COMPITO DI
IDENTIFICAZIONE DEI MODELLI ED ANALISI DEI DATI
25 giugno 2012

Domanda 1. Sia $x_t \in \mathbb{R}^2$, $t \in \mathbb{Z} \cap [0, N]$, la traiettoria di un robot che si muove nel piano ($\mathbb{R}^2$). Un modello che tipicamente viene utilizzato per modellare la traiettoria x_t è la così detta "passeggiata aleatoria integrata"

$$\begin{aligned} x_{t+1} &= x_t + v_t \\ v_{t+1} &= v_t + n_t \end{aligned}$$

dove n_t è un rumore bianco, Gaussiano a media nulla e varianza σ_n^2.

Si assuma siano disponibili misure $y_t = x_t + e_t$ dove $\{e_t\}_{t \in \mathbb{Z}}$ è un rumore bianco Gaussiano, a media zero e varianza σ^2, indipendente da $\{n_t\}_{t \in \mathbb{Z}}$.

1. Si descriva un procedimento per calcolare la covarianza $\mathbb{E}[x_t x_s^\top]$, $t \geq 0$, $s \geq 0$, del processo x_t assumendo che x_0 e v_0 siano variabili Gaussiane indipendenti, indipendenti da $\{n_t, e_t; t \geq 0\}$, a media 0 e varianze $\sigma_{x_0}^2$ e $\sigma_{v_0}^2$ rispettivamente. Dire se il processo y_t è stazionario.

2. Si scrivano le equazioni del filtro di Kalman per la stima della posizione x_t e della velocità v_t.

3. Si discuta come tarare i rumori di misura e modello, avendo a disposizione degli esperimenti di "test".

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

$$y(t) = \sum_{k=0}^{\infty} \theta_k^o u(t-k) + e_0(t)$$

dove $e_0(t)$ è un rumore bianco, a media nulla e varianza σ^2, indipendente dal processo u. Il processo di "ingresso" $u(t)$ si assume bianco, a media nulla e varianza unitaria.

Per fare "identificazione", si utilizzi il criterio di minimizzazione dell'errore di predizione (PEM) ed il modello

$$y(t) = \sum_{k=0}^{m} \theta_k u(t-k) + e(t)$$

con $e(t)$ bianco.

1. Si calcoli la varianza del processo $y(t)$; si enunci una condizione sufficiente sui coefficienti θ_k^o affinchè il processo $y(k)$ sia stazionario e a varianza finita.

2. Assumendo siano disponibili dati $y(t), u(t)$, $t \in [1, N]$, si imposti il problema di identificazione; si trovi lo stimatore $\underline{\hat{\theta}}_N$ del parametro vettoriale $\underline{\theta} := [\theta_1, ...\theta_m]$, scrivendo chiaramente quale funzione dei dati si deve minimizzare.

3. Si dica se lo stimatore $\hat{\theta}_N$ PEM converge e se si, dove.

Domanda 3. Si consideri il seguente scenario: alla stazione ferroviaria sono aperti due sportelli (A e B). Il tempo di attesa T tra un cliente ed il seccessivo si può modellare con una variabile aleatoria esponenziale con parametro λ:

$$f_T(t) = \lambda e^{-\lambda t} \quad t \geq 0.$$

Il parametro λ è legato alla media della variabile aleatoria dalla relazione

$$\mathbb{E}T = \frac{1}{\lambda}.$$

Uno sportello fornisce il servizio ad una velocità mediamente doppia dell'altro e il tempo di attesa medio dello sportello più veloce è 1 minuto.

Supponendo di aver osservato tre clienti successivi allo sportello A, di aver misurato i loro tempi di attesa $t_{A,1} = 0.1\ min$, $t_{A,2} = 1.5\ min$ e $t_{A,3} = 3\ min$ e supponendo ci siano N_A persone in coda allo sportello A ed N_B persone in coda allo sportello B, dire quale coda si sceglierebbe e perchè.

Domanda 4. Si voglia modellare la forza F esercitata da una molla, in funzione dello spostamento x dalla sua posizione di equilibrio, con una relazione del tipo

$$F(x) = \alpha_1 x + \alpha_3 x^3.$$

Si assuma siano disponibili misure rumorose

$$z_i := F(x_i) + e_i = \alpha_1 x_i + \alpha_3 x_i^3 + e_i \quad i = 1, .., N$$

dove i rumori di misura e_i, $i = 1, .., N$ sono indipendenti, Gaussiani, a media nulla e vrianza σ_e^2.

1. Si ricavi uno stimatore di α_1, α_3
2. Si ricavi l'espressione per la varianza dell'errore di stima $\tilde{\alpha}_i = \alpha_i - \hat{\alpha}_i$, $i = 1$ e $i = 3$.
3. Si illustri come si può ottenere una versione ricorsiva (rispetto alle misure z_i) di questo stimatore.

SOLUZIONI

Domanda 1.[8 Punti]

1. È immediato verificare che x_t e v_t hanno media nulla per ogni t. Si definisca

$$A := \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad C := \begin{bmatrix} 1 & 0 \end{bmatrix} \quad G := \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$Q := \begin{bmatrix} 0 & 0 \\ 0 & \sigma_n^2 \end{bmatrix} \quad R := \sigma_e^2 \quad S = \begin{bmatrix} 0 & 0 \end{bmatrix}$$

   e $z_t := [x_t^\top \ v_t^\top]^\top$. Utilizzando la relazione:

$$z_{t+k} = A^k z_t + \sum_{i=0}^{k-1} A^i G n_{t+k-1-i}$$

   si ottiene, per $t \geq s$ (usando il fatto che tutte le grandezze sono a media nulla e che z_s è scorrelato da n_t per $t \geq s$)

$$\mathbb{E} z_t z_s^\top = A^{t-s} \mathbb{E} z_s z_s^\top$$

   e

$$\mathbb{E} z_s z_s^\top = A \mathbb{E} z_0 z_0^\top + \sum \sum_{i=0}^{k-1} A^i Q \left(A^i\right)^\top$$

   Data la struttura di A si vede immediatamente che la varianza di stato $Var\{z_t\}$ diverge per $t \to \infty$ e quindi, certamente, il processo non può essere stazionario (al secondo ordine)

2. Basta utilizzare le matrici A, C, Q, R, S definite sopra nelle equazioni del filtro di Kalman-ricavate a lezione

3. Il tuning dei rumori si può fare utlizzando il test del periodogramma cumulato. Si vedano gli appunti delle lezioni.

Domanda 2.[8 Punti]

1. La varianza del processo di uscita y_t si ottiene da (si noti che dalle ipotesi discende che $\mathbb{E} y_t = 0$)

$$\begin{array}{rcl} Var\{y_t\} & = & \mathbb{E} \sum_{k=1}^{\infty} \sum_{h=1}^{\infty} \theta_k^0 u_{t-k} u_{t-k} \theta_h^0 + \sigma_e^2 \\ & = & \mathbb{E} \sum_{k=1}^{\infty} \left(\theta_k^0\right)^2 + \sigma^2 \end{array}$$

   Quindi la condizione affinchè y_t abbia varianza finita è $\sum_{k=1}^{\infty} \left(\theta_k^0\right)^2 < \infty$. È immediato verificare che sotto le stesse condizioni la covarianza $\mathbb{E} y_t y_s$ dipende solo dalla differenza $t - s$ e quindi il processo è anche stazionario

2. Il modello scelto per fare identificazione è lineare nei parametri e si può scrivere nella forma

$$y(t) = \phi^\top(t)\theta + e(t)$$

   dove $\phi^\top(t) := [u(t)\, u(t-1) \ \dots \ u(t-m)]$; e $\theta^\top = [\theta_0, ..., \theta_m]$. Utilizzando il metodo di minimizzazione dell'errore di predizione

$$\begin{array}{rcl} \hat{\theta}_N & := & \arg\min_\theta \frac{1}{N} \sum_{t=1}^{N} (y(t) - \hat{y}_\theta(t|t-1))^2 \\ & = & \arg\min_\theta \frac{1}{N} \sum_{t=1}^{N} (y(t) - \phi^\top(t)\theta)^2 \end{array}$$

   la cui soluzione è data da:

$$\hat{\theta}_N := \left[\sum_{k=1}^{N} \phi(t)\phi^\top(t) \right]^{-1} \sum_{k=1}^{N} \phi(t) y(t)$$

3. Lo stimatore PEM converge, in ipotesi di ergodicità, all'insieme dei punti di minimo del funzionale

$$J(\theta) := \mathbb{E}(y(t) - \phi^\top(t)\theta)^2 = \sum_{k=0}^{m} \left(\theta_k^0 - \theta_k\right)^2 + \sum_{k=m+1}^{\infty} \left(\theta_k^0\right)^2 + \sigma^2$$

poichè il secondo ed il terzo termine non dipendono da θ, occorre e basta minimizzare il primo termine. Chiaramente questo funzionale ha un solo punto di minimo che si ottiene scegliendo $\theta_k = \theta_k^0$, $k = 0, .., m$. Quindi $\hat{\theta}_{N,k}$ (la $k-$esima componente di $\hat{\theta}_N$) converge a θ_k^0, $k = 0, .., m$.

Domanda 3.[8 Punti] Introduciamo una variabile discreta $C \in \{v, \ell\}$ dove dove i pedici $_v$ e $_\ell$ sono le iniziali di "veloce" e "lento" rispettivamente. Poichè la coda più veloce ha media del tempo di servizio uguale ad 1, mentre la coda più lenta ha il tempo di attesa medio doppio, si ha che $\lambda_v = 1$ e $\lambda_l = 1/2$. Dato che i tempi di attesa di clienti successivi sono indipendenti se condizionati dalla conoscenza della coda scelta (lenta o veloce) abbiamo:

$$\begin{array}{rcl} f_{T_1,T_2,T_3|C}(t_1, t_2, t_3|C = v) & = & \prod_{i=1}^{3}, f_{T_i|C}(t_i|v) = e^{-\sum_{i=1}^{3} t_i} \\ f_{T_1,T_2,T_3|C}(t_1, t_2, t_3|C = \ell) & = & \prod_{i=1}^{3}, f_{T_i|C}(t_i|\ell) = \frac{1}{2}e^{-\frac{1}{2}\sum_{i=1}^{3} t_i} \end{array}$$

Utilizzando la regola di Bayes abbiamo:

$$P(v|T_1, T_2, T_3) = \frac{f_{T_1,T_2,T_3}(t_1, t_2, t_3|C = v)p_v}{f_{T_1,T_2,T_3|C}(t_1, t_2, t_3|C = v)p_v + f_{T_1,T_2,T_3|C}(t_1, t_2, t_3|C = \ell)p_\ell}$$

and $P(\ell|T_1, T_2, T_3) = 1 - P(v|T_1, T_2, T_3)$. Prima di procedere dobbiamo specificare le probabilità *a priori* p_v e $p_{ell} = 1 - p_v$; qui faremo l'ipotesi che $p_v = p_\ell = 1/2$. Con questi dati si ottiene:

$$P(v|T_1, T_2, T_3) = 0.4451 \quad P(\ell|T_1, T_2, T_3) = 0.5549$$

Quindi è più probabile che la coda A sia la coda lenta. Di conseguenza si sceglierà la coda B se $N_B < 2 * N_A$, altrimenti si sceglierà la coda A.

Assumendo che la classificazione lenta/veloce sia corretta, si osservi che il tempo di attesa per la coda A sarà

$$T_A = \sum_{i=1}^{N_A} T_{A,i}$$

mentre

$$T_B = \sum_{i=1}^{N_B} T_{B,i}$$

Le medie e varianze di questi tempi di attesa sono rispettivamente

$$ET_A = \frac{N_A}{\lambda_\ell} = 2N_A \quad Var\{T_A\} = \frac{N_A}{\lambda_\ell^2} = 4N_A$$

$$ET_B = \frac{N_B}{\lambda_v} = N_B \quad Var\{T_B\} = \frac{N_B}{\lambda_v^2} = N_B$$

Si osservi che potrebbe essere conveniente scegliere la coda B anche se N_B è leggermente maggiore di $2N_A$ perchè la deviazione standard del tempo T_B soddisfa

$$Var^{1/2}\{T_B\} < Var^{1/2}\{T_A\} \quad N_B < 4N_A$$

Una valutazione precisa richiede di confrontare le densità di T_A e T_B (sono variabili "Erlang"), ma questo era fuori dagli scopi dell'esercizio.

Domanda 4.[8 Punti]

1. Il modello si può scrivere nella forma

$$z_i = \phi_i^\top \theta + e_i$$

con $\phi_i^\top := [x_i \ x_i^3]$ e $\theta := [\alpha_1 \ \alpha_3]^\top$.

Sotto l'ipotesi di rumore Gaussiano, lo stimatore a massima verosimiglianza di θ conicide con lo stimatore dei minimi quadrati (si vedano gli appunti) e ha la forma:

$$\hat{\theta}_N = \left[\sum_{i=1}^{N} \phi_i \phi_i^\top\right]^{-1} \sum_{i=1}^{N} \phi_i z_i$$

2. Sostituendo $z_i = \phi_i^\top \theta + e_i$ nell'equazione dello stimatore si ottiene

$$\hat{\theta}_N = \theta + \left[\sum_{i=1}^{N} \phi_i \phi_i^\top\right]^{-1} \sum_{i=1}^{N} \phi_i e_i$$

da cui

$$\tilde{\theta}_N := \theta - \hat{\theta}_N = -\left[\sum_{i=1}^{N} \phi_i \phi_i^\top\right]^{-1} \sum_{i=1}^{N} \phi_i e_i$$

e la sua varianza si calcola immediatamente sfruttando il fatto che gli e_i sono indipendenti, a media nulla e varianza σ_e^2:

$$\begin{aligned} Var\{\tilde{\theta}_N\} &= \mathbb{E}\left\{\left[\sum_{i=1}^N \phi_i \phi_i^\top\right]^{-1} \sum_{i=1}^N \phi_i e_i \sum_{j=1}^N e_j \phi_j^\top \left[\sum_{i=1}^N \phi_i \phi_i^\top\right]^{-1}\right\} \\ &= \sigma_e^2 \left\{\left[\sum_{i=1}^N \phi_i \phi_i^\top\right]^{-1} \sum_{i=1}^N \phi_i \phi_i^\top \left[\sum_{i=1}^N \phi_i \phi_i^\top\right]^{-1}\right\} \\ &= \sigma_e^2 \left[\sum_{i=1}^N \phi_i \phi_i^\top\right]^{-1} \end{aligned}$$

3. La versione ricorsiva si ottiene utilizzando le equazioni dei minimi quadrati ricorsivi (RLS) o, quivalentemente, utilizzando le equazioni del filtro di Kalman per il modello

$$\begin{aligned} \theta_{i+1} &= \theta_i \\ z_i &= \phi_i^\top \theta_i + e_i \end{aligned}$$

Basta utilizzare le equazioni dinamiche del filtro (si vedano gli appunti) con

$$A = I \quad C = \phi_i^\top \quad Q = 0 \quad S = 0 \quad R = \sigma_e^2$$

(Nel compito è richiesto di SCRIVERE le equazioni non solo di dire "si usano i minimi quadrati ricorsivi"). Non vengono riportate in questa soluzione perchè sono disponibili negli appunti.