COMPITO DI
IDENTIFICAZIONE DEI MODELLI ED ANALISI DEI DATI
12 luglio 2013

Tutte le risposte vanno adeguatamente giustificate. Si scriva i modo chiaro e ordinato. Non saranno corretti compiti disordinati.

Domanda 1. Si enuncino e si dimostrino la regola di Bayes ed il Teorema della Probabilità totale.

Domanda 2. Siano X_i, $i = 1, .., N$ variabili aleatorie indipendenti ed identicamente distribuite con densità

$$f_\theta(x) := c(\theta)e^{-\theta x}\delta_{-1}(x) \quad \theta > 0$$

1. Si ricavi l'espressione di $c(\theta)$ in modo che $f_\theta(x)$ sia una densità di probabilità.
2. Si scriva lo stimatore a massima verosimiglianza di θ basato sulle misure $x_1, ..., x_N$.
3. Si scriva lo stimatore a massima verosimiglianza della probabilità

   $$P_a := P[X > a]$$

   dove $X \sim f_\theta(x)$

Domanda 3. Sia $y(t)$ un processo stazionario, a media nulla e spettro

$$\Phi_y(z) := \frac{(z + 1/4)(z + 4)}{(z - 2)(1/2 - z)}$$

1. Si trovi una funzione $H(z)$ tale che $y(t)$ si possa rappresentare nella forma

   $$y(t) = H(z)\epsilon(t)$$

   dove $\epsilon(t)$ è bianco, a media nulla e varianza unitaria.
2. Si trovi l'espressione del predittore a due passi $\hat{y}(t|t - 2)$.

Domanda 4. Si consideri il modello di misura

$$Y = h(X) + E \quad Y \in \mathbb{R}, \quad X \in \mathbb{R} \quad E \in \mathbb{R}$$

dove $h(\cdot)$ è una funzione nota, X è una variabile aleatoria Gaussiana a media 0 e varianza σ_x^2, ed E è una variabile aleatoria Gaussiana a media zero e varianza σ_e^2. Si disponga di una misura y di Y. Si vuole calcolare lo stimatore a minima varianza $\hat{X}(Y)$ di X dato Y.

1. Si dica come si può calcolare, in principio, lo stimatore $\hat{X}(Y)$.
2. Si assuma che x_i, $i = 1, .., N$ siano campioni indipendenti ed identicamente distribuiti di $X \sim \mathcal{N}(0, \sigma_x^2)$, cioè N possibili realizzazioni indipendenti della variabile aleatoria X. Si dimostri che, definendo $w_i := \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(y-h(x_i))^2} = f_{Y|X}(y|x_i)$,

   $$\frac{\frac{1}{N}\sum_{i=1}^N x_i w_i}{\frac{1}{N}\sum_{i=1}^N w_i} \simeq \mathbb{E}[X|Y = y]$$

Suggerimento: si usi il fatto che $f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x) f_X(x)}{f_Y(y)}$ *e che, se* x_i *sono indipendenti e identicamente distribuite secondo* $f_X(x)$,

$$\frac{1}{N} \sum_{i=1}^{N} g(x_i) \simeq \int g(x) f_X(x)\, dx = \mathbb{E}_f g(x)$$

SOLUZIONI

Domanda 1.[8 Punti]
Si vedano gli appunti del corso per enunciato e dimostrazione
Domanda 2.[8 Punti]

1. Affinchè $f_\theta(x)$ sia una densità di probabilità deve essere non negativa ed avere integrale unitario, quindi:

$$\int_{-\infty}^{\infty} c(\theta)e^{-\theta x}\delta_{-1}(x)\,dx = \int_{0}^{\infty} c(\theta)e^{-\theta x}\,dx = \frac{c(\theta)}{\theta} = 1$$

   da cui $c(\theta) = \theta$.

2. Lo stimatore a massima verosimiglianza si ottiene da:

$$\begin{aligned} \hat{\theta}_{ML} &:= \arg\max_\theta f_\theta(x_1, ..., x_N) \\ &= \arg\max_\theta \prod_{i=1}^{N} f_\theta(x_i) \\ &= \arg\max_\theta \theta^N e^{-\theta \sum_{i=1}^{N} x_i} \end{aligned}$$

   facendo la derivata rispetto a θ di $\log f_\theta(x_1, .., x_N)$ (il logaritmo è una funzione monotona) ed uguagliando a zero si ottiene

$$\frac{\partial f_\theta(x_1, .., x_N)}{\partial\theta} = N\frac{1}{\theta} - \sum_{i=1}^{N} x_i = 0$$

   che fornisce la soluzione

$$\hat{\theta}_{ML} = \frac{1}{\frac{1}{N}\sum_{i=1}^{N} x_i}$$

3. La probabilità cercata si scrive nella forma

$$P_a = P_\theta[X > a] = 1 - P_\theta[X \le a] = 1 - F_\theta(a) = 1 - \int_0^a \theta e^{-\theta a}\,dx = e^{-\theta a}$$

   Quindo lo stimatore a massima verosimigliana di P_a si ottiene, per il principio di invarianza, da

$$\hat{P}_a := e^{-\hat{\theta}_{ML} a}$$

Domanda 3.[8 Punti]

1. Lo spettro[1] si può scrivere nella forma:

$$\begin{aligned} \Phi_y(z) &= \frac{(z+1/4)(z+4)}{(z-2)(1/2-z)} \\ &= \frac{(z+1/4)}{(z-1/2)}\frac{(z+4)}{(2-z)} \\ &= \frac{\sqrt{2}(z+1/4)}{(z-1/2)}\frac{(z+4)}{\sqrt{2}(2-z)} \\ &= H(z)H(z^{-1}) \end{aligned}$$

   con $H(z) := \frac{\sqrt{2}(z+1/4)}{(z-1/2)}$ causale e causalmente invertibile (polo in 1/2 e zero in $-1/4$)

[1] Nel testo era presente un errore di segno nel denominatore dello spettro. È evidente che quello NON era uno spettro valido perchè NON era una funzione positiva sul cerchio unitario $z = e^{j\theta}$.

2. Il predittore a due passi si trova utilizzando la formula

$$\hat{y}(t|t-2) = \left(H(z) - h(0) - h(1)z^{-1}\right) H^{-1}(z)y(t)$$

dove $h(0) = H(\infty) = \sqrt{2}$ e $h(1) = lim_{z\to\infty} z\,(H(z) - h(0)) = \frac{3\sqrt{2}}{4}$ da cui:

$$H(z) - h(0) - h(1)z^{-1} = \frac{\sqrt{2}(z+1/4)}{(z-1/2)} - \sqrt{2} - \frac{3\sqrt{2}}{4}z^{-1} = \frac{3\sqrt{2}}{8}\frac{1}{z(z-1/2)}$$

e quindi:

$$\hat{y}(t|t-2) = \frac{3\sqrt{2}}{8}\frac{1}{z(z-1/2)}\frac{(z-1/2)}{\sqrt{2}(z+1/4)}y(t) = \frac{3}{8}\frac{1}{z(z+1/4)}y(t)$$

In termini di equazione alle differenze:

$$\hat{y}(t|t-2) + \frac{1}{4}\hat{y}(t-1|t-3) = \frac{3}{8}y(t-2)$$

Domanda 4.[8 Punti]

1. È noto dalla teoria (si vedano gli appunti) che lo stimatore a Minima Varianza è la media condizionata, quindi, in principio è ragionevole utilizzare come stimatore

$$\hat{X}(Y) := \mathbb{E}[X|Y]$$

Purtroppo, quando le variabili aleatorie in gioco non sono congiuntamente Gaussiane (come in questo caso dato che $h(X)$, può essere una funzione non lineare) la media condizionata è difficile da calcolare.

Il punto successivo mostra come sia possibile ottenere un'approssimazione di $\mathbb{E}[X|Y]$ in questo caso.

2. Si noti che, usando la regola di Bayes per il calcolo della densità condizionata,

$$\begin{aligned}\mathbb{E}[X|Y=y] &= \int x f_{X|Y}(x|y)\,dx \\ &= \int x \frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)}\,dx \\ &= \frac{1}{f_Y(y)}\int x f_{Y|X}(y|x)f_X(x)\,dx\end{aligned}$$

Si noti ora che, usando il suggerimento,

$$\int x f_{Y|X}(y|x)f_X(x)\,dx = \mathbb{E}_f\left[x f_{Y|X}(y|x)\right] \simeq \frac{1}{N}\sum_{i=1}^{N} x_i f_{Y|X}(y|x_i)$$

e

$$f_Y(y) = \int f_{Y|X}(y|x)f_X(x)\,dx = \mathbb{E}_f\left[f_{Y|X}(y|x)\right] \simeq \frac{1}{N}\sum_{i=1}^{N} f_{Y|X}(y|x_i)$$

Di conseguenza, definendo

$$w_i := f_{Y|X}(y|x_i)$$

si ottiene

$$\mathbb{E}[X|Y=y] = \frac{1}{f_Y(y)}\int x f_{Y|X}(y|x)f_X(x)\,dx \simeq \frac{\frac{1}{N}\sum_{i=1}^{N} x_i f_{Y|X}(y|x_i)}{\frac{1}{N}\sum_{i=1}^{N} f_{Y|X}(y|x_i)} = \frac{\frac{1}{N}\sum_{i=1}^{N} x_i w_i}{\frac{1}{N}\sum_{i=1}^{N} w_i}$$