Processamento Estatístico de Sinais: Lista 3

Problema 1 (CRLB for biased estimators) Seja $\mathbf{x} \in \Re^n$ uma amostra do vetor aleatório $\mathbf{X}$ com função densidade de probabilidade estritamente positiva, $p(\mathbf{x} \mid \underline{\theta})$, parametrizada pelo vetor determinístico $\underline{\theta} \in \Re^p$. Seja $\hat{\underline{\theta}}(\mathbf{x})$ uma estimativa *polarizada* de $\underline{\theta}$ tal que

$$E\left[\hat{\underline{\theta}}(\mathbf{X})\right] = \mathbf{g}(\underline{\theta}) \ . \tag{1}$$

a) Mostre que, se $\ln p(\mathbf{x} \mid \underline{\theta})$ é contínua e diferenciável em relação a todas as componentes θ_i, $i = 1, \ldots, p$, do vetor $\underline{\theta}$, então

$$E\left[\mathbf{s}(\underline{\theta},\, \mathbf{X})\, (\hat{\underline{\theta}}(\mathbf{X}) - \mathbf{g}(\underline{\theta}))^T\right] = \frac{\partial}{\partial\, \underline{\theta}}\, \mathbf{g}^T(\underline{\theta}) \tag{2}$$

onde $\mathbf{s}(\underline{\theta},\, \mathbf{x})$ é o gradiente de $\ln\, p(\mathbf{x} \mid \underline{\theta})$ (*score function*).

b) Defina a matriz de informação de Fisher

$$\mathbf{J}(\underline{\theta}) = E\left[\mathbf{s}(\underline{\theta},\, \mathbf{X})\ \mathbf{s}(\underline{\theta},\, \mathbf{X})^T\right] \ . \tag{3}$$

Admitindo-se que a matriz $\mathbf{J}(\underline{\theta})$ é inversível, mostre que a matriz

$$\mathbf{C}(\underline{\theta}) - \left[\frac{\partial}{\partial\, \underline{\theta}}\, \mathbf{g}^T(\underline{\theta})\right]^T \mathbf{J}^{-1}(\underline{\theta}) \left[\frac{\partial}{\partial\, \underline{\theta}}\, \mathbf{g}^T(\underline{\theta})\right] \tag{4}$$

é *positiva semidefinida* onde

$$\mathbf{C}(\underline{\theta}) = E\left[(\hat{\underline{\theta}}(\mathbf{X}) - \mathbf{g}(\underline{\theta}))\ (\hat{\underline{\theta}}(\mathbf{X}) - \mathbf{g}(\underline{\theta}))^T\right] \tag{5}$$

é a matriz de covariância da estimativa $\hat{\underline{\theta}}(\mathbf{X})$.

Problema 2 (Weighted Least Squares) Seja $\mathbf{W} \in \Re^{n \times n}$ uma matriz simétrica conhecida de posto n e, $\mathbf{H} \in \Re^{n \times p}$, $n \geq p$, também conhecida e de posto p.

(a) Dado $\mathbf{y} \in \Re^n$, obtenha $\theta = \hat{\theta}_{WLS} \in \Re^p$ que minimize

$$\varepsilon = (\mathbf{y} - \mathbf{H}\theta)^T\, \mathbf{W}\, (\mathbf{y} - \mathbf{H}\theta) \ . \tag{6}$$

(b) Assuma que $\mathbf{y}$ é um vetor gaussiano com função densidade de probabilidade (fdp) $N(\mathbf{H}\theta,\, \mathbf{R})$, $\mathbf{R} > 0$. Mostre que, se tomarmos $\mathbf{W} = \mathbf{R}^{-1}$,

$$p(\hat{\theta}_{WLS}) = N\left[\theta,\, (\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\right] \ . \tag{7}$$

Problema 3 Seja $\mathbf{y} \in \Re^n$ um vetor aleatório observado tal que

$$\mathbf{y} = \mathbf{H}\,\theta + \mathbf{n} \tag{8}$$

onde $\theta \in \Re^P$ é um vetor desconhecido de parâmetros determinísticos; $\mathbf{H} \in \Re^{n \times p}$, $n \geq p$, é uma matriz conhecida de posto p e, $\mathbf{n}$ é um vetor gaussiano com fdp $N(\mathbf{0}, \mathbf{R})$, $\mathbf{R} > 0$.

(a) Obtenha a estimativa ML do vetor desconhecido θ, ou seja, ache

$$\hat{\theta}_{ML} = \arg\max_{\theta} p(\mathbf{y} \mid \theta) \ . \tag{9}$$

Compare esse resultado com a estimativa WLS obtida no problema anterior.

(b) Calcule a matriz de informação de Fisher

$$\mathbf{J}(\theta) = E\left[(\frac{\partial}{\partial\,\theta} \ln\, p(\mathbf{y} \mid \theta))(\frac{\partial}{\partial\,\theta} \ln\, p(\mathbf{y} \mid \theta))^T \right] \tag{10}$$

e verifique que

$$\mathbf{J}(\theta)(\hat{\theta}_{ML} - \theta) = \frac{\partial}{\partial\,\theta} \ln p(\mathbf{y} \mid \theta) \ . \tag{11}$$

c) Verifique que $E\hat{\theta}_{ML} = \theta$ e, usando (11), conclua que $\hat{\theta}_{ML}$ é uma estimativa eficiente. Qual é a matriz de covariância de $\hat{\theta}_{ML}$?

Problema 4 Suponha que seja observada uma seqüência de vetores i.i.d., $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$, $\mathbf{x}_i \in \Re^n$, onde $p(\mathbf{x}_i) = N(\mathbf{m}, \mathbf{R})$, $1 \leq i \leq m$.

Assumindo-se o vetor $\mathbf{m}$ conhecido e a matriz $\mathbf{R}$ desconhecida, mostre que a estimativa ML de $\mathbf{R}$ é dada por

$$\hat{\mathbf{R}}_{ML} = \mathbf{S} = \frac{1}{M} \sum_{i=1}^{M} (\mathbf{x}_i - \mathbf{m})\,(\mathbf{x}_i - \mathbf{m})^T \ . \tag{12}$$

A matriz $\mathbf{S}$ é também conhecida como *sample covariance.*

Problema 5 (Sequential Least Squares) Seja $\theta \in \Re^p$ um vetor de parâmetros desconhecidos e seja $\{y_k\}$ uma seqüência observada tal que

$$y_k = \mathbf{c}_k^T\,\theta + n_k \qquad k = 1, 2, \ldots \tag{13}$$

Escrevendo-se (13) para $1 \leq k \leq t$, obtém-se o sistema de equações

$$\mathbf{y}_t = \mathbf{H}_t\theta + \mathbf{n}_t \tag{14}$$

onde $\mathbf{y}_t = [y_1 \dots y_t]^T$, $\mathbf{n}_t = [n_1 \dots n_t]^T$ e $\mathbf{H}_t$ é uma matriz t x p cujas linhas são dadas pelos vetores linha $\mathbf{c}_k^T$, $1 \leq k \leq t$.

a) Seja $t_0 > p$ tal que o posto da matriz H_{t_0} é igual a p. Mostre que, se $\mathbf{n}_t$ é uma realização de um vetor aleatório gaussiano com função densidade de probabilidade $N(\mathbf{0}, \mathbf{I})$ então a estimativa ML do vetor θ dada as observações $\mathbf{y}_{t_0}$ coincide com a estimativa de mínimos quadrados determinística

$$\hat{\theta}_{t_0} = \min_{\theta} \, || \, \mathbf{y}_{t_0} - \mathbf{H}_{t_0}\theta \, ||^2 \quad .$$

b) Verifique que $\hat{\theta}_{t_0}$ satisfaz a equação

$$\mathbf{P}_{t_0}^{-1}\hat{\theta}_{t_0} = \mathbf{H}_{t_0}^T \mathbf{y}_{t_0} \tag{15}$$

com $\mathbf{P}_{t_0}^{-1} = \mathbf{H}_{t_0}^T \mathbf{H}_{t_0}$.

b) Mostre que, para $t > t_0$, a matriz $\mathbf{P}_t$ pode ser calculada recursivamente através da expressão

$$\mathbf{P}_t = \mathbf{P}_{t-1} - \frac{\mathbf{P}_{t-1}\mathbf{c}_t\mathbf{c}_t^T\mathbf{P}_{t-1}}{1 + \mathbf{c}_t^T\mathbf{P}_{t-1}\mathbf{c}_t} \quad . \tag{16}$$

Sugestão: Verifique que

$$\mathbf{H}_t^T\mathbf{H}_t = \mathbf{H}_{t-1}^T\mathbf{H}_{t-1} + \mathbf{c}_t \, \mathbf{c}_t^T \tag{17}$$

e use o lema de inversão de matrizes, segundo o qual, para $\mathbf{\Sigma}$ e $\mathbf{R}$ inversíveis, tem-se

$$(\mathbf{\Sigma}^{-1} + \mathbf{C}\mathbf{R}^{-1}\mathbf{C}^T)^{-1} = \mathbf{\Sigma} - \mathbf{\Sigma}\,\mathbf{C}(\mathbf{C}^T\mathbf{\Sigma}\mathbf{C} + \mathbf{R})^{-1}\mathbf{C}^T\mathbf{\Sigma} \ . \tag{18}$$

c) Mostre que, para $t > t_0$, a estimativa $\hat{\theta}_t$ pode ser então obtida recursivamente pela expressão

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \mathbf{k}_t(\mathbf{y}_t - \mathbf{c}_t^T\hat{\theta}_{t-1}) \tag{19}$$

onde

$$\mathbf{k}_t = \frac{\mathbf{P}_{t-1}\mathbf{c}_t}{1 + \mathbf{c}_t^T\mathbf{P}_{t-1}\mathbf{c}_t} \quad . \tag{20}$$

Observação: Os resultados (16) e (19) obtidos nesse exercício fornecem uma interpretação alternativa para o conhecido algoritmo RLS usado em filtragem adaptativa.