Processamento Estatístico de Sinais: Lista 4

Problema 1 Sejam $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_M$, M vetores aleatórios de dimensão N, independentes e identicamente distribuídos com função densidade de probabilidade normal $\mathcal{N}(\mathbf{m}, \mathbf{R})$.

a) Assumindo que a matriz $\mathbf{R}$ é conhecida e o vetor $\mathbf{m}$ é desconhecido e parametrizado por um vetor $\underline{\theta} \in \Re^p$, mostre que a matriz de informação de Fisher usada no cálculo do limite inferior de Cramer-Rao para o erro quadrático médio de qualquer estimativa sem viés dos parâmetros θ_i, $i = 1, \ldots, p$, é dada por $\mathbf{J} = [J_{k,l}]_{1 \leq k,l \leq p}$, onde

$$J_{k,l} = M \, \frac{\partial \mathbf{m}^T}{\partial \theta_k} \, \mathbf{R}^{-1} \frac{\partial \mathbf{m}}{\partial \theta_l} \; .$$

b) Assumindo a seguir que $\mathbf{m}$ é conhecido e a matriz $\mathbf{R}$ é desconhecida e parametrizada pelo vetor $\underline{\theta}$, mostre que

$$J_{k,l} = \frac{M}{2} \mathrm{tr} \left[\mathbf{R}^{-1} \, \frac{\partial \mathbf{R}}{\partial \theta_k} \, \mathbf{R}^{-1} \, \frac{\partial \mathbf{R}}{\partial \theta_l} \right] \; .$$

Problema 2 (Rank One Signal) Observa-se uma amostra da seqüência iid de vetores aleatórios $(\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_M)$ onde

$$Y_k(n) = A \, e^{j\omega n} + V_k(n) \qquad 0 \leq n \leq N-1, \qquad 1 \leq k \leq M.$$

Na equação acima, ω é um parâmetro real determinístico no intervalo $[-\pi, \pi]$ e A é uma variável aleatória complexa de média nula e variância σ_1^2. O ruído de observação $\mathbf{V}_k$ é gaussiano com $E\left[V_k(n)\right] = 0$ e $E\left[V_k(n) V_k^*(m)\right] = \sigma^2 \delta(n-m)$, $0 \leq n, m \leq N-1$. Assuma ainda que $E\left[A \, V_k^*(n)\right] = 0$.

a) Considerando que o parâmetro ω é conhecido, mostre que as estimativas ML de σ_1^2 e σ^2 dada uma seqüência observada $(\mathbf{y}_1, \ldots, \mathbf{y}_M)$ satisfazem a equação

$$(\mathbf{e}_1^H \mathbf{e}_1) \widehat{\sigma_1^2} + \widehat{\sigma^2} = \frac{\mathbf{e}_1^H \, \mathbf{S} \, \mathbf{e}_1}{\mathbf{e}_1^H \mathbf{e}_1}$$

onde H denota o hermitiano de um vetor, $\mathbf{e}_1 = [1 \; \exp(-j\omega) \ldots \; \exp(-j(N-1)\omega)]^H$, e S é a "matriz de covariância amostral" dada por

$$\mathbf{S} = \frac{1}{M} \sum_{i=1}^{M} \mathbf{y}_i \mathbf{y}_i^H .$$

b) Calcule o limite inferior de Cramér-Rao para o erro quadrático médio de estimação dos parâmetros

σ_1^2 e σ^2.

Problema 3 (Least Squares Fitting of Structured Covariance Matrices)

Seja $\mathbf{R} \in \Re^{n \times n}$ uma matriz de covariância parametrizada por um vetor de parâmetros desconhecidos $\theta \in \Re^p$. Seja $\mathbf{S}$ uma estimativa *não-estruturada* de $\mathbf{R}$ obtida a partir de dados experimentais. Por exemplo, tome $\mathbf{S}$ igual à "sample covariance" obtida a partir da amostra $(\mathbf{x}_1, \ldots, \mathbf{x}_M)$ com $\mathbf{x}_i$, $\mathbf{x}_i \in \Re^n$, $1 \leq i \leq M$, uma seqüência i.i.d. onde $p(\mathbf{x}_i) = N(\mathbf{0}, \mathbf{R})$.

Usando-se um critério de mínimos quadrados, deseja-se obter a melhor aproximação estruturada $\hat{\mathbf{R}} = \hat{\mathbf{R}}(\theta)$ para a covariância estimada $\mathbf{S}$, ou seja, deseja-se determinar $\mathbf{R}(\theta)$ que minimize

$$\varepsilon = \mathrm{tr}\left[(\mathbf{S} - \mathbf{R}(\theta))^T (\mathbf{S} - \mathbf{R}(\theta))\right] = \sum_{i=1}^{n} \sum_{j=1}^{n} \left[S_{ij} - R_{ij}(\theta)\right]^2 \tag{1}$$

onde "tr" denota o operador traço.

a) Mostre que as estimativas de mínimos quadrados dos parâmetros θ_k, $1 \leq k \leq p$, são soluções do sistema de equações

$$\mathrm{tr}\left\{[\mathbf{S} - \mathbf{R}(\theta)]^T \, \frac{\partial \mathbf{R}(\theta)}{\partial \theta_k}\right\} = 0 \ . \tag{2}$$

b) Assuma agora que $\mathbf{R}$ é uma matriz *Toeplitz* simétrica com estrutura

$$\mathbf{R} = \sum_{i=0}^{n-1} r_i \mathbf{Q}_i \tag{3}$$

onde $\mathbf{Q}_0 = \mathbf{I}$ e, para $1 \leq k \leq n-1$, $Q_k(i,\, j) = 1$ se $\mid i - j \mid= 1$ e, igual a zero caso contrário. Usando o resultado do item (a), mostre que as estimativas de mínimos quadrados dos parâmetros r_i, $0 \leq i \leq n-1$, são dadas por

$$\hat{r}_0 = \frac{1}{n} \sum_{i=1}^{n} S_{ii} \tag{4}$$

$$\hat{r}_k = \frac{1}{2(n-k)} \sum_{j=k+1}^{n} S_{j-k,\, j} + \sum_{i=k+1}^{n} S_{i,\, i-k} \qquad k \neq 0 \ . \tag{5}$$

Interprete esse resultado.

c) Considere agora a situação em que a matriz $\mathbf{R}$ desejada é uma matriz real n x n de posto $p < n$, com estrutura

$$\mathbf{R} = \sum_{i=1}^{p} \sigma_i^2 \mathbf{u}_i \mathbf{u}_i^T \tag{6}$$

onde $\{\mathbf{u}_i\}$, $1 \leq i \leq p$, é um conjunto de p vetores conhecidos, linearmente independentes e de dimensão n x 1. Deseja-se obter a melhor aproximação de mínimos quadrados

$$\hat{\mathbf{R}} = \sum_{i=1}^{n} \widehat{\sigma^2}_i \mathbf{u}_i \, \mathbf{u}_i^T \tag{7}$$

para a estimativa não-estruturada (posto cheio) $\mathbf{S}$ da matriz de covariância. Usando o resultado do item (a), mostre que as estimativas de mínimos quadrados $\widehat{\sigma^2}$ são a solução do sistema linear

$$\begin{bmatrix} (\mathbf{u}_1^T\mathbf{u}_1)^2 & \dots & (\mathbf{u}_1^T\mathbf{u}_p)^2 \\ \vdots & & \vdots \\ (\mathbf{u}_p^T\mathbf{u}_1)^2 & \dots & (\mathbf{u}_p^T\mathbf{u}_p)^2 \end{bmatrix} \begin{bmatrix} \widehat{\sigma^2}_1 \\ \vdots \\ \widehat{\sigma^2}_p \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^T\,\mathbf{S}\,\mathbf{u}_1 \\ \vdots \\ \mathbf{u}_p^T\,\mathbf{S}\,\mathbf{u}_p \end{bmatrix}. \tag{8}$$

d) Na seqüência, assuma que a estrutura desejada da matriz $\mathbf{R}$ é dada por

$$\mathbf{R} = \mathbf{U}\mathbf{\Sigma}^2\mathbf{U}^T \tag{9}$$

onde

$$\mathbf{\Sigma}^2 = \mathbf{\Sigma}_p^2 + \sigma_n^2\mathbf{I} \tag{10}$$

com $\mathbf{\Sigma}_p^2 = \mathrm{diag}(\sigma_1^2,\, \sigma_2^2, \dots,\, \sigma_p^2,\, 0, \dots, 0)$ e $\mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}$.

Seja $\mathbf{S}$ a matriz de covariância estimada não-estruturada e admita conhecida a matriz $\mathbf{U}$ na estrutura desejada $\mathbf{R}$. Mostre que, no caso particular em que $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_p^2 = \sigma_s^2$, as estimativas de mínimos quadrados $\widehat{\sigma_s^2}$ e $\widehat{\sigma_n}^2$ são dadas por

$$\begin{aligned} \widehat{\sigma_n^2} &= \frac{1}{N-p}\,\mathrm{tr}\,[(\mathbf{I}-\mathbf{P})\mathbf{S}] &(11)\\ \widehat{\sigma_s^2} &= \frac{1}{p}\,\mathrm{tr}\,[\mathbf{P}\,\mathbf{S}] - \widehat{\sigma_n^2} &(12) \end{aligned}$$

onde

$$\mathbf{P} = \mathbf{U}_p\,\mathbf{U}_p^T \tag{13}$$

é o operador de projeção ortogonal sobre o subespaço gerado pelas p primeiras colunas da matriz $\mathbf{U}$.

Problema 4 (Constrained Least Squares) Seja $\mathbf{y} \in \Re^n$ um vetor observado. Deseja-se obter o vetor $\theta \in \Re^p$ que minimize

$$\varepsilon = (\mathbf{y} - \mathbf{H}\theta)^T(\mathbf{y} - \mathbf{H}\theta) \tag{14}$$

e seja simultaneamente sujeito à restrição

$$\mathbf{C}^T\theta = \mathbf{0} \tag{15}$$

onde são conhecidos: $\mathbf{H} \in \Re^{n\times p}$, $n \geq p$, com posto p; $\mathbf{C}^T \in \Re^{r\times p}$, $r \leq p$, com posto r.

(a) Mostre que a solução do problema de minimização com restrições é dada por

$$\hat{\theta}_{CLS} = \mathbf{P}\,\hat{\theta}_{LS} \tag{16}$$

onde

$$\begin{aligned} \hat{\theta}_{LS} &= (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{y} &(17)\\ \mathbf{P} &= \mathbf{I} - (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{C}\left[\mathbf{C}^T(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{C}\right]^{-1}\mathbf{C}^T\,. &(18) \end{aligned}$$

Dica Calcule e iguale a zero o gradiente em relação a θ de

$$J = (\mathbf{y} - \mathbf{H}\theta)^T(\mathbf{y} - \mathbf{H}\theta) + 2\underline{\lambda}^T(\mathbf{C}^T\theta)$$

onde $\underline{\lambda}$ é o vetor de multiplicadores de Lagrange.

(b) Verifique que

$$\widehat{\mathbf{x}}_{CLS} = \mathbf{H}\,\widehat{\theta}_{CLS} = (\mathbf{P}_H - \mathbf{P}_C)\,\mathbf{y} \tag{19}$$

onde

$$\begin{aligned} \mathbf{P}_H &= \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T & (20) \\ \mathbf{P}_C &= \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{C}\left[\mathbf{C}^T(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{C}\right]^{-1}\mathbf{C}^T(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\ . & (21) \end{aligned}$$

(c) Verifique que $\mathbf{P}_C$ é o operador de projeção ortogonal de posto r construído a partir de $\mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{C}$. Conclua ainda que

$$\mathbf{P}_H\mathbf{P}_C\mathbf{P}_H = \mathbf{P}_H\mathbf{P}_C = \mathbf{P}_C\mathbf{P}_H = \mathbf{P}_C \tag{22}$$

e, portanto,

$$\widehat{\mathbf{x}}_{CLS} = (\mathbf{I} - \mathbf{P}_C)\widehat{\mathbf{x}}_{LS} = \mathbf{P}_H(\mathbf{y} - \mathbf{P}_C\widehat{\mathbf{x}}_{LS})\ . \tag{23}$$

Interprete esse resultado geometricamente.