Capitolo 6

Stima di segnali stazionari: la teoria di Wiener-Kolmogorov

per gli studenti del corso di
Stima e identificazione

Luigi Chisci, 28 Febbraio 2019

Stimatore di Wiener non causale

Si considera la stima di un segnale a tempo-discreto X_t sulla base di un segnale osservato, anch'esso a tempo-discreto, Y_t. Si assume che entrambi i segnali siano congiuntamente stazionari, vale a dire il segnale $S_t \triangleq [X_t^T, Y_t^T]^T$ ha spettro

$$\Phi_S(z) = \begin{bmatrix} \Phi_X(z) & \Phi_{XY}(z) \\ \Phi_{YX}(z) & \Phi_Y(z) \end{bmatrix}$$

e si cerca la migliore, nel senso del minimo errore quadratico medio (MMSE), stima lineare $\widehat{X}_t$ di X_t basata su Y_t, ovvero

$$\widehat{X}_t \triangleq \widehat{G}(z) Y_t \tag{1.1}$$

con $\widehat{G}(z) = \sum_{k=-\infty}^{\infty} \hat{g}_k z^{-k}$ BIBO-stabile, cioè $\sum_{k=-\infty}^{\infty} \|\hat{g}_k\| < \infty$, tale che l'errore quadratico medio $E\left[\widetilde{X}_t \widetilde{X}_t^T\right]$, $\widetilde{X}_t = X_t - \widehat{X}_t = X_t - \widehat{G}(z)Y_t$, sia minimo.

Il problema posto di stima lineare MMSE di un segnale stazionario è stato originariamente formulato e risolto da N. Wiener nel 1933 per segnali a tempo-continuo (analogici) e successivamente esteso da A.N. Kolmogorov nel 1941 al caso di segnali a tempo-discreto (digitali). Nel presente contesto verrà trattato esclusivamente il caso tempo-discreto, di maggior interesse pratico per le odierne tecnologie digitali di elaborazione dei segnali.

In primo luogo si osserva che l'errore quadratico medio da minimizzare rispetto alla scelta del filtro $G(z)$ è

$$E\left[\widetilde{X}_t \widetilde{X}_t^T\right] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi_{\widetilde{X}}(\omega)\, d\omega \tag{1.2}$$

con

$$
\begin{aligned}
\varphi_{\widetilde{X}}(\omega) &= \varphi_{X-GY,X-GY}(\omega) \\
&= \varphi_X(\omega) - G(e^{j\omega})\varphi_{YX}(\omega) - \varphi_{XY}(\omega)G^T(e^{-j\omega}) + G(e^{j\omega})\varphi_Y(\omega)G^T(e^{-j\omega}) \\
&= \left[G(e^{j\omega}) - \varphi_{YX}(\omega)\varphi_Y^{-1}(\omega)\right]\ \varphi_Y(\omega)\ \left[\cdots\right]^T + \varphi_X(\omega) - \varphi_{XY}(\omega)\varphi_Y^{-1}(\omega)\varphi_{YX}(\omega)
\end{aligned}
\tag{1.3}
$$

Poiché la minimizzazione dell'errore quadratico medio (1.2) equivale alla minimizzazione, per ogni $\omega \in [-\pi, \pi]$, della densità spettrale $\varphi_{\widetilde{X}}(\omega)$ del segnale errore, da (1.3) si evince che, per ogni ω, la risposta in frequenza del filtro ottimo $\widehat{G}(z)$ deve soddisfare

$$\widehat{G}(e^{j\omega}) = \varphi_{XY}(\omega)\ \varphi_Y^{-1}(\omega) \tag{1.4}$$

o, equivalentemente, il filtro ottimo di Wiener non causale è dato da

$$\widehat{G}(z) = \Phi_{XY}(z)\Phi_Y^{-1}(z) \tag{1.5}$$

a cui è associato uno spettro del segnale errore

$$\Phi_{\widetilde{X}}(z) = \Phi_X(z) - \Phi_{XY}(z)\Phi_Y^{-1}(z)\Phi_{YX}(z) \tag{1.6}$$

Riassumendo i precedenti sviluppi, vale il seguente risultato.

Teorema 1 (**Filtro di Wiener-Kolmogorov causale**) - La stima lineare MMSE del segnale tempo-discreto X_t basata sul segnale osservato Y_t è data da (1.1) dove il filtro ottimo $\widehat{G}(z)$ è dato da (1.5) ed il corrispondente errore quadratico medio minimo (MMSE) è dato da (1.2) con $\varphi_{\widetilde{X}}(\omega) = \Phi_{\widetilde{X}}(e^{j\omega})$ e spettro del segnale errore $\Phi_{\widetilde{X}}(\cdot)$ definito in (1.6).

Se Y_t ha uno spettro razionale (il suo spettro effettivo può sempre essere approssimato con arbitrario livello di accuratezza come tale), si può operare la fattorizzazione spettrale

$$\Phi_Y(z) \ = \ H(z)\Sigma_e H^T(z^{-1}) \tag{1.7}$$

ed il filtro di Wiener-Kolmogorov assume la forma

$$\widehat{G}(z) \ = \ \Phi_{XY}(z)\left[H(z)\Sigma_e H^T(z^{-1})\right]^{-1}$$

Nel caso particolare di segnali X_t e Y_t scalari, il filtro di Wiener assume la forma

$$\widehat{G}(z) = \frac{\Phi_{XY}(z)}{\Phi_Y(z)} \ = \ \frac{\Phi_{XY}(z)}{H(z)H(z^{-1})\sigma_e^2} \tag{1.8}$$

dove $\Phi_Y(z) = H(z)H(z^{-1})\sigma_e^2$ è la fattorizzazione spettrale canonica del processo Y_t. Effettuando la fattorizzazione spettrale canonica del segnale errore

$$\Phi_{\widetilde{X}}(z) = \Phi_X(z) - \Phi_{XY}(z)\Phi_Y^{-1}(z)\Phi_{YX}(z) = L(z)L(z^{-1})\sigma_\epsilon^2$$

si può determinare l'errore quadratico medio minimo

$$MMSE \triangleq E\left[\widetilde{X}_t^2\right] = R_{\widetilde{X}}(0)$$

risolvendo le equazioni di Yule-Walker per il processo ARMA

$$\widetilde{X}_t = L(z)\epsilon_t, \qquad \epsilon_t = wn\left(0, \sigma_\epsilon^2\right)$$

Come si vedrà successivamente con alcuni esempi, il filtro di Wiener $\widehat{G}(z)$ fornito da (1.5) risulta, di norma, *non causale*.

Applicazioni della stima alla Wiener

Si esaminano adesso alcune applicazioni classiche della teoria della stima di Wiener-Kolmogorov, con riferimento a segnali X_t e Y_t scalari. In particolare, si considerano i seguenti problemi di *filtraggio*, *deconvoluzione* e *predizione*.

Filtraggio

Si considera la seguente relazione

$$Y_t = X_t + V_t \tag{1.9}$$

dove Y_t è il segnale rumoroso osservato, X_t il segnale da stimare e V_t il rumore di misura additivo, che si assume scorrelato con X_t, i.e. $X_t \perp V_t$. Questa situazione si presenta in molti problemi di telecomunicazione nei quali Y_t rappresenta il segnale rumoroso ricevuto, X_t il segnale utile da stimare e V_t il rumore introdotto dal ricevitore: l'obiettivo è quello di estrarre il segnale utile X_t dal segnale rumoroso ricevuto Y_t. Sfruttando la scorrelazione di X_t e V_t, risulta che $\Phi_{XY}(z) = \Phi_{X,X+V}(z) = \Phi_X(z) + \Phi_{XV}(z) = \Phi_X(z)$ e $\Phi_Y(z) = \Phi_X(z) + \Phi_V(z)$, da cui il filtro MMSE di Wiener assume la forma

$$\widehat{G}_f(z) = \frac{\Phi_X(z)}{\Phi_X(z) + \Phi_V(z)} = \frac{\Phi_X(z)}{H(z)H(z^{-1})\sigma_e^2} \tag{1.10}$$

dove $(\sigma_e^2, H(z))$ fornisce la fattorizzazione spettrale canonica di $\Phi_Y(z) = \Phi_X(z) + \Phi_V(z)$. Inoltre, lo spettro del segnale errore è dato da

$$\Phi_{\widetilde{X}}(z) = \frac{\Phi_X(z)\Phi_V(z)}{\Phi_X(z) + \Phi_V(z)} = L(z)L(z^{-1})\sigma_\epsilon^2$$

Si noti come la formula dello spettro del segnale errore, in funzione degli spettri del segnale e del rumore, ricordi la formula del parallelo di due resistenze (impedenze) nel senso che

$\Phi_{\widetilde{X}}(z) = \Phi_X(z) \parallel \Phi_V(z)$.

Esempio 1 - Si consideri il seguente problema di filtraggio di un segnale autoregressivo X_t

$$\begin{aligned}
Y_t &= \underbrace{\frac{1}{1-az^{-1}} W_t}_{X_t} + V_t \\
W_t &= wn\left(0, \sigma_w^2\right) \perp V_t = wn\left(0, \sigma_v^2\right) \\
a &= \tfrac{3}{5}, \sigma_w^2 = \tfrac{41}{50}, \sigma_v^2 = 1
\end{aligned} \tag{1.11}$$

Si ha:

$$\Phi_{XY}(z) = \Phi_X(z) = \frac{\sigma_w^2}{(1-az^{-1})(1-az)} \tag{1.12}$$

$$\Phi_Y(z) = \Phi_X(z) + \Phi_V(z) = \frac{\sigma_w^2}{(1-az^{-1})(1-az)} + \sigma_v^2 = \underbrace{\frac{1-cz^{-1}}{1-az^{-1}}}_{H(z)} \underbrace{\frac{1-cz}{1-az}}_{H(z^{-1})} \sigma_e^2 \tag{1.13}$$

dove c, $|c| < 1$, e $\sigma_e^2 > 0$ risolvono il problema di fattorizzazione spettrale

$$\sigma_w^2 + (1-az^{-1})(1-az)\sigma_v^2 = (1-cz^{-1})(1-cz)\sigma_e^2$$

Con i valori numerici di a, σ_w^2 e σ_v^2 indicati in (1.11), si ottiene:

$$c = \frac{3}{10}, \sigma_e^2 = 2$$

da cui si deduce il filtro ottimo

$$\widehat{G}_f(z) = \frac{\Phi_{XY}(z)}{\Phi_Y(z)} = \frac{\sigma_w^2/\sigma_e^2}{(1-cz^{-1})(1-cz)} = \frac{\frac{41}{100}}{\left(1-\frac{3}{10}z^{-1}\right)\left(1-\frac{3}{10}z\right)} \tag{1.14}$$

e lo spettro del segnale errore

$$\Phi_{\widetilde{X}}(z) = \frac{\Phi_X(z)\Phi_V(z)}{\Phi_X(z)+\Phi_V(z)} = \frac{\sigma_w^2\sigma_v^2/\sigma_e^2}{(1-cz^{-1})(1-cz)} = \frac{\frac{41}{100}}{\left(1-\frac{3}{10}z^{-1}\right)\left(1-\frac{3}{10}z\right)}$$

Risolvendo le equazioni di Yule-Walker per il modello autoregressivo $\tilde{x}_t - \frac{3}{10}\tilde{x}_{t-1} = \epsilon_t$ con $\epsilon_t = wn\left(0, \frac{41}{100}\right)$ si determina l'errore quadratico medio del filtro MMSE, cioè

$$MMSE \triangleq E\left[\widetilde{X}_t^2\right] = R_{\widetilde{X}}(0) = \frac{41}{91}.$$

Deconvoluzione

Per deconvoluzione si intende la stima del segnale di ingresso X_t ad un sistema dinamico di funzione di trasferimento nota, $F(z)$, dall'osservazione dell'uscita Y_t affetta da rumore di osservazione additivo V_t. Il problema è rappresentato dalla seguente relazione

$$Y_t \;=\; S_t + V_t = F(z)X_t + V_t \tag{1.15}$$

dove solitamente si ipotizza che X_t e V_t siano scorrelati. Pertanto si ha:

$$\begin{aligned}
\Phi_{XY}(z) &= \Phi_{X,FX+V}(z) &&= F(z)^{-1}\Phi_X(z) \\
\Phi_Y(z) &= \Phi_{FX+V,FX+V}(z) &&= F(z)F(z^{-1})\Phi_X(z) + \Phi_V(z)
\end{aligned}$$

da cui si deduce la funzione di trasferimento del deconvolutore MMSE di Wiener

$$\begin{aligned}
\widehat{G}_d(z) &= \frac{F(z)^{-1}\Phi_X(z)}{F(z)F(z^{-1})\Phi_X(z)+\Phi_V(z)} &&= \frac{F(z^{-1})\Phi_X(z)}{H(z)H(z^{-1})\sigma_e^2} \\
&= \frac{1}{F(z)}\frac{F(z)F(z)^{-1}\Phi_X(z)}{F(z)F(z^{-1})\Phi_X(z)+\Phi_V(z)} &&= F^{-1}(z)\frac{\Phi_S(z)}{\Phi_S(z)+\Phi_V(z)}
\end{aligned} \tag{1.16}$$

dove $(\sigma_e^2, H(z))$ fornisce la fattorizzazione spettrale canonica di $\Phi_Y(z) = F(z)F(z^{-1})\Phi_X(z) + \Phi_V(z)$. Si noti da (1.16) che la funzione di trasferimento ottima del deconvolutore è ottenuta da quella del filtro ottimo di $S_t = F(z)X_t$ basato su Y_t, i.e. $\widehat{G}_{SY}(z) = \Phi_S(z)\left(\Phi_S(z)+\Phi_Y(z)\right)^{-1}$, mediante moltiplicazione per l'inverso $F^{-1}(z)$ del filtro $F(z)$.

Esempio 2 - Si consideri il seguente problema di deconvoluzione di un segnale autoregressivo X_t all'ingresso di un filtro FIR $F(z)$:

$$\begin{aligned}
Y_t \;&=\; \underbrace{\left(1+\beta z^{-1}\right)}_{F(z)}\; \underbrace{\frac{1}{1+az^{-1}}\,W_t}_{X_t} \;+ V_t \\
&W_t = wn\left(0,\sigma_w^2\right) \perp V_t = wn\left(0,\sigma_v^2\right) \\
&a = \tfrac{3}{4}, \beta = -2, \sigma_w^2 = \tfrac{9}{22}, \sigma_v^2 = \tfrac{56}{11}
\end{aligned} \tag{1.17}$$

Si ha:

$$\Phi_{XY}(z) \;=\; F(z^{-1})\Phi_X(z) = \frac{(1+\beta z)\,\sigma_w^2}{(1+az^{-1})(1+az)} \tag{1.18}$$

$$\begin{aligned}
\Phi_Y(z) &= F(z)F(z^{-1})\Phi_X(z) + \Phi_V(z) = \frac{(1+\beta z)\,(1+\beta z^{-1})\,\sigma_w^2}{(1+az^{-1})(1+az)} + \sigma_v^2 \\
&= \underbrace{\frac{1+cz^{-1}}{1+az^{-1}}}_{H(z)}\underbrace{\frac{1+cz}{1+az}}_{H(z^{-1})}\sigma_e^2
\end{aligned} \tag{1.19}$$

dove c, $|c| < 1$, e $\sigma_e^2 > 0$ risolvono il problema di fattorizzazione spettrale

$$(1 + \beta z)\left(1 + \beta z^{-1}\right)\sigma_w^2 + (1 + az^{-1})(1 + az)\sigma_v^2 \;=\; (1 + cz^{-1})(1 + cz)\sigma_e^2$$

Con i valori numerici di a, β, σ_w^2 e σ_v^2 indicati in (1.17), si ottiene:

$$c = \frac{1}{3}, \sigma_e^2 = 9$$

da cui si deduce il deconvolutore ottimo

$$\widehat{G}_d(z) = \frac{\Phi_{XY}(z)}{\Phi_Y(z)} = \frac{(1 + \beta z)\sigma_w^2/\sigma_e^2}{(1 + cz^{-1})(1 + cz)} = \frac{(1 - 2z)\frac{1}{22}}{\left(1 + \frac{1}{3}z^{-1}\right)\left(1 + \frac{1}{3}z\right)} \tag{1.20}$$

e lo spettro del segnale errore

$$\Phi_{\widetilde{X}}(z) = \frac{\Phi_X(z)\Phi_V(z)}{F(z)F\left(z^{-1}\right)\Phi_X(z) + \Phi_V(z)} = \frac{\sigma_w^2\sigma_v^2/\sigma_e^2}{\left(1 + cz^{-1}\right)\left(1 + cz\right)} = \frac{\frac{28}{121}}{\left(1 + \frac{1}{3}z^{-1}\right)\left(1 + \frac{1}{3}z\right)}$$

Risolvendo le equazioni di Yule-Walker per il modello autoregressivo $\tilde{x}_t + \frac{1}{3}\tilde{x}_{t-1} = \epsilon_t$ con $\epsilon_t = wn\left(0, \frac{28}{121}\right)$ si determina l'errore quadratico medio del deconvolutore MMSE, cioè

$$MMSE \triangleq E\left[\widetilde{X}_t^2\right] \;=\; R_{\widetilde{X}}(0) \;=\; \frac{63}{242}.$$

Predizione

Si desidera prevedere con $T > 0$ passi di anticipo il segnale stazionario Y_t, ovvero stimare il segnale $X_t \triangleq Y_{t+T} = z^T Y_t$ dall'osservazione del segnale Y_t. La teoria di Wiener-Kolmogorov fornisce, senza imporre la causalità, il seguente predittore MMSE a T passi

$$\widehat{G}_{p,T}(z) \;=\; \frac{\Phi_{XY}(z)}{\Phi_Y(z)} \;=\; \frac{\Phi_{z^T Y,Y}(z)}{\Phi_{Y(z)}} \;=\; \frac{z^T\Phi_Y(z)}{\Phi_Y(z)} \;=\; z^T.$$

Pertanto la stima ottima di Y_{t+T}, senza il vincolo della causalità, risulta banalmente

$$\widehat{Y}_{t+T} = \widehat{G}_{p,T}(z)Y_t = z^T Y_t = Y_{t+T}$$

che non è, tuttavia, di alcuna utilità pratica in quanto per stimare il valore di Y_{t+T} si richiede di dover attendere l'istante $t + T$. Il problema della predizione di un segnale va dunque affrontato imponendo a priori la causalità del filtro di predizione $\widehat{G}_{p,T}(z)$.

Decomposizione causale-anticausale

Ogni operatore lineare, tempo-invariante, definito sullo spazio delle successioni limitate è univocamente caratterizzato dalla funzione di trasferimento

$$G(z) \;=\; \sum_{k=-\infty}^{\infty} g_k z^{-k} \tag{1.21}$$

soddisfacente la condizione di BIBO stabilità

$$\sum_{k=-\infty}^{\infty} |g_k| < \infty \tag{1.22}$$

L'applicazione dell'operatore $G(z)$ al segnale U_t produce il segnale

$$Y_t \;=\; G(z)U_t \;=\; \sum_{k=-\infty}^{\infty} g_k U_{t-k} \tag{1.23}$$

che, in virtù della condizione (1.22), garantisce la limitatezza della successione di uscita Y_t se la successione di ingresso U_t è limitata.

Definizione - L'operatore lineare definito da (1.21)-(1.22) dicesi;

- **causale** se $g_k = 0$ per $k < 0$, cioè $G(z) = \sum_{k=0}^{\infty} g_k z^{-k}$;

- **strettamente causale** se $g_k = 0$ per $k \leq 0$, cioè $G(z) = \sum_{k=1}^{\infty} g_k z^{-k}$;

- **anticausale** se $g_k = 0$ per $k > 0$, cioè $G(z) = \sum_{k=0}^{\infty} g_{-k} z^{k}$;

- **strettamente anticausale** se $g_k = 0$ per $k \geq 0$, cioè $G(z) = \sum_{k=1}^{\infty} g_{-k} z^{k}$.

Ogni operatore lineare $G(z)$ può, quindi, essere decomposto in parte *causale* $G_+(z)$ e parte *strettamente anticausale* $G_-(z)$, ovvero

$$G(z) \;=\; G_+(z) + G_-(z), \qquad G_+(z) = \sum_{k=0}^{\infty} g_k z^{-k}, \qquad G_-(z) = \sum_{k=1}^{\infty} g_{-k} z^{k}$$

a cui corrisponde la decomposizione di (1.23) come

$$
\begin{aligned}
Y_t &= \underbrace{G_+(z)U_t}_{Y_t^+} + \underbrace{G_-(z)U_t}_{Y_t^-} \\
Y_t^+ &= G_+(z)U_t = \sum_{k=0}^{\infty} g_k \, U_{t-k} \\
Y_t^- &= G_-(z)U_t = \sum_{k=1}^{\infty} g_{-k} \, U_{t+k}
\end{aligned}
$$

Risulta pertanto banale decomporre un operatore lineare in parti causale e strettamente anticausale se si conoscono i coefficienti della risposta all'impulso. Ci si chiede come si debba operare, viceversa, se si conoscono i coefficienti della funzione di trasferimento. In primo luogo si osserva che $G(z)$ non ha poli in $|z| \geq 1$ se G è un operatore causale. Infatti, per ogni z tale che $|z| \geq 1$

$$
\left| \sum_{k=0}^{\infty} g_k z^{-k} \right| \leq \sum_{k=0}^{\infty} |g_k||z|^{-k} \leq \sum_{k=0}^{\infty} |g_k| < \infty
$$

Analogamente, $G(z)$ non ha poli in $|z| \leq 1$ se G è un operatore anticausale. Infatti, per ogni z tale che $|z| \leq 1$

$$
\left| \sum_{k=0}^{\infty} g_{-k} z^{k} \right| \leq \sum_{k=0}^{\infty} |g_{-k}||z|^{k} \leq \sum_{k=0}^{\infty} |g_{-k}| < \infty
$$

Pertanto vale il seguente risultato.

Teorema 2 - L'operatore lineare associato alla funzione di trasferimento razionale $G(z)$ è

- **causale** se i poli di $G(z)$ sono interni al cerchio di raggio unitario;

- **anticausale** se i poli di $G(z)$ sono esterni al cerchio di raggio unitario.

Dal teorema consegue che ogni operatore $G(z)$ razionale, privo di poli sul cerchio unitario, può essere univocamente decomposto come somma di una parte causale $G_+(z)$ e di una parte strettamente anticausale $G_-(z)$. Un modo sistematico di procedere (vedi paragrafo 7.3.6 a pag. 191 di [1]) è quello di operare una sorta di decomposizione in fratti parziali della funzione razionale $G(z)$ e di separare i fratti parziali con poli in $|z| < 1$ da quelli con poli in $|z| > 1$.

Nel seguito, a titolo esemplificativo, si illustra la decomposizione causale-anticausale con riferimento al filtro ottimo (1.14) ed al deconvolutore ottimo (1.20)

Esempio 3. Decomposizione causale-anticausale del filtro ottimo (1.14) - Si nota anzitutto che la funzione di trasferimento $\widehat{G}_f(z)$ in (1.14) ha un polo in $z = \frac{3}{10}$, interno al cerchio unitario, ed uno in $z = \frac{10}{3}$, esterno al cerchio unitario. Questo suggerisce la seguente decomposizione in fratti parziali:

$$\widehat{G}_f(z) \;=\; \underbrace{\frac{A}{1 - cz^{-1}}}_{\widehat{G}_f^+(z)} + \underbrace{\frac{Bz}{1 - cz}}_{\widehat{G}_f^-(z)} \tag{1.24}$$

dove il fattore z a numeratore della seconda frazione è stato introdotto per garantire la stretta anticausalità di $\widehat{G}_f^-(z)$. Le costanti A e B sono semplicemente ottenute eguagliando (1.24) a (1.14):

$$A = \frac{\sigma_w^2/\sigma_e^2}{1 - c^2} = \frac{41}{91}, \qquad B = \frac{c\sigma_w^2/\sigma_e^2}{1 - c^2} = \frac{123}{910}$$

Grazie a questa decomposizione, il filtraggio non causale del segnale Y_t può essere implementato nel seguente modo

$$\widehat{X}_t \;=\; \widehat{X}_t^+ + \widehat{X}_t^-$$

dove il termine $\widehat{X}_t^+$ è determinato, a partire da una condizione iniziale X_0^+, mediante la ricorsione in avanti

$$\widehat{X}_t^+ = c\widehat{X}_{t-1}^+ + AY_t \;=\; \frac{3}{10}\widehat{X}_{t-1}^+ + \frac{41}{91}\,Y_t,$$

mentre $\widehat{X}_t^-$ è determinato, a partire dalla condizione terminale X_N^+, mediante la ricorsione all'indietro

$$\widehat{X}_t^- = c\widehat{X}_{t+1}^- + BY_{t+1} \;=\; \frac{3}{10}\widehat{X}_{t+1}^- + \frac{123}{910}\,Y_{t+1}$$

Esempio 4. Decomposizione causale-anticausale del deconvolutore ottimo (1.20) - Procedendo in modo del tutto analogo al precedente esempio, si ha:

$$\widehat{G}_d(z) \;=\; \underbrace{\frac{A}{1 + cz^{-1}}}_{\widehat{G}_f^+(z)} + \underbrace{\frac{Bz}{1 + cz}}_{\widehat{G}_f^-(z)} \tag{1.25}$$

dove, eguagliando (1.25) a (1.20), le costanti A e B risultano:

$$A = \frac{1 - \beta c}{1 - c^2}\;\frac{\sigma_w^2}{\sigma_e^2} = \frac{15}{176}, \qquad B = \frac{\beta - c}{1 - c^2}\;\frac{\sigma_w^2}{\sigma_e^2} = -\frac{21}{176}$$

Grazie a questa decomposizione, la deconvoluzione non causale del segnale Y_t può essere implementata nel seguente modo

$$\widehat{X}_t \;=\; \widehat{X}_t^+ + \widehat{X}_t^-$$

dove il termine $\widehat{X}_t^+$ è determinato, a partire da una condizione iniziale X_0^+, mediante la ricorsione in avanti

$$\widehat{X}_t^+ = -c\widehat{X}_{t-1}^+ + AY_t \;=\; -\frac{1}{3}\widehat{X}_{t-1}^+ + \frac{15}{176}\,Y_t,$$

mentre $\widehat{X}_t^-$ è determinato, a partire dalla condizione terminale X_N^+, mediante la ricorsione all'indietro

$$\widehat{X}_t^- = -c\widehat{X}_{t+1}^- + BY_{t+1} \;=\; -\frac{1}{3}\widehat{X}_{t+1}^- - \frac{21}{176}\,Y_{t+1}$$

Predizione MMSE di processi ARMA

Si vuole affrontare il problema della predizione ottima, nel senso MMSE, a $T > 0$ passi in avanti di un processo ARMA

$$y_t \;=\; H(z)e_t \;=\; \sum_{k=0}^{\infty} h_k e_{t-k}$$

Si ricorda che un tale processo è ottenuto filtrando un rumore bianco $e_t = wn(0, \sigma_e^2)$ mediante una funzione di trasferimento

$$H(z) \;=\; \frac{C(z)}{A(z)} \;=\; \frac{1 + c_1 z^{-1} + \cdots + c_m z^{-m}}{1 + a_1 z^{-1} + \cdots + a_n z^{-n}} \;=\; \sum_{k=0}^{\infty} h_k z^{-k}$$

avente poli in $|z| < 1$ e zeri in $|z| \leq 1$. L'obiettivo è quello di determinare un filtro causale $\widehat{G}_T(z)$ tale che l'errore di predizione $\tilde{y}_{t+T} \stackrel{\triangle}{=} y_{t+T} - \hat{y}_{t+T} = y_{t+T} - \widehat{G}_T(z)y_t$ sia minimo in media quadratica, ovvero sia minimo $E\left[\tilde{y}_{t+T}^2\right]$. A tale proposito si osserva che:

$$\begin{aligned} y_{t+T} \;&=\; H(z)e_{t+T} = \sum_{k=0}^{\infty} h_k\, e_{t+T-k} \\ &=\; \underbrace{e_{t+T} + h_1 e_{t+T-1} + \cdots + h_{T-1}e_{t+1}}_{H_T^-(z)e_{t+T}} \;+\; \underbrace{\sum_{k=0}^{\infty} h_{T+k}e_{t-k}}_{z^{-T}H_T^+(z)e_{t+T}} \end{aligned} \tag{1.26}$$

dove il primo termine $y_{t+T}^- \stackrel{\triangle}{=} H_T^-(z)e_{t+T}$ è combinazione lineare delle innovazioni future $e_{t+1}, \ldots, e_{t+T}$ mentre il secondo termine $y_{t+T}^+ \stackrel{\triangle}{=} z^{-T}H_T^+(z)e_{t+T} = H_T^+(z)e_t$ è combinazione delle innovazioni presente e passate $e_t, e_{t-1}, \ldots$. In virtù della bianchezza di e_t, i due termini y_{t+T}^- e y_{t+T}^+ risultano incorrelati. Pertanto

$$\tilde{y}_{t+T} \stackrel{\triangle}{=} y_{t+T} - \hat{y}_{t+T} \;=\; y_{t+T}^- + \left(y_{t+T}^+ - \hat{y}_{t+T}\right)$$

dove $y^+_{t+T} - \hat{y}_{t+T}$ è combinazione lineare di $e_t, e_{t-1}, \ldots$ e, di conseguenza, è incorrelato con y^-_{t+T}. Pertanto, si ha:

$$E\left[\tilde{y}^2_{t+T}\right] \;=\; E\left[\left(y^-_{t+T}\right)^2\right] \;+\; E\left[\left(y^+_{t+T} - \hat{y}_{t+T}\right)^2\right]$$

da cui si deduce la predizione MMSE a T passi

$$\hat{y}_{t+T} \;=\; y^+_{t+T} \;=\; H^+_T(z)e_t \;=\; H^+_T(z)H^{-1}(z)y_t \tag{1.27}$$

ed il corrispondente predittore ottimo

$$\widehat{G}_T(z) \;=\; H^+_T(z)H^{-1}(z). \tag{1.28}$$

L'errore quadratico medio corrispondente a questa scelta è

$$E\left[\tilde{y}^2_{t+T}\right] \;=\; E\left[\left(y^-_{t+T}\right)^2\right] \;=\; E\left[\left(\sum_{k=0}^{T-1} h_k e_{t+T-k}\right)^2\right] \;=\; \left[\sum_{k=0}^{T-1} h^2_k\right]\sigma^2_e \tag{1.29}$$

Si noti da (1.26) che per calcolare la funzione di trasferimento del predittore MMSE (1.28) a T passi, occorre decomporre la funzione di trasferimento $H(z) = C(z)/A(z)$ del processo ARMA nel seguente modo:

$$H(z) \;=\; H^-_T(z) + z^{-T}H^+_T(z) \tag{1.30}$$

con la condizione che $H^-_T(z)$ sia un polinomio, in z^{-1}, di grado $T-1$, vale a dire

$$H^-_T(z) \;=\; 1 + h_1 z^{-1} + \cdots + h_{T-1}z^{-(T-1)}$$

dove $h_0 = 1, h_1, \cdots, h_{T-1}$ non sono altro che i primi T coefficienti della risposta impulsiva del filtro $H(z)$. Per effettuare questa decomposizione, si può procedere nella divisione lunga fra il numeratore $C(z)$ ed il denominatore $A(z)$ della funzione di trasferimento $H(z)$. Indicando con $Q_T(z)$ il quoziente di tale divisione arrestato dopo T passi (ovvero in modo che il polinomio $Q_T(z)$ abbia grado $T-1$) e con $z^{-T}R_T(z)$ il relativo resto, vale la seguente relazione:

$$C(z) \;=\; Q_T(z)A(z) + z^{-T}R_T(z) \tag{1.31}$$

Dividendo ambo i membri di (1.31) per $A(z)$, si ha

$$H(z) \;=\; \frac{C(z)}{A(z)} \;=\; Q_T(z) + z^{-T}\frac{R_T(z)}{A(z)}$$

da cui, per confronto con (1.30), si deduce che:

$$H^-_T(z) = Q_T(z), \qquad H^+_T(z) = \frac{R_T(z)}{A(z)} \tag{1.32}$$

Sostituendo (1.32) in (1.28), si ottiene la funzione di trasferimento del predittore MMSE a T passi:

$$\widehat{G}_T(z) \; = \; \frac{H_T^+(z)}{H(z)} \; = \; \frac{R_T(z)/A(z)}{C(z)/A(z)} \; = \; \frac{R_T(z)}{C(z)} \tag{1.33}$$

Si noti che, qualunque sia il numero di passi in avanti T della predizione, il denominatore del predittore ottimo coincide con il numeratore $C(z)$ del modello ARMA. Il predittore MMSE risulta BIBO stabile, e quindi applicabile in pratica, se e solo se $C(z)$ non ha radici in $|z| = 1$ o, equivalentemente, la densità spettrale $\varphi_Y(\omega)$ del segnale di cui si vuole effettuare la predizione non si annulla per nessun valore di ω. Viceversa il numeratore $R_T(z)$ del predittore ottimo dipende dal numero di passi T e viene determinato moltiplicando il resto della divisione (1.31) per z^T. Il resto $Q_T(z)$ della divisione serve invece per valutare l'entità dell'errore di predizione. Risulta infatti che l'errore di predizione coincide con il seguente processo a media mobile (MA) di ordine $T - 1$

$$\tilde{y}_{t+T} = y_{t+T}^- = H_T^-(z) e_{t+T} = Q_T(z) e_{t+T} \tag{1.34}$$

per cui l'errore quadratico medio di predizione viene ottenuto, come indicato anche in (1.29), moltiplicando la varianza σ_e^2 del rumore bianco e_t per la somma dei quadrati dei coefficienti del polinomio $Q_T(z)$. Si noti che, per sua natura, il procedimento di calcolo del predittore MMSE è ricorsivo rispetto al numero di passi di predizione, nel senso che per determinare il predittore ottimo a T passi, si devono necessariamente determinare in sequenza i predittori ottimi a $1, 2, \ldots, T - 1$ passi. Per quantificare l'efficacia della predizione si introduce il seguente *guadagno di predizione*

$$\eta_T \; \triangleq \; \frac{E\left[\tilde{y}_{t+T}^2\right]}{E\left[y_{t+T}^2\right]} \; = \; \frac{R_{\tilde{Y}}(0)}{R_Y(0)} \; = \; \frac{\sum_{k=0}^{T-1} h_k^2}{\sum_{k=0}^{\infty} h_k^2} \in (0, 1) \tag{1.35}$$

definito come rapporto fra l'errore quadratico medio di predizione fornito dal predittore ottimo e l'errore quadratico medio di predizione fornito dal predittore media $\hat{y}_{t+T} = 0$. Per definizione, si pone $\eta_0 = 0$. Si noti da (1.35) che η_T è una successione monotona crescente da 0 a 1 per numero di passi di predizione T che varia da 0 a ∞. Ovviamente, la predizione a T passi risulta tanto più efficace quanto più η_T è piccolo. Ad esempio un guadagno di predizione $\eta_T = 0.05$ indica che l'errore quadratico medio fornito dal predittore MMSE è 20 volte più piccolo di quello fornito dal predittore banale $\hat{y}_{t+T} = 0$ che predice il segnale con la sua media. In ogni caso, poiché

$$\lim_{T \to \infty} \eta_T = 1$$

esisterà sempre un valore sufficientemente elevato di T oltre il quale non è conveniente ricorrere alla predizione ottima.

La teoria della predizione, finora considerata per segnali ARMA a media nulla, è banalmente estendibile a segnali di media arbitraria $\overline{y}_t$, eventualmente tempo-variante, della forma

$$y_t \;=\; \overline{y}_t + H(z)e_t$$

per i quali la predizione ottima a T passi è ottenibile tramite

$$\hat{y}_{t+T} \;=\; \overline{y}_{t+T} + \widehat{G}_T(z)\,(y_t - \overline{y}_t)$$

dove $\widehat{G}_T(z)$ è il predittore ottimo determinato per la parte ARMA $H(z)e_t$ del segnale y_t.

Esempio 5 - A titolo esemplificativo, si illustra la determinazione dei predittori ottimi a T passi per il processo ARMA del prim'ordine

$$y_t + ay_{t-1} \;=\; e_t + ce_{t-1}$$

con $|a| < 1$ e $|c| < 1$. Si noti che

$$\frac{1+cz^{-1}}{1+az^{-1}} \;=\; \sum_{k=0}^{\infty} h_k z^{-k} \text{ con } h_k = \begin{cases} 1, & k=0 \\ (c-a)(-a)^{k-1}, & k \geq 1 \end{cases}$$

Inoltre, risulta

$$R_T(z) = (c-a)(-a)^{T-1} \;\Longrightarrow\; \widehat{G}_T(z) = \frac{R_T(z)}{C(z)} = \frac{(c-a)(-a)^{T-1}}{1+cz^{-1}}$$

Pertanto, il predittore ottimo a T passi è definito dalla seguente equazione alle differenze

$$\hat{y}_{t+T|t} \;+ c\,\hat{y}_{t+T-1|t-1} \;=\; (c-a)(-a)^{T-1}\,y_t$$

Il corrispondente guadagno di predizione è

$$\eta_T = \frac{\displaystyle\sum_{k=0}^{T-1} h_k^2}{\displaystyle\sum_{k=0}^{\infty} h_k^2} \;=\; \frac{1+c^2-2ac-(c-a)^2 a^{2(T-1)}}{1+c^2-2ac}$$

La tabella riporta, ad esempio, i valori del guadagno η_T per $T = 1, 2, 3, 4, 5$ passi e coefficienti $c = -0.9$, $a = 0.8$.

T	η_T
1	0.06
2	0.24
3	0.38
4	0.50
5	0.60

Tabella 1.1: Guadagno di predizione per il processo ARMA $y_t + 0.8y_{t-1} = e_t - 0.9e_{t-1}$.

Esempio 6 - Per un generico processo ARMA di ordine n, il predittore MMSE ad un passo ha funzione di trasferimento

$$\widehat{G}_1(z) = \frac{z\,[C(z) - A(z)]}{C(z)} = \frac{(c_1 - a_1) + (c_2 - a_2)z^{-1} + \cdots + (c_n - a_n)z^{-(n-1)}}{1 + c_1 z^{-1} + \cdots + c_n z^{-n}}$$

ed è implementato dalla seguente equazione alle differenze

$$\hat{y}_{t+1|t} + c_1\hat{y}_{t|t-1} + \cdots + c_n\hat{y}_{t+1-n|t-n} = (c_1 - a_1)y_t + \cdots + (c_n - a_n)y_{t+1-n}$$

Il corrispondente MMSE di predizione ad un passo è dato da $E\left[\tilde{y}^2_{t+1|t}\right] = \sigma^2_e$.

Per gli sviluppi del paragrafo successivo, si fa notare che la predizione ottima di un rumore bianco $y_t = e_t$, i.e. un processo ARMA con $A(z) = C(z) = 1$, è nulla, i.e. $\hat{y}_{t+T} = \hat{e}_{t+T} = 0$, qualunque sia il numero di passi $T > 0$, con errore quadratico medio σ^2_e.

Filtro di Wiener causale

Si vuole determinare la stima ottima causale del segnale X_t basata sul segnale osservato Y_t, i.e.

$$\widehat{X}^c_t = \widehat{G}_c(z)Y_t, \qquad \widehat{G}_c(z) = \sum_{k=0}^{\infty} \hat{g}^c_k z^{-k}$$

in modo che $E\left[\left(X_t - \widehat{X}^c_t\right)\left(X_t - \widehat{X}^c_t\right)^T\right]$ sia minimo. A tale proposito, si utilizzerà il seguente risultato.

Lemma 1 (**Principio di ortogonalità della stima lineare MMSE**) - Sia $\widehat{X}_t = \widehat{G}(z)Y_t$ la stima ottima MMSE non causale di X_t basata su Y_t e sia $\widetilde{X}_t = X_t - \widehat{X}_t$ il corrispondente errore di stima. Tale errore di stima risulta incorrelato con ogni segnale S_t ottenuto filtrando linearmente Y_t, cioè

$$\Phi_{\widetilde{X},GY}(z) = 0$$

qualunque sia il filtro $G(z)$.

Dimostrazione - Si ha:

$$\begin{aligned}
\Phi_{\widetilde{X},GY}(z) &= \Phi_{X-\widehat{G}Y,GY}(z) \\
&= \Phi_{X,GY}(z) - \Phi_{\widehat{G}Y,GY}(z) \\
&= \Phi_{XY}(z)G^T(z^{-1}) - \widehat{G}(z)\Phi_Y(z)G^T(z^{-1}) \\
&= \left[\Phi_{XY}(z) - \widehat{G}(z)\Phi_Y(z)\right] G^T(z^{-1}) \\
&= \left[\Phi_{XY}(z) - \Phi_{XY}(z)\Phi_Y^{-1}(z)\Phi_Y(z)\right] G^T(z^{-1}) \\
&= 0
\end{aligned}$$

qualunque sia il filtro $G(z)$ come volevasi dimostrare. □

Si noti, quindi, che

$$\widetilde{X}_t^c \triangleq X_t - \widehat{X}_t^c = X_t - \widehat{X}_t + \widehat{X}_t - \widehat{X}_t^c = \widetilde{X}_t + \underbrace{\widehat{X}_t - \widehat{X}_t^c}_{\delta\widehat{X}_t} \tag{1.36}$$

con $\delta\widehat{X}_t \triangleq \widehat{X}_t - \widehat{X}_t^c = \left[\widehat{G}(z) - \widehat{G}_c(z)\right] Y_t \perp \widetilde{X}_t$, in virtù del precedente lemma. Indicando con $\hat{g}_k$ la risposta impulsiva del filtro di Wiener non causale $\widehat{G}(z)$, si ha

$$\widehat{X}_t = \widehat{G}(z)Y_t = \sum_{k=0}^{\infty} \hat{g}_k\, y_{t-k} + \sum_{k=1}^{\infty} \hat{g}_{-k}\, y_{t+k} \tag{1.37}$$

Sostituendo in (1.37) la decomposizione:

$$y_{t+k} = \hat{y}_{t+k|t} + \tilde{y}_{t+k|t}$$

dove $\hat{y}_{t+k|t}$ è la predizione ottima a k passi del segnale y effettuata all'istante t e $\tilde{y}_{t+k|t}$ il relativo errore di predizione, si ottiene

$$\widehat{X}_t = \underbrace{\sum_{k=0}^{\infty} \hat{g}_k\, y_{t-k} + \sum_{k=1}^{\infty} \hat{g}_{-k}\, \hat{y}_{t+k|t}}_{\widehat{X}_t^p} + \underbrace{\sum_{k=1}^{\infty} \hat{g}_{-k}\, \tilde{y}_{t+k|t}}_{\widehat{X}_t^f} \tag{1.38}$$

Ricordando (1.34), il termine $\widetilde{X}_t^f$ è combinazione lineare delle innovazioni future $e_{t+1}, e_{t+2}, \ldots$ e, quindi, è incorrelato con $\widehat{X}_t^p - \widehat{X}_t^c$ che, viceversa, è combinazione lineare delle innovazioni presente e passate $e_t, e_{t-1}, \ldots$. Pertanto, da (1.36) e (1.38) si ha:

$$\begin{aligned} E\left[\widetilde{X}_t^c\left(\widetilde{X}_t^c\right)^T\right] &= E\left[\widetilde{X}_t\widetilde{X}_t^T\right]+E\left\{\left[\widehat{X}_t^f+\left(\widehat{X}_t^p-\widehat{X}_t^c\right)\right]\left[\widehat{X}_t^f+\left(\widehat{X}_t^p-\widehat{X}_t^c\right)\right]^T\right\} \\ &= E\left[\widetilde{X}_t\widetilde{X}_t^T\right]+E\left[\widehat{X}_t^f\left(\widehat{X}_t^f\right)^T\right]+E\left[\left(\widehat{X}_t^p-\widehat{X}_t^c\right)\left(\widehat{X}_t^p-\widehat{X}_t^c\right)^T\right] \end{aligned} \tag{1.39}$$

dove si è sfruttata l'incorrelazione fra $\widetilde{X}_t$ e $\delta\widehat{X}_t \triangleq \widehat{X}_t - \widehat{X}_t^c$ e, successivamente, quella fra $\widehat{X}_t^f$ e $\widehat{X}_t^p - \widehat{X}_t^c$. Poiché i primi due termini in (1.39) non dipendono da $\widehat{X}_t^c$, la scelta ottima di quest'ultimo si ottiene azzerando il terzo termine, per cui

$$\widehat{X}_t^c \;=\; \widehat{X}_t^p \;=\; \sum_{k=0}^{\infty}\hat{g}_k\, y_{t-k} + \sum_{k=1}^{\infty}\hat{g}_{-k}\,\hat{y}_{t+k|t} \tag{1.40}$$

Da (1.40) si evince che la stima ottima causale di X_t si ottiene applicando la parte causale $\widehat{G}_+(z)$ del filtro di Wiener non causale al segnale osservato nonché la parte strettamente anticausale del medesimo filtro alla successione delle predizioni ottime $\hat{y}_{t+k|t}$. Ricordando che la predizione ottima a k passi si ottiene mediante $\hat{y}_{t+k|t} = \widehat{G}_k(z)y_t$, dove $\widehat{G}_k(z)$ è appunto la funzione di trasferimento del predittore ottimo a k passi fornita da (1.33), si può ottenere in linea di principio la funzione di trasferimento del filtro ottimo causale di X_t basato su Y_t mediante

$$\widehat{G}_c(z) \;=\; \widehat{G}_+(z) \;+\; \sum_{k=1}^{\infty}\hat{g}_{-k}\,\widehat{G}_k(z)$$

Si noti, tuttavia, come la suddetta formula non sia utilizzabile in pratica in quanto richiede, per calcolare il filtro causale ottimo $\widehat{G}_c(z)$, la combinazione lineare di infinite funzioni di trasferimento $\widehat{G}_k(z)$ dei vari predittori ottimi, per k che va da 1 a ∞. La necessità di questa somma di infinite funzioni di trasferimento deriva dalla presenza delle predizioni $\hat{y}_{t+k|t}$. Si può pensare allora di ricavare la stima ottima causale di X_t ricorrendo al segnale delle innovazioni e_t, ottenibile dal segnale osservato y_t mediante l'applicazione a quest'ultimo del filtro sbiancante $H^{-1}(z)$ (i.e. $e_t = H^{-1}(z)Y_t$), in quanto tale segnale, essendo bianco, è caratterizzzato da predizioni $\hat{e}_{t+k|t} = 0$.

Facendo riferimento alla fattorizzazione spettrale (1.7), si considera, dunque, la stima del segnale X_t sulla base del segnale e_t. Dalla teoria precedentemente sviluppata, risulta

$$\widehat{X}_t \;=\; \widehat{G}_{XE}(z)\, e_t$$

con

$$\widehat{G}_{XE}(z) = \Phi_{XE}(z)\Phi_E^{-1}(z) = \Phi_{X,H^{-1}Y}(z)\Sigma_e^{-1} = \Phi_{XY}(z)H^{-T}(z^{-1})\Sigma_e^{-1} \;=\; \sum_{k=-\infty}^{\infty}\gamma_k z^{-k}$$

Indicando con $[G(z)]_+$ la parte causale del filtro $G(z)$, si ha

$$\begin{aligned}
\widehat{X}_t^c &= [G_{XE}(z)]_+ \, e_t + \sum_{k=1}^{\infty} \gamma_k \, \hat{e}_{t+k|t} \\
&= \left[\Phi_{XY}(z) H^{-T}(z^{-1})\right]_+ \Sigma_e^{-1} \, H^{-1}(z) \, Y_t \\
&= \widehat{G}_c(z) Y_t
\end{aligned}$$

per cui il filtro di Wiener causale può essere determinato mediante la seguente formula

$$\widehat{G}_c(z) = \left[\Phi_{XY}(z) H^{-T}(z^{-1})\right]_+ \Sigma_e^{-1} \, H^{-1}(z) \tag{1.41}$$

Si noti che, in generale, il filtro ottimo causale differisce dalla parte causale del filtro ottimo non causale; infatti

$$\widehat{G}_c(z) = \left[\Phi_{XY}(z) H^{-T}(z^{-1})\right]_+ \Sigma_e^{-1} \, H^{-1}(z) \;\neq\; \left[\Phi_{XY}(z) H^{-T}(z^{-1}) \Sigma_e^{-1} \, H^{-1}(z)\right]_+ = \widehat{G}_+(z)$$

Nel caso di segnali X_t e Y_t scalari, il filtro di Wiener causale assume la forma:

$$\widehat{G}_c(z) = \left[\frac{\Phi_{XY}(z)}{H(z^{-1})}\right]_+ \frac{1}{H(z)\sigma_e^2} \tag{1.42}$$

Esempio 7 - Si consideri il problema di filtraggio dell'esempio 1. Sostituendo (1.12) e (1.13) in (1.42), si ottiene il filtro ottimo causale

$$\widehat{G}_c(z) = \left[\frac{\sigma_w^2}{(1 - az^{-1})(1 - cz)}\right]_+ \frac{(1 - az^{-1})}{\sigma_e^2 (1 - cz^{-1})}$$

Effettuando la decomposizione causale-anticausale

$$\frac{\sigma_w^2}{(1 - az^{-1})(1 - cz)} = \underbrace{\frac{A'}{1 - az^{-1}}}_{\text{causale}} + \underbrace{\frac{B'z}{1 - cz}}_{\text{s. anticausale}}$$

con

$$A' = \frac{\sigma_w^2}{1 - ac}, \quad B' = \frac{c\sigma_w^2}{1 - ac}$$

si ha

$$\widehat{G}_c(z) = \frac{A'\sigma_e^{-2}}{1 - cz^{-1}} = \frac{\sigma_w^2 \sigma_e^{-2} (1 - ac)^{-1}}{1 - cz^{-1}} = \frac{\frac{1}{2}}{1 - \frac{3}{10} z^{-1}}$$

Pertanto, il filtro causale ottimo è implementato dalla seguente ricorsione:

$$\widehat{X}_t^c = c\widehat{X}_{t-1}^c + A'\sigma_e^{-2} \, Y_t = \frac{3}{10} \widehat{X}_{t-1}^c + \frac{1}{2} Y_t$$

Esempio 8 - Si consideri il problema di deconvoluzione dell'esempio 2. Sostituendo (1.18) e (1.19) in (1.42), si ottiene il deconvolutore ottimo causale

$$\widehat{G}_c(z) \;=\; \left[\frac{(1+\beta z)\sigma_w^2}{(1+az^{-1})(1+cz)}\right]_+ \frac{(1+az^{-1})}{\sigma_e^2(1+cz^{-1})}$$

Effettuando la decomposizione causale-anticausale

$$\frac{(1+\beta z)\sigma_w^2}{(1+az^{-1})(1+cz)} \;=\; \underbrace{\frac{A'}{1+az^{-1}}}_{\text{causale}} + \underbrace{\frac{B'z}{1+cz}}_{\text{s. anticausale}}$$

con

$$A' = \frac{(1-a\beta)\sigma_w^2}{1-ac}, \quad B' = \frac{(\beta-c)\sigma_w^2}{1-ac}$$

si ha

$$\widehat{G}_c(z) \;=\; \frac{A'\sigma_e^{-2}}{1+cz^{-1}} \;=\; \frac{(1-a\beta)\sigma_w^2\sigma_e^{-2}(1-ac)^{-1}}{1+cz^{-1}} \;=\; \frac{\frac{5}{33}}{1+\frac{1}{3}z^{-1}}$$

Pertanto, il deconvolutore causale ottimo è implementato dalla seguente ricorsione:

$$\widehat{X}_t^c \;=\; -c\widehat{X}_{t-1}^c + A'\sigma_e^{-2}\,Y_t \;=\; -\frac{1}{3}\widehat{X}_{t-1}^c \;+\; \frac{5}{33}Y_t$$

Bibliografia

[1] T. Sodestrom: *Discrete-time stochastic systems: estimation and control*, Prentice Hall, 1994.