Primera Tarea IPD-431

07 de Mayo de 2012

Generalidades

- La tarea es de carácter *individual*.
- Su informe debe estar escrito en algún procesador de texto. (Se recomienda usar LaTeX.) El formato específico del informe queda a su criterio.
- *Justifique en forma cuidadosa* cada paso en sus desarrollos (use referencias si es necesario).
- La notación utilizada en los problemas es la misma utilizada en clases.
- Los enunciados se han escrito de buena fe. Si Ud. cree que hay errores, por favor hágalo notar enviando un e-mail a `eduardo.silva@usm.cl`
- **Fecha de entrega (límite):** 23 de Mayo de 2012, 9:45hrs.

Problema 1 (Ejercicio 2.3 en [2]; 10 puntos)

Suponga que X_1 y X_2 son variables aleatorias conjuntamente Gaussianas, con densidad

$$f_{X_1 X_2} = \mathcal{N}\left(\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}1 & 2\rho\\ 2\rho & 4\end{bmatrix}\right). \tag{1}$$

1. Calcule las densidades marginales de X_1 y X_2, y la densidad de probabilidad condicional de X_2, dado X_1.
2. ¿Para qué valores de ρ se tiene una matriz de covarianza (conjunta) positiva definida?
3. Grafique algunas curvas de nivel de $f_{X_1 X_2}$. Muestre gráficamente que, si $\rho \approx -1$, entonces $x_2 \approx 4 - 2x_1$. ¿Cómo se comparan sus conclusiones con la respuesta dada en el Punto 2?

Problema 2 (50 puntos)

Considere una variable aleatoria X de dimensión n_X y densidad f_X. La entropía de X se define como [1, Capítulo 8]

$$h(X) \triangleq -\mathcal{E}\left\{\log f_X(X)\right\}, \tag{2}$$

donde la base del logaritmo es, usualmente, 2 ó e. En la definición anterior se usa la convención de que $0 \log 0 = 0$, y se supone que la integral existe.

Ejercicio 1 *Suponga que b es un vector en $\mathbb{R}^{n_X}$ y que A es una matriz en $\mathbb{R}^{n_X \times n_X}$. Demuestre que*

$$h(AX + b) = h(X) + \log\left|\det(A)\right|. \tag{3}$$

De forma similar, dadas dos variables aleatorias conjuntamente distribuidas X e Y, se define la entropía condicional de X, dado Y, como [1, Capítulo 8]

$$h(X|Y) \triangleq -\mathcal{E}\left\{\log f_{X|Y}(X)\right\}. \tag{4}$$

Ejercicio 2 *Suponga que X e Y son variables aleatorias conjuntamente distribuidas. Demuestre la siguiente cadena de igualdades:*

$$h(X) - h(X|Y) = h(Y) - h(Y|X) = h(X) + h(Y) - h(X,Y), \tag{5}$$

donde $h(X,Y)$ ha de interpretarse como $h(Z)$ con $Z \triangleq [X^T \;\; Y^T]^T$.

Ejercicio 3 *Suponga que $X_1, \ldots, X_n$ son variables aleatorias conjuntamente distribuidas. Demuestre que*

$$h(X_1, \ldots, X_n) = \sum_{i=1}^{n} h(X_i | X_1, \ldots, X_{i-1}), \tag{6}$$

donde se ha usado la extensión obvia de la notación introducida en el ejercicio anterior.

Ejercicio 4 *Suponga que X e Y son conjuntamente Gaussianas. Demuestre que*

$$\det\left(\mathcal{E}\left\{(X - \hat{X}(Y))(X - \hat{X}(Y))^T\right\}\right) \geq \frac{1}{(2\pi e)^{n_X}} e^{2h(X|Y)}, \tag{7}$$

*para cualquier función $\hat{X}$ suficientemente decente. Use la desigualdad anterior para dar una interpretación a la entropía condicional de una variable, dada otra. (**Ayuda:** Calcule primero la entropía de una variable aleatoria Gaussiana.)*

Problema 3 (40 puntos)

Considere una variable aleatoria de segundo orden Y. Suponga que $\mu_Y = X\theta$ y $P_Y = \sigma^2 I_{n_Y}$, donde X es una matriz y θ un vector, ambos determinísticos. Suponga que $X^T X$ es no singular y defina

$$\hat{\theta} \triangleq \left(X^T X\right)^{-1} X^T Y. \tag{8}$$

Ejercicio 5 *Demuestre que $\mathcal{E}\left\{\hat{\theta}\right\} = \theta$ y que $P_{\hat{\theta}} = (X^T X)^{-1}\sigma^2$.*

Ejercicio 6 *Suponga que $\bar{\theta} \triangleq CY$ con C tal que $\mathcal{E}\left\{\bar{\theta}\right\} = \theta$. Demuestre que $P_{\bar{\theta}} \geq P_{\hat{\theta}}$.*

Al probar lo anterior, Ud. habrá demostrado que $\hat{\theta}$ es el *estimador lineal de mínima varianza y sin sesgo de θ, dadas las mediciones de Y*. (Que el estimador sea lineal y sin sesgo significa que el estimador es de la forma MY, con $\mathcal{E}\{MY\} = \theta$ (las estimaciones correspondientes son My).)

Ejercicio 7 *Considere un sistema lineal e invariante en el tiempo modelado a través de la relación*

$$Y(k) = \left(\sum_{i=0}^{n} \theta_i U(k-i)\right) + E(k), \quad k \in \mathbb{N}_0, \tag{9}$$

*donde $n \in \mathbb{N}$, $\theta_i \in \mathbb{R}$, U es una entrada manipulable, Y es la salida, y E es una perturbación (i.e., una entrada no manipulable). Todas las señales se suponen escalares. Suponga que $U(i) = 0$ para $i < 0$ y que E es una secuencia de ruido blanco de varianza unitaria y media cero, i.e., E es tal que $P_{E(i)E(j)} = 0$ para todo $i \neq j$, $\mu_{E(i)} = 0$ y $P_{E(i)} = 1$ para todo i. Se cuenta con mediciones del par $(u(i), y(i))$ para $i \in \{0, \ldots, N\}$, $N >> n$. Halle el estimador lineal de mínima varianza y sin sesgo del vector de parámetros $\theta \triangleq [\theta_0 \;\; \ldots \;\; \theta_n]^T$, dadas las mediciones de Y. Ilustre su solución vía simulaciones, considerando un sistema lineal sencillo (i.e., suponiendo $n = 2$ ó $n = 3$). (**Ayuda:** En Matlab, (i) `randn(N,1)` genera una secuencia de ruido blanco de dimensión $N \times 1$ con las propiedades enunciadas más arriba, (ii) los comandos `\` ó `/` permiten calcular inversas en forma implícita y razonablemente bien condicionada.)*

Referencias

[1] T.M. Cover and J.A. Thomas. *Elements of Information Theory.* John Wiley and Sons, Inc., 2nd edition, 2006.

[2] T. Söderström. *Discrete-time stochastic systems.* Springer, second edition, 2002.