EUREKOM 07/12/2023
SSP Statistical Signal Processing

# Homework 2

**Due: 21/12/2023** (HW2 to be turned in by email to Zilu.Zhao@eurecom.fr)

Homework policy: the homework is individual. Students can discuss with fellow students to try to find the main structure of the solution for a problem, especially if they are totally stuck at the beginning of the problem. However, you should work out the details yourself and write down in your own words only what you understand yourself.

## Bayesian Parameter Estimation

### 1. On the Beneficial Bias of MMSE Estimation

Consider the Bayesian linear model $Y = H\,\theta + V$ with $\theta \sim \mathcal{N}(0, C_{\theta\theta})$ and $V \sim \mathcal{N}(0, C_{VV})$ independent (we consider here $m_\theta = 0$ for simplicity).

(a) The LMMSE estimator is $\widehat{\theta}_{LMMSE} = C_{\theta Y} C_{YY}^{-1} Y = (C_{\theta\theta}^{-1} + H^T C_{VV}^{-1} H)^{-1} H^T C_{VV}^{-1} Y$.
What are the unconstrained (non-linear) MMSE and the MAP estimators?

they are the same?

(b) What is the error covariance matrix ?

$$R_{\widetilde{\theta}\widetilde{\theta}}^{LMMSE} = E_\theta E_{Y|\theta}\, \widetilde{\theta}_{LMMSE} \widetilde{\theta}_{LMMSE}^T \tag{1}$$

(c) The conditional bias of an estimator $\widehat{\theta}$ is $b_{\widehat{\theta}}(\theta) = E_{Y|\theta}\widehat{\theta}(Y) - \theta$.
The BLUE estimator is the LMMSE estimator under the constraint of conditional unbiasedness. So $b_{BLUE}(\theta) = 0$.
What is $\widehat{\theta}_{BLUE}$ in terms of the quantities appearing in the Bayesian linear model considered here?
Is there another classical deterministic estimator that equals $\widehat{\theta}_{BLUE}$ in this case?

(d) What is the error covariance matrix ?

$$R_{\widetilde{\theta}\widetilde{\theta}}^{BLUE} = E_\theta E_{Y|\theta}\, \widetilde{\theta}_{BLUE} \widetilde{\theta}_{BLUE}^T \tag{2}$$

(e) Show that $R_{\widetilde{\theta}\widetilde{\theta}}^{LMMSE} \leq R_{\widetilde{\theta}\widetilde{\theta}}^{BLUE}$.
Note that this is true in spite of $\widehat{\theta}_{LMMSE}$ being (conditionally) biased and $\widehat{\theta}_{BLUE}$ being unbiased.

(f) Returning to $\widehat{\theta}_{LMMSE}$, what is the bias $b_{LMMSE}(\theta)$?

(g) Show that

$$R_{\widetilde{\theta}\widetilde{\theta}} = \underbrace{E_\theta\, b_{\widehat{\theta}}(\theta)\, b_{\widehat{\theta}}^T(\theta)}_{\text{(bias)}^2} + \underbrace{E_\theta E_{Y|\theta}\,(\widehat{\theta} - E_{Y|\theta}\widehat{\theta})(\widehat{\theta} - E_{Y|\theta}\widehat{\theta})^T}_{\text{variance}} \tag{3}$$

(h) Compute $E_\theta\, b_{LMMSE}(\theta)\, b_{LMMSE}^T(\theta)$.

(i) Compute $E_\theta E_{Y|\theta}\,(\widehat{\theta}_{LMMSE} - E_{Y|\theta}\widehat{\theta}_{LMMSE})(\widehat{\theta}_{LMMSE} - E_{Y|\theta}\widehat{\theta}_{LMMSE})^T$.

Note that the sum of the positive definite matrices in (h) and (i) yields $R_{\widetilde{\theta}\widetilde{\theta}}^{LMMSE}$, for which (e) holds. Hence, in spite of the the fact that LMMSE introduces a (conditional) bias, it allows to reduce the variance so much that the sum of variance and squared bias gets lower than the variance in the unbiased case.

## Deterministic Parameter Estimation

**2. ML Estimation of Roundtrip Delay Distribution.**

Assume that for the roundtrip delay in a computer network, as considered in the homework, we now consider a truncated exponential distribution:

$$f(y|\lambda,\alpha,\beta) = \left\{ \begin{array}{ll} 0 & , y < \alpha \\ \gamma\, e^{-\lambda y} & , \alpha \le y \le \beta \\ 0 & , \beta < y \end{array} \right\} = \gamma\, e^{-\lambda y}\, 1_{[\alpha,\beta]}(y) \tag{4}$$

where $\gamma$ is a normalization constant and

$$1_{\mathcal{A}}(y) = \left\{ \begin{array}{ll} 1 & , \; y \in \mathcal{A} \\ 0 & , \; y \notin \mathcal{A} \end{array} \right.$$

is the indicator function for the set $\mathcal{A}$.

(a) Determine the normalization constant $\gamma$ as a function of $\lambda$, $\alpha$ and $\beta$.
In what follows, you substitute $\gamma$ in $f(y|\lambda,\alpha,\beta)$ by this function of $\lambda$, $\alpha$ and $\beta$.

(b) We now collect $n$ i.i.d. measurements $y_i$ into the vector $Y$. Assume for the moment that $\lambda > 0$ is a given constant.
Find the likelihood function $l(\alpha,\beta|Y,\lambda)$ for $\alpha$ and $\beta$ given $Y$ and $\lambda$.
Note that $1_{[\alpha,\beta]}(y) = 1_{[\alpha,\infty)}(y)\; 1_{(-\infty,\beta]}(y)$.

(c) Maximize this likelihood function to determine the Maximum Likelihood (ML) estimate of $\alpha$ and $\beta$ on the basis of (for given) $Y$ and $\lambda$.

(d) Consider now also $\lambda$ as unknown and determine its ML estimate.

**In what follows**, consider the special case in which $\alpha = 0$, $\beta = \infty$, i.e. the untruncated exponential distribution case. In this case, $\lambda$ is the only remaining parameter in the distribution $f(y|\lambda)$.

(e) Determine the mean $m_y = \mathrm{E}\; y$ and the variance $\sigma_y^2 = \mathrm{E}\; y^2 - (\,\mathrm{E}\; y)^2$ as a function of $\lambda$.

(f) Determine the log likelihood function from $n$ i.i.d. measurements $Y$, $L(\lambda|Y) = \ln f(Y|\lambda)$.

(g) Determine the Maximum Likelihood (ML) estimate $\widehat{\lambda}_{ML}$ and express it as a function of the sample mean $\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$.

(h) Show that in this case the ML estimate can be interpreted as an application of the method of moments.

(i) Asymptotically ($n \gg 1$), the estimation error $\overline{y} - m_y$ will be very small. Hence develop $\widehat{\lambda}_{ML}$ up to first order in $\overline{y} - m_y$. From this asymptotic expression of $\widehat{\lambda}_{ML}$, obtain the asymptotic mean $m_{\widehat{\lambda}_{ML}}$ of $\widehat{\lambda}_{ML}$ and asymptotic variance (for large but finite $n$) $\sigma^2_{\widetilde{\lambda}_{ML}}$ of $\widetilde{\lambda}_{ML} = \lambda - \widehat{\lambda}_{ML}$. Express both $m_{\widehat{\lambda}_{ML}}$ and $\sigma^2_{\widetilde{\lambda}_{ML}}$ in terms of $\lambda$.
Is $\widehat{\lambda}_{ML}$ asymptotically unbiased?

(j) Determine the Fisher Information and the Cramer-Rao bound (CRB) for any unbiased estimator $\widehat{\lambda}$.
Is ML asymptotically efficient in this case?