Esercizi sull'identificazione

1 Esercizio 1

Supponiamo di avere infiniti dati presi dal processo stocastico stazionario:

$$\mathcal{S} \;:\; y(t) = 2 \cdot e(t) + \frac{1}{2} \cdot e(t-1) + e(t-2) \qquad\qquad e(t) \sim WN(0,1)$$

e i tre modelli:

$$\begin{aligned}
\mathcal{M}_1(a) &: \; y(t) = a \cdot y(t-1) + \eta(t) && \eta(t) \sim WN\left(0, \lambda^2\right) \\
\mathcal{M}_2(a,b) &: \; y(t) = a \cdot y(t-1) + b \cdot y(t-2) + \eta(t) && \eta(t) \sim WN\left(0, \lambda^2\right) \\
\mathcal{M}_3(a,b) &: \; y(t) = \eta(t) + a \cdot \eta(t-1) + b \cdot \eta(t-2) && \eta(t) \sim WN\left(0, \lambda^2\right)
\end{aligned}$$

Identificare il miglior modello in ogni classe usando la metodologia PEM.

1.1 Note preliminari

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:

$$J_N(\boldsymbol{\vartheta}) = \frac{1}{N} \sum_{t=1}^{N} \left(y(t) - \hat{y}(t|t-1; \boldsymbol{\vartheta})\right)^2$$

In questo caso, abbiamo a disposizione infiniti campioni. Quindi, non abbiamo bisogno di stimare la varianza dell'errore di predizione, ma possiamo usare la sua definizione direttamente:

$$J(\boldsymbol{\vartheta}) = \mathbb{E}\left[\left(y(t) - \hat{y}(t|t-1; \boldsymbol{\vartheta})\right)^2\right]$$

In questo caso, le misure $y(t)$ sono campionate dal processo $\mathcal{S}$ mentre il predittore è quello calcolato con il modello che si vuole identificare.

1.2 Modello $\mathcal{M}_1(a)$

Iniziamo calcolando il predittore a un passo di $\mathcal{M}_1(a)$:

$$\hat{y}(t|t-1, a) = a \cdot y(t-1)$$

NOTA: il sistema è un AR in forma canonica e quindi il predittore a un passo corrisponde alla forma ricorsiva del processo senza il rumore.

Quindi la cifra di merito diventa:

$$\begin{aligned}
J(a) &= \mathbb{E}\left[\left(y(t) - \hat{y}(t|t-1; a)\right)^2\right] \\
&= \mathbb{E}\left[\left(y(t) - ay(t-1)\right)^2\right]
\end{aligned}$$

dove $y(t)$ e $y(t-1)$ sono campioni del processo $\mathcal{S}$. Con qualche passaggio:

$$\begin{aligned}
J(a) &= \mathbb{E}\left[y(t)^2 - 2ay(t)\,y(t-1) + a^2 y(t-1)^2\right] \\
&= \underbrace{\mathbb{E}\left[y(t)^2\right]}_{\gamma_{yy}(0)} - 2a \underbrace{\mathbb{E}\left[y(t)\,y(t-1)\right]}_{\gamma_{yy}(1)} + a^2 \underbrace{\mathbb{E}\left[y(t-1)^2\right]}_{\gamma_{yy}(0)} \\
&= \gamma_{yy}(0) - 2a\gamma_{yy}(1) + a^2\gamma_{yy}(0) \\
&= a^2\gamma_{yy}(0) - 2a\gamma_{yy}(1) + \gamma_{yy}(0)
\end{aligned}$$

Intuitivamente, la funzione di costo $J(a)$ è sempre positiva (poichè dalla teoria $J(a) = Var\left[\varepsilon_1(t)\right] > 0$). Inoltre, $J(a)$ per modelli *AR* è sempre lineare nei parametri, perciò $J(a)$ sarà sempre una parabola (oppure un paraboloide nel caso in cui i parametri da

identificare siano più di uno). Grazie a queste due considerazioni deduciamo che $\frac{d}{da}J(a) = 0$ è sempre un punto di minimo. Ciò non vale per modelli *MA* e *ARMA*, quindi bisognerà sempre verificare che $\frac{d}{da}J(a) = 0$ sia un punto di minimo.
Per trovarlo possiamo derivare la funzione di costo:

$$\begin{aligned}\frac{d}{da}J(a) &= \frac{d}{da}\left(a^2\gamma_{yy}(0) - 2a\gamma_{yy}(1) + \gamma_{yy}(0)\right)\\ &= 2a\gamma_{yy}(0) - 2\gamma_{yy}(1)\end{aligned}$$

e trovare il suo zero:

$$\begin{aligned}\left.\frac{d}{da}J(a)\right|_{a=\hat{a}} &= 0\\ 2\hat{a}\gamma_{yy}(0) - 2\gamma_{yy}(1) &= 0\\ 2\hat{a}\gamma_{yy}(0) &= 2\gamma_{yy}(1)\\ \hat{a} &= \frac{\gamma_{yy}(1)}{\gamma_{yy}(0)}\end{aligned}$$

Per ricavare la stima è necessario calcolare la varianza del processo e la funzione di auto-correlazione a un passo. Fortunatamente, il processo $\mathcal{S}$ è un MA(n):

$$\begin{aligned}\gamma_{yy}(0) &= \lambda_e^2 \cdot \sum_{i=0}^{n} c_i^2\\ &= 2^2 + \left(\frac{1}{2}\right)^2 + (1)^2 = 4 + \frac{1}{4} + 1 = \frac{21}{4}\end{aligned}$$

$$\begin{aligned}\gamma_{yy}(\tau) &= \lambda_e^2 \cdot \sum_{i=0}^{n-\tau} c_i c_{i-\tau}\\ \gamma_{yy}(1) &= 2 \cdot \frac{1}{2} + \frac{1}{2} \cdot 1 = 1 + \frac{1}{2} = \frac{3}{2}\end{aligned}$$

infine:

$$\begin{aligned}\hat{a} &= \frac{\gamma_{yy}(1)}{\gamma_{yy}(0)}\\ &= \frac{\frac{3}{2}}{\frac{21}{4}} = \frac{3}{2} \cdot \frac{4}{21} = \frac{2}{7} \simeq 0.286\end{aligned}$$

per finire il processo d'identificazione è necessario stimare anche la varianza del rumore. Dalla teoria sappiamo che:

$$\varepsilon_1(t) = y(t) - \hat{y}(t|t-1; \hat{a}) = \eta(t)$$

quindi:

$$\begin{aligned}\widehat{\lambda}^2 = Var[\eta(t)] &= Var[y(t) - \hat{y}(t|t-1, \hat{a})] = \mathbb{E}\left[(y(t) - \hat{a}y(t-1))^2\right]\\ &= J(\hat{a}) = \hat{a}^2\gamma_{yy}(0) - 2\hat{a}\gamma_{yy}(1) + \gamma_{yy}(0)\\ &= \left(\frac{2}{7}\right)^2 \cdot \frac{21}{4} - 2 \cdot \frac{\cancel{2}}{7} \cdot \frac{3}{\cancel{2}} + \frac{21}{4} = \frac{135}{28} \simeq 4.82\end{aligned}$$

quindi il sistema identificato è:

$$y(t) = \frac{2}{7}y(t-1) + \eta(t) \qquad \eta(t) \sim WN\left(0, \frac{135}{28}\right)$$

che è un processo stocastico stazionario in forma canonica.

1.3 Modello $\mathcal{M}_2\,(a, b)$

Il procedimento è molto simile. Iniziamo calcolando il predittore a un passo di $\mathcal{M}_2\,(a, b)$:

$$\hat{y}\,(t|t-1; a, b) = a \cdot y\,(t-1) + b \cdot y\,(t-2)$$

NOTA: il sistema è un AR in forma canonica e quindi il predittore a un passo corrisponde alla forma ricorsiva del processo senza il rumore.
Quindi, la cifra di merito diventa:

$$J\,(a, b) = \mathbb{E}\left[(y\,(t) - \hat{y}\,(t|t-1; a, b))^2\right] = \mathbb{E}\left[(y\,(t) - ay\,(t-1) - by\,(t-2))^2\right]$$

dove $y\,(t)$, $y\,(t-1)$ e $y\,(t-2)$ sono campioni del processo $\mathcal{S}$. Con qualche passaggio:

$$\begin{aligned}
J\,(a, b) &= \mathbb{E}\left[y\,(t)^2 + a^2 y\,(t-1)^2 + b^2 y\,(t-2)^2 - 2ay\,(t)\,y\,(t-1) - 2by\,(t)\,y\,(t-2) + 2aby\,(t-1)\,y\,(t-2)\right] \\
&= \mathbb{E}\left[y\,(t)^2\right] + a^2\mathbb{E}\left[y\,(t-1)^2\right] + b^2\mathbb{E}\left[y\,(t-2)^2\right] - 2a\mathbb{E}\left[y\,(t)\,y\,(t-1)\right] - 2b\mathbb{E}\left[y\,(t)\,y\,(t-2)\right] + 2ab\mathbb{E}\left[y\,(t-1)\,y\,(t-2)\right] \\
&= \gamma_{yy}\,(0) + a^2\gamma_{yy}\,(0) + b^2\gamma_{yy}\,(0) - 2a\gamma_{yy}\,(1) - 2b\gamma_{yy}\,(2) + 2ab\gamma_{yy}\,(1)
\end{aligned}$$

dato che (come visto prima):

$$\begin{aligned}
\gamma_{yy}\,(0) &= \frac{21}{4} \\
\gamma_{yy}\,(1) &= \frac{3}{2}
\end{aligned}$$

e che:

$$\gamma_{yy}\,(2) = c_0 \cdot c_{-2} = 2 \cdot 1 = 2$$

si ottiene:

$$\begin{aligned}
J\,(a, b) &= \frac{21}{4} + a^2\frac{21}{4} + b^2\frac{21}{4} - 2a\frac{3}{2} - 2b2 + 2ab\frac{3}{2} \\
&= a^2\frac{21}{4} + b^2\frac{21}{4} + 3ab - 3a - 4b + \frac{21}{4}
\end{aligned}$$

Come nella sezione precedente, la funzione di costo $J\,(a)$ è sempre positiva (poichè dalla teoria $J\,(a) = Var\left[\varepsilon_1\,(t)\right] > 0$). Inoltre, $J\,(a)$ per modelli *AR* è sempre lineare nei parametri, perciò $J\,(a)$ sarà sempre un paraboloide. Quindi ha un minimo unico. Per trovarlo possiamo derivare la funzione di costo:

$$\begin{aligned}
\frac{\partial}{\partial a} J\,(a, b) &= \frac{\partial}{\partial a}\left(a^2\frac{21}{4} + b^2\frac{21}{4} + 3ab - 3a - 4b + \frac{21}{4}\right) \\
&= \frac{21}{2}a + 3b - 3
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial b} J\,(a, b) &= \frac{\partial}{\partial b}\left(a^2\frac{21}{4} + b^2\frac{21}{4} + 3ab - 3a - 4b + \frac{21}{4}\right) \\
&= \frac{21}{2}b + 3a - 4
\end{aligned}$$

e trovare il suo zero:

$$\begin{cases}
\frac{21}{2}\hat{a} + 3\hat{b} - 3 = 0 \\
\frac{21}{2}\hat{b} + 3\hat{a} - 4 = 0
\end{cases}$$

questo è un sistema lineare a due incognite e due equazioni:

$$\begin{cases} \frac{21}{2}\hat{a} + 3\hat{b} = 3 \\ \frac{21}{2}\hat{b} + 3\hat{a} = 4 \end{cases}$$

$$\begin{cases} \hat{a} = \frac{2}{21} \cdot 3 - \frac{2}{21} \cdot 3\hat{b} \\ - \end{cases}$$

$$\begin{cases} \hat{a} = \frac{2}{7} - \frac{2}{7}\hat{b} \\ \frac{21}{2}\hat{b} + 3\left(\frac{2}{7} - \frac{2}{7}\hat{b}\right) = 4 \end{cases}$$

$$\begin{cases} - \\ \frac{135}{14}\hat{b} = \frac{22}{7} \end{cases}$$

$$\begin{cases} - \\ \hat{b} = \frac{14}{135} \cdot \frac{22}{7} \end{cases}$$

$$\begin{cases} - \\ \hat{b} = \frac{44}{135} \end{cases}$$

$$\begin{cases} \hat{a} = \frac{26}{135} \\ \hat{b} = \frac{44}{135} \end{cases}$$

per finire il processo d'identificazione è necessario stimare anche la varianza del rumore, quindi:

$$\begin{aligned} \widehat{\lambda}^2 &= J\left(\hat{a}, \hat{b}\right) \\ &= \hat{a}^2 \frac{21}{4} + \hat{b}^2 \frac{21}{4} + 3\hat{a}\hat{b} - 3\hat{a} - 4\hat{b} + \frac{21}{4} \\ &= \left(\frac{26}{135}\right)^2 \cdot \frac{21}{4} + \left(\frac{44}{135}\right)^2 \cdot \frac{21}{4} + 3 \cdot \left(\frac{26}{135}\right)\left(\frac{44}{135}\right) - 3 \cdot \left(\frac{26}{135}\right) - 4 \cdot \left(\frac{44}{135}\right) + \frac{21}{4} \\ &= \frac{2307}{540} \end{aligned}$$

quindi il sistema identificato è:

$$y(t) = \frac{26}{135} \cdot y(t-1) + \frac{44}{135} \cdot y(t-2) + \eta(t) \qquad \eta(t) \sim WN\left(0, \frac{2307}{540}\right)$$

per controllare che il processo identificato sia effettivamente stazionario è necessario calcolarne i poli:

$$\begin{aligned} y(t) &= \frac{26}{135} \cdot y(t) z^{-1} + \frac{44}{135} \cdot y(t) z^{-2} + \eta(t) \\ y(t)\left(1 - \frac{26}{135}z^{-1} - \frac{44}{135}z^{-2}\right) &= \eta(t) \\ y(t) &= \frac{1}{1 - \frac{26}{135}z^{-1} - \frac{44}{135}z^{-2}} \cdot \eta(t) \end{aligned}$$

$$y(t) = \frac{z^2}{z^2 - \frac{26}{135}z - \frac{44}{135}} \cdot \eta(t)$$

che ha poli in:

$$\begin{aligned} p_{1,2} &= \frac{\frac{26}{135} \pm \sqrt{\left(\frac{26}{135}\right)^2 + 4 \cdot \frac{44}{135}}}{2} \\ &= \frac{13}{135} \pm \frac{\sqrt{6109}}{135} \end{aligned}$$

quindi:

$$\begin{aligned} p_1 &= \frac{13}{135} + \frac{\sqrt{6109}}{135} \simeq 0.67 < 1 \\ p_2 &= \frac{13}{135} - \frac{\sqrt{6109}}{135} \simeq -0.48 > -1 \end{aligned}$$

quindi il processo è stazionario.

1.4 Modello $\mathcal{M}_3(a, b)$

Prima di calcolare il predittore è possibile notare che questo modello è un $MA(2)$. Lo stesso tipo di modello di quello usato per generare i dati. Dalla teoria, sappiamo che 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:

$$\lim_{N \to \infty} \hat{\vartheta}_N = \vartheta$$

L'unico problema è che $\mathcal{S}$ ha $c_0 \neq 1$ mentre il modello $\mathcal{M}_3(a, b)$ ha $c_0 = 1$ perchè è considerato in forma canonica. Per risolvere il problema è necessario portare $\mathcal{S}$ in forma canonica.
Per farlo conviene ricavare la funzione di trasferimento:

$$\begin{aligned} y(t) &= 2 \cdot e(t) + \frac{1}{2} \cdot e(t) z^{-1} + e(t) z^{-2} \\ &= \left(2 + \frac{1}{2}z^{-1} + z^{-2}\right) e(t) \\ &= \frac{2z^2 + \frac{1}{2}z + 1}{z^2} e(t) \end{aligned}$$

scomponendo il numeratore:

$$\begin{aligned} z_{1,2} &= \frac{-\frac{1}{2} \pm \sqrt{\frac{1}{4} - 4 \cdot 2 \cdot 1}}{4} \\ &= -\frac{1}{8} \pm j \cdot \frac{\sqrt{31}}{8} \end{aligned}$$

$$\begin{aligned} \left|z_{1,2}\right| &= \sqrt{\left(-\frac{1}{8}\right)^2 + \left(\frac{\sqrt{31}}{8}\right)^2} \\ &= \sqrt{\frac{1}{64} + \frac{31}{64}} \end{aligned}$$

$$= \sqrt{\frac{32}{64}}$$
$$= \frac{1}{\sqrt{2}} < 1$$

quindi basta raccogliere il 2 per rendere il polinomio monico:

$$\begin{aligned} y(t) &= \frac{2z^2 + \frac{1}{2}z + 1}{z^2} e(t) \\ &= \frac{z^2 + \frac{1}{4}z + \frac{1}{2}}{z^2} \left(\underbrace{2 \cdot e(t)}_{w(t)} \right) \end{aligned}$$

dove:

$$w(t) \sim WN(0,4)$$

quindi:

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

quindi il modello identificato è:

$$\mathcal{M}_3\left(\hat{a}, \hat{b}\right) \ : \ y(t) = \eta(t) + \hat{a} \cdot \eta(t-1) + \hat{b} \cdot \eta(t-2) \qquad \eta(t) \sim WN(0,4)$$
$$\mathcal{M}_3\left(\frac{1}{4}, \frac{1}{2}\right) \ : \ y(t) = \eta(t) + \frac{1}{4} \cdot \eta(t-1) + \frac{1}{2} \cdot \eta(t-2) \qquad \eta(t) \sim WN(0,4)$$

2 Esercizio 2

Supponiamo di avere a disposizione il dataset

t	1	2	3	4
$y(t)$	-1	2	0	1

e i seguenti modelli:

$$\begin{aligned} \mathcal{M}_1(a) &: y(t) = a \cdot y(t-1) + \eta(t) && \eta(t) \sim WN\left(0, \lambda^2\right) \\ \mathcal{M}_2(a,b) &: y(t) = a \cdot y(t-1) + b \cdot y(t-2) + \eta(t) && \eta(t) \sim WN\left(0, \lambda^2\right) \\ \mathcal{M}_3(a) &: y(t) = \eta(t) + a \cdot \eta(t-1) && \eta(t) \sim WN\left(0, \lambda^2\right) \end{aligned}$$

Identificare il miglior modello in ogni classe usando la metodologia PEM, con i dati a disposizione.

2.1 Modello $\mathcal{M}_1(a)$

2.1.1 Stima di a - metodo diretto

Iniziamo calcolando il predittore a un passo di $\mathcal{M}_1(a)$:

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

NOTA: il sistema è un AR in forma canonica e quindi il predittore a un passo corrisponde alla forma ricorsiva del processo senza il rumore.
Quindi la cifra di merito diventa:

$$J_4(a) = \frac{1}{4}\sum_{t=1}^{4}\left(y(t) - a \cdot y(t-1)\right)^2$$

dato che non conosciamo $y(0)$ è necessario troncare la sommatoria:

$$J_4(a) = \frac{1}{3}\sum_{t=2}^{4}\left(y(t) - a \cdot y(t-1)\right)^2$$

dove è possibile sostiture i valori noti:

$$\begin{aligned} J_4(a) &= \frac{1}{3} \cdot \left[(2 - a \cdot (-1))^2 + (0 - a \cdot 2)^2 + (1 - a \cdot (0))^2\right] \\ &= \frac{1}{3} \cdot \left[\left(4 + 4a + a^2\right) + \left(4a^2\right) + (1)\right] \\ &= \frac{1}{3} \cdot \left[5a^2 + 4a + 5\right] \\ &= \frac{5a^2 + 4a + 5}{3} \end{aligned}$$

che ha minimo in

$$\begin{aligned} \frac{d}{da}J_4(\hat{a}) = \frac{10\hat{a} + 4}{3} &= 0 \\ \hat{a} &= -\frac{2}{5} \end{aligned}$$

2.1.2 Stima di a - formule note

Ricordando la teoria, la stima si può trovare risolvendo il sistema lineare:

$$\Phi^\top \cdot \Phi \cdot \hat{a} = \Phi^\top \cdot \boldsymbol{y}$$

dove:

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

quindi:

$$\begin{aligned} \begin{bmatrix} -1 & 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} \cdot \hat{a} = \begin{bmatrix} -1 & 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \\ (1+4+0) \cdot \hat{a} = -2+0+0 \\ \hat{a} = -\frac{2}{5} \end{aligned}$$

2.1.3 Stima di λ^2

La varianza del rumore può essere ricavata calcolando la cifra di merito coi coefficienti stimati:

$$J_4(a) = \frac{5a^2+4a+5}{3}$$

e quindi:

$$\begin{aligned} \widehat{\lambda^2} &= J(\hat{a}) = J_4\left(-\frac{2}{5}\right) \\ &= \frac{5\left(-\frac{2}{5}\right)^2 + 4\left(-\frac{2}{5}\right) + 5}{3} \\ &= \frac{5\frac{4}{25} - \frac{8}{5} + 5}{3} = \frac{\frac{4}{5} - \frac{8}{5} + \frac{25}{5}}{3} \\ &= \frac{21}{15} = \frac{7}{5} \end{aligned}$$

quindi il sistema identificato è:

$$y(t) = -\frac{2}{5} \cdot y(t-1) + \eta(t) \qquad \eta(t) \sim WN\left(0, \frac{7}{5}\right)$$

che è un processo stocastico stazionario.

2.2 Modello $\mathcal{M}_2(a,b)$

2.2.1 Stima di a e b - metodo diretto

Iniziamo calcolando il predittore a un passo di $\mathcal{M}_2(a,b)$:

$$\hat{y}(t|t-1;a,b) = a \cdot y(t-1) + b \cdot y(t-2)$$

NOTA: il sistema è un AR in forma canonica e quindi il predittore a un passo corrisponde alla forma ricorsiva del processo senza il

rumore.
Quindi la cifra di merito diventa:

$$J_4(a,b) = \frac{1}{4}\sum_{t=1}^{4}\left(y(t) - a\cdot y(t-1) - b\cdot y(t-2)\right)^2$$

dato che non conosciamo $y(0)$ e $y(-1)$ è necessario troncare la sommatoria:

$$J_4(a,b) = \frac{1}{2}\sum_{t=3}^{4}\left(y(t) - a\cdot y(t-1) - b\cdot y(t-2)\right)^2$$

dove è possibile sostiture i valori noti:

$$\begin{aligned} J_4(a,b) &= \frac{1}{2}\cdot\left[(0 - a\cdot(2) - b\cdot(-1))^2 + (1 - a\cdot 0 - b\cdot 2)^2\right] \\ &= \frac{1}{2}\cdot\left[\left(4a^2 + b^2 - 4ab\right) + \left(1 + 4b^2 - 4b\right)\right] \\ &= \frac{4a^2 + 5b^2 - 4ab - 4b + 1}{2} \end{aligned}$$

dove è possibile ricavare le due derivate parziali:

$$\begin{aligned} \frac{\partial}{\partial a}J_4(a,b) &= \frac{8a - 4b}{2} = 4a - 2b \\ \frac{\partial}{\partial b}J_4(a,b) &= \frac{10b - 4a - 4}{2} = 5b - 2a - 2 \end{aligned}$$

infine la stima si può trovare risolvendo il sistema lineare:

$$\begin{cases} 4\hat{a} - 2\hat{b} = 0 \\ 5\hat{b} - 2\hat{a} - 2 = 0 \end{cases} = \begin{cases} \hat{b} = 2\hat{a} \\ - \end{cases} = \begin{cases} - \\ 5(2\hat{a}) - 2\hat{a} - 2 = 0 \end{cases} = \begin{cases} - \\ 8\hat{a} = 2 \end{cases} = \begin{cases} \hat{b} = \dfrac{1}{2} \\ \hat{a} = \dfrac{1}{4} \end{cases}$$

2.2.2 Stima di a e b - formule dei minimi quadrati

Ricordando la teoria, la stima si può trovare risolvendo il sistema lineare:

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

dove:

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

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

quindi:

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

$$\begin{aligned}\Phi^\top \cdot \boldsymbol{y} &= \begin{bmatrix} 2 & 0 \\ -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 2 \end{bmatrix}\end{aligned}$$

quindi il sistema lineare è:

$$\begin{aligned}\begin{bmatrix} 4 & -2 \\ -2 & 5 \end{bmatrix} \cdot \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} &= \begin{bmatrix} 0 \\ 2 \end{bmatrix} \\ \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} &= \begin{bmatrix} 4 & -2 \\ -2 & 5 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \\ \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} &= \frac{1}{4 \cdot 5 - (-2) \cdot (-2)} \cdot \begin{bmatrix} 5 & 2 \\ 2 & 4 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \\ \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} &= \frac{1}{16} \cdot \begin{bmatrix} 4 \\ 8 \end{bmatrix} \\ \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} &= \begin{bmatrix} \frac{1}{4} \\ \frac{1}{2} \end{bmatrix}\end{aligned}$$

2.2.3 Stima di λ^2

La varianza del rumore può essere ricavata calcolando la cifra di merito coi coefficienti stimati:

$$J_4\left(a, b\right) = \frac{4a^2 + 5b^2 - 4ab - 4b + 1}{2}$$

e quindi:

$$\begin{aligned}J_4\left(\frac{1}{4}, \frac{1}{2}\right) &= \frac{4\left(\frac{1}{4}\right)^2 + 5\left(\frac{1}{2}\right)^2 - 4\left(\frac{1}{4}\right)\left(\frac{1}{2}\right) - 4\left(\frac{1}{2}\right) + 1}{2} \\ &= \frac{\frac{1}{4} + \frac{5}{4} - \frac{1}{2} - 2 + 1}{2} = \frac{1 - 2 + 1}{2} \\ &= 0\end{aligned}$$

quindi stiamo facendo quasi sicuramente over-fitting.
NOTA: Dato che la nostra cifra di costo è una sommatoria di due elementi e abbiamo due parametri il sistema identificato fitta perfettamente i dati. Questo avviene tutte le volte in cui il numero di dati è uguale al numero di parametri ed è una situazione da evitare.
Quindi il processo identificato è:

$$y\left(t\right) = \frac{1}{4} \cdot y\left(t-1\right) + \frac{1}{2} \cdot y\left(t-2\right)$$

che è un processo senza rumore perchè la varianza identificata è 0.

2.3 Modello $\mathcal{M}_3\left(a\right)$

In questo caso il modello non è un *AR* e quindi non è possibile usare le formule dei minimi quadrati. Tuttavia, possiamo utilizzare il metodo diretto. Per prima cosa è necessario ricavare il predittore del modello. Supponendo che il modello sia in forma canonica abbiamo:

$$\begin{aligned}y\left(t\right) &= \eta\left(t\right) + a \cdot \eta\left(t-1\right) \\ &= \eta\left(t\right) \cdot \left(1 + a \cdot z^{-1}\right)\end{aligned}$$

quindi abbiamo:

$$A(z) = 1$$
$$C(z) = 1 + a \cdot z^{-1}$$

quindi:

$$\hat{y}(t|t-1;a) = \frac{R_1(z)}{C(z)} \cdot y(t)$$
$$\hat{y}(t|t-1;a) = \frac{C(z) - A(z)}{C(z)} \cdot y(t)$$
$$\hat{y}(t|t-1;a) = \frac{a \cdot z^{-1}}{1 + a \cdot z^{-1}} \cdot y(t)$$

di conseguenza la funzione di costo diventa:

$$\begin{aligned} J_4(a) &= \frac{1}{4} \cdot \sum_{t=1}^{4} \left(y(t) - \hat{y}(t|t-1;a)\right)^2 \\ &= \frac{1}{4} \cdot \sum_{t=1}^{4} \left(\varepsilon_1(t;a)\right)^2 \end{aligned}$$

in questo caso calcolare il predittore non è banale perchè è ricorsivo. Quindi conviene calcolare l'errore di predizione ricorsivamente. Infatti, possiamo notare che:

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

quindi:

$$\varepsilon_1(t;a) \cdot \left(1 + a \cdot z^{-1}\right) = y(t)$$
$$\varepsilon_1(t;a) + a \cdot \varepsilon_1(t-1;a) = y(t)$$
$$\varepsilon_1(t;a) = y(t) - a \cdot \varepsilon_1(t-1;a)$$

di conseguenza, per poter calcolare l'errore di predizione è necessario inizializzarlo in qualche modo. Esistono varie opzioni:

- Si inizializza il predittore ponendolo uguale al predittore banale:

$$\hat{y}(1|0) = m_y \simeq \frac{1}{4} \cdot \sum_{t=1}^{4} y(t)$$
$$\varepsilon_1(1;a) = y(1) - \hat{y}(1|0)$$

- Si inizializza il predittore al primo instante noto:

$$\hat{y}(1|0) = y(1)$$
$$\varepsilon_1(1;a) = y(1) - \hat{y}(1|0)$$

- Si inizializza il predittore a 0:

$$
\begin{aligned}
\hat{y}(1|0) &= 0 \\
\varepsilon_1(1;a) &= y(1) - \hat{y}(1|0)
\end{aligned}
$$

e si taglia la sommatoria in modo che parta da 2. In modo analogo a come si fa con gli ARX.
Tuttavia, il predittore è stabile e quindi l'effetto dell'inizializzazione decresce esponenzialmente fino ad arrivare a 0. Quindi in uno scenario pratico con grosse quantità di dati questa scelta non riperquote molto la stima. In questo esercizio conviene usare la seconda opzione in modo da semplificare i calcoli.

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

quindi la cifra di merito diventa:

$$
\begin{aligned}
J_4(a) &= \frac{1}{3} \cdot \sum_{t=2}^{4} (\varepsilon_1(t;a))^2 \\
&= \frac{1}{3} \cdot \left[(2)^2 + (-2a)^2 + \left(2a^2+1\right)^2\right] \\
&= \frac{1}{3} \cdot \left[4 + 4a^2 + 4a^4 + 1 + 4a^2\right] \\
&= \frac{4a^4 + 8a^2 + 5}{3}
\end{aligned}
$$

per trovare il minimo è necessario calcolare la derivata:

$$
\frac{d}{da} J_4(a) = \frac{16a^3 + 16a}{3}
$$

che ha radici in:

$$
\begin{aligned}
\frac{d}{da} J_4(a) &= 0 \\
\frac{16}{3} \cdot \hat{a}\left(\hat{a}^2 + 1\right) &= 0
\end{aligned}
$$

quindi:

$$
\begin{aligned}
\hat{a} &= 0 \\
\hat{a} &= \pm j
\end{aligned}
$$

dato che a è reale, allora l'unica soluzione è:

$$
\hat{a} = 0
$$

bisogna controllare che questo sia effettivamente un minimo, dato che è un MA. Per farlo conviene calcolare la derivata seconda di $J_4(a)$

$$
\frac{d^2}{da^2} J_4(a) = \frac{48a^2 + 16}{3}
$$

e valutarla nel valore identificato:

$$\left.\frac{d^2}{da^2} J_4(a)\right|_{a=\hat{a}} = \frac{16}{3} > 0$$

e quindi è un minimo.

2.3.1 Stima di λ^2

La varianza del rumore può essere ricavata calcolando la cifra di merito coi coefficienti stimati:

$$J_4(\hat{a}) = \frac{4\hat{a}^4 + 8\hat{a}^2 + 5}{3}$$

e quindi:

$$J_4(0) = \frac{5}{3}$$

Quindi il processo identificato è:

$$y(t) = \eta(t) \qquad \eta(t) \sim WN\left(0, \frac{5}{3}\right)$$

che è un processo senza dinamica poichè $\hat{a} = 0$.

3 Esercizio 3

Consideriamo di avere infiniti dati presi dal processo stocastico stazionario a media nulla con la seguente funzione di autocovarianza:

$$\gamma_{yy}(\tau) = \begin{cases} 4 & \text{se } \tau = 0 \\ -2 & \text{se } |\tau| = 1 \\ 1 & \text{se } |\tau| = 2 \\ 0 & \text{se } |\tau| > 2 \end{cases}$$

Identificare il miglior modello per le seguenti classi usando la metodologia PEM:

$$\begin{aligned} \mathcal{M}_1(a) &: y(t) = \eta_1(t) + a \cdot y(t-1) & \eta_1 \sim WN\left(0, \lambda_1^2\right) \\ \mathcal{M}_2(b,c) &: y(t) = \eta_2(t) + b \cdot y(t-2) + c \cdot y(t-3) & \eta_2 \sim WN\left(0, \lambda_2^2\right) \end{aligned}$$

3.1 Identificazione di $\mathcal{M}_1(a)$

Il predittore del modello $\mathcal{M}_1$ è

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

con la seguente cifra di merito PEM

$$\begin{aligned} J(a) &= \mathbb{E}\left[(y(t) - a \cdot y(t-1))^2\right] \\ &= \mathbb{E}\left[y(t)^2 - 2 \cdot a \cdot y(t) \cdot y(t-1) + a^2 \cdot y(t-1)^2\right] \\ &= \mathbb{E}\left[y(t)^2\right] - 2 \cdot a \cdot \mathbb{E}\left[y(t) \cdot y(t-1)\right] + a^2 \cdot \mathbb{E}\left[y(t-1)^2\right] \\ &= \gamma_{yy}(0) - 2 \cdot a \cdot \gamma_{yy}(1) + a^2 \cdot \gamma_{yy}(0) \\ &= 4 - 2 \cdot a \cdot (-2) + a^2 \cdot 4 \\ &= 4 \cdot a^2 + 4 \cdot a + 4 \end{aligned}$$

Calcoliamo quindi la derivata per trovare il minimo è:

$$\begin{aligned} \frac{d}{da} J(a) &= \frac{d}{da}\left(4 \cdot a^2 + 4 \cdot a + 4\right) \\ &= 8 \cdot a + 4 \end{aligned}$$

e trovandone le radici

$$\begin{aligned} \frac{d}{da} J(a) &= 0 \\ 8 \cdot \hat{a} + 4 &= 0 \\ \hat{a} &= -\frac{1}{2} \end{aligned}$$

Si calcola la derivata seconda:

$$\left.\frac{d^2}{da^2} J(a)\right|_{a=\hat{a}} = 8$$

Poichè la derivata seconda è > 0 il punto trovato è un minimo.
La stima della varianza dell'errore di predizione λ^2 può essere fatta valutando la funzione obiettivo in $\hat{a}$:

$$\hat{\lambda}^2 = J(\hat{a})$$

$$
\begin{aligned}
&= 4 \cdot \frac{1}{4} + 4 \cdot \left(-\frac{1}{2}\right) + 4 = 1 - 2 + 4 \\
&= 3
\end{aligned}
$$

quindi il modello identificato è

$$y(t) = \eta_1(t) - \frac{1}{2} \cdot y(t-1), \qquad \eta_1(t) \sim WN(0,3)$$

3.2 Identificazione di $\mathcal{M}_2(b,c)$

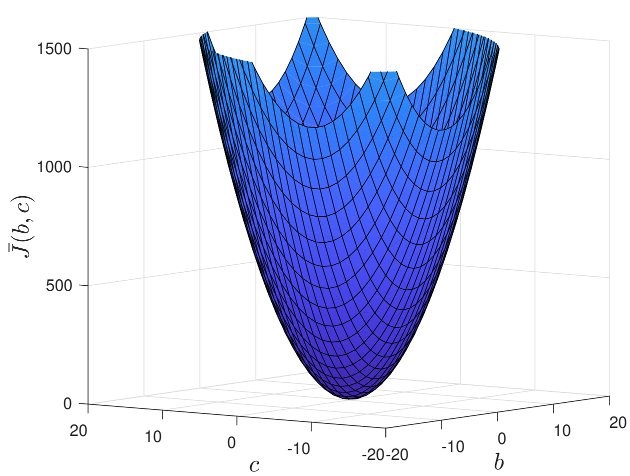
Il predittore del modello $\mathcal{M}_2$ è:

$$\hat{y}(t|t-1;b,c) = b \cdot y(t-2) + c \cdot y(t-3)$$

con la seguente cifra di merito PEM

$$
\begin{aligned}
J(b,c) &= \mathbb{E}\left[(y(t) - b \cdot y(t-2) - c \cdot y(t-3))^2\right] \\
&= \mathbb{E}\left[y(t)^2 + b^2 \cdot y(t-2)^2 + c^2 \cdot y(t-3)^2 - 2 \cdot b \cdot y(t) \cdot y(t-2) - 2 \cdot c \cdot y(t) \cdot y(t-3) + 2 \cdot b \cdot c \cdot y(t-2) \cdot y(t-3)\right] \\
&= \mathbb{E}\left[y(t)^2\right] + b^2 \cdot \mathbb{E}\left[y(t-2)^2\right] + c^2 \cdot \mathbb{E}\left[y(t-3)^2\right] - \\
&\quad - 2 \cdot b \cdot \mathbb{E}\left[y(t) \cdot y(t-2)\right] - 2 \cdot c \cdot \mathbb{E}\left[y(t) \cdot y(t-3)\right] + 2 \cdot b \cdot c \cdot \mathbb{E}\left[y(t-2) \cdot y(t-3)\right] \\
&= \gamma_{yy}(0) + b^2 \cdot \gamma_{yy}(0) + c^2 \cdot \gamma_{yy}(0) - 2 \cdot b \cdot \gamma_{yy}(2) - 2 \cdot c \cdot \gamma_{yy}(3) + 2 \cdot b \cdot c \cdot \gamma_{yy}(1) \\
&= 4 + 4 \cdot b^2 + 4 \cdot c^2 - 2 \cdot b \cdot 1 - 2 \cdot c \cdot 0 + 2 \cdot b \cdot c \cdot (-2) \\
&= 4 \cdot b^2 - 4 \cdot b \cdot c + 4 \cdot c^2 - 2 \cdot b + 4
\end{aligned}
$$

La funzione di costo $J(b,c)$ è sempre positiva (poichè dalla teoria $J(b,c) = Var[\varepsilon_1(t)] > 0$). Inoltre, $J(b,c)$ per modelli AR è sempre lineare nei parametri, perciò $J(b,c)$ sarà un paraboloide dato che i parametri da identificare siano più di uno.



Calcoliamo quindi il gradiente per trovare i minimi.

$$\nabla J(b,c) = \begin{bmatrix} \frac{\partial}{\partial b} J(b,c) \\ \frac{\partial}{\partial c} J(b,c) \end{bmatrix}$$

$$
\begin{aligned}
\frac{\partial}{\partial b} J(b,c) &= \frac{\partial}{\partial b}\left(4 \cdot b^2 - 4 \cdot b \cdot c + 4 \cdot c^2 - 2 \cdot b + 4\right) \\
&= 8 \cdot b - 4 \cdot c + 0 - 2 + 0 \\
&= 8 \cdot b - 4 \cdot c - 2
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial}{\partial c} J(b, c) &= \frac{\partial}{\partial c}\left(4 \cdot b^2 - 4 \cdot b \cdot c + 4 \cdot c^2 - 2 \cdot b + 4\right) \\
&= 0 - 4 \cdot b + 8 \cdot c - 0 + 0 \\
&= -4 \cdot b + 8 \cdot c
\end{aligned}
$$

e trovandone le radici. Si ottiene il seguente sistema lineare

$$
\begin{cases} 8 \cdot \hat{b} - 4 \cdot \hat{c} - 2 = 0 \\ -4 \cdot \hat{b} + 8 \cdot \hat{c} = 0 \end{cases}
= \begin{cases} 4 \cdot \hat{b} - 2 \cdot \hat{c} = 1 \\ -\hat{b} + 2 \cdot \hat{c} = 0 \end{cases}
= \begin{cases} 4 \cdot \hat{b} - 2 \cdot \hat{c} = 1 \\ 2 \cdot \hat{c} = b \end{cases}
= \begin{cases} 4 \cdot \hat{b} - \hat{b} = 1 \\ 2 \cdot \hat{c} = \hat{b} \end{cases}
= \begin{cases} \hat{b} = \frac{1}{3} \\ 2 \cdot \hat{c} = \frac{1}{3} \end{cases}
$$

quindi

$$
\hat{b} = \frac{1}{3}, \qquad \hat{c} = \frac{1}{6}
$$

La stima della varianza dell'errore di predizione λ^2 può essere fatta valutando la funzione obiettivo in $\hat{b}$ e $\hat{c}$

$$
\begin{aligned}
\hat{\lambda}^2 = J\left(\hat{b}, \hat{c}\right) &= 4 \cdot \hat{b}^2 - 4 \cdot \hat{b} \cdot \hat{c} + 4 \cdot \hat{c}^2 - 2 \cdot \hat{b} + 4 \\
&= 4 \cdot \left(\frac{1}{3}\right)^2 - 4 \cdot \frac{1}{3} \cdot \frac{1}{6} + 4 \cdot \left(\frac{1}{6}\right)^2 - 2 \cdot \frac{1}{3} + 4 = \frac{4}{9} - \frac{4}{18} + \frac{4}{36} - \frac{2}{3} + 4 \\
&= \frac{4}{9} - \frac{2}{9} + \frac{1}{9} - \frac{6}{9} + \frac{36}{9} = \frac{33}{9} = \frac{11}{3}
\end{aligned}
$$

quindi il modello identificato è

$$
y(t) = \eta_2(t) + \frac{1}{3} \cdot y(t-2) + \frac{1}{6} \cdot y(t-3), \qquad \eta_2(t) \sim WN\left(0, \frac{11}{3}\right)
$$

4 Esercizio 4

Supponiamo di avere i dati:

$$u(t) = \begin{cases} 0 & \text{se } t \text{ è pari} \\ 2 & \text{se } t \text{ è dispari} \end{cases} \qquad t = 1, \dots, 100$$

$$y(t) = \begin{cases} 1 & \text{se } t \text{ è pari} \\ -1 & \text{se } t \text{ è dispari} \end{cases} \qquad t = 1, \dots, 100$$

Identificare il miglior modello per le seguenti classi usando la metodologia PEM:

$$\mathcal{M}_1(a,b) \;:\; y(t) = a \cdot y(t-1) + b \cdot u(t-1) + \eta(t) \qquad \eta(t) \sim WN\left(0, \lambda^2\right)$$
$$\mathcal{M}_2(a,b) \;:\; y(t) = a \cdot u(t-1) + b \cdot u(t-2) + \eta(t) \qquad \eta(t) \sim WN\left(0, \lambda^2\right)$$

4.1 Identificazione di $\mathcal{M}_1(a,b)$

Il predittore è:

$$\hat{y}(t|t-1) = a \cdot y(t-1) + b \cdot u(t-1)$$

quindi la funzione di costo diventa:

$$J_{100}(a,b) = \frac{1}{100} \sum_{t=1}^{100} \left(y(t) - a \cdot y(t-1) - b \cdot u(t-1)\right)^2$$

dato che $y(0)$ non è noto:

$$J_{100}(a,b) = \frac{1}{99} \sum_{t=2}^{100} \left(y(t) - a \cdot y(t-1) - b \cdot u(t-1)\right)^2$$

dato che i dati sono divisi tra pari e dispari conviene separare la sommatoria in due parti:

$$\begin{aligned}
J_{100}(a,b) &= \frac{1}{99}\left[\sum_{t=2, t \text{ pari}}^{100} \left(y(t) - a \cdot y(t-1) - b \cdot u(t-1)\right)^2 + \sum_{t=3, t \text{ dispari}}^{99} \left(y(t) - a \cdot y(t-1) - b \cdot u(t-1)\right)^2\right] \\
&= \frac{1}{99}\left[\sum_{t=1}^{50} \left(y(2t) - a \cdot y(2t-1) - b \cdot u(2t-1)\right)^2 + \sum_{t=1}^{49} \left(y(2t+1) - a \cdot y(2t+1-1) - b \cdot u(2t+1-1)\right)^2\right] \\
&= \frac{1}{99}\left[\sum_{t=1}^{50} \left(y(2t) - a \cdot y(2t-1) - b \cdot u(2t-1)\right)^2 + \sum_{t=1}^{49} \left(y(2t+1) - a \cdot y(2t) - b \cdot u(2t)\right)^2\right] \\
&= \frac{1}{99}\left[\sum_{t=1}^{50} \left(1 - a \cdot (-1) - b \cdot 2\right)^2 + \sum_{t=1}^{49} \left((-1) - a \cdot 1 - b \cdot 0\right)^2\right] \\
&= \frac{1}{99}\left[\sum_{t=1}^{50} (1 + a - 2b)^2 + \sum_{t=1}^{49} (-1 - a)^2\right] \\
&= \frac{1}{99}\left[\sum_{t=1}^{50} \left(1 + a^2 + 4b^2 + 2a - 4b - 4ab\right) + \sum_{t=1}^{49} \left(1 + a^2 + 2a\right)\right] \\
&= \frac{1}{99}\left[50 \cdot \left(1 + a^2 + 4b^2 + 2a - 4b - 4ab\right) + 49 \cdot \left(1 + a^2 + 2a\right)\right] \\
&= \frac{1}{99}\left[a^2 \cdot (50 + 49) + b^2 \cdot 200 + a \cdot (100 + 98) + b \cdot (-200) + ab\,(-200) + (50 + 49)\right] \\
&= \frac{1}{99}\left[99a^2 + 200b^2 + 198a - 200b - 200ab + 99\right] \\
&= \frac{99a^2 + 200b^2 + 198a - 200b - 200ab + 99}{99}
\end{aligned}$$

Poichè il modello è un ARX valgono le medesime considerazioni dei modelli AR (lineare nei parametri e funzione di costo positiva), quindi si procede nel medesimo modo:

$$\frac{\partial}{\partial a} J_{100}(a,b) = \frac{198a + 198 - 200b}{99}$$
$$\frac{\partial}{\partial b} J_{100}(a,b) = \frac{400b - 200 - 200a}{99}$$

quindi il minimo si trova:

$$\begin{cases} \frac{198\hat{a} + 198 - 200\hat{b}}{\cancel{99}} = 0 \\ \frac{400\hat{b} - 200 - 200\hat{a}}{\cancel{99}} = 0 \end{cases}$$

$$\begin{cases} \hat{a} = \frac{200\hat{b} - 198}{198} \\ 400\hat{b} - 200 - 200\hat{a} = 0 \end{cases}$$

$$\begin{cases} - \\ 400\hat{b} - 200 - 200\left(\frac{200\hat{b} - 198}{198}\right) = 0 \end{cases}$$

$$\begin{cases} - \\ \left(400 - \frac{400}{198}\right)\hat{b} = 200 - 200 \end{cases}$$

$$\begin{cases} - \\ \hat{b} = 0 \end{cases}$$

$$\begin{cases} \hat{a} = \frac{200 \cdot 0 - 198}{198} \\ \hat{b} = 0 \end{cases}$$

$$\begin{cases} \hat{a} = -1 \\ \hat{b} = 0 \end{cases}$$

infine:

$$\begin{aligned} \widehat{\lambda}^2 &= J_{100}\left(\hat{a}, \hat{b}\right) = J_{100}(-1, 0) \\ &= \frac{99(-1)^2 + 200(0)^2 + 198(-1) - 200(0) - 200(-1)(0) + 99}{99} \\ &= \frac{99 - 198 + 99}{99} \\ &= 0 \end{aligned}$$

$\widehat{\lambda}^2 = 0$ quando si ha overfitting oppure sistema deterministico. In questo caso è un sistema deterministico ed identifichiamo:

$$y(t) = -y(t-1)$$

che ha senso. L'uscita oscilla tra $+1$ e -1.

4.2 Identificazione di $\mathcal{M}_2(a, b)$

Il predittore è:

$$\hat{y}(t|t-1) = a \cdot u(t-1) + b \cdot u(t-2)$$

quindi la funzione di costo diventa:

$$J_{100}(a, b) = \frac{1}{100} \sum_{t=1}^{100} (y(t) - a \cdot u(t-1) - b \cdot u(t-2))^2$$

dato che $u(t)$ con $t \leq 0$ non sono noti:

$$J_{100}(a, b) = \frac{1}{98} \sum_{t=3}^{100} (y(t) - a \cdot u(t-1) - b \cdot u(t-2))^2$$

come prima conviene la sommatoria in due parti:

$$\begin{aligned}
J_{100}(a, b) &= \frac{1}{98} \left[\sum_{t=3, t \text{ dispari}}^{99} (y(t) - a \cdot u(t-1) - b \cdot u(t-2))^2 + \sum_{t=4, t \text{ pari}}^{100} (y(t) - a \cdot u(t-1) - b \cdot u(t-2))^2 \right] \\
&= \frac{1}{98} \left[\sum_{t=1}^{49} (y(2t) - a \cdot u(2t-1) - b \cdot u(2t-2))^2 + \sum_{t=1}^{49} (y(2t+1) - a \cdot u(2t+1-1) - b \cdot u(2t+1-2))^2 \right] \\
&= \frac{1}{98} \left[\sum_{t=1}^{49} (y(2t) - a \cdot u(2t-1) - b \cdot u(2t-2))^2 + \sum_{t=1}^{49} (y(2t+1) - a \cdot u(2t) - b \cdot u(2t-1))^2 \right] \\
&= \frac{1}{98} \left[\sum_{t=1}^{49} (1 - a \cdot 2 - b \cdot 0)^2 + \sum_{t=1}^{49} (-1 - a \cdot 0 - b \cdot 2)^2 \right] \\
&= \frac{1}{98} \left[\sum_{t=1}^{49} (1 - 2a)^2 + \sum_{t=1}^{49} (-1 - 2b)^2 \right] \\
&= \frac{1}{98} \left[\sum_{t=1}^{50} \left(1 + 4a^2 - 4a\right) + \sum_{t=1}^{49} \left(1 + 4b^2 + 4b\right) \right] \\
&= \frac{1}{98} \left[49 \left(1 + 4a^2 - 4a\right) + 49 \left(1 + 4b^2 + 4b\right) \right] \\
&= \frac{1}{98} \left[196a^2 + 196b^2 - 196a + 196b + 98 \right] \\
&= \frac{196a^2 + 196b^2 - 196a + 196b + 98}{98}
\end{aligned}$$

Poichè il modello è un ARX valgono le medesime considerazioni dei modelli AR (lineare nei parametri e funzione di costo positiva), quindi si procede nel medesimo modo:

$$\frac{\partial}{\partial a} J_{100}(a, b) = \frac{392a - 196}{98}$$
$$\frac{\partial}{\partial b} J_{100}(a, b) = \frac{392b + 196}{98}$$

Quindi il minimo si trova:

$$\begin{cases} \dfrac{392\hat{a} - 196}{\cancel{98}} = 0 \\ \dfrac{392\hat{b} + 196}{\cancel{98}} = 0 \end{cases}$$

$$\begin{cases} \hat{a} = \frac{196}{392} = \frac{1}{2} \\ \hat{b} = -\frac{196}{392} = -\frac{1}{2} \end{cases}$$

infine:

$$\begin{aligned} \widehat{\lambda}^2 &= J_{100}\left(\hat{a}, \hat{b}\right) = J_{100}\left(\frac{1}{2}, -\frac{1}{2}\right) \\ &= \frac{196\left(\frac{1}{2}\right)^2 + 196\left(-\frac{1}{2}\right)^2 - 196\left(\frac{1}{2}\right) + 196\left(-\frac{1}{2}\right) + 98}{98} \\ &= \frac{(49 + 49 - 98) + (-98 + 98)}{98} = 0 \end{aligned}$$

$\widehat{\lambda}^2 = 0$ quindi, analogamente a prima, o si ha overfitting oppure il sistema è deterministico.
In questo caso il sistema è deterministico e corrisponde a:

$$y\left(t\right) = \frac{1}{2} \cdot u\left(t-1\right) - \frac{1}{2} \cdot u\left(t-2\right)$$

5 Esercizio 5

Consideriamo di avere infiniti dati presi dal processo stocastico stazionario con funzione di autocovarianza definita nel seguente modo:

$$\gamma_{yy}(\tau) = \begin{cases} 5 & se \ \tau = 0 \\ 3 & se \ |\tau| = 1 \\ 1 & se \ |\tau| = 2 \\ 0 & se \ |\tau| > 2 \end{cases}$$

Identificare il miglior modello per la seguente classe usando la metodologia PEM:

$$\mathcal{M}_1(a) \ : \ y(t) = (1-a) \cdot y(t-1) + a \cdot y(t-2) + \eta(t) \qquad \eta(t) \sim WN\left(0, \lambda^2\right)$$

5.1 Identificazione di $\mathcal{M}_1(a)$

In questo caso il predittore è:

$$\hat{y}(t|t-1) = (1-a) \cdot y(t-1) + a \cdot y(t-2)$$

NOTA: il sistema è un AR in forma canonica e quindi il predittore a un passo corrisponde alla forma ricorsiva del processo senza il rumore.

Quindi la cifra di merito diventa:

$$\begin{aligned} J(a) &= \mathbb{E}\left[(y(t) - \hat{y}(t|t-1; a))^2\right] \\ &= \mathbb{E}\left[(y(t) - (1-a) \cdot y(t-1) - a \cdot y(t-2))^2\right] \end{aligned}$$

dove $y(t)$ e $y(t-1)$ sono campioni del processo $\mathcal{S}$. Con qualche passaggio:

$$\begin{aligned} J(a) &= \mathbb{E}\left[y(t)^2\right] + (1-a)^2 \cdot \mathbb{E}\left[y(t-1)^2\right] + a^2 \cdot \mathbb{E}\left[y(t-2)^2\right] \\ &\quad - 2(1-a) \cdot \mathbb{E}\left[y(t) \cdot y(t-1)\right] - 2a \cdot \mathbb{E}\left[y(t) \cdot y(t-2)\right] + 2a \cdot (1-a) \cdot \mathbb{E}\left[y(t-1) \cdot y(t-2)\right] \\ &= \gamma_{yy}(0) + (1-a)^2 \cdot \gamma_{yy}(0) + a^2 \cdot \gamma_{yy}(0) - 2(1-a) \cdot \gamma_{yy}(1) - 2a \cdot \gamma_{yy}(2) + 2\left(a - a^2\right) \cdot \gamma_{yy}(1) \\ &= 5 + 5\left(1 - 2a + a^2\right) + 5a^2 - 6(1-a) - 2a + 6\left(a - a^2\right) \\ &= (5 + 5 - 6)\, a^2 + (-10 + 6 - 2 + 6)\, a + (5 + 5 - 6) \\ &= 4a^2 + 4 \end{aligned}$$

che ha minimo in:

$$\frac{d}{da} J(a) = 8a$$

$$\begin{aligned} 8\hat{a} &= 0 \\ \hat{a} &= 0 \end{aligned}$$

Poichè, la funzione di costo $J(a)$ è sempre positiva (poichè dalla teoria $J(a) = Var[\varepsilon_1(t)] > 0$) ed $J(a)$ per modelli AR è sempre lineare nei parametri, perciò $J(a)$ sarà una parabola. Mentre la varianza del rumore diventa:

$$J(\hat{a}) = 4 \cdot 0 + 4 = 4$$

quindi il modello identificato è:

$$\begin{aligned} y(t) &= (1-0) \cdot y(t-1) + 0 \cdot y(t-2) + \eta(t) \qquad & \eta(t) \sim WN(0, 4) \\ y(t) &= y(t-1) + \eta(t) \qquad & \eta(t) \sim WN(0, 4) \end{aligned}$$

Tuttavia il modello stimato è un random walk ($a_0 = 1$) e quindi il sistema non è asintoticamente stabile, ma solo stabile.

6 Esercizio 6

Supponiamo di avere il processo stocastico stazionario:

$$\mathcal{S} \ : \ y(t) = 0.25 \cdot y(t-2) + 2 \cdot e(t) \qquad e(t) \sim WN(0,1)$$

e la classe:

$$\mathcal{M}(a) \ : \ y(t) = (0.5 - a) \cdot y(t-1) + 0.5 \cdot a \cdot y(t-2) + \eta(t) \qquad \eta(t) \sim WN\left(0, \lambda^2\right)$$

Identificare il miglior modello per la classe usando la metodologia PEM, assumendo che si abbiano infiniti dati a disposizione ($N \longrightarrow \infty$).

6.1 Modello $\mathcal{M}(a)$

Il modello $\mathcal{S}$ è un $AR(2)$, verifichiamo che sia in forma canonica:

$$y(t) = \frac{2}{1 - 0.25z^{-2}} e(t)$$

$$C(z) = 2$$

$$A(z) = 1 - 0.25z^{-2}$$

Il modello non è in forma canonica visto che $C(z)$ non è monico, si definisce quindi:

$$\begin{aligned} \tilde{e}(t) &= 2e(t) \\ \tilde{e}(t) &\sim WN(0,4) \end{aligned}$$

Quindi la rappresentazione canonica di $\mathcal{S}$ è:

$$y(t) = \frac{1}{1 - 0.25z^{-2}} \tilde{e}(t)$$

Si noti che anche il modello $\mathcal{M}(a)$ è un $AR(2)$, quindi si potrebbe applicare la teoria dell'identificazione PEM, ovvero: l'identificatore PEM è asintoticamente corretto se la classe dei modelli contiene esattamente il modello usato per generare i dati (vedi 1.4).
Si calcola il predittore ad un passo di $\mathcal{M}(a)$:

$$\hat{y}(t \mid t-1; a) = (0.5 - a)\, y(t-1) + 0.5ay(t-2)$$

Quindi si calcola l'errore di predizione ad un passo:

$$\begin{aligned} \varepsilon_1(t;a) &= y(t) - \hat{y}(t \mid t-1; a) = y(t) - (0.5-a)\, y(t-1) + 0.5ay(t-2) \\ &= \left(1 - (0.5-a)\, z^{-1} - 0.5az^{-2}\right) y(t) \end{aligned}$$

Sostituendo $y(t) = \frac{1}{1-0.25z^{-2}} \tilde{e}(t)$ in $\varepsilon(t;a)$, si ottiene:

$$\varepsilon_1(t;a) = \frac{\left(1 - (0.5-a)\, z^{-1} - 0.5az^{-2}\right)}{1 - 0.25z^{-2}} \tilde{e}(t) = F(z;a)\, \tilde{e}(t)$$

Si noti che la varianza di $\varepsilon_1(t;a)$ è minimo quando $\varepsilon_1(t;a) = \tilde{e}(t)$. Questo è vero se il denominatore e il numeratore di $F(z;a)$

sono uguali (ovvero per $F(z;a) = 1$).
Si procede con risolvere il seguente sistema di 2 equazioni e una incognita:

$$\begin{cases} -(0.5 - \hat{a}) = 0 \\ 0.5\hat{a} = 0.25 \end{cases}$$

Il sistema ha un'unica solizione: $\hat{a} = 0.5$.
Si calcola ora la varianza del rumore:

$$\widehat{\lambda}^2 = Var\left[\tilde{e}(t)\right] = Var\left[\varepsilon_1(t;\hat{a})\right] = \mathbb{E}\left[\left(\frac{\left(1 - (0.5 - 0.5)z^{-1} - 0.5 \cdot 0.5z^{-2}\right)}{1 - 0.25z^{-2}}\tilde{e}(t)\right)^2\right]$$

$$\widehat{\lambda}^2 = \mathbb{E}\left[(\tilde{e}(t))^2\right]$$

$$\widehat{\lambda}^2 = 4$$

Quindi il modello identificato è:

$$y(t) = 0.25y(t-2) + \eta(t) \qquad \eta(t) \sim WN(0,4)$$

7 Esercizio 7

Si identifichi il miglior modello della seguente classe usando la metodologia PEM:

$$\mathcal{M}(a) \ : \ y(t) = \eta(t) + a \cdot \eta(t-1) \qquad \eta(t) \sim WN\left(0, \lambda^2\right)$$

avendo a disposizione tre misurazioni acquisite del sistema $\mathcal{S}$:

$$\begin{aligned} y(1) &= 2 \\ y(2) &= 0 \\ y(3) &= 2 \end{aligned}$$

7.1 Modello $\mathcal{M}(a)$

Supponendo che il modello sia in forma canonica abbiamo :

$$\begin{aligned} y(t) &= \eta(t) + a \cdot \eta(t-1) \\ &= \eta(t) \cdot \left(1 + a \cdot z^{-1}\right) \end{aligned}$$

quindi abbiamo:

$$\begin{aligned} A(z) &= 1 \\ C(z) &= 1 + a \cdot z^{-1} \end{aligned}$$

Il modello è in forma canonica se $|a| < 1$.
Calcoliamo il predittore ad un passo:

$$\begin{aligned} \hat{y}(t|t-1;a) &= \frac{R_1(z)}{C(z)} \cdot y(t) \\ \hat{y}(t|t-1;a) &= \frac{C(z) - A(z)}{C(z)} \cdot y(t) \\ \hat{y}(t|t-1;a) &= \frac{az^{-1}}{1 + az^{-1}} \cdot y(t) \\ \hat{y}(t|t-1;a) &= -a\hat{y}(t-1|t-2;a) + ay(t-1) \end{aligned}$$

Si calcola la seguente tabella con i dati a disposizione (si decide di inizializzare $y(0) = 0$ e $\hat{y}(0|-1) = 0$):

t	$y(t)$	$\hat{y}(t\|t-1;a)$
0	0	0
1	2	0
2	0	$2a$
3	2	$-2a^2$

Si calcola la funzione di costo (non troncando la sommatoria, o troncandola supponendo $y(0)$ come parte del dataset a disposizione):

$$\begin{aligned} J_3(a) &= \frac{1}{3} \sum_{t=1}^{3} \left(y(t) - \hat{y}(t|t-1)\right)^2 = \frac{1}{3}\left[(2-0)^2 + (0-2a)^2 + \left(2+2a^2\right)^2\right] \\ &= \frac{1}{3}\left[4 + 4a^2 + 4 + 8a^2 + 4a^4\right] \end{aligned}$$

$$= \frac{4}{3}\left[a^4 + 3a^2 + 2\right]$$

Si calcola la derivata della funzione di costo:

$$\frac{d}{da} J_3(a) = \frac{4}{3}\left[4a^3 + 6a\right]$$

Si calcola il minimo:

$$\frac{8}{3}\left[2\hat{a}^3 + 3\hat{a}\right] = 0$$
$$\hat{a}\left(2\hat{a}^2 + 3\right) = 0$$

L'equazione ammette tre soluzioni:

$$\hat{a}_1 = 0$$
$$\hat{a}_2 = j\sqrt{\frac{3}{2}}$$
$$\hat{a}_3 = -j\sqrt{\frac{3}{2}}$$

Si ricorda che $\hat{a} \in \mathbb{R}$, quindi si devono scartare le soluzioni $\hat{a}_2$ e $\hat{a}_3$. L'unica soluzione fatttibile è di conseguenza $\hat{a}_1$. Dato che è un MA si deve verificare che la soluzione è un minimo
Si analizza la derivata seconda di $J_3(a)$:

$$\frac{d^2}{da^2} J_3(a) = \frac{4}{3}\left[12a^2 + 6\right]$$

Con $\hat{a}_1 = 0$:

$$\frac{24}{3} > 0$$

Quindi $\hat{a}_1$ è un punto di minimo.
Si calcola ora la varianza del rumore

$$\widehat{\lambda}^2 = \bar{J}(\hat{a}) = \frac{4}{3}[2]$$
$$\widehat{\lambda}^2 = \frac{8}{3}$$

Il sistema identificato è:

$$y(t) = \eta(t) \qquad \eta(t) \sim WN\left(0, \frac{8}{3}\right)$$

che è un processo senza dinamica.

8 Esercizio 8

Si consideri il sistema:

$$\mathcal{S}_1 \ : \ y(t) = e(t) \qquad\qquad e(t) \sim WN(0,1)$$

Identificare il miglior modello per la seguente classe usando la metodologia PEM:

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

Valutare cosa succede se il sistema fosse:

$$\mathcal{S}_2 \ : \ y(t) = e(t) + 2e(t-1) \qquad\qquad e(t) \sim WN(0,1)$$

8.1 Modello $\mathcal{M}(a)$ con sistema $\mathcal{S}_1$:

Il modello $\mathcal{M}(a)$ è un ARMA:

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

$$C(z) = 1 + \frac{1}{2}z^{-1}$$

$$A(z;a) = 1 + az^{-1}$$

Il sistema è in forma canonica se $|a| < 1$.
Si calcola il predittore ad un passo:

$$\begin{aligned}
\hat{y}(t|t-1;a) &= \frac{R_1(z)}{C(z)} \cdot y(t) \\
&= \frac{C(z) - A(z;a)}{C(z)} y(t) \\
&= \frac{1 + \frac{1}{2}z^{-1} - 1 - az^{-1}}{1 + \frac{1}{2}z^{-1}} y(t) \\
&= \frac{\left(\frac{1}{2} - a\right) z^{-1}}{1 + \frac{1}{2}z^{-1}} y(t)
\end{aligned}$$

L'errore di predizione è:

$$\begin{aligned}
\varepsilon_1(t;a) = y(t) - \hat{y}(t|t-1;a) &= y(t) - \frac{\left(\frac{1}{2} - a\right) z^{-1}}{1 + \frac{1}{2}z^{-1}} y(t) \\
&= \frac{1 + \frac{1}{2}z^{-1} - \frac{1}{2}z^{-1} + az^{-1}}{1 + \frac{1}{2}z^{-1}} y(t) = \frac{1 + az^{-1}}{1 + \frac{1}{2}z^{-1}} y(t)
\end{aligned}$$

Dato che vale $y(t) = e(t)$, si ha che:

$$\varepsilon_1(t;a) = \frac{1 + az^{-1}}{1 + \frac{1}{2}z^{-1}} e(t)$$

Si noti che la varianza di $\varepsilon_1(t;a)$ è minima quando il valore di $\hat{a}$ è tale che $\varepsilon_1(t;a) = e(t)$.
Quindi:

$$\hat{a} = \frac{1}{2}$$

Si calcola ora la varianza del rumore:

$$\begin{aligned}
\widehat{\lambda}^2 &= Var\left[e(t)\right] = Var\left[\varepsilon_1(t;\hat{a})\right] = Var\left[\frac{1+\frac{1}{2}z^{-1}}{1+\frac{1}{2}z^{-1}}e(t)\right] \\
\widehat{\lambda}^2 &= \mathbb{E}\left[(e(t))^2\right] \\
\widehat{\lambda}^2 &= 1
\end{aligned}$$

Il sistema identificato è:

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

8.2 Modello $\mathcal{M}(a)$ con sistema $\mathcal{S}_2$:

Dato che vale $y(t) = e(t) + 2e(t-1)$, l'errore di predizione diventa:

$$\begin{aligned}
\varepsilon_1(t;a) &= \frac{1+az^{-1}}{1+\frac{1}{2}z^{-1}}y(t) \\
&= \frac{1+az^{-1}}{1+\frac{1}{2}z^{-1}}\left(1+2z^{-1}\right)e(t) \\
&= 2\cdot \underbrace{\frac{1}{2}\left(\frac{1+2z^{-1}}{1+\frac{1}{2}z^{-1}}\right)}_{filtro\ passa-tutto}\left(1+az^{-1}\right)e(t) \\
&= 2\left(1+az^{-1}\right)e(t) \\
&= 2e(t) + 2ae(t-1)
\end{aligned}$$

Quindi la cifra di merito diventa:

$$\begin{aligned}
J(a) &= \mathbb{E}\left[(2e(t) + 2ae(t-1))^2\right] \\
&= 4\underbrace{\mathbb{E}\left[e(t)^2\right]}_{\gamma_{ee}(0)} + 4a^2\underbrace{\mathbb{E}\left[e(t-1)^2\right]}_{\gamma_{ee}(0)} + 8a\underbrace{\mathbb{E}\left[e(t)\,e(t-1)\right]}_{\gamma_{ee}(1)} \\
&= 4\cdot 1 + 4a^2\cdot 1 + 8a\cdot 0 \\
&= 4a^2 + 4
\end{aligned}$$

Si calcola la derivata della funzione di costo:

$$\frac{d}{da}J(a) = 8a$$

Si calcola il minimo:

$$\hat{a} = 0$$

Si calcola ora la varianza del rumore:

$$\widehat{\lambda}^2 = J(\hat{a}) = 4 \cdot 0^2 + 4$$
$$\widehat{\lambda}^2 = 4$$

Il sistema identificato è:

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