underline{x} poi le prime due e così via, fino ad utilizzare tutte le N componenti. Il punto chiave è che, passando dallo stimatore di lunghezza m a quello di lunghezza $m + 1$, utilizziamo, per migliorare la stima, l'errore di predizione "all'indietro" e cioè la differenza tra il valore effettivo della $m + 1$ -esima componente di \underline{x} e la sua stima in base alle prime m componenti di \underline{x} che già sono state utilizzate per stimare z . In questo modo arricchiamo la stima fatta in precedenza con l'uso di una nuova variabile casuale che da un lato è **incorrelata con tutti** i dati utilizzati finora (e quindi porta nuova informazione) ma d'altra parte può essere correlata con la variabile casuale da stimare. Quindi, l'errore di stima non può che decrescere e la minimizzazione dell'errore quadratico medio può essere fatta progressivamente su di una variabile

23.2 IL GRADIENTE CONIUGATO

alla volta senza distruggere l'ottimizzazione fatta con le variabili precedenti. Inoltre, la determinazione dello stimatore ottimo termina in N passi, perchè terminerà quando abbiamo utilizzato tutte le N variabili casuali, tra loro incorrelate, con cui possiamo rappresentare il vettore casuale $\underline{\mathbf{x}}$.

Nel caso generale del gradiente coniugato il ragionamento va come segue: la ricerca dello stimatore ottimo $\underline{\mathbf{a}}_z$ parte dalla determinazione di un primo stimatore di prova $\underline{\mathbf{a}}_0$. La prima "direzione" $\underline{\mathbf{D}}_0$ può essere trovata con il criterio della massima pendenza e quindi calcolando il gradiente dell'errore quadratico medio:

$$\underline{\mathbf{D}}_0 = \nabla_{\underline{\mathbf{a}}} \left(\frac{\underline{\mathbf{a}}^* \underline{\mathbf{R}} \underline{\mathbf{a}}}{2} - \underline{\mathbf{b}}^* \underline{\mathbf{a}} \right) \Big|_{\underline{\mathbf{a}}=\underline{\mathbf{a}}_0} = \underline{\mathbf{R}} \underline{\mathbf{a}}_0 - \underline{\mathbf{b}} \quad (23.26)$$

Poi, si cerca la lunghezza del passo minimizzando l'errore quadratico medio e quindi imponendo che la variabile casuale che corrisponde all'errore residuo sia incorrelata rispetto al dato e cioè la variabile casuale $\underline{\mathbf{D}}_0^* \underline{\mathbf{x}}$ pertanto si calcola il numero α_0 tale che:

$$\alpha_0 = - \frac{(\underline{\mathbf{R}} \underline{\mathbf{a}}_0 - \underline{\mathbf{b}})^* \underline{\mathbf{D}}_0}{\underline{\mathbf{D}}_0^* \underline{\mathbf{R}} \underline{\mathbf{D}}_0} \quad (23.27)$$

Il nuovo stimatore $\underline{\mathbf{a}}_1$ è:

$$\underline{\mathbf{a}}_1 = \underline{\mathbf{a}}_0 + \alpha_0 \underline{\mathbf{D}}_0 \quad (23.28)$$

Ora un altro passo; vogliamo minimizzare ancora l'errore quadratico medio, muovendoci nella direzione $\underline{\mathbf{D}}_1$; come nella ricorsione di Levinson, vogliamo che la variabile casuale $\underline{\mathbf{D}}_1^* \underline{\mathbf{x}}$ sia incorrelata con la variabile casuale che abbiamo usato prima, per utilizzare **solamente** dell' "informazione nuova". Per accelerare la minimizzazione dell'errore, la nuova direzione $\underline{\mathbf{D}}_1$ sarà vicina al gradiente in $\underline{\mathbf{a}}_1$, e cioè $\underline{\mathbf{g}}_1 = \underline{\mathbf{R}} \underline{\mathbf{a}}_1 - \underline{\mathbf{b}}$, per quanto possibile; deve anche però essere incorrelata (ortogonale coniugata) rispetto alla prima direzione presa $\underline{\mathbf{D}}_0$. Quindi:

$$\underline{\mathbf{D}}_1 = \underline{\mathbf{R}} \underline{\mathbf{a}}_1 - \underline{\mathbf{b}} + \beta_1 \underline{\mathbf{D}}_0 = \underline{\mathbf{g}}_1 + \beta_1 \underline{\mathbf{D}}_0 \quad (23.29)$$

dove β_1 si trova imponendo l'ortogonalità coniugata tra $\underline{\mathbf{D}}_1$ e $\underline{\mathbf{D}}_0$.

$$\beta_1 = - \frac{\underline{\mathbf{g}}_1^* \underline{\mathbf{R}} \underline{\mathbf{D}}_0}{\underline{\mathbf{D}}_0^* \underline{\mathbf{R}} \underline{\mathbf{D}}_0} \quad (23.30)$$

L'algoritmo procede aggiornando lo stimatore $\underline{\mathbf{a}}_1$ spingendosi nella nuova direzione $\underline{\mathbf{D}}_1$ con l'opportuna lunghezza del passo; ciò avviene ortogonalizzando l'errore di stima residuo rispetto al nuovo dato e cioè la variabile casuale $\underline{\mathbf{D}}_1^* \underline{\mathbf{x}}$ e quindi imponendo l'ortogonalità tra il gradiente $\underline{\mathbf{g}}_2$ nella nuova posizione $\underline{\mathbf{a}}_2$ e $\underline{\mathbf{D}}_1$:

$$\underline{\mathbf{a}}_2 = \underline{\mathbf{a}}_1 + \alpha_1 \underline{\mathbf{D}}_1 \quad (23.31)$$

$$\alpha_1 = - \frac{\underline{\mathbf{g}}_1^* \underline{\mathbf{D}}_1}{\underline{\mathbf{D}}_1^* \underline{\mathbf{R}} \underline{\mathbf{D}}_1} \quad (23.32)$$

A questo punto dobbiamo individuare un'altra direzione $\underline{\mathbf{D}}_2$ e cioè un'altra variabile casuale $\underline{\mathbf{D}}_2 \underline{\mathbf{x}}$ con cui fare un'ulteriore stima del residuo. Per quanto detto scegliamo questa direzione con il solito criterio di stare vicini al gradiente in $\underline{\mathbf{a}}_2$ pur mantenendo l'ortogonalità coniugata con la direzione $\underline{\mathbf{D}}_1$ risulta allora:

$$\underline{\mathbf{D}}_2 = \underline{\mathbf{R}} \underline{\mathbf{a}}_2 - \underline{\mathbf{b}} + \beta_2 \underline{\mathbf{D}}_1 \quad (23.33)$$

La variabile β_2 viene scelta con il criterio di imporre che la nuova variabile casuale $\underline{\mathbf{D}}_2^* \underline{\mathbf{x}}$ sia incorrelata con quella usata precedentemente e cioè $\underline{\mathbf{D}}_1^* \underline{\mathbf{x}}$. Quindi:

$$\beta_2 = - \frac{\underline{\mathbf{g}}_2^* \underline{\mathbf{R}} \underline{\mathbf{D}}_1}{\underline{\mathbf{D}}_1^* \underline{\mathbf{R}} \underline{\mathbf{D}}_1} \quad (23.34)$$

Se si desidera ancora decrescere il valore quadratico medio dell'errore di stima, ci dobbiamo comunque muovere nella direzione del gradiente dell'errore rispetto al vettore $\underline{\mathbf{a}}$. Pur così facendo, ci si muove verso una direzione ortogonale coniugata rispetto a tutte le direzioni precedenti (aggiungere una variabile casuale incorrelata con le variabili casuali finora utilizzate) altrimenti le minimizzazioni precedenti avrebbero potuto essere migliorate portando ad errori inferiori. Il nuovo vettore perciò non può puntare verso una delle direzioni già utilizzate.

Il residuo è **incorrelato** con le componenti per cui è stata fatta la minimizzazione e pertanto punta in direzione ortogonale coniugata ad esse; lo spazio dei parametri si riduce di una dimensione ad ogni passo fino a vuotarsi in N passi. L'algoritmo quindi prosegue e termina in N passi, con la determinazione dello stimatore ottimo $\underline{\mathbf{a}}_z$. Si osservi che già nei primi passi si era trovata una buona approssimazione dello stimatore; ciò è tanto più vero quanto maggiore è la dispersione degli autovalori della matrice $\underline{\mathbf{R}}$.

23.3 Gradiente e gradiente coniugato

Sulla base delle osservazioni statistiche fatte nella sezione precedente si comprende come mai il metodo del gradiente coniugato sia assai piú efficiente di quello della massima pendenza. Anzitutto, si ha convergenza all'ottimo in N passi e non in infiniti come nella massima pendenza. Infatti, l'ottimizzazione è progressivamente fatta rispetto a variabili tra loro incorrelate, fino a che tutte le componenti del vettore \underline{x} siano perfettamente descritte. Ancora, la convergenza è la piú veloce possibile, e cioè in pratica dopo pochi passi le componenti di \underline{a}_z relative agli autovettori della matrice \mathbf{R} con autovalori piú elevati sono determinate con buona cura. Successivamente sono determinate tutte le altre componenti, fino alla soluzione completa cui si giunge in N passi, come detto.

Il paragone classico che viene fatto per spiegare la tecnica del gradiente coniugato è quello relativo ad uno sciatore in un canalone. Con la tecnica del gradiente, lo sciatore parte lontano dal fondo del canalone e si avvia lungo la direzione della linea di massimo pendio; poi, prosegue sempre dritto (traversando il canalone) fino a quasi risalire il pendio; si ferma cioè quando gli sci sono tangenti ad una curva di livello (Il punto di partenza è in alto a destra nella fig 23.1; attenzione, le curve di livello **non** sono equispaziate). Poi riparte per il tratto successivo, riorientando gli sci nella direzione di massimo pendio, e quindi in una direzione ortogonale a quella precedente. Ciò lo porta a ritraversare il canalone fino a portarsi un poco piú in basso sul versante opposto; in altre parole lo sciatore fa numerosissimi passi brevi. Questa situazione corrisponde ad una forma quadratica caratterizzata da un forte dispersione degli autovalori (piccolo nella direzione del canalone, grandi nelle direzioni normali ad esso).

Con la tecnica del gradiente coniugato, lo sciatore è forzato ad allontanarsi dalla direzione di massimo pendio procedendo verso una direzione "nuova" rispetto all'ultima direzione presa ed anche a tutte le precedenti. Pertanto, invece di disporre gli sci lungo la direzione di massimo pendio, vi rinuncia e imbocca il canalone facendo passi piú lunghi, anche se in una direzione in cui il pendio è meno ripido. Nel caso della fig. 23.2, si arriva al minimo di quota in due passi, coerentemente con il fatto che se la minimizzazione è fatta con il gradiente coniugato, si arriva al minimo in N passi nel caso di una forma quadratica N - dimensionale). Il primo passo è lo stesso che per il metodo del gradiente. Con il gradiente co-

CAPITOLO 23 TECNICHE DI OTTIMIZZAZIONE

niugato, quindi, la direzione di discesa si allontana progressivamente da quelle di massima concavità per esplorare le direzioni dove le concavità sono minori. Peraltro, le successive ottimizzazioni **non sono in contrasto tra loro** e progressivamente portano alla soluzione. Negli ultimi passi, lo sciatore è nel "campetto" vicino al traguardo e anche se si ferma, non accade nulla di grave. In altre parole, la presenza di errori di stima della matrice di covarianza portano comunque a errori nel posizionamento del minimo globale (*pozzanghera vicino al traguardo*). Peraltro, anche se lo sciatore non arriva esattamente al traguardo (il processo termina prima di N passi), la perdita di livello è minima.