Statistical signal processing
5CTA0

Estimation theory - parameter transformation

Parameter transformation

- Interested in transformed parameter $\alpha = h(\theta)$
- Example: estimate in squared magnitude A^2 rather than amplitude A
- Knowing the Fisher information for the parameter θ and the mapping $\alpha = h(\theta)$ allows to find the Cramér-Rao lower bound for $\hat{\alpha}$.
- Maximum likelihood estimate is invariant under parameter transform

CRLB for transformed parameters

$$E[\hat{\alpha}] = \alpha = h(\theta) \quad \Rightarrow \quad E[(\hat{\alpha}-\alpha)] = 0$$

$$\int (\underbrace{\hat{\alpha}}_{g(\underline{x})} - \underbrace{\alpha}_{\alpha = h(\theta)})\, p(\underline{x};\alpha)\, d\underline{x} = \int (\hat{\alpha} - h(\theta))\, p(\underline{x};\theta)\, d\underline{x} = 0$$

$$p(\underline{x};\theta)\frac{\partial}{\partial\theta}\ln p(\underline{x};\theta)$$

$$-\int \frac{\partial}{\partial\theta} h(\theta)\, p(\underline{x};\theta)\, d\underline{x} + \int (\hat{\alpha} - h(\theta)) \overbrace{\frac{\partial}{\partial\theta} p(\underline{x};\theta)}\, dx = 0$$

$$\left(\int (\hat{\alpha} - h(\theta))\, p(\underline{x};\theta)\frac{\partial}{\partial\theta}\ln p(\underline{x};\theta)\, dx\right)^2 = \left(\frac{d}{d\theta} h(\theta)\right)^2$$

$$\left(\int \underbrace{(\hat{\alpha} - h(\theta))\sqrt{p(\underline{x};\theta)}}\, \underbrace{\frac{\partial}{\partial\theta}\ln p(\underline{x};\theta)\sqrt{p(\underline{x};\theta)}}\, dx\right)^2 = \left(\frac{d}{d\theta} h(\theta)\right)^2$$

CRLB for transformed parameters

$$\underbrace{\int (\hat{\alpha} - h(\theta))\, p(\underline{x};\theta)\, d\underline{x}}_{Var[\hat{\alpha}]} \underbrace{\int \left(\frac{\partial}{\partial\theta} \ln p(\underline{x};\theta)\right)^2 p(\underline{x};\theta)\, d\underline{x}}_{I(\theta)} \geq \left(\frac{d}{d\theta} h(\theta)\right)^2$$

$$Var[\hat{\alpha}] \geq \frac{\left(\frac{d}{d\theta} h(\theta)\right)^2}{I(\theta)}$$

Transformed parameter - efficiency

- Efficiency is preserved in a linear (affine) transformation: $\alpha = a\theta + b$
- Nonlinear transform do not preserve efficiency

$$\bar{x} = \frac{1}{N} \sum_{n=0}^{N-1} x_n = \hat{A} \qquad A^2 = ?$$

$$E[\bar{x}^2] = E^2[\bar{x}] + Var[\bar{x}] = A^2 + \frac{\sigma^2}{N} \neq A^2$$

CRLB for transformed vector parameters

- $\boldsymbol{\theta} \in \mathbb{R}^p$ and $\boldsymbol{\alpha} \in \mathbb{R}^r$, $h : \mathbb{R}^p \to \mathbb{R}^r$
- $$\mathbf{a}\left(\mathbf{C}_{\hat{\boldsymbol{\alpha}}} - \frac{\partial h(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \mathbf{I}^{-1}(\boldsymbol{\theta}) \frac{\partial h(\boldsymbol{\theta})^T}{\partial \boldsymbol{\theta}}\right) \mathbf{a} \geq 0, \qquad \mathbf{a} \neq \mathbf{0}$$
- Jacobian matrix
$$\frac{\partial h(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{\partial}{\partial \theta_1} h_1(\boldsymbol{\theta}) & \frac{\partial}{\partial \theta_2} h_1(\boldsymbol{\theta}) & \cdots & \frac{\partial}{\partial \theta_p} h_1(\boldsymbol{\theta}) \\ \frac{\partial}{\partial \theta_1} h_2(\boldsymbol{\theta}) & \frac{\partial}{\partial \theta_2} h_2(\boldsymbol{\theta}) & \cdots & \frac{\partial}{\partial \theta_p} h_2(\boldsymbol{\theta}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial \theta_1} h_r(\boldsymbol{\theta}) & \frac{\partial}{\partial \theta_2} h_r(\boldsymbol{\theta}) & \cdots & \frac{\partial}{\partial \theta_p} h_r(\boldsymbol{\theta}) \end{bmatrix}.$$

Maximum likelihood estimation - transformation invariance

- Suppose tha $\hat{\theta}_{\mathrm{ML}}$ is the maximum likelihood estimator of the parameters θ. Furthermore, suppose that $h(\theta)$ is a, not necessarily one-to-one, scalar function of the parameter. Then, $h(\hat{\theta}_{\mathrm{ML}})$ is the maximum likelihood estimator of $\alpha = h(\theta)$.

$$\{\theta : \alpha = h(\theta)\}$$

$$p_T(\underline{x}; \alpha) = \max_{\{\theta : \alpha = h(\theta)\}} p(\underline{x}; \theta) \leq \max_{\theta} p(\underline{x}; \theta) = p(\underline{x}; \hat{\theta}_{ML})$$

$$p(\underline{x}; \hat{\theta}_{ML}) = \max_{\{\theta : h(\hat{\theta}_{ML}) = h(\theta)\}} p(\underline{x}; \theta) = p_T(\underline{x}, h(\hat{\theta}_{ML}))$$