IMAD esercitazioni

Silviu Filote

May 2023

Contents

1 Teoria I

Defizione: un processo stocastico $v(t,s)$ a tempo discreto é una successione infinita di variabili casuali, definite a partire dallo stesso esperimento causale s e ordinate secondo un indice temporale $t \in \mathbb{N}$

$$v(1,s), v(2,s), \dots, v(N,s) \qquad N \in \mathbb{N}$$

Osservazioni:

- **Fissato un esito** $s = \bar{s}$, si ottine una realizzazione $v(t,\bar{s})$ del processo stocastico, ovvero una serie di valori deterministici nel tempo (un segnale)
- **Fissato un istante temporale** $t = \bar{t}$, si **ottine la variabile casuale** $v(\bar{t},s)$, ovvero la variabile casuale al tempo $\bar{t}$
- **Fissati** $s = \bar{s}$ e $t = \bar{t}$, si ottiene un numero $v(\bar{t},\bar{s})$

Osservazione: un processo é completamente caratterizzato dal punto di vista probabilistico dalla distribuzione di probabilità congiunta di infinite v.c. Per cui, spesso ci si limita a considerare solo valore atteso e funzione di covarianza di un processo stocastico.

Definizione: un processo stocastico $v(t)$ si dice **stazionario in senso forte** se $\forall n \in \mathbb{N}$, scelti $t_1, t_2, \dots, t_n$ istanti di tempo, le caratteristiche probabilistiche della $n-$upla $v(t_1), v(t_2), \dots, v(t_n)$ sono sono uguali a quelle della $n-$upla $v(t_1+\tau), v(t_2+\tau), \dots, v(t_n+\tau)$ con $\forall \tau \in \mathbb{N}$

Stesse caratteristiche probabilistiche:

$$v(t_1), v(t_2), \dots, v(t_n) \equiv v(t_1+\tau), v(t_2+\tau), \dots, v(t_n+\tau) \quad \forall \tau \in \mathbb{N}$$

Definizione: un processo stocastico $v(t)$ si dice **stazionario in senso debole** se:

- $\mathbb{E}[v(t)] = m_v(t) = m \qquad \forall t$
- $\gamma_{vv}(t_1,t_2) = \gamma_{vv}(t_3,t_4)$ nel caso in cui $|t_4 - t_3| = |t_2 - t_1| = \tau \qquad \Rightarrow \qquad \gamma_{vv}(\tau)$

L' autocovarianza dipende solo da ***LAG*** τ
e non dagli specifici valori di t_1, t_2, t_3, t_4

Definzione: due processi stocastici stazionari $v_1(t)$ e $v_2(t)$ si dicono **equivalenti** se hanno lo stesso valore atteso m e stessa funzione di autocovarianza.$\gamma(\tau)$.

Teorema: Dato un processo stocastico $y(t)$ ottenuto dalla risposta a regime di un sistema dinamico $H(z)$ alimentato da un processo stocastico $e(t)$, si ha che $y(t)$ é un **processo stazionario in senso debole** se e solo se:

- il sistema $H(z)$ é asintoticamente stabile
- il proceso $e(t)$ é stazionario in senso debole

Definizione: un sistema dinamico LTI a tempo discreto si dice **asintoticamente stabile** se i suoi poli sono in modulo minore di 1:

$$z^2 - 0.3z - 0.1 \qquad \textit{Poli: } z_1 = 0.5 \;\; z_2 = -.02$$

$$|z_1| < 1 \textit{ and } |z_2| < 1 \rightarrow \textit{ Sistema asintoticamente stabile}$$

Proprietà della funzione di autocovarianza di un pss

1. $\gamma_{yy}(0) = \mathbb{E}\big[\big(y(t) - m\big)^2\big] \geqslant 0 \qquad$ (varianza del processo)
2. $\gamma_{yy}(\tau) = \gamma_{yy}(-\tau) \qquad$ (funzione pari)
3. $|\gamma_{yy}(\tau)| \leqslant \gamma_{yy}(0), \forall \tau \qquad$ (funzione limitata)

$\gamma_{yy}(0)$ o é costante (processo non dimentica) oppure
tende a diminure con τ (processo dimentica)

Rumore bianco (white noise)

Un pss $e(t) \sim WN(\mu, \lambda^2)$ é detto rumore bianco se:

$$1)\quad \mathbb{E}[e(t)] = \mu = m$$

$$2)\quad \gamma_{ee}(0) = Var\big[e(t)\big] = \mathbb{E}\bigg[\big(e(t)-\mu\big)^2\bigg] = \lambda^2 \qquad \forall t$$

$$2)\quad \gamma_{ee}(\tau) = \mathbb{E}\bigg[\big(e(t)-\mu\big)\cdot\big(e(t+\tau)-\mu\big)\bigg] = 0 \qquad \forall t, \forall \tau \neq 0$$

$$3)\quad \Gamma_{ee} = \sum_{\tau - \infty}^{+\infty} \gamma_{ee}(\tau)\cdot e^{-j\omega\tau} = \lambda^2 \cdot e^{j\omega 0} = \lambda^2 \qquad \Phi(z) = \lambda^2$$

4) Il termine che esplodo é solo y(t), posso ricarvalo tramite la proprietá di stazionarietá, y(t-1), y(t - k) non li esplodo

$$\begin{aligned}\mathbb{E}\big[\widetilde{\eta}(t-1)\cdot\widetilde{y}(t-1)\big] &= \mathbb{E}\big[\widetilde{\eta}(t)\cdot\widetilde{y}(t)\big] = \textit{stazionarietá}\\ &= \mathbb{E}\big[\widetilde{\eta}(t)\cdot\big(\dots\big)\big] \neq 0\end{aligned}$$

la stazionarietá implica dipendenza da τ e non dai t

$$y(t-1)\cdot\eta(t-1) = y(t)\cdot\eta(t)$$

$$\eta(t-2)\cdot y(t-1) = \eta(t-1)\cdot y(t)$$

$$5)\quad e_1(t) \perp e_2(t) \Rightarrow w_1(t) \perp w_2(t) \qquad w_1(t) = W_1(z)\cdot e_1(t) \qquad w_2(t) = W_2(z)\cdot e_2(t)$$

NB: *Il rumore bianco è stazionario per definizione*

Nota: spesso, considereremo processi stocastici stazionari a media nulla. Infatti, il valore della media del processo non modifica la sua funzione di autocovarianza (in altri termini, si dice che non modifica le **caratteristiche spettrali** del processo).

Teorema: Se:

1. $v(t,s)$ é un pss
2. $\big|\mathbb{E}_s\big[v(t,s)\big]\big| < \infty$ (ovvero se la media esiste finita)

Allora il limite $\langle v(t,s)\rangle$ **converge quasi certamente.** Le conseguenze del teorema sono che la media temporale converge alla media di insieme o verticale del processo:

$$\mathbb{E}_s\big[\langle v(t,s)\rangle\big] = \mathbb{E}_s\big[v(t,s)\big] = m$$

$$\mathbb{E}_s\big[\langle v(t,s)\cdot v(t+\tau,s)\rangle\big] = R_{vv}(\tau)$$

Definzione: il processo stocastico $v(t,s)$ é detto **ergodico** se:

1. $v(t,s)$ é stazionario
2. Per $N \to +\infty$, i momenti temporali convergono quasi certamente ai rispettivi momenti di insieme

Definzione: il processo stocastico $v(t,s)$ é detto **ergodico nella media** (proprietà più debole) se:

$$\lim_{N\to+\infty} \langle v(t,s)\rangle_N = m \quad \textit{q.c. = quasi certamente}$$

Teorema: Sia $v(t,s)$ un **pss in senso debole**. Allora se:

1. $|\gamma_{vv}(0)| < +\infty$ (la varianza esiste finita)
2. $\lim_{\tau\to+\infty} \gamma_{vv}(\tau) = 0$ (la funzione di autocovarianza tende a zero)

Si ha che $v(t,s)$ é **ergodico nella media**.

Teorema: Sia $v(t,s)$ **stazionario e Gaussiano**. Allora se:

1. $|\gamma_{vv}(0)| < +\infty$ (la varianza esiste finita)
2. $\lim\limits_{\tau \to +\infty} \gamma_{vv}(\tau) = 0$ (la funzione di autocovarianza tende a zero)

Si ha che $v(t,s)$ é **ergodico**.

Sistemi dinamici LTI discreti stocastici

Riassumendo, possiamo rappresentare un sistema dinamico lineare LTI come:

1) Spazio di stato

2) Funzione di trasferimento

$$\begin{cases} x_1(t+1) = 0.1x_1(t) + 0.4x_2(t) + u(t) \\ x_2(t+1) = 0.3x_2(t) + 0.2x_2(t) \\ \quad y(t) = 3x_1(t) + x_2(t) \end{cases} \quad \underset{\text{Realizzazione}}{\overset{C(zI_n - A)^{-1}B + D}{\rightleftarrows}} \quad G(z) = \frac{3z^{-1} - 0.3z^{-2}}{1 - 0.3z^{-1} - 0.1z^{-2}}$$

3) Forma ricorsiva (o di filtraggio)

$$y(t) = 0.3y(t-1) + 0.1y(t-2) + 3u(t-1) - 0.3u(t-2)$$

Definizione: Un sistema dinamico LTI a tempo discreto si dice **asintoticamente stabile** se i suoi poli sono in modulo minore di 1

$$z^2 - 0.3z - 0.1 \qquad Poli: \ \ z_1 = 0.5 \ \ z_2 = -0.2$$

$$|z_1| < 1 \ \ e \ \ |z_2| < 1 \quad \Rightarrow \quad Sistema\ asintoticamente\ stabile$$

Definzione: ***Zeri*** sono le radici del numeratore, i ***Poli*** sono le radici del denominatore.

Densitá spettrale di potenza

$$\Gamma_{vv}(\omega) \equiv \mathcal{F}\big[\gamma_{vv}(\tau)\big] = \sum_{\tau=-\infty}^{+\infty} \gamma_{vv}(\tau) \cdot e^{-j\omega\tau}$$

$$\Phi_{vv}(z) \equiv \mathcal{Z}\big[\gamma(\tau)\big] = \sum_{\tau=-\infty}^{+\infty} \gamma_{vv}(\tau) \cdot z^{-t}$$

$$\Gamma_{vv}(\omega) = \Phi_{vv}(e^{j\omega})$$

VALORE ATTESO

Il valore atteso di $y(t)$ non dipende da t

$$\mathbb{E}\big[y(t)\big] = \sum_{i=0}^{+\infty} g(i) \cdot \mathbb{E}\big[u(t-i)\big] = G(1) \cdot m_u$$

Teorema:

$$\Gamma_{yy}(\omega) = |G(e^{j\omega})|^2 \cdot \Gamma_{uu}(\omega)$$

$$\Phi_{yy}(z) = G(z) \cdot G(z^{-1}) \cdot \Phi_{uu}(z)$$

Piú segnali in ingresso dove $y(t) = w_1(t) + w_2(t)$

$$\Gamma_{yy}(\omega) = \Gamma_{w_1w_1}(\omega) + \Gamma_{w_2w_2}(\omega) = |W_1(e^{j\omega})|^2 \cdot \Gamma_{e_1e_1}(\omega) + |W_2(e^{j\omega})|^2 \cdot \Gamma_{e_2e_2}(\omega)$$

$$= \sum_{\tau=-\infty}^{+\infty} \gamma_{w_1w_1}(\tau) \cdot e^{-j\omega\tau} + \sum_{\tau=-\infty}^{+\infty} \gamma_{w_2w_2}(\tau) \cdot e^{-j\omega\tau}$$

Famiglie di modelli stocastici

I processi stocastici che si ottengono filtrando un rumore bianco tramite un filtro asintoticamente stabile $H(z) = \frac{C(z)}{A(z)}$ sono detti **processi a spettro razionale**, dove $C(z)$ e $A(z)$ sono polinomi a coefficienti reali nella variabile z o z^{-1}.

Modelli di processi stocastici per **serie temporali:**

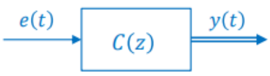
•*MA (Moving Average)* • *AR (AutoRegressive)* • *ARMA (AutoRegressive Moving Average)*

Modelli di processi stocastici per **sistemi dinamici** (quindi con ingresso $u(t)$ noto)

•*ARX (AutoRegressive with eXogenous input)* •*ARMAX* •*OE (Output Error)* •*BJ (Box-Jenkins)*

Modelli Moving Average (MA)

Sapendo che $e(t) \sim WN(\mu, \lambda^2)$ avremo:



$$y(t) = \sum_{i=0}^{n_c} c_i \cdot e(t-i) = C(z) \cdot e(t)$$

$$C(z) = \frac{y(t)}{e(t)} = \frac{z^{n_c} c_0 + \cdots + c_{n_c}}{z^{n_c}} \qquad \textit{n_c poli in 0, MA(n_c) sono sempre stazionari}$$

Valore atteso:

$$\begin{aligned} m_y = \mathbb{E}\big[y(t)\big] &= \mathbb{E}\big[c_0 \cdot e(t) + c_1 \cdot e(t-1) + \ldots + c_{n_c} \cdot e(t-n_c)\big] \\ &= c_0\mathbb{E}\big[e(t)\big] + c_1\mathbb{E}\big[e(t-1)\big] + \cdots + c_{n_c}\mathbb{E}\big[e(t-n_c)\big] \\ &= c_0 \cdot \mu + c_1 \cdot \mu + \cdots + c_{n_c} \cdot \mu = \mu \cdot \sum_{i=0}^{n_c} c_i \end{aligned}$$

Funzione di autocovarianza:

$$\begin{aligned} \gamma_{yy}(0) = \mathbb{E}\big[\big(y(t) - m_y\big)^2\big] = \mathbb{E}\big[y(t)^2\big] &= \mathbb{E}\big[\big(c_0 e(t) + c_1 e(t-1) + \cdots + c_{n_c} e(t-n_c)\big)^2\big] \\ &= c_0^2 \lambda_{ee}(0) + c_1^2 \lambda_{ee}(0) + \cdots + c_{n_c}^2 \lambda_{ee}(0) = \lambda_{ee}^2 \cdot \sum_{i=0}^{n_c} c_i^2 \end{aligned}$$

$$\begin{aligned} \gamma_{yy}(1) = \mathbb{E}\big[\big(y(t) - m_y\big)\big(y(t-1) - m_y\big)\big] &= \mathbb{E}\big[y(t)y(t-1)\big] \\ &= \lambda_{ee}^2 \cdot (c_0 c_1 + c_1 c_2 + \cdots + c_{n_c-1} c_{n_c}) \end{aligned}$$

$$\gamma_{yy}(0) = \lambda_{ee}^2 \cdot \sum_{i=0}^{n_c} c_i^2 \qquad \gamma_{yy}(\tau) = \lambda_{ee}^2 \cdot \sum_{i=0}^{n_c-\tau} c_i \cdot c_{i+\tau}$$

$$\gamma_{yy}(1) = \lambda_{ee}^2 \cdot (c_0 c_1 + c_1 c_2 + \ldots + c_{n_c-1} c_{n_c})$$

$$\gamma_{yy}(2) = \lambda_{ee}^2 \cdot (c_0 c_2 + c_1 c_3 + \ldots + c_{n_c-3} c_{n_c})$$

$$\gamma_{yy}(\tau) = 0, \forall \tau > n_c \quad \Rightarrow \quad \textit{il processo ha memoria finita}$$

Densitá spettrale di potenza:

$$\Gamma_{yy}(\omega) = \sum_{\tau=-\infty}^{+\infty} \gamma_{yy}(\tau) \cdot e^{-j\omega\tau} = \sum_{\tau=-n_c}^{+n_c} \gamma_{yy}(\tau) \cdot e^{-j\omega\tau}$$

$$\Gamma_{yy}(\omega) = |H(e^{j\omega})|^2 \cdot \Gamma_{ee}(\omega) = |C(e^{j\omega})|^2 \cdot \Gamma_{ee}(\omega)$$

Modelli AutoRegressive (AR)

Sapendo che $e(t) \sim WN(\mu, \lambda^2)$ avremo:

$$y(t) = \sum_{i=1}^{n_a} a_i \cdot y(t-i) + e(t) = \frac{1}{A(z)} \cdot e(t)$$

$$A(z) = \frac{z^{n_a}}{z^{n_a} - a_1 z^{n_a - 1} - \cdots - a_{n_1}} \qquad \textit{n_a poli, $AR(n_a)$ é stazionario se asint. stabile}$$

Valore atteso:

$$\mathbb{E}[e(t)] \neq 0 \;\; \Rightarrow \;\; \mathbb{E}[y(t)] \neq 0 \;\; \Rightarrow \textit{depolarizzo } e(t) \textit{ e } y(t)$$

$$\begin{aligned} m_y = \mathbb{E}\big[y(t)\big] &= \mathbb{E}\big[a_1 y(t-1) + a_2 y(t-2) + \cdots + a_{n_a} y(t-n_a) + e(t)\big] \\ &= a_1 \mathbb{E}\big[y(t-1)\big] + \cdots + a_{n_a} \mathbb{E}\big[y(t-n_{n_a})\big] + \mu \\ &= a_1 m_y + \cdots + a_{n_a} m_y + \mu \end{aligned}$$

$$(1 - a_1 - \cdots - a_{n_a}) m_y = \mu$$

$$m_y = \mathbb{E}\big[y(t)\big] = \frac{\mu}{1 - a_1 - \cdots - a_{n_a}}$$

Funzione di autocovarianza:

$$\begin{aligned} \gamma_{yy}(0) = \mathbb{E}\Big[\big(y(t) - m_y\big)^2\Big] = \mathbb{E}\big[y(t)^2\big] &= \mathbb{E}\Big[\big(a_1 y(t-1) + \cdots + a_{n_a} y(t-n_a) + e(t)\big)^2\Big] \\ &= a_1^2 \mathbb{E}\big[y(t-1)^2\big] + \cdots + \mathbb{E}\big[e(t)^2\big] + \cancel{2a_1 \mathbb{E}\big[y(t-1)e(t)\big]} \\ &= a_1^2 \gamma_{yy}(0) + \cdots + \lambda_{ee}^2 \end{aligned}$$

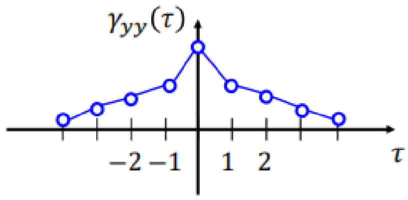
NB: *$y(t-1)$ dipende solo da $e(t-1), e(t-2), \ldots$*
ecco perché viene cancellato il termine

Densitá spettrale di potenza:

$$\Gamma_{yy}(\omega) = |H(e^{j\omega})|^2 \cdot \Gamma_{ee}(\omega) = |1/A(e^{j\omega})|^2 \cdot \Gamma_{ee}(\omega)$$

Osservazioni:

- Il processo AR(1)

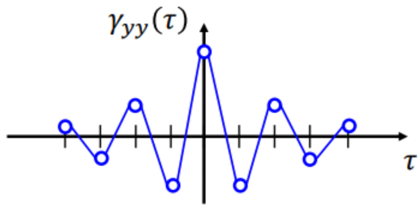
  

  $\bar{y}(t) = a_1 \cdot \bar{y}(t-1) + e(t), \quad e(t) \sim WN(0, \lambda^2), \quad$ $\mathbf{0 < a_1 < 1}$

  ha la funzione di autocovarianza $\gamma_{yy}(\tau) > 0$, $\forall \tau \Rightarrow$ **sará decrescente ma non raggiunge mai lo 0**.

  Le **realizzazioni** del processo **variano lentamente e sono smooth**, poiché le variabili casuali sono **correlate positivamente** fra loro. In media, le realizzazioni **non cambiano segno** da un istante al successivo. Le **componenti a bassa frequenza dominano** nella densità spettrale di potenza

- Il processo AR(1)

  $\bar{y}(t) = a_1 \cdot \bar{y}(t-1) + e(t), \quad e(t) \sim WN(0, \lambda^2), \quad$ $\mathbf{-1 < a_1 < 0}$

  ha la funzione di autocovarianza **che cambia di segno** ad ogni τ, in modo alternato $\Rightarrow$ **sará decrescente ma non raggiunge mai lo 0**.
  Le realizzazioni del processo **variano velocemente e sono nervose**, poiché le variabili casuali sono **correlate negativamente** fra loro. In media, le realizzazioni cambiano segno da un istante al successivo e le componenti ad **alta frequenza dominano** nella densità spettrale di potenza

- Per i processi $AR(n_a)$ possiamo osservare un comportamento analogo guardando la **funzione di autocorrelazione parziale (PACF)** $\gamma_{yy}^{PAR}(\tau)$. La PACF é tale che

$$\gamma_{yy}^{PAR}(\tau) = 0, \qquad \forall \tau > n_a$$

2 Teoria II

Proprietá: Ogni **processo stocastico stazionario** $y(t)$ puó essere scritto come:

$$y(t) = \bar{y}(t) + y_p(t)$$

- $y_p(t)$: processo stocastico stazionario **completamente predicibile**
- $\bar{y}(t)$: parte **puramente stocastica,** tale che:

$$MA(\infty) \qquad \bar{y}(t) = \sum_{i=0}^{+\infty} c_i \cdot e(t-i), \qquad e(t) \sim WN(0, \lambda^2) \qquad \sum_{i=0}^{+\infty} c_i^2 < \infty$$

- $\bar{y}(t)$ e $y_p(t)$: sono **incorrelati**

Forma canonica - ipotesi di lavoro:

- Anche eventuali **componenti di non stazionarietà** come stagionalità o trend devono essere stimate e rimosse dai dati per ottenere solo la parte stocastica del processo
- La parte puramente stocastica $\bar{y}(t)$ ovvero il processo $MA(\infty)$ puó essere approssimata da processi a spettro razionale
- Supporremo di lavorare con processi $y(t)$ generati in uscita da un sistema dinamico lineare asintoticamente stabile con funzione di trasferimento $H(z)$ razionale fratta, alimentata da rumore bianco $e(t) \sim WN(0, \lambda^2)$
$$H(z) = \frac{C(z)}{A(z)}$$
- Il problema della fattorizzazione spettrale consiste nel trovare tutte le coppie $\{H(z), \lambda^2\} \Rightarrow$ per processi a spettro razionale, esistono infiniti fattori spettrali
- Ai fini della predizione ci interessa un solo fattore spettrale $\Rightarrow$ **fattore spettrale canonico**

Teorema della fattorizzazione spettrale: dato un processo stocastico stazionario a spettro razionale, esiste **un solo fattore spettrale** $\{\widetilde{H}(z), \widetilde{\lambda}^2\}$, detto **fattore spettrale canonico**, dove $\widetilde{H}(z) = \frac{C(z)}{A(z)}$ tale che:

1. $C(z)$ e $A(z)$ hanno lo **stesso grado** (grado relativo nullo)
2. $C(z)$ e $A(z)$ sono **coprimi** (non ci son fattori in comune)
3. $C(z)$ e $A(z)$ sono **monici** (il coefficiente del termine di grado massimo è 1)
4. $C(z)$ e $A(z)$ hanno **radici interne al cerchio unitario**

Spettro razionale della parte stocastica $e(t)$, consiste di trovare:
- funzione di trasferimento $\Rightarrow H(z)$
- varianza di λ^2 di $e(t)$
se manca una della due porre = 1

$$H(z) = \frac{C(z)}{A(z)}$$

Filtro passa-tutto e forma canonica

Il filtro passa-tutto è un filtro di ordine 1 definito come:

$$T(z) = \frac{1}{a} \cdot \frac{z+a}{z+\frac{1}{a}} \qquad a \neq 0, a \in \mathbb{R}$$

$$zero : z = -a \qquad polo : z = -\frac{1}{a}$$

Lo zero é reciproco del polo

Il fattore moltiplicativo é come il polo

Inoltre la densitá spettrale di potenza $\Gamma_{yy}(\omega)$ di un processo $y(t)$ in uscita dal passa tutto $T(z)$ alimentato da un generico processo stazionario in ingresso $v(t)$ é uguale a:

$$\Gamma_{yy}(\omega) = |T(e^{j\omega})|^2 \cdot \Gamma_{vv}(\omega)$$

Il filtro passa-tutto non modifica il modulo delle frequenze nella densità spettrale di potenza dell'ingresso. Quindi, si ha che:

$$\Gamma_{yy}(\omega) = \Gamma_{vv}(\omega)$$

Esempio: Consideriamo il seguente processo ARMA(1,1)

$$y(t) = \frac{z+2}{z-\frac{1}{3}} \cdot e(t-2) \qquad e(t) \sim WN(0,1)$$

Applicando il teorema della fattorizzazione spettrale:

Applichiamo un filtro passa-tutto pechè $H(z)$ non è canonico, abbiamo uno zero fuori dal cerchio unitario

$$y(t) = \frac{\cancel{z+2}}{z-\frac{1}{3}} \cdot 2\frac{z+\frac{1}{2}}{\cancel{z+2}} \cdot e(t-2)$$

$$y(t) = \frac{z+\frac{1}{2}}{z-\frac{1}{3}} \cdot \underbrace{2e(t-2)}_{\eta(t)\sim WN(0,4)}$$

$$y(t) = \frac{1+\frac{1}{2}z^{-1}}{1-\frac{1}{3}z^{-1}} \cdot \eta(t) \qquad \eta(t) \sim WN(2\cdot 0, 2^2 \cdot 1) = WN(0,4)$$

Predittore ottimo lineare

Predizione: stimare il dato al tempo t avendo a disposizione dati fino al tempo $t-k$. Equivalentemente, stimare il dato al tempo $t+k$ avendo a disposizione dati fino al tempo t. Indichiamo il **predittore** come:

$$\hat{y}(t|t-k) \quad \textit{oppure} \quad \hat{y}(t+k|t)$$

Vogliamo però trovare il **predittore lineare ottimo** dai dati, ovvero quello che minimizza il seguente criterio **Mean Squared Error (MSE):**

$$\mathbb{E}\big[\varepsilon_k(t)^2\big] = \mathbb{E}\Big[\big(y(t) - \hat{y}(t|t-k)\big)^2\Big]$$

Definizione: un predittore (lineare) ottimo se:

- $\mathbb{E}\big[\varepsilon_k(t)\big] = \mathbb{E}\big[y(t) - \hat{y}(t|t-k)\big] = 0 \Rightarrow$ il valore atteso dell'errore di predizione è nullo.

  Quindi il processo $y(t)$ e il predittore $\hat{y}(t|t-1)$ hanno lo stesso valore atteso

- $\mathbb{E}\big[\hat{y}(t|t-k) \cdot \varepsilon_k(t)\big] = 0 \Rightarrow$ il predittore e l'errore di predizione **sono incorrelati** quindi il predittore ha utilizzato tutta l'informazione disponibile.
- $Var\big[\varepsilon_k(t)^2\big]$ minima

Scomposizione del processo come:

$$y(t) = \hat{y}(t|t-k) + \varepsilon_k(t)$$

Errore di predizione:

$$\varepsilon_k(t) = y(t) - \hat{y}(t|t-k)$$

- $y(t) = \hat{y}(t|t-k)$ é la **parte predicibile** al tempo $t-k$
- $\varepsilon_k(t)$ é la **parte impredicibile** al tempo $t-k$

Predittori ottimi

- Il filtro sbiancante viene scritto come

$$y(t) = \widehat{y}(t|t-1) + e(t) \qquad \Rightarrow \qquad e(t) = \frac{1}{\widehat{y}(t|t-1)} \cdot y(t)$$

- **Il predittore ottimo dal rumore** viene scritto in funzione di $e(t)$
- **Il predittore ottimo dai dati** viene scritto in funzione di $y(t)$ e bisogna sostituire $e(t)$ con il filtro sbiancante

Predizione

Metodo della lunga divisione k passi:
C(z): dividendo - A(z): divisore - R(z): resto - E(z): quoziente
$z^{-k}\widetilde{R} = k$ *colpi di lunga divisione*

$$H(z) = \frac{C(z)}{A(z)} \quad \Rightarrow \quad \underbrace{C(z)/A(z)}_{divido} \quad \Rightarrow \quad H(z) = E(z) + \frac{z^{-k}R(z)}{A(z)}$$

$$\begin{aligned} y(t) &= H(z) \cdot \eta(t) = \frac{C(z)}{A(z)} \cdot \eta(t) \\ &= \left(E(z) + \frac{z^{-k}\widetilde{R}(z)}{A(z)}\right) \cdot \eta(t) = \underbrace{E(z) \cdot \eta(t)}_{non\ pred} + \underbrace{\frac{\widetilde{R}(z)}{A(z)} \cdot \eta(t-k)}_{pred} \end{aligned}$$

Filtro sbiancante

$$y(t) = \frac{C(z)}{A(z)} \cdot \eta(t) \qquad \Rightarrow \qquad \eta(t) = \frac{A(z)}{C(z)} \cdot y(t)$$

Il predittore é:

$$\widehat{y}(t|t-k) = \frac{z^{-k}\widetilde{R}(z)}{A(z)} \cdot \eta(t) = \frac{\widetilde{R}(z)}{A(z)} \cdot \eta(t-k)$$

$$\begin{aligned} \widehat{y}(t|t-k) &= \frac{z^{-k}\widetilde{R}(z)}{A(z)} \cdot \eta(t) \\ &= \frac{z^{-k}\widetilde{R}(z)}{A(z)} \cdot \underbrace{\frac{A(z)}{C(z)} \cdot y(t)}_{f.sbiancante} = \frac{\widetilde{R}(z)}{C(z)} \cdot y(t-k) \end{aligned}$$

Errore di predizione:

$$\varepsilon_k(t) = y(t) - \widehat{y}(t|t-k) = E(z) \cdot e(t)$$

Secondo metodo: formule note
predittore ad un passo $z^{-1} \Rightarrow k = 1$

$$R(z) = C(z) - A(z)$$

$$E(z) = 1$$

Predittore ottimo per processi $MA(n_c)$

$$y(t) = e(t) + c_1 e(t-1) + c_2 e(t-2) + \dots + c_{n_c} e(t-n_c) \qquad e(t) \sim WN(0, \lambda^2)$$
$$y(t) = C(z) \cdot e(t) = H(z) \cdot e(t)$$

Predittore ad un passo:

$$\begin{aligned} y(t) &= \varepsilon_1(t) + \widehat{y}(t|t-1) \\ &= e(t) + c_1 e(t-1) + \ldots + c_{n_c} e(t-n_c) \end{aligned}$$

Filtro sbiancante:

$$e(t) = \frac{1}{1 + c_1 z^{-1} + c_2 z^{-2} + \dots + c_{n_c} e^{-n_c}} \cdot y(t)$$

Predittore ottimo dal rumore:

$$\begin{aligned} \widehat{y}(t|t-1) &= c_1 e(t-1) + \ldots + c_{n_c} e(t-n_c) \\ &= \left[c_1 z^{-1} + c_2 z^{-2} + \dots + c_{n_c} e^{-n_c}\right] \cdot e(t) \end{aligned}$$

Predittore ottimo dai dati:

$$\widehat{y}(t|t-1) = \frac{c_1 e(t-1) + \ldots + c_{n_c} e(t-n_c)}{1 + c_1 e(t-1) + \ldots + c_{n_c} e(t-n_c)} \cdot y(t)$$

Parte impredicibile:

$$\varepsilon_1(t) = y(t) - \widehat{y}(t|t-1) = e(t)$$

Predittore a k passi:

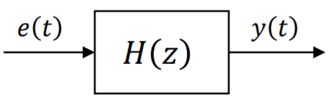
$$\begin{aligned} y(t) &= \varepsilon_k(t) + \widehat{y}(t|t-k) \\ &= e(t) + c_1 e(t-1) \ldots + c_{k-1} e(t-k+1) + c_k e(t-k) + \ldots + c_{n_c} e(t-n_c) \end{aligned}$$

$$\begin{aligned} \widehat{y}(t|t-k) &= c_k e(t-k) + \ldots + c_{n_c} e(t-n_c) \\ &= \left[c_k z^{-k} + \ldots + c_{n_c} z^{-n_c}\right] \cdot e(t) \end{aligned}$$

$$\begin{aligned} \varepsilon_k(t) &= y(t) - \widehat{y}(t|t-k) \\ &= e(t) + c_1 e(t-1) + \ldots + c_{k-1} e\big(t-k+1) \end{aligned}$$

NB: *la varianza $\varepsilon_k(t)$ aumenta con l'orizzonte di predizione, fino a diventare uguale alla varianza del processo $y(t)$*

$$MA(n_c) \rightarrow Var\big[\varepsilon_{n_c+1}(t)\big] = Var\big[y(t)\big]$$



$H(z) = C(z)/A(z)$ **fdt razionale fratta**

Predittore ottimo per processi $ARMA(n_a, n_c)$

$$y(t) = \frac{C(z)}{A(z)} \cdot e(t) = H(z) \cdot e(t) \qquad e(t) \sim WN(0, \lambda^2)$$

Lunga divisione:

$C(z)/A(z) \rightarrow k$ *colpi di lunga divisione = previsione a* k *passi*

$$\begin{aligned} y(t) = \frac{C(z)}{A(z)} \cdot e(t) &= \left(E(z) + \frac{z^{-k}\widetilde{R}(z)}{A(z)} \right) \cdot e(t) \\ &= \color{red}{E(z) \cdot e(t)} + \color{blue}{\frac{\widetilde{R}(z)}{A(z)} \cdot e(t-k)} \end{aligned}$$

filtro sbiancante:

$$e(t) = \frac{A(z)}{C(z)} \cdot y(t)$$

Predittore a k *passi:*

$$\begin{aligned} \widehat{y}(t|t-k) &= \frac{\widetilde{R}(z)}{A(z)} \cdot e(t-k) \qquad \textit{predittore ottimo dal rumore} \\ &= \frac{z^{-k}\widetilde{R}(z)}{A(z)} \cdot e(t) = \frac{z^{-k}\widetilde{R}(z)}{A(z)} \cdot \frac{A(z)}{C(z)} \cdot y(t) \\ &= \frac{\widetilde{R}(z)}{C(z)} \cdot y(t-k) \qquad \textit{predittore ottimo dai dati} \end{aligned}$$

L'errore di predizione

$$\varepsilon_k(t) = y(t) - \widehat{y}(t|t-k) = E(z) \cdot e(t)$$

Caso particolare: predizione ad un passo $k = 1$

$$R(z) = C(z) - A(z)$$

$$E(z) = 1$$

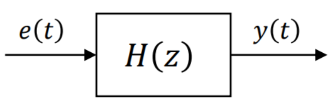
$$\widehat{y}(t|t-1) = \frac{R(z)}{C(z)} \cdot y(t) = \frac{C(z) - A(z)}{C(z)} \cdot y(t)$$

$$\varepsilon_1(t) = E(z) \cdot e(t) = e(t)$$

Qualità del predittore

$$ESR = \frac{Var\big[\varepsilon_k(t)\big]}{Var\big[y(t)\big]}$$

% Var spiegata = *1 - ESR*
ci fornisce la percentuale di varianza del
processo che è stata catturata dal predittore



$H(z) = C(z)/A(z)$ **fdt razionale fratta**

Predittore ottimo per processi $ARMAX(n_a, n_c, n_b, k)$

$$y(t) = \frac{B(z)}{A(z)} \cdot u(t-k) + \frac{C(z)}{A(z)} \cdot e(t) \qquad e(t) \sim WN(0, \lambda^2)$$

Lunga divisione:

$C(z)/A(z) \rightarrow$ *k colpi di lunga divisione = previsione a k passi*

$$y(t) = \frac{B(z)}{A(z)} \cdot u(t-k) + \frac{\widetilde{R}(z)}{A(z)} \cdot e(t-k) + E(z) \cdot e(t)$$

$$\widehat{y}(t|t-k) = \frac{B(z)}{A(z)} \cdot u(t-k) + \frac{\widetilde{R}(z)}{A(z)} \cdot e(t-k)$$

$$e(t) = \frac{A(z)}{C(z)} \cdot y(t) - \frac{B(z)}{C(z)} \cdot u(t-k)$$

$$\widehat{y}(t|t-k) = \frac{\widetilde{R}(z)}{C(z)} \cdot y(t-k) + \frac{B(z) \cdot E(z)}{C(z)} \cdot u(t-k)$$

$$\varepsilon_k(t) = y(t) - \widehat{y}(t|t-k) = E(z) \cdot e(t)$$

Caso particolare: predizione ad un passo $k=1$:

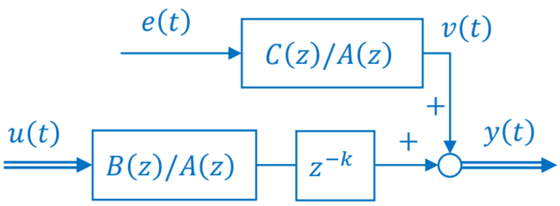
$$R(z) = C(z) - A(z)$$

$$E(z) = 1$$

$$\begin{aligned}\widehat{y}(t|t-1) &= \frac{\widetilde{R}(z)}{C(z)} \cdot y(t) + \frac{B(z) \cdot E(z)}{C(z)} \cdot u(t-1) \\ &= \frac{C(z) - A(z)}{C(z)} \cdot y(t) + \frac{B(z)}{C(z)} \cdot u(t-1)\end{aligned}$$

$$\varepsilon_1(t) = E(z) \cdot e(t) = e(t)$$

NB: *ARMAX ha un ingresso esogeno $u(t-k)$ quindi sensato fare una previsione a k passi, in modo che l'ingresso riesca ad influenzare l'uscita. Dunque il $u(t-k)$ con k passi di lunga divisione, raccogliendo nella FDT_u puó essere uguale dopo la lunga divisione*



Qualità del predittore

$$ESR = \frac{Var\big[\varepsilon_k(t)\big]}{Var\left[\frac{C(z)}{A(z)} \cdot e(t)\right]}$$

% Var spiegata = *1 - ESR*

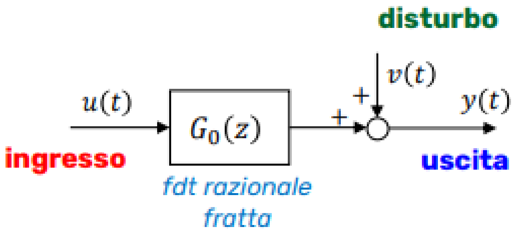
Con il predittore ARMAX, la varianza di $\varepsilon_k(t)$ è data solo dalla parte ARMA, in quanto unica parte stocastica del modello.

3 Teoria III

Criterio di identificazione: è necessario decidere come confrontare il modello con i dati. Ciò si traduce spesso nella definizione di una **funzione di costo da minimizzare**.

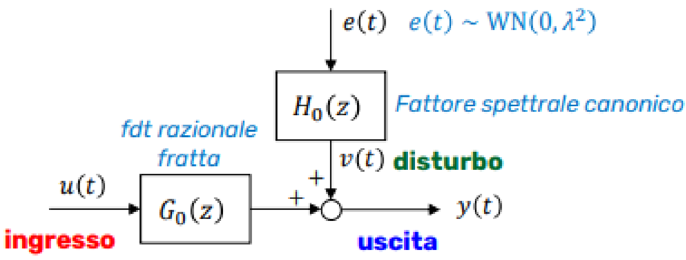
Ipotesi di lavoro 1: i dati sono generati da un sistema LTI SISO con uscita rumorosa

- I parametri da stimare sono i coefficienti del numeratore e del denominatore di $G_0(z)$
- Il disturbo $v(t)$ modella rumore di misura, ingresso non misurabili



Ipotesi di lavoro 2: il disturbo $v(t)$ è modellizzabile come un processo stocastico stazionario a spettro razionale, indipendente da $u(t)$. In questo caso, vogliamo sia stimare:

- i coefficienti del numeratore e denominatore $G_0(z)$
- i coefficienti del numeratore e denominatore $H_0(z)$
- anche stimare λ^2



Il modello più generale che usiamo per stimare un sistema dinamico è dato da:

$$y(t) = G(z, \boldsymbol{\theta}) \cdot u(t) + H(z, \boldsymbol{\theta}) \cdot e(t) \qquad e(t) \sim WN(0, \lambda^2)$$

- $H(z, \boldsymbol{\theta})$: **fattore spettrale canonico**, numeratore e denominatore monici, coprimi, di uguale grado, poli e zeri nel cerchio unitario
- $G(z, \boldsymbol{\theta})$: **funzione di trasferimento**, che descrive l'effetto dell'ingresso $u(t)$, misurabile o noto sull'uscita $y(t)$

Proprietá di $G(z, \boldsymbol{\theta})$:

- Spesso si ipotizza che $G(z, \boldsymbol{\theta})$ sia **strettamente propria**, ovvero che il grado del numeratore é minore del grado del denominatore $\Rightarrow$ fdt asintoticamente stabile e questo fa si che vi sia un ritardo $k \neq 0$ tra ingresso e uscita
- $G(z, \boldsymbol{\theta})$ puó avere **zeri fuori dal cerchio unitari** o numeratore e denominatore **non monici**
- $G(z, \boldsymbol{\theta})$ rappresenta un sistema fisico, mentre $H(z, \boldsymbol{\theta})$ ed $e(t)$ **non esistono nella realtà (sono solo costrutti matematici)**

Metodi a minimizzazione dell'errore di predizione (PEM)

- Considero come residuo $\epsilon_{\boldsymbol{\theta}}(t)$ da minimizzare l'errore di predizione a un passo $\varepsilon_1(t;\boldsymbol{\theta})$
- $\boldsymbol{\Theta}$ *é l'insieme dei valori ammissibili di* $\boldsymbol{\theta}$
- $\boldsymbol{\theta}^0$ é il theta ottimo
- L'identificazione PEM consiste nel trovare il modello che minimizza la varianza dell'errore di predizione ad un passo del modello sui dati a disposizione

$$\widehat{\boldsymbol{\theta}}_N = arg\ \underset{\boldsymbol{\theta}\in\Theta}{min} J_N(\boldsymbol{\theta})$$

$$J_N(\boldsymbol{\theta}) = \frac{1}{N}\cdot\sum_{t=1}^{N}\varepsilon_1(t;\boldsymbol{\theta})^2 = \frac{1}{N}\cdot\sum_{t=1}^{N}\left(y(t)-\widehat{y}(t|t-1;\boldsymbol{\theta})\right)^2$$

$$J_N(\boldsymbol{\theta}) = \frac{1}{N}\cdot\sum_{t=1}^{N}\left(y(t)-\boldsymbol{\varphi}^T(t)\cdot\boldsymbol{\theta}\right)^2$$

I metodi di stima basati sulla minimizzazione dell'errore di predizione prendono il nome di ***Prediction Error Methods (PEM)***

- Per valori iniziali $t = 1, 2, \ldots$ il predittore potrebbe non avere dati disponibili. Si usa quindi un'**inizializzazione convenzionale** $\Rightarrow$ ipotizzo che i valori passati di $y(\cdot)$ siano nulli. L'inizializzazione non è un problema in quanto il predittore è stabile

In alternativa, scarto quei dati iniziali che non hanno una predizione associata

- Abbiamo già in parte visto che l'errore di predizione a un passo $\varepsilon_1(t;\boldsymbol{\theta})$ gode di interessanti proprietá, che ci permettono di **distinguere** $\boldsymbol{\theta}^0$ da un qualsiasi $\boldsymbol{\theta}$
  1. Dato $\boldsymbol{\theta}$ e i dati $\{u(1),\ldots,u(N)\}$ e $\{y(1),\ldots,y(N)\}$ $\Rightarrow$ é sempre possibile calcolare $\varepsilon_1(t;\boldsymbol{\theta})$
  2. Se $\exists\boldsymbol{\theta} = \boldsymbol{\theta}^0$ $t.c$ $G_0(z) = G(z,\boldsymbol{\theta}^0)$ e $H_0(z) = H(z,\boldsymbol{\theta}^0)$, abbiamo che $\varepsilon_1(t;\boldsymbol{\theta}^0) = e(t)$ $\Rightarrow$ ovvero ci permette di capire se il modello è buono
  3. $\varepsilon_1(t;\boldsymbol{\theta}^0) \neq e(t)$ $\Rightarrow$ per qualsiasi $\boldsymbol{\theta}\neq\boldsymbol{\theta}^0$ (supponendo di avere un ingresso adeguato)
  4. $\boldsymbol{\theta}^0$ minimizza la varianza dell'errore di predizione a un passo
- **NB:** se ipotizzo che:
  - $\mathcal{S} = \mathcal{M}(\boldsymbol{\theta}^0)$
  - $e(t) \sim WN$ Gaussiano

  $\Rightarrow$ lo stimatore PEM è circa uguale allo stimatore a massima verosimiglianza.

Predittore ottimo ad un passo per sistemi ingresso\uscita

$$y(t) = G(z, \boldsymbol{\theta}) \cdot u(t) + H(z, \boldsymbol{\theta}) \cdot e(t)$$
$$e(t) = H^{-1}(z, \boldsymbol{\theta}) \cdot \big[y(t) - G(z, \boldsymbol{\theta}) \cdot u(t)\big]$$

$$\begin{aligned}
y(t) &= G(z, \boldsymbol{\theta}) \cdot u(t) + H(z, \boldsymbol{\theta}) \cdot e(t) \\
&= G(z, \boldsymbol{\theta}) \cdot u(t) + H(z, \boldsymbol{\theta}) \cdot e(t) - e(t) + e(t) && \textit{sommo e tolgo e(t)} \\
&= G(z, \boldsymbol{\theta}) \cdot u(t) + \big[H(z, \boldsymbol{\theta}) - 1\big] \cdot e(t) + e(t) \\
&= G(z, \boldsymbol{\theta}) \cdot u(t) + \big[H(z, \boldsymbol{\theta}) - 1\big] \cdot H^{-1}(z, \boldsymbol{\theta}) \cdot \big[y(t) - G(z, \boldsymbol{\theta}) \cdot u(t)\big] + e(t) \\
&= G(z, \boldsymbol{\theta}) \cdot u(t) + \big[1 - H^{-1}(z, \boldsymbol{\theta})\big] \cdot \big[y(t) - G(z, \boldsymbol{\theta}) \cdot u(t)\big] + e(t) \\
&= G(z, \boldsymbol{\theta}) \cdot u(t) + \big[1 - H^{-1}(z, \boldsymbol{\theta})\big] \cdot \big[y(t) - G(z, \boldsymbol{\theta}) \cdot u(t)\big] + e(t) \\
&= G(z, \boldsymbol{\theta}) \cdot u(t) - G(z, \boldsymbol{\theta}) \cdot u(t) + H^{-1}(z, \boldsymbol{\theta}) \cdot G(z, \boldsymbol{\theta}) \cdot u(t) + \big[1 - H^{-1}(z, \boldsymbol{\theta})\big] \cdot y(t) + e(t) \\
&= H^{-1}(z, \boldsymbol{\theta}) \cdot G(z, \boldsymbol{\theta}) \cdot u(t) + \big[1 - H^{-1}(z, \boldsymbol{\theta})\big] \cdot y(t) + e(t)
\end{aligned}$$

$$\begin{aligned}
\widehat{\mathcal{M}}(\boldsymbol{\theta}) : \quad \hat{y}(t|t-1; \boldsymbol{\theta}) &= H^{-1}(z, \boldsymbol{\theta}) \cdot G(z, \boldsymbol{\theta}) \cdot u(t) + \big[1 - H^{-1}(z, \boldsymbol{\theta})\big] \cdot y(t) \\
\varepsilon_1(t; \boldsymbol{\theta}) &= e(t) = H^{-1}(z, \boldsymbol{\theta}) \cdot \big[y(t) - G(z, \boldsymbol{\theta}) \cdot u(t)\big]
\end{aligned}$$

$$\begin{aligned}
\varepsilon_1(t; \boldsymbol{\theta}) &= H^{-1}(z, \boldsymbol{\theta}) \cdot \big[y(t) - G(z, \boldsymbol{\theta}) \cdot u(t)\big] \\
&= H^{-1}(z, \boldsymbol{\theta}) \cdot \big[G(z, \boldsymbol{\theta}) \cdot u(t) + H(z, \boldsymbol{\theta}) \cdot e(t) - G(z, \boldsymbol{\theta}) \cdot u(t)\big] \\
&= e(t)
\end{aligned}$$

Se $\exists \boldsymbol{\theta} = \boldsymbol{\theta}^0$ tale che $G_0(z) = G(z, \boldsymbol{\theta}^0)$ e $H_0(z) = H(z, \boldsymbol{\theta}^0)$ ovvero, se il ***sistema vero appartiene alla famiglia di modelli scelta****, otteniamo che*

$$\varepsilon_1(t; \boldsymbol{\theta}^0) = e(t)$$

Quindi, il valore $\boldsymbol{\theta}^0$:

- *É l'unico valore che rende $\varepsilon_1(t; \boldsymbol{\theta}^0) = e(t)$*
- *Minimizza la varianza dell'errore di predizione a un passo*

$\Rightarrow$ *$\varepsilon_1(t)$ è un buon indicatore della bontà di un modello dinamico. Useremo questa proprietà per trovare un criterio di identificazione dei modelli dinamici*

Identificazione PEM di modelli ARX

$$\mathcal{M}(\boldsymbol{\theta}): \quad y(t) = \frac{B(z,\boldsymbol{\theta})}{A(z,\boldsymbol{\theta})} \cdot u(t-1) + \frac{1}{A(z,\boldsymbol{\theta})} \cdot e(t) \qquad e(t) \sim WN(0,\lambda^2)$$

$$B(z) = b_0 + b_1 z^{-1} + \cdots + b_{n_b} z^{-n_b}$$
$$A(z) = 1 - a_1 z^{-1} - \cdots - a_{n_a} z^{-n_a}$$

$$\begin{aligned}
\widehat{\mathcal{M}}(\boldsymbol{\theta}): \quad \widehat{y}(t|t-1;\boldsymbol{\theta}) &= H^{-1}(z,\boldsymbol{\theta}) \cdot G(z,\boldsymbol{\theta}) \cdot u(t) + \left[1 - H^{-1}(z,\boldsymbol{\theta})\right] \cdot y(t) \\
&= B(z,\boldsymbol{\theta}) \cdot u(t-1) + \left[1 - A(z,\boldsymbol{\theta})\right] \cdot y(t) \\
&= (b_0 + b_1 z^{-1} + \cdots + b_{n_b} z^{-n_b}) \cdot u(t-1) + (a_1 z^{-1} + \cdots + a_{n_a} z^{-n_a}) \cdot y(t)
\end{aligned}$$

$$\widehat{y}(t|t-1;\boldsymbol{\theta}) = \ b_0 u(t-1) + \cdots + b_{n_b} u(t-n_b-1) + a_1 y(t-1) + \cdots + a_{n_a} u(t-n_a)$$

$$\underset{d\times 1}{\boldsymbol{\theta}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n_a} \\ b_0 \\ b_1 \\ \vdots \\ b_{n_b} \end{bmatrix} \qquad \underset{d\times 1}{\boldsymbol{\varphi}(t)} = \begin{bmatrix} y(t-1) \\ y(t-2) \\ \vdots \\ y(t-n_a) \\ u(t-1) \\ u(t-2) \\ \vdots \\ u(t-n_b-1) \end{bmatrix} \qquad d = n_a + n_b + 1$$

Possiamo scrivere:

$$\mathcal{M}(\boldsymbol{\theta}): \qquad y(t) = \boldsymbol{\varphi}^T(t) \cdot \boldsymbol{\theta} + e(t)$$
$$\widehat{\mathcal{M}}(\boldsymbol{\theta}): \qquad \widehat{y}(t|t-1;\boldsymbol{\theta}) = \boldsymbol{\varphi}^T(t) \cdot \boldsymbol{\theta}$$

Il predittore è lineare nei parametri $\boldsymbol{\theta}$

Per trovare la stima PEM minimizziamo la funzione di costo:

$$J_N(\boldsymbol{\theta}) = \frac{1}{N} \sum_{t=1}^{N} \varepsilon_1(t;\boldsymbol{\theta})^2 = \frac{1}{N} \sum_{t=1}^{N} \Big(y(t) - \widehat{y}(t|t-1;\boldsymbol{\theta})\Big)^2$$

$$J_N(\boldsymbol{\theta}) = \frac{1}{N} \sum_{t=1}^{N} \Big(y(t) - \boldsymbol{\varphi}^T(t) \cdot \boldsymbol{\theta}\Big)^2$$

La soluzione è analoga alla stima
a minimi quadrati di un modello lineare statico!

$$\widehat{\boldsymbol{\theta}}_N = \left[\sum_{t=1}^{N} \boldsymbol{\varphi}(t) \cdot \boldsymbol{\varphi}^T(t)\right]^{-1} \cdot \left[\sum_{t=1}^{N} \boldsymbol{\varphi}(t) \cdot y(t)\right]$$

Se $S(N) = \sum_{t=1}^{N} \boldsymbol{\varphi}(t) \cdot \boldsymbol{\varphi}^T(t)$ *é invertibile* $\Rightarrow$ *allora la soluzione* $\widehat{\boldsymbol{\theta}}_N$ *é* ***unica***
in quanto la funzione di costo é convessa

Come per la regressione lineare, posso esprimere il modello ARX in forma matriciale:

$$\underset{N\times d}{\boldsymbol{\Phi}} = \begin{bmatrix} \boldsymbol{\varphi}^T(1) \\ \boldsymbol{\varphi}^T(2) \\ \vdots \\ \boldsymbol{\varphi}^T(N) \end{bmatrix} \qquad \underset{d\times 1}{\boldsymbol{\theta}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n_a} \\ b_0 \\ b_1 \\ \vdots \\ b_{n_b} \end{bmatrix} \qquad \underset{N\times 1}{\boldsymbol{Y}} = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix} \qquad \underset{N\times 1}{\boldsymbol{E}} = \begin{bmatrix} e(1) \\ e(2) \\ \vdots \\ e(N) \end{bmatrix}$$

$$\boldsymbol{Y} = \boldsymbol{\Phi}\boldsymbol{\theta} + \boldsymbol{E} \qquad \Rightarrow \qquad \widehat{\boldsymbol{\theta}} = \left(\boldsymbol{\Phi}^T\boldsymbol{\Phi}\right)^{-1}\boldsymbol{\Phi}^T\boldsymbol{Y}$$

Identificazione PEM di modelli ARMAX

$$\mathcal{M}(\boldsymbol{\theta}): \quad y(t) = \frac{B(z,\boldsymbol{\theta})}{A(z,\boldsymbol{\theta})} \cdot u(t-1) + \frac{C(z,\boldsymbol{\theta})}{A(z,\boldsymbol{\theta})} \cdot e(t) \qquad e(t) \sim WN(0,\lambda^2)$$

$$B(z) = b_0 + b_1 z^{-1} + \cdots + b_{n_b} z^{-n_b}$$

$$A(z) = 1 - a_1 z^{-1} - \cdots + a_{n_a} z^{-n_a}$$

$$C(z) = 1 + c_1 z^{-1} + \cdots + c_{n_c} z^{-n_c}$$

Il vettore dei parametri, in questo caso, è:

$$\boldsymbol{\theta} = \begin{bmatrix} a_1 \dots a_{n_a} & b_0 b_1 \dots b_{n_b} & c_1 \dots c_{n_c} \end{bmatrix}^T$$

$$n_a + n_b + 1 + n_c \times 1 = d \times 1$$

Utilizzando l'espressione generica:

$$\begin{aligned} \varepsilon_1(t;\boldsymbol{\theta}) &= H^{-1}(z,\boldsymbol{\theta}) \cdot \left[y(t) - G(z,\boldsymbol{\theta}) \cdot u(t)\right] \\ &= \frac{A(z)}{C(z)} \cdot \left[y(t) - \frac{B(z)}{A(z)} \cdot u(t-1)\right] \\ &= \frac{A(z)}{C(z)} y(t) - \frac{B(z)}{C(z)} u(t-1) \end{aligned}$$

Funzione di costo:

$$J_N(\boldsymbol{\theta}) = \frac{1}{N}\sum_{t=1}^{N} \varepsilon_1(t;\boldsymbol{\theta})^2 = \frac{1}{N}\sum_{t=1}^{N}\left[\frac{A(z,\boldsymbol{\theta})}{C(z,\boldsymbol{\theta})} y(t) - \frac{B(z,\boldsymbol{\theta})}{C(z,\boldsymbol{\theta})} u(t-1)\right]^2$$

Osservazioni:

- Dato che ho $C(z,\boldsymbol{\theta})$ al denominatore, non é piú convessa $\Rightarrow$ avró dei **minimi locali**
- Per la risoluzione del problema di minimizzazione, devo utilizzare dei **metodi iterativi (Newton)**

4 Esercitazione 1

4.1 Classificazione

Modelli	F. trasferimento	Note
$MA(n_c)$	$y(t) = C(z) \cdot e(t)$	$\gamma_{yy}(\tau) = 0, \forall \tau > n_c$
$AR(n_a)$	$y(t) = \frac{1}{A(z)} \cdot e(t)$	$\gamma_{yy}^{PAR}(\tau) = 0, \forall \tau > n_a$
$ARMA(n_a, n_c)$	$y(t) = \frac{C(z)}{A(z)} \cdot e(t)$	
$ARX(n_a, n_b, k)$	$y(t) = \frac{B(z)}{A(z)} \cdot u(t-k) + \frac{1}{A(z)} \cdot e(t)$	
$ARMAX(n_a, n_c, n_b, k)$	$y(t) = \frac{B(z)}{A(z)} \cdot u(t-k) + \frac{C(z)}{A(z)} \cdot e(t)$	
$FIR(n_b, k)$	$y(t) = B(z) \cdot u(t-k) + e(t)$	
$OE(n_b, n_f, k)$	$y(t) = \frac{B(z)}{F(z)} \cdot u(t-k) + e(t)$	
$BJ(n_c, n_b, n_d, n_f, k)$	$y(t) = \frac{B(z)}{F(z)} \cdot u(t-k) + \frac{C(z)}{D(z)} \cdot e(t)$	

La componente **stocastica** $H(z)$ dei processi é riassumibile come:

$$AR(n_a \text{ poli}) \rightarrow H(z) = \frac{1}{A(z)}$$
$$MA(n_c \text{ zeri}) \rightarrow H(z) = C(z)$$
$$ARMA(n_a \text{ poli}, n_c \text{ zeri}) \rightarrow H(z) = \frac{C(z)}{A(z)}$$
$$ARX, ARMAX \rightarrow \text{ingresso esegeno noto } u(t)$$

Forma ricorsiva dei modelli:

$$MA(n_c) \rightarrow y(t) = \sum_{i=0}^{n_c} c_i \cdot e(t-i)$$

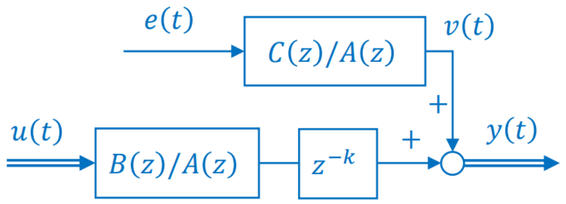
$$AR(n_a) \rightarrow y(t) = \sum_{i=1}^{n_a} a_i \cdot y(t-i) + e(t)$$

$$ARMA(n_a, n_c) \rightarrow y(t) = \sum_{i=1}^{n_a} a_i \cdot y(t-i) + \sum_{i=0}^{n_c} c_i \cdot e(t-i)$$

$$ARX(n_a, n_b, k) \rightarrow y(t) = \sum_{i=1}^{n_a} a_i \cdot y(t-i) + e(t) + \sum_{i=0}^{n_b} b_i \cdot u(t-k-i)$$

$$ARMAX(n_a, n_c, n_b, k) \rightarrow y(t) = \sum_{i=1}^{n_a} a_i \cdot y(t-i) + \sum_{i=0}^{n_b} b_i \cdot u(t-k-i) + \sum_{i=0}^{n_c} c_i \cdot e(t-i)$$

$$FIR(n_b, k) \rightarrow y(t) = \sum_{i=0}^{n_b} b_i \cdot u(t-k-i) + e(t)$$



Classificazione:

- Forma ricorsiva • Schema a blocchi • Spazio di stato / rappresentazione di stato

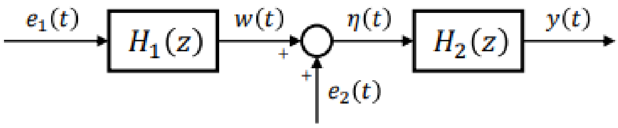
Classificazione forma ricorsiva

$$y(t-k) = z^{-k} \cdot y(t)$$

$$y(t) = \frac{1}{3}y(t-3) + u(t-1) + 3u(t-2) + e(t) + \frac{1}{2}e(t-1)$$

$$y(t) = \frac{1+3z^{-1}}{1-\frac{1}{3}z^{-3}} \cdot u(t-1) + \frac{1+\frac{1}{2}z^{-1}}{1-\frac{1}{3}z^{-3}} \cdot e(t)$$

Classificazione schema a blocchi



$$e_1(t) \sim WN(0,1) \quad e_2(t) \sim WN(1,1) \quad e_1(t) \perp e_2(t)$$

$$H_1(z) = \frac{5z+25}{5z+1} \qquad H_2(z) = \frac{3z+12}{6z^2+5z+1}$$

$$\begin{aligned} w(t) = H_1(z) \cdot e_1(t) &= \frac{5z+25}{5z+1} \cdot e_1(t) \\ &= 5 \cdot \underbrace{\frac{1}{5} \cdot \frac{z+5}{z+\frac{1}{5}}}_{f.\ passa\ tutto} \cdot e_1(t) = 5 \cdot e_1(t) \end{aligned}$$

$$\begin{aligned} \eta(t) &= w(t) + e_2(t) = 5e_1(t) + e_2(t) \\ \mathbb{E}[\eta(t)] = m_\eta &= 5 \cdot \mathbb{E}\big[e_1(t)\big] + \mathbb{E}\big[e_2(t)\big] \\ &= 5 \cdot 0 + 1 = 1 \end{aligned}$$

$$\begin{aligned} \gamma_{\eta\eta}(0) = \mathbb{E}\big[\tilde{\eta}(t)^2\big] &= \mathbb{E}\Big[\big(5e_1(t) + e_2(t)\big)^2\Big] \\ &= 25 \cdot \mathbb{E}\big[e_1(t)^2\big] + \cancel{10 \cdot \mathbb{E}\big[e_1(t) \cdot e_2(t)\big]} + \mathbb{E}\big[e_2(t)^2\big] \\ &= 25 \cdot 1 + 1 = 26 \end{aligned}$$

$$\Rightarrow \qquad ARMA(2,1) \qquad y(t) = \frac{3z+12}{6z^2+5z+1} \cdot \eta(t) \qquad \eta(t) \sim WN(1,26)$$

Spazio di stato / rappresentazione di stato

$$\begin{cases} z \cdot x_1(t) &= 0.8x_1(t) + u(t) + e(t) \\ z \cdot x_2(t) &= x_1(t) + 0.5x_2(t) + 4e(t) \\ y(t) &= x_2(t) \end{cases}$$

$$\begin{cases} (z-0.8) \cdot x_1(t) &= u(t) + e(t) \\ (z-0.5) \cdot x_2(t) &= x_1(t) + 4e(t) \\ y(t) &= x_2(t) \end{cases}$$

$$\begin{cases} x_1(t) &= \frac{1}{z-0.8}u(t) + \frac{1}{z-0.8}e(t) \\ x_2(t) &= \frac{1}{z-0.5}x_1(t) + \frac{4}{z-0.5}e(t) \\ y(t) &= x_2(t) \end{cases}$$

$$\begin{cases} x_1(t) &= \frac{1}{z-0.8}u(t) + \frac{1}{z-0.8}e(t) \\ x_2(t) &= \frac{1}{(z-0.5)(z-0.8)}u(t) + \frac{1}{(z-0.5)(z-0.8)}e(t) + \frac{4}{z-0.5}e(t) \\ y(t) &= x_2(t) \end{cases}$$

$$\begin{cases} x_1(t) &= \frac{1}{z-0.8}u(t) + \frac{1}{z-0.8}e(t) \\ x_2(t) &= \frac{1}{(z-0.5)(z-0.8)}u(t) + \frac{1+4(z-0.8)}{(z-0.5)(z-0.8)}e(t) \\ y(t) &= x_2(t) \end{cases}$$

$$\begin{cases} x_1(t) &= \frac{1}{z-0.8}u(t) + \frac{1}{z-0.8}e(t) \\ x_2(t) &= \frac{1}{(z-0.5)(z-0.8)}u(t) + \frac{4z-2.2}{(z-0.5)(z-0.8)}e(t) \\ y(t) &= x_2(t) \end{cases}$$

$$y(t) = x_2(t) = \frac{1}{1-1.3z^{-1}+0.4z^{-2}}u(t-2) + \frac{4z^{-1}-2.2z^{-2}}{1-1.3z^{-1}+0.4z^{-2}}e(t)$$

4.2 Stazionarietá

- Secondo il teorema dobbiamo verificare che $H(z)$ é asitoticamente stabile e $e(t)$ sia stazionario in senso debole $\rightarrow$ processo $y(t)$ stazionario in senso debole
- Trasformare la funzione da forma ricorsiva a forma dinamica $\rightarrow$ raccolgliendo i segnali e si moltiplicandoli per il ritardo $e(t-1) = e(t) \cdot z^{-1}$
- Ricavare la funzione di trasferimento del processo stocastico con **potenze positive** moltiplicando numeratore e denominatore per la potenza presente piú grande
- controllare che i $|poli| < 1$ di $H(z)$
- $e(t)$ é un processo stazionario per definizione

$$y(t) = \underbrace{\frac{3z^2 - \frac{9}{4}z - \frac{15}{8}}{z^2 - \frac{2}{3}z}}_{H_1(z) = as.\ stabile} \cdot e(t)$$

Esempio:

- $H_1(z)$ é asintoticamente stabile
- $e(t)$ processo stazionario per definizione

$\Rightarrow y(t)$é un processo stazionario

4.3 Media del processo

Proprietá della media:

- Proprietá 1: $\mathbb{E}\big[y(t)\big] = \mathbb{E}\big[e(t) + \frac{1}{2}y(t-1)\big] = \mathbb{E}\big[e(t)\big] + \frac{1}{2}\mathbb{E}\big[y(t-1)\big]$
- Proprietá 2: $\mathbb{E}\big[y(t)\big] = \mathbb{E}\big[10 + \frac{1}{2}y(t-1)\big] = 10 + \frac{1}{2}\mathbb{E}\big[y(t-1)\big]$

Essendo un processo stazionario la media non dipende da t, allora la media per istanti temporali diversi (es: $t+1$, $t-1$) sono sempre uguale a μ

$$m_y = \mathbb{E}\big[y(t)\big] = \mathbb{E}\big[y(t-1)\big] = \mathbb{E}\big[y(t-2)\big] = \dots \qquad \rightarrow \textbf{\textit{esplodo y(t)}}$$

Dato che $y(t)$ é l'uscita a regime del sistema dinamico $H(z)$, é possibile ricavare la media usado il guadagno statico di $H(z)$ (usando forma dinamica)

$$m_y = \mathbb{E}\big[y(t)\big] = \sum_{i=0}^{+\infty} g(i) \cdot \mathbb{E}\big[u(t-i)\big] = H(1) \cdot m_e$$

$$m_y = H(1) \cdot m_e = H(1) \cdot m_u$$

$$m_y = H(1) \cdot m_{ingresso}$$

4.4 Funzione di auto-covarianza

Proprietá:

- Proprietà 1: $Var\big[y(t)\big] = \mathbb{E}\big[\big(y(t) - m_y\big)^2\big]$
- Proprietá 2: $Var\big[y(t-1) + \frac{1}{5}\big] = Var\big[y(t-1)\big] + 0$
- Proprietá 3: $Var\big[b \cdot y(t-1) + e(t)\big] = b^2 \cdot Var\big[y(t-1)\big] + Var\big[e(t)\big]$

NB: *se $m_y \neq 0 \Rightarrow$ e $m_e \neq 0$ depolarizzare, ossia rendere la media = 0 per semplificare i calcoli:*

Usare la forma ricorsiva di $y(t)$ per calcolarla ***esplodere soltato y(t)***

$$\gamma_{yy}(0) = Var\big[y(t)\big] = \mathbb{E}\big[\big(y(t) - m_y\big)^2\big] \overset{m_y=0}{=} \mathbb{E}\big[\big(y(t)\big)^2\big]$$

$$\gamma_{yy}(\tau) = \mathbb{E}\big[\big(y(t) - m_y\big) \cdot \big(y(t-\tau) - m_y\big)\big] \overset{m_y=0}{=} \mathbb{E}\big[y(t) \cdot y(t-\tau)\big]$$

NB: stazionario $\mathbb{E}\big[\tilde{\eta}(t-1) \cdot \tilde{y}(t-1)\big] = \mathbb{E}\big[\tilde{\eta}(t) \cdot \tilde{y}(t)\big] = \mathbb{E}\big[\tilde{\eta}(t) \cdot (\dots)\big] \neq 0$

NB: white noise $\gamma_{ee}(\tau) = \mathbb{E}\big[e(t)e(t-\tau)\big] = 0, \ \forall \tau \neq 0$

4.5 Densitá spettrale di potenza

Proprietá:

- **Reale:** dato che $\gamma_{vv}(\tau)$ é pari, i termini immaginari del tipo $\pm j \cdot sin(\omega)$ si elidono
- **Positiva:** $\Gamma_{vv}(\omega) \geqslant 0, \ \forall \omega \in \mathbb{R}$
- **Pari:** $\Gamma_{vv}(\omega) = \Gamma_{vv}(-\omega), \ \forall \omega \in \mathbb{R}$
- **Periodica di periodo 2π:** $\Gamma_{vv}(\omega) = \Gamma_{vv}(\omega + k \cdot 2\pi), \ \forall \omega \in \mathbb{R}, \forall k \in \mathbb{Z}$
- Ci basta valutare $\Gamma_{vv}(\omega)$ tra $[0, \pi]$

Nota bene:

- $e^{jk\omega} + e^{-jk\omega} = 2cos(k\omega)$
- $\Gamma_{ee}(\omega) = \lambda^2$
- $\left|1 + \frac{1}{2}e^{-j\omega}\right|^2 = \left(1 + \frac{1}{2}e^{-j\omega}\right) \cdot \left(1 + \frac{1}{2}e^{j\omega}\right) = \left(1 + \frac{1}{2}e^{-j\omega} + \frac{1}{2}e^{j\omega} + \frac{1}{4}\right) = \frac{5}{4} + 3\cos(\omega)$
- $\left|\dfrac{2 + 3e^{-j\omega} - 2e^{-2j\omega}}{1 + 4e^{-j\omega}}\right|^2 = \dfrac{|2 + 3e^{-j\omega} - 2e^{-2j\omega}|^2}{|1 + 4e^{-j\omega}|^2}$

Metodo diretto:
Applicabile solo al MA(n_c) perché sappiamo che é simmetrica e quindi dobbiamo partire con τ da $-n_c$ e arrivare a n_c, ma per altri processi esempio AR dobbiamo usare la rappresentazione dinamica.

$$\Gamma_{yy}(\omega) \equiv \mathcal{F}|\gamma_{yy}(\tau)| = \sum_{\tau=-\infty}^{+\infty} \gamma_{yy}(\tau) \cdot e^{-j\omega\tau}$$

Sfruttando la rappresentazione dinamica:
*con **esponenti negativi***

$$\Gamma_{yy}(\omega) = |H(e^{j\omega})|^2 \cdot \Gamma_{uu}(\omega)$$
$$\Gamma_{uu}(\omega) = \Gamma_{ee}(\omega) \sim WN = \lambda_{ee}^2$$

***NB:** per disegnare la densitá di potenza é una funzione pari e simmetrica, la valutiamo in $\{0, \frac{\pi}{2}, \pi\}$*

Piú segnali in ingresso dove $y(t) = w_1(t) + w_2(t)$

$$\begin{aligned}\Gamma_{yy}(\omega) = \Gamma_{w_1w_1}(\omega) + \Gamma_{w_2w_2}(\omega) &= |W_1(e^{j\omega})|^2 \cdot \Gamma_{e_1e_1}(\omega) + |W_2(e^{j\omega})|^2 \cdot \Gamma_{e_2e_2}(\omega) \\ &= \sum_{\tau=-\infty}^{+\infty} \gamma_{w_1w_1}(\omega) \cdot e^{-j\omega\tau} + \sum_{\tau=-\infty}^{+\infty} \gamma_{w_2w_2}(\omega) \cdot e^{-j\omega\tau}\end{aligned}$$

4.6 Depolarizzazione:

Effettuare la depolarizzazione di $y(t)$ e $e(t)$ prima del calcolo della autocovarianza $\gamma_{yy}(\tau)$ del processo per semplificare i calcoli.

$$y(t) = e(t-1) + 2e(t-2) + e(t-4) + 2e(t-5) + 1 \qquad e(t) \sim WN(1,1)$$

$$\begin{cases} \tilde{y} = y(t) - m_y = y(t) - 7 \\ \tilde{e}(t) = e(t) - m_e = e(t) - 1 \end{cases} \qquad \begin{cases} y(t) = \tilde{y} + 7 \\ e(t) = \tilde{e}(t) + 1 \end{cases}$$
$$\Rightarrow \qquad 2e(t) = 2\big(\tilde{e}(t) + 1\big)$$

$$\tilde{y}(t) + 7 = \tilde{e}(t-1) + 1 + 2\tilde{e}(t-2) + 2 + \tilde{e}(t-4) + 1 + 2\tilde{e}(t-5) + 2 + 1$$
$$\tilde{y}(t) = \tilde{e}(t-1) + 2\tilde{e}(t) + \tilde{e}(t-4) + 2\tilde{e}(t-5) + (-7 + 1 + 2 + 1 + 2 + 1)$$

$$\tilde{y}(t) = \tilde{e}(t-1) + 2\tilde{e}(t) + \tilde{e}(t-4) + 2\tilde{e}(t-5) \qquad \tilde{e}(t) \sim WN(0,1)$$

$$\gamma_{\tilde{y}\tilde{y}}(0) = \mathbb{E}\big[\tilde{y}(t)^2\big] = \mathbb{E}\Big[\big(\tilde{e}(t-1) + 2\tilde{e}(t) + \tilde{e}(t-4) + 2\tilde{e}(t-5)\big)^2\Big] = 1 + 4 + 1 + 4 = 10$$

5 Esercitazione 2

5.1 Forma canonica

- Verificare che $H(z) = \frac{C(z)}{A(z)}$

A(z) e C(z): monici, stesso grado, radici interne, coprimi
coprimi e radici interne ⇒ scomposizione esponenti pos.
monici ⇒ raccogliere esponente piú grande esponenti neg.

- Ricavare la forma dinamica se l'equazione é espresa in forma ricorsiva
- Si scompone l'equazione dinamica ***della componente stocastica*** $H(z)$ per vedere se ci sono semplificazioni → si usano **le potenze positive**, moltiplico per esponente piú grande
- Se possediamo zeri/poli fuori dal cerchio unitario ⇒ applicare **filtro passa-tutto**

Filtro passa tutto:

$$T(z) = \frac{1}{a} \cdot \frac{z+a}{z+\frac{1}{a}} \qquad a \neq 0, a \in \mathbb{R}$$

Si ha un zero fuori dal cerchio unitario:

$$y(t) = \frac{3 \cdot \cancel{\left(z - \frac{5}{4}\right)} \cdot \left(z + \frac{1}{2}\right)}{z \cdot \left(z - \frac{2}{3}\right)} \cdot e(t) \cdot -\frac{5}{4} \frac{\left(z - \frac{4}{5}\right)}{\cancel{\left(z - \frac{5}{4}\right)}} = \frac{\left(z + \frac{1}{2}\right) \cdot \left(z - \frac{4}{5}\right)}{z \cdot \left(z - \frac{2}{3}\right)} \cdot \underbrace{\left(- \frac{15}{4} \cdot e(t) \right)}_{\eta(t)}$$

Bisogna ora sistemare la componente stocastica:

$$e(t) \sim WN(0, 2)$$

$$e(t) \sim WN\left(0 \cdot \left(-\frac{15}{4}\right), 2 \cdot \left(-\frac{15}{4}\right)^2\right)$$

$$\eta(t) \sim WN\left(0, \frac{225}{8}\right)$$

- Si raccoglie l'esponente piú grande dell'equazione ⇒ e si porta tutto in con **potenze negative**
- Quando numeratore e denomitore possiedono gradi diversi ⇒ aggiungere fattore di **ritardio e anticipo (entrambi perché l'equazione non deve cambiare matematicamente)** $\{z^k, z^{-k}\}$ in modo tale da avere lo stesso grado si al numeratore che al denominatore, il fattore rimanente viene aggiunto al segnale $u(t), e(t), \eta(t)$ che moltiplica la Fdt

- C(z) non é monica
- C(z) e A(z) non hanno lo stesso grado

$$\begin{aligned} H(z) \cdot e(t) = \frac{C(z)}{A(z)} &= \frac{4z^{-1} - 2.2z^{-2}}{1 - 1.3z^{-1} + 0.4z^{-2}} \cdot e(t) \\ &= \frac{4z^{-1} - 2.2z^{-2}}{1 - 1.3z^{-1} + 0.4z^{-2}} \cdot z \cdot z^{-1} \cdot \frac{4}{4} \cdot e(t) \\ &= \frac{1 - 0.55z^{-1}}{1 - 1.3z^{-1} + 0.4z^{-2}} \cdot 4e(t-1) \\ &= \frac{1 - 0.55z^{-1}}{1 - 1.3z^{-1} + 0.4z^{-2}} \cdot \eta(t) \end{aligned}$$

NB: $\quad 4z^{-1} - 2.2z^{-2} = z^{-1}(4z^{-2} - 2.2z^{-3})$

$$\eta(t) \sim WN(0, 1 \cdot 4^2)$$

- se ho una costante aggiungerla al segnale di uscita $\eta(t)$ o $e(t)$ che modificherá la media e la varianza del WN creando **un nuovo segnale**. **Fare questo passaggio sono quando ho giá la forma canonica.**

$$e(t) \sim WN(1, 1)$$

$$5 \cdot e(t) \sim WN(1 \cdot 5, 1 \cdot 5^2) = WN(5, 25)$$

$$\eta(t) \sim WN(5, 25)$$

5.2 Predittore ad un passo

- $H(z)$ proviene dalla forma canonica, verificare prima i 4 punti!!
- Dato un processo stocastico generale $y(t)$ dove:

$$y(t) = \ldots + H(z)\cdot d(t) \qquad d(t) \sim WN(m_d, \lambda^2)$$

- Se $\mathbb{E}\big[\varepsilon_1(t)\big] = \mathbb{E}\big[E(z)\cdot d(t)\big] \neq 0$, il predittore non é corretto e si deve trasformare il processo in un ARMAX, aggiungendo in ingresso un segnale $u(t)$ costante che mi porti la media a 0

$$y(t) = H(z)\cdot d(t) \qquad d(t) \sim WN(2, 104) \quad \Rightarrow \quad m_d = 2 \neq 0$$

Aggiungo in ingresso un segnale $u(t)$ che é una costante, dove
$u(t) = u(t-1) = u(t-2) = \ldots = m_d = 2$

$$y(t) = H(z)\cdot \Big(\widetilde{d}(t) + u(t)\Big)$$

Sostituire solo alla fine u(t) con n

$$\widetilde{d}(t) = d(t) - m_d = d(t) - u(t) = d(t) - 2 \quad \Rightarrow \quad \widetilde{d}(t) \sim WN(0, 104)$$
$$d(t) = \widetilde{d}(t) + u(t)$$

$$\begin{aligned} ARMAX = y(t) &= H(z)\cdot u(t) + H(z)\cdot \widetilde{d}(t) \\ &= \frac{B(z)}{A(z)}\cdot u(t) + \frac{C(z)}{A(z)}\cdot \widetilde{d}(t) \qquad B(z) = C(z) \end{aligned}$$

- Effttuo k colpi di lunga divisione sulla **componente stocastica**:

$H(z)$ é la stessa della forma canonica → potenze negative

$$H(z) = \frac{C(z)}{A(z)} \rightarrow \underbrace{C(z)/A(z)}_{divido}$$

Dopo aver effettuato la lunga divisione H(z) diventa:

$$\begin{aligned} H(z) &= E(z) + \frac{R(z)}{A(z)} = E(z) + \frac{z^{-k}\widetilde{R}(z)}{A(z)} \\ y(t) &= \ldots + H(z)\cdot d(t) \\ &= \ldots + \frac{\widetilde{R}(z)}{A(z)}\cdot d(t-k) + E(z)\cdot d(t) \end{aligned}$$

Predittore a k passi: dal rumore:

$$\widehat{y}(t|t-k) = \ldots + \frac{\widetilde{R}(z)}{A(z)}\cdot d(t-k)$$

Predittore a k passi: dai dati + applicando filtro sbiancante:

$$y(t) = \ldots + H(z)\cdot d(t) \quad \Rightarrow \quad d(t) = \ldots$$
$$\widehat{y}(t|t-k) = \ldots + \frac{\widetilde{R}(z)}{C(z)}\cdot y(t-k)$$

Errore di predizione a k passi:

$$\varepsilon_k(t) = y(t) - \widehat{y}(t|t-k) = E(z)\cdot e(t)$$

Predizione a k = 1

$$R(z) = C(z) - A(z)$$
$$E(z) = 1$$
$$\widehat{y}(t|t-1) = \ldots + \frac{\widetilde{R}(z)}{C(z)}\cdot y(t-k) = \ldots + \frac{C(z)-A(z)}{C(z)}\cdot y(t)$$
$$\varepsilon_1(t) = E(z)\cdot e(t) = e(t)$$

5.3 Correttezza del predittore

*Il **predittore lineare ottimo** dai dati é quello che minimizza il seguente criterio **Mean Squared Error (MSE):***

$$var\big[\varepsilon_k(t)\big] = \mathbb{E}\big[\varepsilon_k(t)^2\big] = \mathbb{E}\Big[\big(y(t) - \hat{y}(t|t-k)\big)^2\Big]$$

Affiché il predittore sia ottimo occorre che:

- $\mathbb{E}\big[\varepsilon_k(t)\big] = \mathbb{E}\big[y(t) - \hat{y}(t|t-k)\big] = 0 \Rightarrow$ sia **Corretto**, ossia valore atteso nullo.
- $\mathbb{E}\big[\hat{y}(t|t-k) \cdot \varepsilon_k(t)\big] = 0 \Rightarrow$ il predittore e l'errore di predizione sono **incorrelati**
- $Var\big[\varepsilon_k(t)\big] = \mathbb{E}\big[\varepsilon_k(t)^2\big] = \mathbb{E}\Big[\big(y(t) - \hat{y}(t|t-k)\big)^2\Big]$ minima (la varianza dell'errore del preditttore deve essere uguale o minore alla varianza del processo)

$$\Rightarrow \quad \mathbb{E}\big[\varepsilon_k(t)\big] = \mathbb{E}\big[y(t) - \hat{y}(t|t-k)\big] = \big[E(z) \cdot e(t)\big]$$

***NB:** La media del preditttore deve essere uguale alla media del processo (per ogni preditttore per ogni processo)*

$$\mathbb{E}\big[y(t)\big] \equiv \mathbb{E}\big[y(t|t-k)\big]$$

5.4 Varianza dell'errore di predizione

- Il rumore a istanti di tempo diversi é incorrelato
- **NB:** la varianza dell'errore del preditttore deve essere uguale o minore alla varianza del processo (per ogni preditttore per ogni processo), prendere il processo iniziale e depolarizzare per il calcolo della $\gamma_{yy}(0)$ come in ESE I.

$$Var\big[\varepsilon_k(t)\big] = Var\big[y(t) - \widehat{y}(t|t-k)\big] = Var\big[E(z) \cdot e(t)\big] \leqslant Var\big[y(t)\big]$$

$$\mathbb{E}\big[(\varepsilon_1(t) - m_{\varepsilon_1})^2\big] = \mathbb{E}\big[\varepsilon_1(t)^2\big] = \mathbb{E}\Big[\big(y(t) - \hat{y}(t|t-1)\big)^2\Big]$$

5.5 Calcolare l'ESR del preditttore

Prima calcoliamo la varianza del processo: dove $\widetilde{y}(t)$ é il processo depolarizzato

$$Var\big[y(t)\big] = \mathbb{E}\big[\widetilde{y}(t)^2\big] = \mathbb{E}\Bigg[\bigg(\frac{C(z)}{A(z)} \cdot e(t)\bigg)^2\Bigg] \quad \Rightarrow \quad ESR = \frac{Var\big[\varepsilon_k(t)\big]}{Var\big[y(t)\big]}$$

La percentuale di varianza del processo che é stata catturata dal preditttore é definita come:

$$1 - ESR$$

5.6 Passaggi

1. Classificazione: media - varianza $\rightarrow$ depolarizzazione
2. Calcolo asintotica stabilitá di $H(z)$ $|poli| < 1 \Rightarrow y(t)$ stazionario in senso debole

   ***NB:** filtro passa tutto*

3. Calcolo forma canonica $H(z) = C(z)/A(z)$ e devono rispettare le 4 proprietá

   monici, coprimi, all'interno cerchio unitario, stesso grado

   ***NB:** filtro passa tutto (sistemo $(\eta(t)$ - $(z^k \cdot z^{-k})$*

4. Calcolo del preditttore $\widehat{y}(t|t-k)$ e $\varepsilon_k(t)$ deve essere ottimo:

   valore atteso 0, incorrelato, varianza minima

   ***NB:** ARMAX media $\eta(t) \neq 0$*

5.7 Ricavare il valore della predizione $\widehat{y}(8|7)$

- Possediamo delle informazioni dal testo

$$y(1) = 0.06 \quad y(2) = 0.73 \quad y(3) = -0.63 \quad y(4) = -3.2 \quad y(5) = 0.8 \quad y(6) = 2.92 \quad y(7) = -1.48$$

- Portiamo in forma ricorsiva il predittore che dipende dai **dati**

$$y(t) = -0.3y(t-2) + e(t) - 2e(t-2) \qquad e(t) \sim Wn(0,1)$$

portiamo in forma canonica e applichiamo formule/lunga divisione

$$\widehat{y}(t|t-1) = 0.8y(t-2) + \frac{1}{2}\widehat{y}(t-2|t-3)$$

- Dato che è un'equazione ricorsiva è necessario avere una condizione iniziale. Solitamente si pone la media del processo

$$t = 2 \qquad \rightarrow \qquad \widehat{y}(2|1) = 0.8 \cdot y(0) + \frac{1}{2} \cdot \widehat{y}(0|1)$$

Media del processo stocastico:

$$\widehat{y}(0|-1) = m_y = 0$$

se non si sa la media vera si può usare
lo stimatore campionario, ma in questo caso la media è nota

$$\widehat{y}(0|-1) = \frac{1}{8} \cdot \sum_{t=1}^{8} \frac{y(1) + y(2) + y(3) + \ldots + y(8)}{8}$$

Se non possediamo informazioni ulteriori sui dati possiamo supporre che:

$$y(t) = 0, \forall t \leqslant 0$$

2 metodi: tabella o a mano:

t	$y(t-2)$	$\widehat{y}(t-2|t-3)$	$\widehat{y}(t|t-1)$

$$\widehat{y}(8|7) = -0.8 \cdot y(6) + 0.5 \cdot \widehat{y}(6|5) = -1.2$$
$$\widehat{y}(6|5) = -0.8 \cdot y(4) + 0.5 \cdot \widehat{y}(4|3) = 2.2680$$
$$\widehat{y}(4|3) = -0.8 \cdot y(2) + 0.5 \cdot \widehat{y}(2|1) = -0.5840$$
$$\widehat{y}(2|1) = -0.8 \cdot y(0) + 0.5 \cdot \widehat{y}(0|-1) = m_y = 0$$

Nota bene

- L'ingresso influenza l'uscita $(t-2)$ non avrebbe senso studiare il predittore a un passo bensi partire da $\varepsilon_2(t)$ in poi

$$y(t) = G(z) \cdot u(t-2) + H(z) \cdot e(t)$$

- Se il mio predittore a un passo é $\hat{y}(t|t-1) = 4y(t-3)$ allora $\hat{y}(t|t-1) = \hat{y}(t|t-2) = \hat{y}(t|t-3)$, perché uso le informazioni dei miei dati fino al tempo $t-3$ per predirre al tempo t.

- Quando calcolo la funzione di autocovarianza/varianza del processo, posso considerare tutto il processo depolarizzato. Perché la media del processo non modifica la funzione di autocovarianza.

$$\eta(t) = 5e_1(t) + e_2(t) \qquad e_1(t) \sim Wn(0,1) \qquad e_2(t) \sim Wn(1,1)$$

$$m_\eta = \mathbb{E}\big[\eta(t)\big] = \mathbb{E}\big[5e_1(t) + e_2(t)\big] = 1 \neq 0$$

$$\gamma_{\eta\eta}(0) = \mathbb{E}\Big[\big(\eta(t) - m_\eta\big)^2\Big] = \mathbb{E}\big[\eta(t)\big] = \mathbb{E}\big[25e_1(t) + e_2(t)^2 + \cancel{10e_1(t)e_2(t)}\big] = 25 + 1 = 26$$

$$\eta(t) \sim WN(1, 26)$$

6 Esercitazione 3

Calcolo della funzione di costo

- N osservazioni del processo stocastico stazionario $\mathcal{S}$

$$J_N(\boldsymbol{\theta}) = \frac{1}{N} \cdot \sum_{t=1}^{N} \varepsilon_1(t;\boldsymbol{\theta})^2 = \frac{1}{N} \cdot \sum_{t=1}^{N} \big(y(t) - \widehat{y}(t|t-1;\boldsymbol{\theta})\big)^2$$

- Osservazioni infinite del processo stocastico stazionario $\mathcal{S}$

Ipotesi: *i dati osservati sono corrispondenti ad una unica realizzazione s, ma noi dobbiamo studiare il processo in generale, per cui **grazie all'ipotesi di ergodicitá** per $N \to \infty$ la funzione di costo converge ad unica curva deterministica $\bar{J}(\boldsymbol{\theta})$*

$$J_N(\boldsymbol{\theta}) = \frac{1}{N} \cdot \sum_{t=1}^{N} \varepsilon_1(t;\boldsymbol{\theta},s)^2 \quad \underset{N\to+\infty}{\longrightarrow} \quad \bar{J}(\boldsymbol{\theta}) = \mathbb{E}_s\big[\varepsilon_1(t;\boldsymbol{\theta})^2\big]$$

$$J(\boldsymbol{\theta}) = \mathbb{E}_s\big[\varepsilon_1(t;\boldsymbol{\theta})^2\big] = \mathbb{E}\Big[\big(y(t) - \widehat{y}(t|t-1;\boldsymbol{\theta})\big)^2\Big]$$

- Le misure $y(t)$ sono campionate dal processo $\mathcal{S}$ mentre il predittore é quello calcolato con il modello che si vuole identificare

- Quando $\mathcal{S} \in \mathcal{M}$, ossia il modello usato per generare i dati é lo stesso della famiglia di modelli scelta $\mathcal{M}$, bisogna utlizzare la forma canonica e cercare di far risultare che $\varepsilon_1(t) = e(t)$ ossia che abbia la varianza minima.

$$\mathcal{S}: \qquad y(t) = e(t) \qquad e(t) \sim Wn(0,1)$$

$$\mathcal{M}: \qquad y(t) = -ay(t-1) + \eta(t) + \frac{1}{2} \cdot \eta(t-1) \qquad \eta(t) \sim Wn(0,\lambda^2)$$

Forma canonica: predittore deve dipendere dai dati y(t)

$$y(t) = \frac{1 + \frac{1}{2}z^{-1}}{1 + az^{-1}}\eta(t)$$

$$C(z)/A(z) \quad \to \quad \widehat{y}(t|t-1;a) = \frac{(\frac{1}{2} - a) \cdot z^{-1}}{1 + \frac{1}{2}z^{-1}} \cdot y(t)$$

Calcolo errore di predizione ad un passo:

$$\begin{aligned}
\varepsilon_1(t;a) &= y(t) - \widehat{y}(t|t-1) \\
&= y(t) - \frac{(\frac{1}{2} - a) \cdot z^{-1}}{1 + \frac{1}{2}z^{-1}} \cdot y(t) \\
&= \left(1 - \frac{(\frac{1}{2} - a) \cdot z^{-1}}{1 + \frac{1}{2}z^{-1}}\right) \cdot y(t) \\
&= \frac{1 + az^{-1}}{1 + \frac{1}{2}z^{-1}} \cdot y(t) \qquad y(t) = e(t) \to \mathcal{S} \\
&= \frac{1 + az^{-1}}{1 + \frac{1}{2}z^{-1}} \cdot e(t)
\end{aligned}$$

Si noti che la varianza di $\varepsilon_1(t;a)$ é minima quando $\widehat{a}$ é tale che:

$$\varepsilon_1(t;a) = e(t) \qquad \widehat{a} = \frac{1}{2}$$

$$\widehat{\lambda}^2 = Var\big[\varepsilon_1(t;\widehat{a})\big] = Var\big[e(t)\big]$$

Predittore

- Il predittore dev'essere esplicitato in funzione dei dati, ossia y(t)

$$\mathcal{M}: \qquad y(t) = \eta(t) + a \cdot \eta(t-1)$$

- Applicare la forma canonica → $C(z)/A(z)$

$$y(t) = (1 + a \cdot z^{-1}) \cdot \eta(t)$$

$$\widehat{y}(t|t-1;a) = \frac{a \cdot z^{-1}}{1 + a \cdot z^{-1}} \cdot y(t)$$

- Ci sono due modi per procedere:

$y(1) = -1 \;\; y(2) = 2 \;\; y(3) = 0 \;\; y(4) = 1$

$$J(a) = \frac{1}{4} \cdot \sum_{t=1}^{4} \big(\varepsilon_1(t;a)\big)^2$$

$$\begin{aligned}\varepsilon_1(t;a) &= y(t) - \widehat{y}(t|t-1;a)\\ &= y(t) - \frac{a \cdot z^{-1}}{1 + a \cdot z^{-1}} \cdot y(t)\\ &= \frac{1}{1 + a \cdot z^{-1}} \cdot y(t)\\ &= y(t) - a \cdot \varepsilon_1(t-1;a)\end{aligned}$$

Ricorsivo → inizializzare

$$\varepsilon_1(t;a) = y(t) - a \cdot \underbrace{\varepsilon_1(t-1;a)}$$

1) Media processo

$$\widehat{y}(1|0) = m_y = \frac{1}{4} \sum_{t=1}^{4} y(t)$$

2) Si inizializza il predittore al primo instante noto

$$\widehat{y}(1|0) = y(1)$$

$$\begin{aligned}\Rightarrow \varepsilon_1(1;a) &= y(1) - \widehat{y}(1|0) = 0\\ \varepsilon_1(2;a) &= y(2) - a \cdot \varepsilon_1(1;a) = 2\\ \varepsilon_1(3;a) &= y(3) - a \cdot \varepsilon_1(2;a) = -2a\\ \varepsilon_1(4;a) &= y(4) - a \cdot \varepsilon_1(3;a) = 1 + 2a^2\end{aligned}$$

Tolgo 1 perché l'ho usato per inizializzare il tutto quindi parto da t = 2

$$J(a) = \frac{1}{3} \sum_{t=2}^{4} \varepsilon_1(t)^2 = \frac{1}{3}\left(4 + 4a^2 + 1 + 4a^2 + 4a^4\right)$$

$y(1) = 2 \;\; y(2) = 0 \;\; y(3) = 2$

$$J(a) = \frac{1}{3} \cdot \sum_{t=1}^{3} \big(\varepsilon_1(t;a)\big)^2$$

Esplicito il predittore

$$\begin{aligned}\widehat{y}(t|t-1;a) &= \frac{a \cdot z^{-1}}{1 + a \cdot z^{-1}} \cdot y(t)\\ &= a \cdot y(t-1) - a \cdot \widehat{y}(t-1|t-2;a)\end{aligned}$$

Ricorsivo → inizializzare

$$\widehat{y}(t|t-1;a) = a \cdot y(t-1) - a \cdot \widehat{y}(t-1|t-2;a)$$

Imposto y(0) = 0

$$\widehat{y}(t|t-1;a) = a \cdot y(t-1) - a \cdot \widehat{y}(t-1|t-2;a)$$

$\Rightarrow t = 0$	$\widehat{y}(0\|-1;a) = y(-1) + a0 = 0$
$t = 1$	$\widehat{y}(1\|0;a) = ay(0) - a0 = m_y = 0$
$t = 2$	$\widehat{y}(2\|1;a) = ay(1) - a0 = 2a$
$t = 3$	$\widehat{y}(3\|2;a) = ay(2) - a(2a) = -2a^2$

6.1 Risoluzione definitiva

- Calcolo il predittore $\widehat{y}(t|t-1)$

- Il predittore dev'essere in funzione dei dati y(t) nel caso
non lo fosse applico la risoluzione della forma canonica $C(z)/A(z)$
- se $\mathcal{M} \equiv \mathcal{S}$ applicare la teoria dell'identificazione PEM

- La funzione di costo va a minimizzare la varianza dell'errore di predizione ad un passo $\varepsilon_1(t)$, puó succedere che quest'ultimo dipenda da $\boldsymbol{\theta}$

$$\varepsilon_1(t; \boldsymbol{\theta} = a) = y(t) - a \cdot \varepsilon_1(t-1; a)$$

1) Si inizializza il predittore ponendolo uguale al predittore banale

$$\widehat{y}(1|0) = m_y = \frac{1}{N} \cdot \sum_{t=1}^{N} y(t)$$

2) Si inizializza il predittore al primo instante noto

$$\widehat{y}(1|0) = y(1)$$

$$\Rightarrow \varepsilon_1(1, a) = y(1) - \widehat{y}(1|0)$$

- Calcolo la funzione di costo $J(\boldsymbol{\theta})$

Funzione di costo ha infiniti dati osservati: dipendenza da $\gamma_{yy}(\tau)$

$$J(\boldsymbol{\theta}) = var\big[\varepsilon_1(t; \boldsymbol{\theta})\big] = \mathbb{E}_s\big[\varepsilon_1(t; \boldsymbol{\theta})^2\big] = \mathbb{E}\Big[\big(y(t) - \widehat{y}(t|t-1; \boldsymbol{\theta})\big)^2\Big]$$

Funzione di costo ha finiti dati osservati: dipendenza dai dati oservati $y(i)$

$$J_N(\boldsymbol{\theta}) = \frac{1}{N} \cdot \sum_{t=1}^{N} \varepsilon_1(t; \boldsymbol{\theta})^2 = \frac{1}{N} \cdot \sum_{t=1}^{N} \big(y(t) - \widehat{y}(t|t-1; \boldsymbol{\theta})\big)^2$$

La funzione di costo puó dipendere da:
- funzione di autocovarianza $\gamma_{yy}(k) \rightarrow$ vanno calcolate su $\mathcal{S}$
- da dati finiti osservati $y(1), y(t), \ldots, y(n) \rightarrow$ somma troncata
- NB: nella funzione di costo non esplodo $y(t)$ bensí solo $\widehat{y}(t|t-1)$

- Calcolo il minimo $\boldsymbol{\theta}$, ossia il $\boldsymbol{\theta}^0$ che minimizza $J(\boldsymbol{\theta})$

- Metodo diretto: tutti i processi possono applicarlo

$$\frac{d}{d\boldsymbol{\theta}} J(\boldsymbol{\theta}) \quad \rightarrow \quad \frac{d}{d\boldsymbol{\theta}} J(\boldsymbol{\theta})\bigg|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} = 0 \quad \Rightarrow \quad \widehat{\boldsymbol{\theta}} \textit{ sistemi} \quad \rightarrow \quad \frac{d^2}{d\boldsymbol{\theta}^2} J(\widehat{\boldsymbol{\theta}})\bigg|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} > 0$$

- Metodo lineare: solo AR e ARX

$$\boldsymbol{\Phi} = \textit{parte predittoria} \qquad \boldsymbol{Y} = \textit{parte osservata } y(t)$$

$$\boldsymbol{\Phi}^T \cdot \boldsymbol{\Phi} \cdot \widehat{\boldsymbol{\theta}} = \boldsymbol{\Phi}^T \cdot \boldsymbol{Y}$$

- Stimo $\widehat{\lambda^2}$ di $\mathcal{M}$

$$\widehat{\lambda^2} = J(\boldsymbol{\theta})$$

*Nel caso in cui $\widehat{\lambda^2} = 0 \rightarrow$ **overfitting o sistema deterministico***
Questo avviene tutte le volte in cui il numero di dati è uguale
al numero di parametri ed è una situazione da evitare
***Deterministico** $\rightarrow$ no componente stocastica in $y(t) = y(t-1)$*

- Scrivo il modello finale in funzione dei paramtri $\boldsymbol{\theta}^0$ e di $\widehat{\lambda^2}$ esempio:

$$\Rightarrow \qquad y(t) = \widehat{a} \cdot y(t-1) + \eta(t) \qquad \eta(t) \sim WN(0, \widehat{\lambda^2})$$

- Controllo se il processo del modello $\widehat{\mathcal{M}}$ é stazionario, calcolo i $|poli| < 1$

Modelli AR

$$\mathcal{S}: \quad y(1) = -1 \quad y(2) = 2 \quad y(3) = 0 \quad y(4) = 1$$

$$M_1(a): \qquad y(t) = a \cdot y(t-1) + \eta(t) \qquad \eta(t) \sim WN(0, \lambda^2)$$

$$\widehat{y}(t|t-1; a) = a \cdot y(t-1)$$

Calcolo la funzione di costo:

$$J_N(\boldsymbol{\theta}) = \frac{1}{N} \cdot \sum_{t=1}^{N} \varepsilon_1(t; \boldsymbol{\theta})^2 = \frac{1}{N} \cdot \sum_{t=1}^{N} \big(y(t) - \widehat{y}(t|t-1; \boldsymbol{\theta})\big)^2$$

$$J_N(a) = \frac{1}{4} \cdot \sum_{t=1}^{4} \big(y(t) - \widehat{y}(t|t-1; a)\big)^2 = \frac{1}{4} \cdot \sum_{t=1}^{4} \big(y(t) - a \cdot y(t-1)\big)^2$$

Il dataset che abbiamo parte da $y(1)$ e sostituendo questo nella formula ci mancherebbe $y(0)$ dunque dato che non conosciamo $y(0)$ ***é necessario troncare la somma***. $N-1$

$$\textit{iterazione 1:} \quad J_N(a) = \big(y(1) - y(0)\big)^2 + \dots$$

$$J_N(a) = \frac{1}{3} \cdot \sum_{t=2}^{4} \big(y(t) - a \cdot y(t-1)\big)^2$$

Adesso sostituiamo i dati della tabella nella formula di costo

$$\begin{aligned} J_N(a) &= \frac{1}{3} \cdot \sum_{t=2}^{4} \big(y(t) - a \cdot y(t-1)\big)^2 \\ &= \frac{1}{3} \cdot \left[\Big(2 - a \cdot (-1)\Big)^2 + \Big(0 - a \cdot 2\Big)^2 + \Big(1 - a \cdot (0)\Big)^2 \right] \\ &= \frac{5a^2 + 4a + 5}{3} \end{aligned}$$

I Metodo diretto

$$\frac{d}{d\boldsymbol{\theta}} J(\boldsymbol{\theta}) \quad \rightarrow \quad \frac{d}{d\boldsymbol{\theta}} J(\boldsymbol{\theta}) \Big|_{\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}} = 0 \quad \Rightarrow \quad \widehat{\boldsymbol{\theta}} \textit{ sistemi}$$

- calcolo il gradiente per ogni componente
- e faccio un sistema dove le pongo uguali a 0

$$\boldsymbol{\theta} = \big[a, b\big] \quad \rightarrow \quad \frac{d}{da} J(\boldsymbol{\theta}) \ \ \frac{d}{db} J(\boldsymbol{\theta}) \quad \Rightarrow$$

$$\begin{cases} \frac{d}{d\widehat{a}} J(\boldsymbol{\theta}) = 0 \\ \frac{d}{d\widehat{b}} J(\boldsymbol{\theta}) = 0 \end{cases} \Rightarrow \begin{cases} \widehat{a} \\ \widehat{b} \end{cases}$$

II Metodo dei minimi quadrati

$\boldsymbol{\Phi}$ *salva su ogni riga i valori assunti dalla parte predittoria ad ogni iterazione $J_N(\boldsymbol{\theta})$ utilizzando il dataset di partenza*

$\boldsymbol{Y}$ *ad ogni iterazione di $J_N(\boldsymbol{\theta})$ salva il valore associato a $y(t)$*

$$\boldsymbol{\Phi}^T \cdot \boldsymbol{\Phi} \cdot \widehat{\boldsymbol{\theta}} = \boldsymbol{\Phi}^T \cdot \boldsymbol{Y}$$

$$\boldsymbol{\Phi} = \begin{bmatrix} \widehat{y}(1) \\ \widehat{y}(2) \\ \widehat{y}(3) \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} \qquad \boldsymbol{Y} = \begin{bmatrix} y(2) \\ y(3) \\ y(4) \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \qquad \widehat{\boldsymbol{\theta}} = \begin{bmatrix} \widehat{a} \end{bmatrix}$$

$$\begin{bmatrix} -1 & 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} \cdot \hat{a} = \begin{bmatrix} -1 & 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

Osservazioni:

- Calcolo la funzione di costo: puó essere necesario **troncare la somma**
- se $J(\boldsymbol{\theta}) = Var\big[\varepsilon_1(k)\big] > 0$ sempre positiva
- $J(\boldsymbol{\theta})$ é lineare nei parametri, perció $J(\boldsymbol{\theta})$ sará sempre un paraboloide oppure $J(\boldsymbol{\theta} = a)$ se un solo parametro una parabola $\Rightarrow$ ha un **minimo unico**

Modelli MA

- Non possibile usare le formule dei minimi quadrati $\rightarrow$ solo metodo diretto
- Calcolo la forma canonica di $\mathcal{M}$ $\rightarrow$ voglio il **predittore in funzione dei dati y(t)**
- Effettuare la lunga divisione per trovare il predittore ad un passo dipendente dai dati

$$\widehat{y}(t|t-1;\boldsymbol{\theta}) = \frac{R(z)}{C(Z)} \cdot y(t) = \frac{C(z) - A(z)}{C(z)} \cdot y(t)$$

- Pongo $\varepsilon_1(t)$ in forma dinamica e successivamente in forma ricorsiva

$$\varepsilon_1(t) = y(t) - \widehat{y}(t|t-1;\boldsymbol{\theta})$$

- Se il $\varepsilon_1(t;\boldsymbol{\theta})$ dipende da se stesso a un instante precendente ossia:

$$\varepsilon_1(t;\boldsymbol{\theta}) = y(t) - a \cdot \varepsilon_1(t-1;\boldsymbol{\theta})$$

  - Si inizializza il predittore ponendolo uguale al predittore banale:

$$\widehat{y}(1|0) = m_y \backsimeq \frac{1}{N} \cdot \sum_{t=1}^{N} y(t) \Rightarrow \textit{media dei dati che ho}$$

$$\varepsilon_1(1;\boldsymbol{\theta}) = y(1) - \widehat{y}(1|0)$$

  - Si inizializza il predittore al primo instante noto: e si taglia la sommatoria in modo che parta da 2 $\Rightarrow$ **troncare la somma**, si preferisce utilizzare questo metodo per semplificare i calcoli

$$\widehat{y}(1|0) = y(1)$$

$$\varepsilon_1(1;\boldsymbol{\theta}) = y(1) - \widehat{y}(1|0) = 0$$

- Effettuata l'inizializzazione mi calcolo tutti la funzione di costo:

$$J_N(\boldsymbol{\theta}) = \sum_{t=2}^{N} \Big(\varepsilon_1(t)\Big)^2 = \sum_{t=2}^{N} \Big(y(t) - \widehat{y}(t|t-1;\boldsymbol{\theta})\Big)^2$$

- Eseguo il metodo diretto, calcolando il gradiente e ponendo il gradiente = 0 $\rightarrow$ devo controllare che il minimo sia effettivamente il minimo **calcolando la derivata secondo di** $J_N(\boldsymbol{\theta})$ e verificando che sia > 0

$$\frac{d}{d\boldsymbol{\theta}} J(\boldsymbol{\theta}) \quad \rightarrow \quad \frac{d}{d\boldsymbol{\theta}} J(\boldsymbol{\theta})\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} = 0 \quad \Rightarrow \quad \widehat{\boldsymbol{\theta}}$$

$$\frac{d^2}{d\boldsymbol{\theta}^2} J_N(\boldsymbol{\theta})\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} = K > 0 \quad \Rightarrow \quad \widehat{\boldsymbol{\theta}} = min$$

Note

- $\mathcal{S}$ non é in forma canonica, bensí $\mathcal{M}$ lo é, ma non si sa se stazionario alla fine → **controllare**
- **IDEA:** dati dei dati raccolti vogliamo vedere a che famiglia di modelli appartiene
- **Metodo dei minimi quadrati** solo AR, ARX e sono lineari nei parametri avremo un paraboloide o una parabola in base a $\boldsymbol{\theta} \Rightarrow$ **unico minimo globale** (funzione di costo convessa)
- Se il modello $\mathcal{M}_k$ é lo stesso del processo $\mathcal{S}$, ossia $\mathcal{M}_k \equiv \mathcal{S} \Rightarrow$ dalla teoria l'identificatore PEM è asintoticamente corretto se la classe di modelli contiene esattamente il modello usato per generare i dati. Quindi la stima dei coefficienti corrisponde esattamente a quelli veri:

  $$\mathcal{M}_k \equiv \mathcal{S} \qquad \Rightarrow \qquad \lim_{N\to\infty} \widehat{\boldsymbol{\theta}} = \boldsymbol{\theta}$$

  Puó succere inoltre che i coefficienti sono diversi, ossia:

  $$\mathcal{S} \to c_0 \neq 1$$

  $$\mathcal{M} \to c_0 = 1$$

  *per risolvere il problema portare $\mathcal{S}$ in **forma canonica**.*

  $$y(t)_{\mathcal{S}} = ..$$

  Sostituisco in $y(t) = y(t)_{\mathcal{S}}$ e $\widehat{y}(t|t-1;\boldsymbol{\theta})$ = predittore del modello $\mathcal{M}$

  $$\varepsilon_1(t;\boldsymbol{\theta}) = y(t) - \widehat{y}(t|t-1;\boldsymbol{\theta})$$
  $$\varepsilon_1(t;\boldsymbol{\theta}) = F(z;\boldsymbol{\theta}) \cdot e(t)$$

  ***NB:** si noti che la varianza $\varepsilon_1(t;\boldsymbol{\theta})$ é minima quando $\varepsilon_1(t;\boldsymbol{\theta}) = e(t)$ (predittore ottimo lineare), ossia quando $F(z;\boldsymbol{\theta}) = 1$, Dunque crea il sistema per soddisfare la proprietá*

- Dato $\mathcal{S}$ in forma dataset puó essere necessario **troncare la funzione di costo** $J_N(\boldsymbol{\theta})$
- Calcolo delle matrici

  $$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

  $$riga \cdot colonna = r1 \cdot c1 = 1 \times 1$$

  $$a \cdot 1 + b \cdot 0 = a = 1 \times 1$$

- Invertire le matrici: scambiare gli elementi sulla diagonale principale e cambiando il segno sulla diagonale secondaria

  $$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{det} \cdot \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} = \frac{1}{a_{11} \cdot a_{22} - a_{12} \cdot a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

- La trasposta di una matrice 2×2 si invertono gli elementi sulla diagonale secondaria

  $$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{T} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$$

- Un processo $ARMA$ lo tratto come un MA inizialemente quindi scrivo il processo in modo tale che dipenda solo dai dati $y(t)$
- Quando calcolo i $\gamma_{yy}(0)$ non usare mai la varianza, calcolarlo tramite $\mathbb{E}\big[(y(t)^2 - m_y)\big]$

7 Riassunto svolgimento

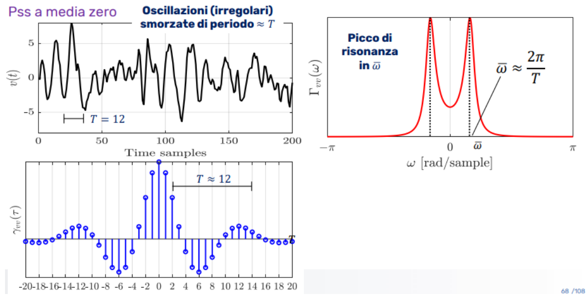
7.1 Esercitazione 1

Svolgimento:

- Verifico stazionaritá processo tramite teorema $\rightarrow$ poli
- Calcolo media del processo $\mathbb{E}[y(t)]$
- Se $\mathbb{E}[y(t)] \neq 0$ e $e(t) \sim Wn(m_e, \lambda^2)$ dove $m_e \neq 0 \Rightarrow$ depolarizzo il processo
- Calcolo l'autocovarianza $\gamma_{yy}(\tau)$ e ulteriori dipendenze
- Calcolo la densitá di potenza del processo $\Gamma_{yy}(\omega)$ e la disegno, sapendo che $\Gamma_{yy}(\omega) \geqslant 0 \ \forall \omega$

Errori:

- Quando calcolo la $\gamma_{yy}(\tau) \overset{m_y=0}{=} \mathbb{E}\big[y(t)y(t-\tau)\big]$ esplodo solo $y(t)$
- Velocizzare il calcolo della $\gamma_{yy}(\tau)$ usando le formule di yule se $AR(1)$ o le formule del $MA(n_c)$
- Depolarizzare prima di calcolare $\gamma_{yy}(\tau)$
- $\gamma_{yy}(\tau)$ puó essere negativa con $\tau > 0$, solo $\gamma_{yy}(0) = var[y(t)] \geqslant 0$
- La pulsazione dominante viene determinata osservando la densitá spettrale di potenza il valore piú grande di $\Gamma_{yy}(\bar{\omega})$ da cui possiamo trovare il periodo del del segnale



$$\overline{w} = \frac{2\pi}{T} = 2\pi f$$

$$\uparrow f, \ \uparrow w, \ \downarrow T$$

- Quando calcolo $\Gamma_{yy}(\omega) = |H(e^{j\omega})|^2 \cdot \Gamma_{ee}(\omega)$, non dimenticarsi di moltiplicare il tutto per $\Gamma_{ee}(\omega)$
- the resonance causes one frequency to prevail on the others. The time behaviour is then a sort of sinuosidal signal dipped in noise. Like in this example

$$\begin{array}{lll} \overline{w} = 0 & \overline{w} = 2\pi f = 0 & f \rightarrow 0 \ \textit{(low frequency)} \\ \overline{w} \rightarrow \infty & \overline{w} = 2\pi f = \infty & f \rightarrow \infty \ \textit{(high frequency)} \\ \overline{w} = k \ \textit{(resonance)} & \overline{w} = 2\pi f = k & \textit{sinusoidal signal + noise} \end{array}$$

7.2 Esercitazione 2

Svolgimento:

- Controllo stazionarietá
- Calcolo base canonica: radici interne cerchio unitario, coprimi, monici, stesso grado $\rightarrow$ filtro passa tutto, e aggiustamenti matematici
- Controllo se $\mathbb{E}[\varepsilon_k] = \mathbb{E}[E(z) \cdot \eta(t)] = 0$ altrimenti portare $m_\eta = 0 \Rightarrow$ metodo ARMAX. Quando uso questo metodo lasciare $u(t)$ simbolico fino alla fine e poi sostituisco quando determino il predittore $\hat{y}(t-k)$
- Calcolo predittore $y(t|t-k)$ e scriverlo in forma ricorsiva $\Rightarrow$ metodo lunga divisione
- Esempio di calcolo $y(5|4)$, avró il predittore (di cui non ho i dati per cui lo inizializzo alla media del processo) $y(\hat{1}|0) = \mathbb{E}[y(t)] = 0$ oppure calcolo $\hat{y}(-1|0) = 0$

Errori:

- Quando calcolo la predizione devo dare una condizione iniziale, dunque $\hat{y}(x|y) = y(x) = m_y$ e devo trascrivere ricorsivamente fino a quando non ho piú alcun campione di dati
- Controllare attentamente se si puó semplificare o meno
- Se viene chiesta la $\mathbb{E}[y(t)] = m_y$ é quella del processo iniziale quindi nel caso devo deopolarizzare (se ARMAX con $u(t)$ devo usare ARMA)
- Ogni volta che depolarizzo/ARMAX creare un nuovo segnale $WN(0, \lambda^2)$
- guardare esercizio 7 ese II di ulteriori esercizi

7.3 Esercitazione 3

Svolgimento: $\mathcal{S} \notin \mathcal{M}$

- Le misure $y(t)$ sono campionate dal processo $\mathcal{S}$ mentre il predittore é quello calcolato con il modello che si vuole identificare
- Calcolo il predittore $\hat{y}(t|t-1)$ in funzione dei dati $y(t)$, nel caso forma canonica $C(z)/A(z)$
- Calcolo $\varepsilon_1(t)$ e vedo se riesco se: $\varepsilon_1(t) = e(t)$ (var min). Se $\mathcal{S} \in \mathcal{M}$ questo é il metodo piú veloce.
- Se $\varepsilon_1(t) \neq e(t)$ uso la solita funzione di costo $J_N(\boldsymbol{\theta})$
- **Metodo diretto:** calcolo i punti stazionari $\frac{\partial}{\partial \boldsymbol{\theta}} J(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = 0$, tiro fuori $\hat{\boldsymbol{\theta}}$ sistemi e controllo se i $\hat{\boldsymbol{\theta}}$ tiratirati fuori hanno $\frac{\partial^2}{\partial \boldsymbol{\theta}^2} J(\hat{\boldsymbol{\theta}})|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} > 0$ allora il punto risulta essere un minimo e ottengo i miei dati stimati $\hat{\boldsymbol{\theta}}$
- **Metodo lineare:** solo i ARX e AR possono utilizzare questo metodo, perché sono lineari nei parametri e dunque hanno un unico minimo globale $\rightarrow$ paraboloide/parabola come funzione di costo
- Ricalcolo la funzione di costo e tiro fuori la varianza: $\hat{\lambda}^2 = J(\hat{\boldsymbol{\theta}})$ e riscrivo $\mathcal{M}(\hat{\boldsymbol{\theta}})$

Svolgimento: $\mathcal{S} \in \mathcal{M}$

- Calcolo il predittore $\hat{y}(t|t-1)$ in forma canonica che dipenda dai dati $y(t) \rightarrow C(z)/A(z)$
- Porto $y(t)_{\mathcal{S}}$ in forma canonica/dinamica e che dipenda da $e(t)$
- Calcolo $\varepsilon_1(t) = y(t) - \hat{y}(t|t-1)$, sostituendo il predittore in forma dinamica al suo interno
- Svolgo i calcoli ottenendo: $\varepsilon_1(t) = F(z) \cdot y(t)$
- Pongo $y(t) = y(t)_{\mathcal{S}}$ ottenendo: $\varepsilon_1(t) = W(z) \cdot e(t)$
- Affinché il predittore abbia varianza minima $\varepsilon_1(t) = e(t)$ $(W(z) = 1)$ e dunque devo eliminare numeratore e denominatore ottenedo i valori esatti di $\hat{\boldsymbol{\theta}}$
- $\hat{\lambda}^2 = \lambda^2_{\mathcal{S}}$ e riscrivo $\mathcal{M}$ con i parametri stimati $\hat{\boldsymbol{\theta}}$

Errori:

- quando calcolo $J(\boldsymbol{\theta}) = \mathbb{E}[\big(y(t) - \hat{y}(t|t-1)\big)^2] = \mathbb{E}[\big(y(t) - (y(t-1) + ay(t-2)\big)^2]$ stare attento al secondo membro all'interno della parentesi che il segno che diventa negativo

8 Note processi stocastici

$$AR \Rightarrow \text{ presenta solo poli}$$
$$MA \Rightarrow \text{ presenta solo zeri}$$
$$ARMA \Rightarrow \text{ presenta poli e zeri}$$

Processo stocastico: $AR(n_a = 2)$

$$y(t) = e(t) + \frac{1}{2}y(t-1)) - \frac{1}{4}y(t-2)$$
$$\mathbb{E}\big[y(t)\big] \Rightarrow m_y = 0 + \frac{1}{2}m_y - \frac{1}{4}m_y$$
$$\gamma_{yy}(0) = \mathbb{E}\Bigg[y(t) \cdot y(t-1)\Bigg] = \mathbb{E}\Bigg[\Bigg(e(t) + \frac{1}{2}y(t-1) - \frac{1}{4}y(t-2)\Bigg) \cdot y(t-1)\Bigg]$$
$$= 1 + \frac{1}{4}\gamma_{yy}(0) + \frac{1}{16}\gamma_{yy}(0) - \frac{1}{4}\gamma_{yy}(1)$$
$$\gamma_{yy}(\tau) = \mathbb{E}\Bigg[y(t) \cdot y(t-\tau)\Bigg] = \mathbb{E}\Bigg[\Bigg(e(t) + \frac{1}{2}y(t-1) - \frac{1}{4}y(t-2)\Bigg) \cdot y(t-\tau)\Bigg]$$

NB: *se ho k incognite in* $\gamma_{yy}(0)$ *devo avere k equazioni e quindi calcolare i successivi* γ_{yy} *fino alla* $\gamma_{yy}(k)$ *e poi mettere a sistema*

Equazioni di Yule-Walker per un AR(1)

$$y(t) = a_1 y(t-1) + e(t) \qquad e(t) \sim WN(0, \lambda^2)$$
$$\begin{cases} \gamma_{yy}(0) = \frac{\lambda^2}{1-a_1^2} & se\ \tau = 0 \\ \gamma_{yy}(\tau) = a_1 \cdot \gamma_{yy}(\tau - 1) & se\ \tau > 0 \end{cases}$$

Si può notare che questi non raggiungeranno mai 0, ma si avvicineranno asintoticame, per il calcolo della densitá spettrale utilizzare la rappresentazione dinamica:

$$\Gamma_{yy}(\omega) = |H(e^{j\omega})|^2 \cdot \Gamma_{ee}(\omega) = |H(e^{j\omega})|^2 \cdot \lambda_{ee}$$

Un processo $MA(n_c)$:

$$\mathbb{E}[\hat{y}(t|t-k)] = m_y, \quad \forall k > n_c$$
$$Var[\varepsilon_k(t)] = \gamma_{yy}(0) = Var[y(t)], \quad \forall k > n_c$$

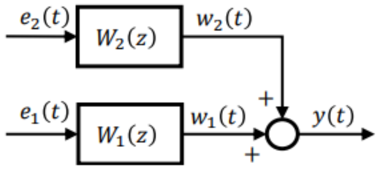
$$\gamma_{yy}(\tau) = 0, \forall \tau > n_c \quad \Rightarrow \text{ il processo ha memoria finita}$$

$$\gamma_{yy}(0) = \lambda_{ee}^2 \cdot \sum_{i=0}^{n_c} c_i^2 \qquad \gamma_{yy}(\tau) = \lambda_{ee}^2 \cdot \sum_{i=0}^{n_c - \tau} c_i \cdot c_{i+\tau}$$
$$\gamma_{yy}(1) = \lambda_{ee}^2 \cdot (c_0 c_1 + c_1 c_2 + \ldots + c_{n_c - 1} c_{n_c})$$
$$\gamma_{yy}(2) = \lambda_{ee}^2 \cdot (c_0 c_2 + c_1 c_3 + \ldots + c_{n_c - 3} c_{n_c})$$

NB: *tutti i processi possono essere visti come un* $MA(\infty)$

$$w_1(t) = W_1(z) \cdot e_1(t)$$
$$w_1(t) = c_{01}e_1(t) + c_{11}e_1(t-1) + c_{21}e_1(t-2) + \ldots$$

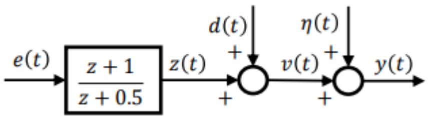
9 Schema a blocchi



$$e_1(t) \perp e_2(t) \Rightarrow w_1(t) \perp w_2(t) \qquad w_1(t) = W_1(z) \cdot e_1(t) \qquad w_2(t) = W_2(z) \cdot e_2(t)$$

$$\begin{aligned}
\gamma_{yy}(\tau) &= \mathbb{E}\Big[y(t) \cdot y(t-\tau)\Big] \\
&= \mathbb{E}\Big[\Big(w_1(t) + w_2(t)\Big) \cdot \Big(w_1(t-\tau) + w_2(t-\tau)\Big)\Big] \\
&= \gamma_{w_1 w_1}(\tau) + \gamma_{w_2 w_2}(\tau) + \mathbb{E}\Big[w_2(t) \cdot w_1(t-\tau)\Big] + \mathbb{E}\Big[w_1(t) \cdot w_2(t-\tau)\Big] \\
&= \gamma_{w_1 w_1}(\tau) + \gamma_{w_2 w_2}(\tau)
\end{aligned}$$

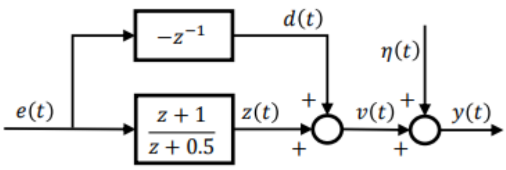
$$\begin{aligned}
\Gamma_{yy}(\omega) &= |W_1(e^{j\omega})|^2 \cdot \Gamma_{w_1 w_1}(\omega) + |W_2(e^{j\omega})|^2 \cdot \Gamma_{w_2 w_2}(\omega) \\
&= |W_1(e^{j\omega})|^2 \cdot \lambda_1^2 + |W_2(e^{j\omega})|^2 \cdot \lambda_2^2
\end{aligned}$$

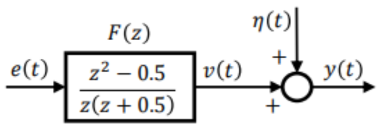


$$d(t) \perp e(t) \perp \eta(t) \sim WN$$

$$\begin{aligned}
\Gamma_{yy}(\omega) &= \Gamma_{\eta\eta}(\omega) + \Gamma_{vv}(\omega) \\
&= \Gamma_{\eta\eta}(\omega) + \Gamma_{dd}(\omega) + \Gamma_{zz}(\omega) = \Gamma_{zz}(\omega) + 2
\end{aligned}$$

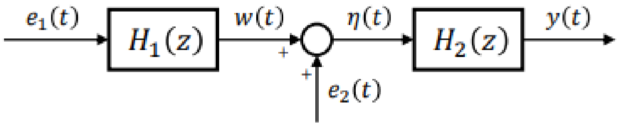
$$\Gamma_{zz}(\omega) = \lambda_e^2 \cdot |W_1(e^{j\omega})|^2$$





$$F(z) = \frac{z+1}{z+0.5} - z^{-1}$$
$$\Gamma_{yy}(\omega) = \Gamma_{\eta\eta}(\omega) + \Gamma_{vv}(\omega) = \Gamma_{vv}(\omega) + 1$$
$$\Gamma_{vv}(\omega) = \lambda_e^2 \cdot |F(e^{j\omega})|^2$$

classificazione:



$$e_1(t) \sim WN(0,1) \quad e_2(t) \sim WN(1,1) \quad e_1(t) \perp e_2(t)$$

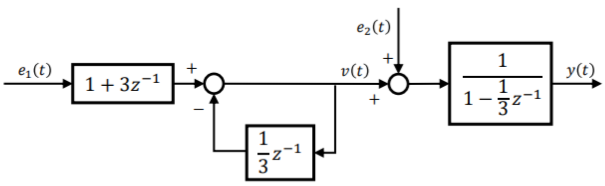
$$H_1(z) = \frac{5z+25}{5z+1} \qquad H_2(z) = \frac{3z+12}{6z^2+5z+1}$$

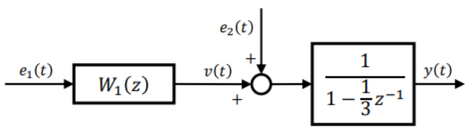
$$\begin{aligned} w(t) = H_1(z) \cdot e_1(t) &= \frac{5z+25}{5z+1} \cdot e_1(t) \\ &= 5 \cdot \underbrace{\frac{1}{5} \cdot \frac{z+5}{z+\frac{1}{5}}}_{f.\ passa\ tutto} \cdot e_1(t) = 5 \cdot e_1(t) \end{aligned}$$

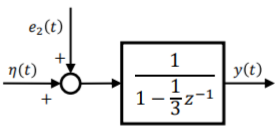
$$\begin{aligned} \eta(t) = w(t) + e_2(t) &= 5e_1(t) + e_2(t) \\ m_\eta &= 5 \cdot \mathbb{E}\big[e_1(t)\big] + \mathbb{E}\big[e_2(t)\big] \\ &= 5 \cdot 0 + 1 = 1 \end{aligned}$$

$$\begin{aligned} \gamma_{\eta\eta}(0) = \mathbb{E}\big[\tilde{\eta}(t)^2\big] &= \mathbb{E}\Big[\big(5e_1(t) + e_2(t)\big)^2\Big] \\ &= 25 \cdot \mathbb{E}\big[e_1(t)^2\big] + \cancel{10 \cdot \mathbb{E}\big[e_1(t) \cdot e_2(t)\big]} + \mathbb{E}\big[e_2(t)^2\big] \\ &= 25 \cdot 1 + 1 = 26 \end{aligned}$$

$$\Rightarrow y(t) = \frac{3z+12}{6z^2+5z+1} \cdot \eta(t) \qquad \eta(t) \sim WN(1, 26)$$





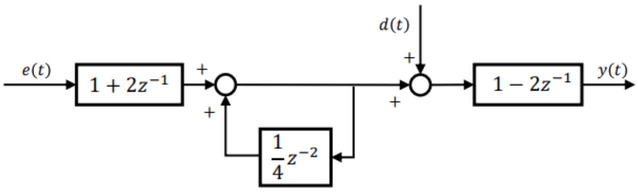


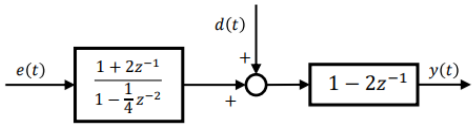
$$W_1 = \frac{v(t)}{e_1(t)} = \left(1 + 3z^{-1}\right) \cdot \left(\frac{1}{1+\frac{1}{3}z^{-1}}\right) = \frac{z+3}{z+\frac{1}{3}}$$

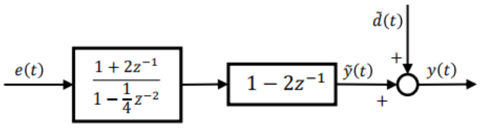
$$v(t) = W_1 \cdot e_1(t) = T(z) \cdot 3 \cdot e_1(t) = 3 \cdot e_1(t) = \eta(t)$$

$$d(t) = \eta(t) + e_2(t)$$

$$m_d = \mathbb{E}\big[d(t)\big] = 4 \qquad \gamma_{dd}(0) = Var\big[d(t)\big] = \mathbb{E}\big[(d(t)-4)^2\big] \qquad d(t) \sim WN\big(m_d, \gamma_{dd}(0)\big)$$







$$d(2) = 2$$

$$W_1(z) = \left(1 + 2z^{-1}\right) \cdot \left(\frac{1}{1 = \frac{1}{4}z^{-2}}\right) = \frac{1 + 2z^{-1}}{1 - \frac{1}{4}z^{-2}}$$

$$W_2(z) = W_1(z) \cdot \left(1 - 2z^{-1}\right) = \frac{(1 + 2z^{-1})(1 - 2z^{-1})}{1 - \frac{1}{4}z^{-2}}$$

Come se fosse uno schema a blocchi: *in serie*

$$\widetilde{d}(t) = \left(1 - 2z^{-1}\right) \cdot d(t) = d(t) - 2d(t-1) = 2 - 4 = -2$$

$$D(z) = 2 \cdot \left(1 - 2z^{-1}\right)$$

$$y(t) = \widetilde{d}(t) + \frac{(1 + 2z^{-1})(1 - 2z^{-1})}{1 - \frac{1}{4}z^{-2}} \cdot e(t) \qquad e(t) \sim WN(0,1) \quad \widetilde{d}(t) = -2$$

$$y(t) = \widetilde{d}(t) + \widetilde{y}(t)$$

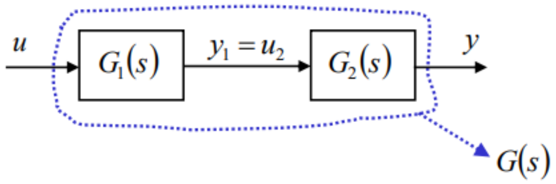
$$\widetilde{y}(t) = \left(\frac{1}{2} \cdot \frac{1 + 2z^{-1}}{1 + \frac{1}{2}z^{-1}}\right) \cdot \left(\frac{1}{2} \cdot \frac{1 - 2z^{-1}}{1 - \frac{1}{2}z^{-1}}\right) \cdot -4e(t)$$
$$= T_1(z) \cdot T_2(z) \cdot -4e(t) = T_1(z) \cdot T_2(z) \cdot \eta(t)$$

$$y(t) = \widetilde{d}(t) + \eta(t) \rightarrow ARMAX$$

$$\widehat{y}(t|t-1) = u(t-1) = -2$$

10 Considerazioni di automatica

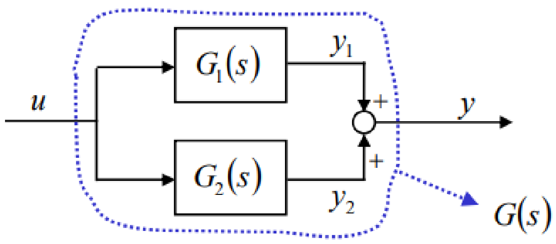
Blocchi in serie:



$$Y(s) = G_2(s) \cdot U_2(s) = G_2(s) \cdot Y_1(s) = G_2(s) \cdot G_1(s) \cdot U(s)$$

$$\frac{uscita}{ingresso} = \frac{Y(s)}{U(s)} = G_2(s) \cdot G_1(s) = G(s)$$

Blocchi in serie:

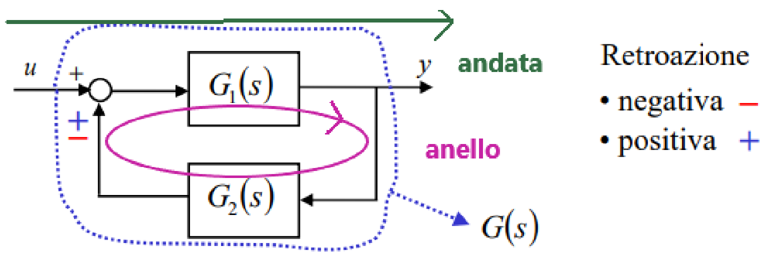


$$Y(s) = Y_1(s) + Y_2(s) = G_1(s) \cdot U(s) + G_2(s) \cdot U(s) = \Big(G_1(s) + G_2(s)\Big) \cdot U(s)$$

$$\frac{uscita}{ingresso} = \frac{Y(s)}{U(s)} = G_1(s) + G_2(s) = G(s)$$

Blocchi in retroazione:

Se la retroazione è negativa - $\rightarrow$ *mettiamo al Denominatore il +*



$$Y(s) = G_1(s)\Big[U(s) \pm G_2(s) \cdot Y(s)\Big] = \Big[1 \mp G_1(s) \cdot G_2(s)\Big] \cdot Y(s) = G_1(s) \cdot U(s)$$

$$\frac{uscita}{ingresso} = \frac{Y(s)}{U(s)} = \frac{Fdt\ andata}{1 \mp\ Fdt\ anello} = \frac{G_1(s)}{1 \mp G_1(s) \cdot G_2(s)} = G(s)$$

11 Robe matematiche

- Applicare la formula del discriminante:

$$6z^2 + 5z + 1 = 0 \qquad \Rightarrow \qquad p_{1,2} = \frac{-5 \pm \sqrt{25 - 4 \cdot 6 \cdot 1}}{12} = \frac{-5 \pm \sqrt{1}}{12}$$

$$p_1 = -\frac{1}{3} \qquad p_2 = -\frac{1}{2}$$

$$6z^2 + 5z + 1 \equiv 6\left(z + \frac{1}{2}\right)\left(z + \frac{1}{3}\right)$$

- Calcolo poli complessi coniugati

$$p_{1,2} = \frac{\frac{10}{3} \pm \sqrt{\frac{100}{9} - 4 \cdot 5 \cdot \frac{10}{9}}}{10} = \frac{\frac{10}{3} \pm \sqrt{-\frac{100}{9}}}{10} = \frac{\frac{10}{3} \pm j\frac{10}{3}}{10}$$
$$= \frac{1}{3} \pm j\frac{1}{3}$$

$$\|p_{1,2}\| = \left\|\frac{1}{3} \pm j\frac{1}{3}\right\| = \sqrt{\frac{1}{9} + \frac{1}{9}} = \frac{\sqrt{2}}{3} < 1$$

$$p_3 = \frac{1}{3} < 1$$

- Calcolo zeri, si raccoglie il 5:

$$z_{1,2} = \frac{\frac{11}{6} \pm \sqrt{\frac{121}{36} - 4 \cdot 1 \cdot \frac{1}{2}}}{2} = \frac{\frac{11}{6} \pm \sqrt{\frac{49}{36}}}{2} = \frac{\frac{11}{6} \pm \frac{7}{6}}{2}$$
$$= \frac{11}{12} \pm \frac{7}{12}$$

$$z_1 = \frac{18}{12} = \frac{3}{2}$$
$$z_2 = \frac{4}{12} = \frac{1}{3}$$

$$w(t) = \frac{\left(z - \frac{3}{2}\right) \cdot \cancel{\left(z - \frac{1}{3}\right)}}{5 \cdot \left(z^2 - \frac{2}{3}z + \frac{2}{9}\right) \cdot \cancel{\left(z - \frac{1}{3}\right)}} \cdot e(t)$$

- Grado massimo numeratore denominatore diversi $\frac{z}{z}$:

$$w(t) = \frac{z^{-1} - \frac{2}{3}z^{-2}}{1 - \frac{2}{3}z^{-1} + \frac{2}{9}z^{-2}} \cdot \frac{z}{z}\left(-\frac{3}{10} \cdot e(t)\right)$$

$$w(t) = \frac{1 - \frac{2}{3}z^{-1}}{1 - \frac{2}{3}z^{-1} + \frac{2}{9}z^{-2}}\left(-\frac{3}{10}z^{-1} \cdot e(t)\right)$$

$$w(t) = \frac{1 - \frac{2}{3}z^{-1}}{1 - \frac{2}{3}z^{-1} + \frac{2}{9}z^{-2}} \cdot \left(-\frac{3}{10} \cdot e(t-1)\right)$$

- Costruzione nuovo segnali, occhi ai segni

$$w(t) = \frac{1-\frac{2}{3}z^{-1}}{1-\frac{2}{3}z^{-1}+\frac{2}{9}z^{-2}} \cdot \left(-\frac{3}{10}\cdot e(t-1)\right)$$

$$\eta(t) = -\frac{3}{10}\cdot e(t-1)$$
$$\eta(t) \sim WN\left(-\frac{3}{10}\cdot 1, \frac{3^2}{10^2}\cdot 1\right) = WN\left(-\frac{3}{10}, \frac{9}{100}\right)$$

$$w(t) = \frac{1-\frac{2}{3}z^{-1}}{1-\frac{2}{3}z^{-1}+\frac{2}{9}z^{-2}} \cdot \eta(t)$$

- classificazione schema a blocchi, igegnoso metodo per calcolo varianza, viene splittato il

$$-4 = -1 + (-3)$$

$$m_e = \mathbb{E}[e(t)] = \mathbb{E}[e_2(t)+\eta(t)] = \mathbb{E}[e_2(t)] + \mathbb{E}[\eta(t)] = 1+3 = 4$$

$$\begin{aligned}
\gamma_{ee}(0) &= \mathbb{E}\left[(e(t)-4)^2\right] = \mathbb{E}\left[(e_2(t)+\eta(t)-4)^2\right] \\
&= \mathbb{E}\left[((e_2(t)-1)+(\eta(t)-3))^2\right] \\
&= \mathbb{E}\left[((\tilde{e}_2(t))+(\tilde{\eta}(t)))^2\right] \\
&= \mathbb{E}\left[(\tilde{e}_2(t))^2\right] + \mathbb{E}\left[(\tilde{\eta}(t))^2\right] + 2\underbrace{\cancel{\mathbb{E}\left[(\tilde{e}_2(t))(\tilde{\eta}(t))\right]}}_{e_2(t)\perp e_1(t)} \\
&= 2+9 = 11
\end{aligned}$$

$$e(t) \sim WN(4, 11)$$

- ***NB:*** quando creo ARMAX devo poi sostituire a $u(t)$ la costante che le ho dato in ingresso

$$\hat{y}(t|t-3) = \left(1+\frac{1}{3}z^{-1}+\frac{1}{9}z^{-2}\right)u(t-3) + \frac{1}{27}y(t-3)$$
$$\hat{y}(t|t-3) = u(t-3) + \frac{1}{3}u(t-4) + \frac{1}{9}u(t-5) + \frac{1}{27}y(t-3)$$

$$\hat{y}(t|t-3) = \left(4+\frac{4}{3}+\frac{4}{9}\right) + \frac{1}{27}y(t-3)$$
$$\hat{y}(t|t-3) = \frac{108+36+12}{27} + \frac{1}{27}y(t-3)$$
$$\hat{y}(t|t-3) = \frac{156}{27} + \frac{1}{27}y(t-3)$$
$$\hat{y}(t|t-3) = \frac{52}{9} + \frac{1}{27}y(t-3)$$

- Depolarizzazione processo

$$y(t) = e(t-1) + 2e(t-2) + e(t-4) + 2e(t-5) + 1, \qquad e(t) \sim WN(1,1)$$

con media:

$$\begin{aligned} m_y = \mathbb{E}[y(t)] &= \mathbb{E}[e(t-1) + 2e(t-2) + e(t-4) + 2e(t-5) + 1] \\ &= \mathbb{E}[e(t-1)] + 2\mathbb{E}[e(t-2)] + \mathbb{E}[e(t-4)] + 2\mathbb{E}[e(t-5)] + 1 \\ &= 1 + 2 + 1 + 2 + 1 = 7 \end{aligned}$$

si depolarizza il processo:

$$\begin{cases} \tilde{y}(t) & = y(t) - m_y = y(t) - 7 \\ \tilde{e}(t) & = e(t) - m_e = e(t) - 1 \end{cases}$$

$$\begin{aligned} \tilde{y}(t) + 7 &= \tilde{e}(t-1) + 1 + 2\tilde{e}(t-2) + 2 + \tilde{e}(t-4) + 1 + 2\tilde{e}(t-5) + 2 + 1 \\ \tilde{y}(t) &= \tilde{e}(t-1) + 2\tilde{e}(t-2) + \tilde{e}(t-4) + 2\tilde{e}(t-5) \end{aligned}$$

- Guardare esercizio 9 ese 2
- Calcolare matrici

$$\Phi^{\top} \cdot \Phi \cdot \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \Phi^{\top} \cdot \boldsymbol{y}$$

$$\Phi = \begin{bmatrix} y(2) & y(1) \\ y(3) & y(2) \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix}$$

$$\boldsymbol{y} = \begin{bmatrix} y(3) \\ y(4) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\begin{aligned} \Phi^{\top} \cdot \Phi &= \begin{bmatrix} 2 & 0 \\ -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 4+0 & -2+0 \\ -2+0 & 1+4 \end{bmatrix} \\ &= \begin{bmatrix} 4 & -2 \\ -2 & 5 \end{bmatrix} \end{aligned}$$