# COMPITO DI
# IDENTIFICAZIONE DEI MODELLI ED ANALISI DEI DATI
# 11 febbraio 2013

**Domanda 1.** Sia $X$ una variabile aleatoria Gaussiana $\mathcal{N}(5,1)$ e siano $Y_1$ e $Y_2$ due "misure rumorose" di $X$ nel senso che $Y_1 = X + W_1$ e $Y_2 = X + W_2$ dove $W_1$ e $W_2$ sono Gaussiane, indipendenti da $X$, a media zero, varianza $\sigma_1^2 = \sigma_2^2 = 1$ e covarianza $\sigma_{12} = 1/2$.

1. Si trovi la densità di $Y := [Y_1\ Y_2]^\top$ e la densità congiunta di $X$ e $Y$.
2. Si calcolino i valori attesi condizionati $\hat{X}_1 = E[X|Y_1]$ e $\hat{X}_{12} = E[X|Y_1, Y_2]$
3. Si descriva come calcolare $P\left[|X - \hat{X}_{12}| > \epsilon | Y_1, Y_2\right]$. Si può dire (senza fare calcoli) che $P\left[|X - \hat{X}_1| > \epsilon | Y_1\right] < P[|X-5| > \epsilon]$? Cosa succede per $P\left[|X - \hat{X}_{12}| > \epsilon | Y_1, Y_2\right]$. Si commenti il risultato.

**Domanda 2.** Si assuma che il processo $y(t)$ sia generato dal modello dinamico

$$y(t) = a_1 y(t-1) + b_1 u(t-1) + e_0(t) \quad |a_1| < 1 \quad t \in \mathbb{Z}$$

dove $e_0(t)$ è un rumore bianco, a media nulla e varianza $\sigma^2$. Il processo di "ingresso" $u(t)$ è descritto dal modello

$$u(t) = v(t) + v(t-1)$$

dove $v(t)$ è un processo stazionario bianco, a media nulla ed indipendente da $e_0(t)$. Per fare "identificazione", si utilizzi il criterio di minimizzazione dell'errore di predizione (PEM) ed il modello

$$y(t) = \theta u(t-1) + e(t) \quad \theta \in \mathbb{R}$$

dove si assume $e(t)$ bianco.

1. Si calcoli la varianza del processo $y(t)$;
2. Assumendo siano disponibili dati $y(t), u(t)$, $t \in [1, N]$, si imposti il problema di identificazione utilizzando il criterio di minimizzazione dell'errore di predizione (PEM); si trovi lo stimatore $\hat{\theta}_N$ del parametro $\theta$, scrivendo chiaramente quale funzione dei dati si deve minimizzare.
3. Si dica se lo stimatore PEM $\hat{\theta}_N$ converge e se è uno stimatore consistente di $b_1$ (nel caso in cui non sia consistente non è necessario calcolare esplicitamente il valore a cui converge)

**Domanda 3.** Si consideri il modello lineare nei parametri

$$y(t) = \sum_{i=1}^{4} \theta_i x^{i-1}(t) + \theta_5 \log(x(t)) + d(t) \tag{1}$$

dove $x(t) > 0$ è una grandezza misurabile e $d(t)$ è un rumore bianco. Si definisca il parametro vettoriale $\theta := [\theta_1\ \theta_2\ \theta_3\ \theta_4\ \theta_5]^\top \in \mathbb{R}^5$.

1. Si scrivano le equazioni del filtro di Kalman per la stima ricorsiva dei parametri $\theta$ nel modello lineare (1).

2. Supponendo che i parametri $\theta$ varino (lentamente) nel tempo, e avendo a disposizione dei dati di “test” $y(t), x(t)$, $t = 1, .., N$, come è opportuno “tarare” il filtro di Kalman?

3. Si descriva, in maniera qualitativa, come si potrebbe utilizzare il filtro di Kalman per rivelare un brusco cambiamento (dovuto, ad esempio, ad un malfunzionamento di un sistema) dei valori $\theta$ (incogniti).

4. Si illustri come si dovrebbe modificare l’algoritmo descritto sopra basato sul filtro di Kalman se $d(t)$, invece di essere bianco, fosse un processo colorato a spettro noto?

**Domanda 4.** Si enunci e si dimostri la regola di Bayes.

# SOLUZIONI

## ESERCIZIO 1 :

1. Le variabili $X, Y_1, Y_2$ soddisfano l'equazione

$$\begin{bmatrix} X \\ Y_1 \\ Y_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}}_{A} \begin{bmatrix} X \\ W_1 \\ W_2 \end{bmatrix}$$

Poiché $X, W_1, W_2$ sono Gaussiane ed indipendenti, il vettore

$$V := \begin{bmatrix} X \\ W_1 \\ W_2 \end{bmatrix}$$ è Gaussiano

con media $\mathbb{E}[V] = \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}$

e matrice varianza

$$\mathrm{Var}\{V\} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1/2 \\ 0 & 1/2 & 1 \end{bmatrix}$$

$\mathrm{Var}\{X\}$

$\sigma_1^2$

$\sigma_2^2$

$\sigma_{12}$

Di conseguenza anche

$$Z := \begin{bmatrix} X \\ Y_1 \\ Y_2 \end{bmatrix} = AV \text{ è un vettore Gaussiano}$$

di medie $\mathbb{E}[Z] = A\,\mathbb{E}[V] = A \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}$

$$= \begin{bmatrix} 5 \\ 5 \\ 5 \end{bmatrix}$$

e matrice Varianza

$$\mathrm{Var}\{Z\} = A\,\mathrm{Var}\{V\}\,A^T = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3/2 \\ 1 & 3/2 & 2 \end{bmatrix}$$

Poiché $Y := \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$ si può

scrivere nella forma:

$$Y = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} Z$$ , anche $Y$ è

un vettore GAUSSIANO con media

$$\mathbb{E}[Y] = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbb{E}[Z] = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$ e

matrice varianza

$$\mathrm{Var}[Y] = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathrm{Var}\{Z\} \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3/2 \\ 3/2 & 2 \end{bmatrix}$$

2. Poiché le variabili in gioco sono congiuntamente GAUSSIANE basta applicare la nota formula

$$\mathbb{E}[X|Y] = \mathrm{cov}(X,Y)\,\mathrm{Var}\{Y\}^{-1}(Y - m_Y) + m_X$$

$$= m_X + \Sigma_{XY} \Sigma_{YY}^{-1} (Y - m_Y)$$

e si ottiene:

$$\hat{X}_1 := \mathbb{E}[X \mid Y_1] = 5 + \frac{1}{2}(Y_1 - 5)$$

$$\hat{X}_{12} := \mathbb{E}[X \mid Y_1, Y_2] =$$

$$= 5 + \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3/2 \\ 3/2 & 2 \end{bmatrix}^{-1} \begin{bmatrix} Y_1 - 5 \\ Y_2 - 5 \end{bmatrix}$$

$$= 5 + \frac{\begin{bmatrix} 1 & 1 \end{bmatrix}}{4 - \frac{9}{4}} \begin{bmatrix} 2 & -3/2 \\ -3/2 & 2 \end{bmatrix} \begin{bmatrix} Y_1 - 5 \\ Y_2 - 5 \end{bmatrix}$$

$$= 5 + \frac{4}{7} \cdot \begin{bmatrix} 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} Y_1 - 5 \\ Y_2 - 5 \end{bmatrix}$$

$$= 5 + \frac{4}{7} \cdot \left[ \frac{Y_1 + Y_2}{2} - 5 \right]$$

3: Basta ricordare che:

$$P\left[|X-\hat{X}_{12}|>\epsilon \mid Y_1, Y_2\right] = \int_{|x-\hat{x}_{12}|>\epsilon} f_{X|Y_1,Y_2}(x/Y_1 Y_2)\,dx$$

dove $f_{X|Y_1,Y_2}$ è la densità di

$X$ condizionata da $Y_1$ e $Y_2$.

Come noto dalla Teoria, $X$ condizionata da $Y_1, Y_2$ è GAUSSIANA con media

$$\mathbb{E}[X|Y_1,Y_2] = \hat{X}_{12}$$

e varianza

$$Var\{X|Y_1,Y_2\} = \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}$$

dove $Y := \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$

DA CUI:

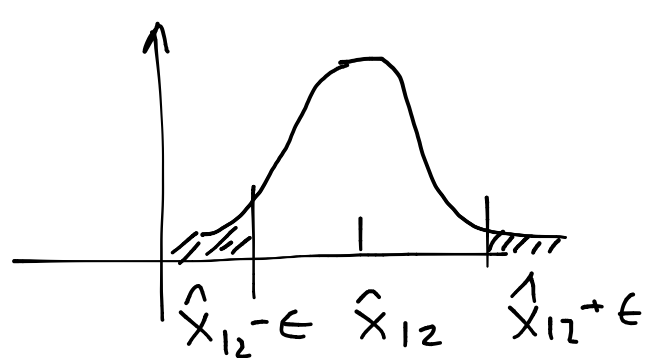
$$\mathrm{Var}\{X|Y_1,Y_2\} = 1 - [1\ 1]\begin{bmatrix}2 & 3/2\\ 3/2 & 2\end{bmatrix}^{-1}\begin{bmatrix}1\\1\end{bmatrix}$$

$$= 1 - \frac{[1\ 1]}{4-\frac{9}{4}}\begin{bmatrix}2 & -3/2\\ -3/2 & 2\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}$$

$$= 1 - \frac{4}{7}\begin{bmatrix}1/2 & 1/2\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}$$

$$= 1 - \frac{4}{7} = \frac{3}{7}$$

Quindi si tratta di calcolare la probabilità che una variabile Gaussiana di varianza 3/7 si discosti dalla sua media più di $\epsilon$.



Poiché vale

$$\mathrm{Var}\{X\} > \mathrm{Var}\{X \mid Y_1\} > \mathrm{Var}\{X \mid Y_1, Y_2\}$$

(il condizionamento riduce l'incertezza)

si ha:

$$P[|X-5| > \epsilon] > P[|X-\hat{X}_1| > \epsilon \mid Y_1] >$$

$$> P[|X-\hat{X}_{12}| > \epsilon \mid Y_1, Y_2]$$

## DOMANDA 2:

**1:**

$$y(t) = \begin{bmatrix} a_1 & b_1 \end{bmatrix} \begin{bmatrix} y(t-1) \\ u(t-1) \end{bmatrix} + e_0(t)$$

$$= \begin{bmatrix} a_1 & b_1 & b_1 \end{bmatrix} \begin{bmatrix} y(t-1) \\ v(t-1) \\ v(t-2) \end{bmatrix} + e_0(t)$$

Si denoti: $\sigma_{yy}(\tau) = \mathbb{E}[y(t+\tau)\,y(t)]$

$\sigma_{yv}(\tau) = \mathbb{E}[y(t+\tau)\,v(t)]$

Si ottiene:

$$\sigma_{yy}(0) = [a_1 \; b_1 \; b_1] \begin{bmatrix} \sigma_{yy}(0) & \sigma_{yv}(0) & \sigma_{yv}(1) \\ \sigma_{yv}(0) & \sigma_{vv}(0) & \sigma_{vv}(1) \\ \sigma_{yv}(1) & \sigma_{vv}(1) & \sigma_{vv}(0) \end{bmatrix} \begin{bmatrix} a_1 \\ b_1 \\ b_1 \end{bmatrix} +$$

$$+ \sigma_{ee}(0)$$

DOVE:

$\sigma_{yv}(0) = 0$ [$v(t)$ è bianco e indipendente da $e_0(t)$, quindi $v(t)$ è scorrelato con $y(s)$ $s \leq t$]

$\sigma_{vv}(0) = \sigma_v^2$ (DA FISSARE) $\boxed{[\sigma_v^2 = 1] \text{ IPOTESI}}$

$\sigma_{vv}(1) = 0$ ($v(t)$ bianco)

$\sigma_{yv}(1) = \mathbb{E}\; y(t)\,v(t-1)$

$$= a_1 \, \mathbb{E}\,\underbrace{\cancel{y(t-1)v(t-1)}}_{=0} + b_1 \underbrace{\mathbb{E}[v(t-1)v(t-1)]}_{=\mathbb{E}[v(t-1)^2]} + \mathbb{E}[\underbrace{\cancel{e_0(t)v(t-1)}}_{=0}] =$$

$$= b_1 \sigma_v^2$$

DA CUI SI OTTIENE $(\sigma_y^2 := \sigma_{yy}(0))$

$$\sigma_y^2 = [a_1 \; b_1 \; b_1] \begin{bmatrix} \sigma_y^2 & 0 & b_1 \\ 0 & 1 & 0 \\ b_1 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ b_1 \\ b_1 \end{bmatrix} + \sigma^2$$

E QUINDI:

$$\sigma_y^2 = a_1^2 \sigma_y^2 + 2 a_1 b_1^2 + 2 b_1^2 + \sigma^2$$

$$\Downarrow$$

$$\sigma_y^2 = \frac{2 a_1 b_1^2 + 2 b_1^2 + \sigma^2}{1 - a_1^2} \qquad \boxed{|a_1| < 1}$$

(STABILITÀ)

**2:** Utilizzando il metodo PEM si deve minimizzare l'errore di predizione

$$\hat{\theta} := \arg\min_{\theta} \frac{1}{N} \cdot \sum_{t=1}^{N} \left(y(t) - \hat{y}_{\theta}(t|t-1)\right)^2$$

DOVE:

$$\hat{y}_{\theta}(t|t-1) = \mathbb{E}_{\theta}\left[y(t) \mid y_t^-\right]$$

$$= \theta\, u(t-1)$$

<u>QUINDI:</u>

$$\hat{\theta}_N = \arg\min_{\theta} \frac{1}{N-1} \cdot \sum_{t=2}^{N} \left(y(t) - \theta\, u(t-1)\right)^2$$

DA CUI SI OTTIENE:

$$\hat{\theta}_N = \left(\sum_{t=2}^{N} u(t-1)^2\right)^{-1} \sum_{t=2}^{N} y(t)\, u(t-1)$$

**3:** Poichè i processi in gioco sono ergodici, lo stimatore $\hat{\theta}_N$ converge per $N \to +\infty$ all'insieme di punti di minimo della funzione

$$J(\theta) := \mathbb{E}\left[\left(y(t) - \hat{y}_\theta(t|t-1)\right)^2\right]$$

TROVIAMO I PUNTI DI MINIMO DI $J(\theta)$

$$J(\theta) = \mathbb{E}\left[\left(a_1 y(t-1) + (b_1 - \theta) u(t-1) + e_0(t)\right)^2\right]$$

$$= \begin{bmatrix} a_1 & (b_1-\theta) & (b_1-\theta) \end{bmatrix} \begin{bmatrix} \sigma_{yy}(0) & \sigma_{yv}(0) & \sigma_{yv}(1) \\ \sigma_{yv}(0) & \sigma_{vv}(0) & \sigma_{vv}(1) \\ \sigma_{yv}(1) & \sigma_{vv}(1) & \sigma_{vv}(0) \end{bmatrix} \begin{bmatrix} a_1 \\ b_1-\theta \\ b_1-\theta \end{bmatrix}$$

$$+ \sigma^2$$

$$= \begin{bmatrix} a & (b_1-\theta) & (b_1-\theta) \end{bmatrix} \begin{bmatrix} \sigma_y^2 & 0 & b_1 \\ 0 & 1 & 0 \\ b_1 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ b_1-\theta \\ b_1-\theta \end{bmatrix} + \sigma^2$$

$$= a_1^2 \sigma_y^2 + 2a_1 b_1 (b_1 - \theta) + 2(b_1-\theta)^2 + \sigma^2$$

<u>CALCOLANDO:</u>

$$\frac{\partial J}{\partial \theta} = -\cancel{2}\, 2 a_1 b_1 - \overset{2}{\cancel{4}}\,(b_1 - \theta) = 0$$

SI OTTIENE CHE PER

$$\theta^* := \frac{b_1 a_1 + 2 b_1}{2} = b_1 \left(1 + \frac{a_1}{2}\right)$$

VALE: $\left.\frac{\partial J}{\partial \theta}\right|_{\theta = \theta^*} = 0$

POICHÈ $\frac{\partial^2 J}{\partial \theta^2} = 4 > 0$

$\theta = \theta^*$ è l'unico punto di minimo di $J(\theta)$

$$\Downarrow$$

$$\boxed{\lim_{N \to +\infty} \hat{\theta}_N = b_1 \left(1 + \frac{a_1}{2}\right)}$$

(Lo stimatore non è consistente perché $\hat{\theta}_N$ non converge a $b_1$)

# DOMANDA 3

**1:** Basta definire:

$$\varphi(t) := \begin{bmatrix} 1 \\ x(t) \\ x^2(t) \\ x^3(t) \\ \log x(t) \end{bmatrix}$$

DA CUI:

$$y(t) = \varphi^T(t)\,\theta + d(t)$$

LE EQUAZIONI DEL KF per la stima ricorsiva di $\theta$ si ottengono applicando il KF (Kalman Filter), si vedano gli appunti, al modello di stato

$$\theta(t+1) = \theta(t)$$

$$y(t) = \varphi^T(t)\,\theta(t) + d(t)$$

**2:** Se i parametri $\theta(t)$ sono tempo varianti si dovrà considerare il modulo

$$\theta(t+1) = \theta(t) + w(t)$$
$$y(t) = \varphi^T(t)\,\theta(t) + d(t)$$

dove la varianza del rumore bianco (e a media nulla) $w(t)$ si può "tarare" cercando di fare in modo che l'errore di predizione

$$y(t) - \varphi^T(t)\,\hat{\theta}(t|t-1)$$

sia "il più bianco possibile"

(e.g. usando il test del Periodogramma cumulato)

3: Per "monitorare" il sistema si può utilizzare l'errore di predizione

$$e(t) := y(t) - \varphi^T(t)\,\hat{\theta}(t|t-1)$$

In presenza di un brusco cambiamento del sistema e(t) aumenterà bruscamente. Verosimilmente e(t) uscirà dalla "fascia" $[-3\sigma_e, +3\sigma_e]$ dove $\sigma_e$ è la deviazione standard di e(t).

4: Se il processo d(t) non è bianco è sufficiente procurarsi un modello di stato per il processo e(t):

$$x(t+1) = A\,x(t) + B\,u(t)$$
$$d(t) = C\,x(t) + D\,u(t)$$

e Applicare l'algoritmo del KF al modello aumentato

$$\begin{bmatrix} \theta(t+1) \\ x(t+1) \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & A \end{bmatrix} \begin{bmatrix} \theta(t) \\ x(t) \end{bmatrix} + \begin{bmatrix} 0 \\ B \end{bmatrix} u(t)$$

$$y(t) = \begin{bmatrix} \varphi^T(t) & C \end{bmatrix} \begin{bmatrix} \theta(t) \\ x(t) \end{bmatrix} + D\,u(t)$$

## DOMANDA 4:

Si vedano gli appunti del corso