Capitolo 11

Filtri di Wiener

La teoria formulata da Norbert Wiener dà i fondamenti sui filtri costruibili da dati, ottimi nel senso dei minimi quadrati. Questi filtri giocano un importante ruolo nelle applicazioni a problemi di predizione lineare, ricostruzione dei segnali, identificazione di sistema, equalizzazione di canale.

Nel primo paragrafo viene formulato e risolto il problema della costruzione del filtro per processi stocastici stazionari a tempo continuo, nel dominio delle frequenze. I risultati vengono applicati a problemi quali la riduzione di rumore additivo ed equalizzazione di canale. Si mostra poi la configurazione di un sistema per la stima dei coefficienti di Wiener nel dominio delle frequenze.

Il secondo paragrafo è dedicato alla derivazione del filtro di Wiener FIR causale per sistemi a tempo discreto; viene discussa la realizzazione di tali filtri per segnali ergodici.

Nel terzo paragrafo, infine, si discute la realizzazione di algoritmi adattativi per il filtro di Wiener e si mostra l'algoritmo LMS.

11.1 Formulazione nel Dominio delle Frequenze

In questo paragrafo deriviamo il filtro di Wiener, formulando il problema per sistemi stazionari a tempo continuo nel dominio delle frequenze.

Il problema può essere posto in questi termini: dati due processi stazionari $X(t)$ e $Y(t)$, determinare il sistema lineare tempo-invariante S tale che la risposta $Z(t) = S(Y(t))$ sia più "vicina possibile" a $X(t)$ (vedi Figura 11.1).

$$Y(t) \longrightarrow \boxed{\text{SISTEMA INCOGNITO}} \longrightarrow Z(t) \approx X(t)$$

Figura 11.1

Allo scopo di riformulare il problema nel dominio delle frequenze, si considerano le trasformate di Fourier $F(\omega)$ e $G(\omega)$, rispettivamente dei processi stocastici $X(t)$ e $Y(t)$; indichiamo inoltre con $W(\omega)$ la funzione di trasferimento (incognita) del sistema S e con $D(\omega)$ la risposta in frequenza del sistema su ingresso $G(\omega)$, così che:

$$D(\omega) = W(w)G(\omega).$$

Fissata una particolare realizzazione del processo $Y(t)$, l'errore $e(\omega)$ tra la risposta $D(\omega)$ e il risultato desiderato $F(\omega)$, a una data frequenza ω, è:

$$e(\omega) = D(\omega) - F(\omega).$$

Una ragionevole nozione di distanza tra i due processi $D(\omega)$ e $F(\omega)$ è l'aspettazione del quadrato del modulo dell'errore, a una data frequenza ω:

$$\text{dist}[D(\omega), F(\omega)] = \mathrm{E}\left[e(\omega)e^*(\omega)\right].$$

Il problema può allora essere riformulato come il seguente problema di ottimizzazione:

dati $F(\omega)$ e $G(\omega)$, per ogni fissato ω determinare $W(\omega)$ che minimizza $\text{dist}[G(\omega)W(\omega), F(\omega)]$.

Ricordiamo che la condizione necessaria di minimo è:

$$\frac{\partial}{\partial W(\omega)}\text{dist}[G(\omega)W(\omega), F(\omega)] = 0, \tag{11.1}$$

dove $\frac{\partial}{\partial W(\omega)}$ è l'operazione di derivata complessa, essendo $W(\omega)$ in generale a valori complessi. Si può calcolare (vedi sotto) che:

$$\frac{\partial}{\partial W(\omega)}\text{dist}[G(\omega)W(\omega), F(\omega)] = 2(W(\omega)S_{YY}(\omega) - S_{XY}(\omega)), \tag{11.2}$$

dove $S_{YY}(\omega) = \mathrm{E}\left[G(\omega)G^*(\omega)\right]$ è lo spettro di potenza di $Y(t)$ e $S_{XY}(\omega) = \mathrm{E}\left[F(\omega)G^*(\omega)\right]$ è lo spettro di potenza incrociato di $X(t)$ e $Y(t)$.

Il filtro di Wiener $W(\omega)$ che minimizza la funzione $\text{dist}[G(\omega)W(\omega), F(\omega)]$ verifica allora:

$$W(\omega)S_{YY}(\omega) - S_{XY}(\omega) = 0.$$

Possiamo quindi concludere:

Fatto 11.1. *Il filtro di Wiener, che minimizza la distanza nel senso del quadrato dell'errore tra la risposta a un processo $Y(t)$ e un processo $X(t)$, ha la funzione di trasferimento $W(\omega)$:*

$$W(\omega) = \frac{S_{XY}(\omega)}{S_{YY}(\omega)}.$$

dove $S_{YY}(\omega)$ è lo spettro di potenza di $Y(t)$ e $S_{XY}(\omega)$ lo spettro di potenza incrociato tra $X(t)$ e $Y(t)$.

Allo scopo di dimostrare la (11.2), richiamiamo brevemente la nozione di derivata complessa. Se $z = a + ib$ e $Z(a,b) = A(a,b) + iB(a,b)$, allora la derivata complessa di Z rispetto a z è:

$$\frac{\partial Z}{\partial z} = \frac{\partial Z}{\partial a} + i\frac{\partial Z}{\partial b}.$$

In particolare vale che:

$$\frac{\partial z}{\partial z} = 0, \qquad \frac{\partial z^*}{\partial z} = 2, \qquad \frac{\partial zz^*}{\partial z} = 2z.$$

Applicando le regole sopra descritte al termine $\frac{\partial}{\partial W(\omega)}\text{dist}[G(\omega)W(\omega), F(\omega)]$, si ottiene:

$$\begin{aligned}
\frac{\partial \mathrm{E}\left[e(\omega)e^*(\omega)\right]}{\partial W(\omega)} &= \frac{\partial}{\partial W(\omega)}\mathrm{E}\left[(G(\omega)W(\omega) - F(\omega))(G^*(\omega)W^*(w) - F^*(\omega))\right] \\
&= \frac{\partial}{\partial W(\omega)}(W(\omega)W^*(\omega)\mathrm{E}\left[G(w)G^*(\omega)\right] - W(\omega)\mathrm{E}\left[G(\omega)F^*(\omega)\right]) \\
&\qquad\qquad - \frac{\partial}{\partial W(\omega)}W^*(\omega)\mathrm{E}\left[G^*(\omega)F(\omega)\right] \\
&= 2W(\omega)\mathrm{E}\left[G(\omega)G^*(\omega)\right] - 2\mathrm{E}\left[F(\omega)G^*(\omega)\right] \\
&= 2(W(\omega)S_{\mathrm{YY}} - S_{\mathrm{XY}}).
\end{aligned}$$

Esempio 11.1.1.

Filtro per la riduzione di rumore additivo. Consideriamo un processo $Y(t)$ ottenuto contaminando un processo $X(t)$ con rumore additivo $N(t)$, a media 0 e scorrelato da $X(t)$:

$$Y(t) = X(t) + N(t).$$

In questo caso $S_{\mathrm{XY}} = S_{\mathrm{XX}}$, mentre $S_{\mathrm{YY}} = S_{\mathrm{XX}} + S_{\mathrm{NN}}$. Il filtro di Wiener W risulta allora:

$$W(\omega) = \frac{S_{\mathrm{XX}}(\omega)}{S_{\mathrm{XX}}(\omega) + S_{\mathrm{NN}}(\omega)}.$$

Dividendo numeratore e denominatore per lo spettro di potenza del rumore $S_{\mathrm{NN}}(\omega)$ e introducendo la variabile $\mathrm{SNR}(\omega) = S_{\mathrm{XX}}(\omega)/S_{\mathrm{NN}}(\omega)$, detta *rapporto segnale rumore* si ottiene infine:

$$W(\omega) = \frac{\mathrm{SNR}(\omega)}{1 + \mathrm{SNR}(\omega)}.$$

Si osservi che:

- $0 \leq W(\omega) \leq 1$,
- se $\mathrm{SNR}(\omega) \approx +\infty$, allora $W(\omega) \approx 1$,
- se $\mathrm{SNR}(\omega) \approx 0$, allora $W(\omega) \approx 0$.

In sostanza, il filtro di Wiener per rumore additivo attenua ogni componente di frequenza in relazione a una stima del rapporto segnale rumore.

Se gli spettri del segnale e del rumore sono completamente separati, allora il filtro di Wiener permette la ricostruzione perfetta del segnale (con combinazioni di filtri passa-banda), altrimenti attenua la potenza del rumore.

Nel caso limite di rumore bianco, il filtro di Wiener risulta:

$$W(\omega) = \frac{S_{XX}(\omega)}{S_{XX}(\omega) + \frac{N_0}{2}}.$$

Esempio 11.1.2.

Equalizzazione di canale. La distorsione introdotta in un canale di comunicazione può essere modellata da una combinazione di un filtro lineare, di cui è nota la funzione di trasferimento, con un rumore additivo (vedi Figura 11.1.2):

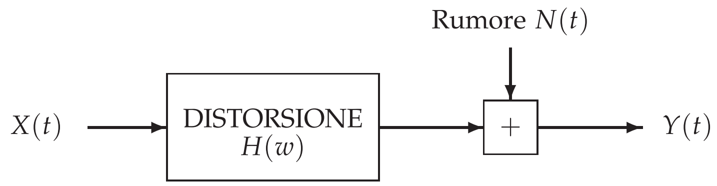

Figura 11.2 Canale di comunicazione con rumore additivo.

Per ottenere una buona ricostruzione di $X(t)$ sulla base dell'osservazione di $Y(t)$, basta applicare a $Y(t)$ il filtro di Wiener $W(\omega)$ con:

$$W(\omega) = \frac{\text{SNR}(\omega)}{|H(\omega)|^2 + \text{SNR}(\omega)}.$$

Esempio 11.1.3.

Implementazione di filtri di Wiener. La implementazione di filtri di Wiener per la riduzione di rumore additivo richiede la conoscenza dello spettro di potenza del segnale e del rumore; il principale problema è che il segnale desiderato è osservato con aggiunta di rumore, e quindi non è direttamente disponibile lo spettro di potenza del segnale. Una soluzione possibile è quella di stimare lo spettro di potenza del segnale più rumore, stimare lo spettro di potenza del rumore per periodi in cui il segnale è disattivato, ottenendo infine per sottrazione lo spettro di potenza del segnale e quindi il filtro W come mostrato in figura 11.3.

L'assunzione che il rumore sia stazionario può essere rilassata a quella che il rumore sia quasi-stazionario, cioè stazionario per periodi sufficientemente lunghi: in tal caso il filtro può essere ricalcolato periodicamente. L'assunzione di quasi-stazionarietà è ragionevole per molti tipi di rumore (ad esempio il rumore creato dal motore di un'automobile, il rumore creato dai computer in un ufficio).

11.2 Filtro di Wiener FIR

In questo paragrafo consideriamo segnali e sistemi a tempo discreto $t \in \{\dots, -1, 0, 1, 2 \dots\}$. Obbiettivo è di derivare un algoritmo per la determinazione del filtro di Wiener S ottimo rispetto a due processi stazionari $X(t)$ e $Y(t)$, sotto l'ipotesi che:

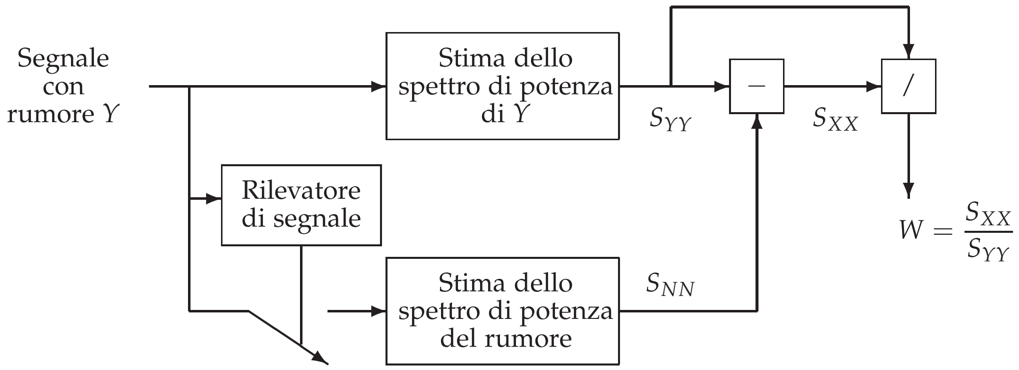

Figura 11.3 Implementazione di filtri di Wiener.

- il filtro sia causale, cioè $w(t) = 0$ per $t < 0$, dove $w(t)$ è la risposta di S all'impulso;
- la risposta $w(t)$ di S all'impulso sia di durata finita, cioè $w(t) = 0$ per $t \geq q$, per un opportuno q (filtro FIR).

Supponiamo che il sistema S abbia un risposta di durata q, con q fissato; nelle ipotesi precedenti esso sarà individuato da un vettore $w = (w_0, \dots, w_{q-1})$ con $w_t = w(t)$ per $0 \leq t < q$.

Dato un processo stazionario $Y(t)$, la risposta $Z(r,t)$ del sistema alla realizzazione $Y(r,t)$ del processo è esprimibile mediante convoluzione $w * Y(t)$:

$$Z(r,t) = \sum_{k=0}^{q-1} w_k Y(r, t-k).$$

Dato un altro processo stazionario $X(t)$ che interpretiamo come segnale desiderato, l'errore $e(t)$ tra $X(t)$ e la risposta $Z(t)$ del filtro w al processo $Y(t)$ è dato da:

$$e = \mathrm{E}\left[(X(t) - w * Y(t))^2\right].$$

Il problema di determinazione del filtro ottimo può allora essere visto come il seguente problema di ottimizzazione:

determinare w che minimizza l'errore $e = \mathrm{E}\left[(X(t) - w * Y(t))^2\right]$.

Condizione necessaria di minimo è:

$$\frac{\partial e}{\partial w_k} = 0 \quad (0 \leq k < q).$$

Calcoliamo le precedenti derivate, denotando con R_{YY} e R_{XY} le funzioni di autocorrelazione di $Y(t)$ e di cross-correlazione di $X(t)$ e $Y(t)$:

$$\begin{aligned}
\frac{\partial e}{\partial w_k} &= \frac{\partial}{\partial w_k} E\left[\left(X(t) - \sum_{j=0}^{q-1} w_j Y(t-j)\right)^2\right] \\
&= E\left[2\left(X(t) - \sum_{j=0}^{q-1} w_j Y(t-j)\right) Y(t-k)\right] \\
&= 2\left(\mathrm{E}\left[X(t)Y(t-k)\right] - \sum_{j=0}^{q-1} w_j \mathrm{E}\left[Y(t-j)Y(t-k)\right]\right) \\
&= 2\left(R_{XY}(k) - \sum_{j=0}^{q-1} w_j R_{YY}(k-j)\right)
\end{aligned}$$

La determinazione del filtro di Wiener w è quindi ridotta alla soluzione del seguente sistema lineare:

$$(*) \qquad R_{XY}(k) - \sum_{j=0}^{q-1} w_j R_{YY}(k-j) = 0 \quad (0 \le k < q).$$

Le informazioni intorno a $X(t)$ e a $Y(t)$ necessarie per poter costruire il filtro di Wiener, che applicato a $Y(t)$ approssima il comportamento desiderato $X(t)$, sono essenzialmente la funzione di autocorrelazione R_{YY} e quella di cross-correlazione R_{XY}.

Ad alto livello, l'algoritmo prevede due passi fondamentali ben distinti:

1. Stima di R_{YY} e R_{XY}.

2. Soluzione del sistema lineare (*).

Riguardo al passo 1., se i processi sono ergodici è sufficiente stimare su una realizzazione la funzione di autocorrelazione temporale di $Y(t)$ e di cross-correlazione temporale di $X(t)$ e $Y(t)$.

Riguardo al passo 2., si osserva che la matrice relativa al sistema lineare (*) ha una struttura altamente regolare, essendo uguali tra loro le componenti sulle diagonali da sinistra a destra (matrice di Toeplitz); la matrice è inoltre simmetrica:

$$\begin{bmatrix}
R_{YY}(0) & R_{YY}(1) & \dots & R_{YY}(q-1) \\
R_{YY}(1) & R_{YY}(0) & \dots & \dots \\
\dots & \dots & \dots & R_{YY}(1) \\
R_{YY}(q-1) & \dots & R_{YY}(1) & R_{YY}(0)
\end{bmatrix}$$

Per risolvere sistemi lineari di questo tipo esistono metodi numerici particolarmente efficienti (esempio: decomposizione di Cholesky).

11.3 Algoritmi Adattativi per Filtri di Wiener: LMS

L'algoritmo che abbiamo precedentemente tratteggiato per il calcolo di un filtro di Wiener consiste di passi ben distinti:

1. acquisizione di un adeguato numero di campioni di una realizzazione dei processi $Y(t)$ e $X(t)$;

2. stima delle funzioni di autocorrelazione e cross-correlazione, che entrano come coefficienti in un sistema lineare;

3. soluzione del sistema lineare, cosa che permette di ottenere il filtro.

Questa impostazione soffre di due principali inconvenienti:

- Il tempo di esecuzione dell'algoritmo può introdurre ritardi indesiderati se si vuole fare uso del filtro in tempo reale.

- Se i processi non sono stazionari, è necessario periodicamente modificare il filtro riapplicando l'algoritmo

Un approccio alternativo consiste nel disegnare un algoritmo adattativo che cerca di stimare in modo ricorsivo i parametri w del filtro. Questo può essere ottenuto da un algoritmo iterativo che per ogni tempo t calcola una stima $w(t)$ dei parametri del filtro come segue:

1. i parametri $w(0)$ sono arbitrari, oppure stabiliti con una tecnica "ad hoc" esterna all'algoritmo;

2. i parametri $w(t+1)$ sono calcolati a partire da $w(t)$, in modo da minimizzare l'errore, cioè la distanza tra il segnale desiderato e l'uscita del filtro con parametri $w(t)$.

Lo schema generale è:

$$w(t+1) = w(t) + \mathit{Aggiornamento}(\text{errore}).$$

Il problema che ora si pone è di come scegliere *Aggiornamento*, data una funzione errore e $w(t)$, in modo che sulla $w(t+1)$ l'errore diminuisca, almeno tendenzialmente.

Una semplice soluzione può essere ottenuta attraverso il gradiente dell'errore: sia $f = f(w)$ una funzione errore le cui variabili sono le componenti di un vettore di parametri w. Poiché l'obiettivo è minimizzare l'errore, se $f(w(t))$ è l'errore al tempo t, i parametri $w(t+1)$ dovranno essere scelti in modo tale da diminuire l'errore: $f(w(t+1)) \leq f(w(t))$.

Se f è una funzione "sufficientemente regolare", un minimo locale di f può essere trovato iterando la regola data nel seguente:

Fatto 11.2. *Per $\eta > 0$ sufficientemente piccolo, sia:*

$$w_k(t+1) = w_k(t) - \eta \frac{\partial f}{\partial w_k}(w(t)), \qquad (0 \leq k < q-1).$$

Allora: $f(w(t+1)) \leq f(w(t))$.

Dimostrazione. Si consideri infatti lo sviluppo di Taylor al primo ordine:

$$\begin{aligned} f(w(t+1)) - f(w(t)) &\approx \sum_{k=0}^{q-1} \frac{\partial f}{\partial w_k}(w(t))(w_k(t+1) - w_k(t)) \\ &= -\eta \sum_{k=0}^{q-1} \left(\frac{\partial f}{\partial w_k}(w(t)) \right)^2 \leq 0. \end{aligned}$$

□

Il metodo qui tratteggiato è una variante della tecnica di ricerca di minimi su base locale, ed è detto *metodo del gradiente*. Il parametro η, detto tasso di adattamento, controlla la stabilità e la velocità di convergenza: valori troppo alti di η rendono il metodo instabile mentre valori troppo bassi lo rendono troppo lento.

Tornando al problema di disegnare un algoritmo adattativo per costruire il filtro di Wiener, consideriamo le realizzazioni $x(t)$ e $y(t)$ di due processi $X(t)$ e $Y(t)$. L'errore al tempo t tra il segnale desiderato $x(t)$ e la risposta del filtro $(w_0, w_1, \dots, w_{q-1})$ al segnale $y(t)$ può essere definito mediante l'errore istantaneo $e(t, w_0, \dots, w_{q-1})$:

$$e = x(t) - \sum_{k=0}^{q-1} w_k y_{t-k}.$$

L'algoritmo LMS (Least Mean Squared) si basa sulla applicazione del metodo del gradiente in cui la funzione errore è il quadrato dell'errore istantaneo.

Poiché il gradiente di $e^2(t)$ è dato da:

$$\frac{\partial e^2(t)}{\partial w_j} = 2e(t) \frac{\partial}{\partial w_j} \left(x(t) - \sum_{k=0}^{q-1} w_k y_{t-k} \right) = -2e(t) y_{t-j},$$

l'algoritmo LMS risulta essere:

Algoritmo LMS: ingresso $x(t)$ e $y(t)$

Per $t = 0$: $\quad w_k(0) = a_k \quad (0 \leq k < q)$

$$\text{Per tutti i } t: \quad \begin{cases} e(t) = x(t) - \sum_{k=0}^{q-1} w_k y_{t-k} \\ w_k(t+1) = w_k(t) + \eta e(t) y_{t-k} \qquad (0 \leq k < q) \end{cases}$$

I coefficienti iniziali a_k del filtro sono scelti a caso o già parzialmente adattati su base di qualche conoscenza preliminare.

Il maggior pregio dell'algoritmo LMS è la sua semplicità realizzativa ed efficienza: una semplice analisi mostra che, sia il tempo di esecuzione di ogni iterazione che lo spazio richiesto sono $O(q)$, dove q è la lunghezza del filtro.

La stabilità e l'adattabilità dell'algoritmo possono essere migliorate introducendo un fattore positivo $\alpha < 1$ come segue:

$$w_k(t+1) = \alpha w_k(t) + \eta e(t) y_{t-k}, \qquad (0 \leq k < q).$$

Il parametro α ha l'effetto di aumentare la stabilità e accelerare l'adattamento del filtro alle variazioni del segnale di ingresso.