Statistical Signal Processing

Homework 2

Student: **Sara Rosato**

Student ID: **64830**

1 Bayesian parameter estimation

- **Point a:**

  Both θ and V are Gaussians, so the estimators will all be equal:
  $\hat{\theta}_{MMSE} = \hat{\theta}_{MAP} = \hat{\theta}_{LMMSE} = C_{\theta Y} C_{YY}^{-1} Y$
  $= (C_{\theta\theta}^{-1} + H^T C_{VV}^{-1} H)^{-1} H^T C_{VV}^{-1} Y$

- **Point b:**

$$
\begin{aligned}
&E_\theta E_{Y|\theta} \left(C_{\theta Y}(C_{YY}^{-1})Y - \theta\right)\left(C_{\theta Y}(C_{YY}^{-1})Y - \theta\right)^T \\
&= E_\theta \left(\theta - \left((C_{\theta\theta})^{-1} + H^T C_{VV}^{-1} H\right)^{-1} H^T C_{VV}^{-1} Y\right)\left(Y^T C_{VV}^{-1} H \left((C_{\theta\theta})^{-1} + H^T C_{VV}^{-1} H\right)^{-1}\right)^T \\
&= E_\theta \left(\left((C_{\theta\theta})^{-1} + H^T C_{VV}^{-1} H\right)^{-1} H^T C_{VV}^{-1} \mathrm{Cov}(\theta, Y) H \left((C_{\theta\theta})^{-1} + H^T C_{VV}^{-1} H\right)^{-1}\right) \\
&= E_\theta \left(\left((C_{\theta\theta})^{-1} + H^T C_{VV}^{-1} H\right)^{-1} H^T C_{VV}^{-1} H C_{\theta\theta} H^T C_{VV}^{-1} H \left((C_{\theta\theta})^{-1} + H^T C_{VV}^{-1} H\right)^{-1}\right) \\
&= \left((C_{\theta\theta})^{-1} + H^T C_{VV}^{-1} H\right)^{-1}
\end{aligned}
$$

- **Point c:**

$$b_{\hat{\theta}}(\theta) = E_{Y|\theta}\hat{\theta}(Y) - \theta \qquad b_{BLUE}(\theta) = 0$$

$$\hat{\theta}_{BLUE} = (H^T C_{VV}^{-1} H)^{-1} H^T C_{VV}^{-1} Y \text{ as seen in the slides}$$

  In this case, $\hat{\theta}_{BLUE}$ is equal to the Maximum Likelihood.

- **Point d:**

$$
\begin{aligned}
&R_{\tilde{\theta}\tilde{\theta}}^{BLUE} = E_\theta E_{Y|\theta} \tilde{\theta}_{BLUE} \tilde{\theta}_{BLUE}^T \\
&\tilde{\theta}_{BLUE} = \hat{\theta}_{BLUE} - \theta
\end{aligned}
$$

  We can use the properties of expectations to get the following results:

$$
\begin{aligned}
&R_{\tilde{\theta}\tilde{\theta}}^{BLUE} = \left(H^\top C_{VV}^{-1} H\right)^{-1} H^\top C_{VV}^{-1} E\left[YY^\top\right] \mathrm{C}_{VV}^{-1} H \left(H^\top \mathrm{C}_{VV}^{-1} H\right)^{-1} \\
&= \left(H^\top C_{VV}^{-1} H\right)^{-1} H^\top \left(C_{VV}^{-1}\right)^2 H \left(H^\top C_{VV}^{-1} H\right)^{-1} \\
&= \left(H^\top C_{VV}^{-1} H\right)^{-1} H^\top I H \left(H^\top C_{VV}^{-1} H\right)^{-1} \\
&= (H^T C_{VV}^{-1} H)^{-1}
\end{aligned}
$$

- **Point e:** The inequality $R_{LMMSE} \leq R_{BLUE}$ holds because $C_{\theta\theta}$ is a positive definite matrix (non-negative).

- **Point f:**

  From what we said in Point a, $\hat{\theta}_{LMMSE} = \hat{\theta}_{MMSE}$, which is an unbiased estimator, so also our LMMSE will be unbiased $\rightarrow$ bias $= 0$

- **Point g:**

$$
\begin{aligned}
&R_{\tilde{\theta}\tilde{\theta}} = E_\theta E_{Y|\theta}[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T] \\
&= E_\theta E_{Y|\theta}[(\hat{\theta} - E_{Y|\theta}\hat{\theta} + E_{Y|\theta}\hat{\theta} - \theta)(\hat{\theta} - E_{Y|\theta}\hat{\theta} + E_{Y|\theta}\hat{\theta} - \theta)^T] \\
&= E_\theta E_{Y|\theta}[(\hat{\theta} - E_{Y|\theta}\hat{\theta})(\hat{\theta} - E_{Y|\theta}\hat{\theta})^T] + E_\theta E_{Y|\theta}[(E_{Y|\theta}\hat{\theta} - \theta)(E_{Y|\theta}\hat{\theta} - \theta)^T] \\
&- E_\theta E_{Y|\theta}[(\hat{\theta} - E_{Y|\theta}\hat{\theta})(E_{Y|\theta}\hat{\theta} - \theta)^T] - E_\theta E_{Y|\theta}[(E_{Y|\theta}\hat{\theta} - \theta)(\hat{\theta} - E_{Y|\theta}\hat{\theta})^T] \\
&= E_\theta E_{Y|\theta}[(\hat{\theta} - E_{Y|\theta}\hat{\theta})(\hat{\theta} - E_{Y|\theta}\hat{\theta})^T] + E_\theta[(E_{Y|\theta}\hat{\theta} - \theta)(E_{Y|\theta}\hat{\theta} - \theta)^T] \\
&= E_\theta b_{\hat{\theta}}(\theta) b_{\hat{\theta}}^T(\theta) + E_\theta E_{Y|\theta}(\hat{\theta} - E_{Y|\theta}\hat{\theta})(\hat{\theta} - E_{Y|\theta}\hat{\theta})^T
\end{aligned}
$$

- **Point h:**

$$E_\theta(b_{LMMSE}(\theta) b_{LMMSE}^T(\theta)) = E_\theta[(E_{Y|\theta}\hat{\theta} - \theta)(E_{Y|\theta}\hat{\theta} - \theta)^T] = C_{\theta\theta} - C_{\theta Y} C_{YY}^{-1} C_{Y\theta}$$

  (as shown in point b)

- **Point i:**

  From point g:

  $R_{\tilde{\theta}\tilde{\theta}}^{LMMSE} = E_\theta[b_{LMMSE}(\theta)b_{LMMSE}^T(\theta)] + E_\theta E_{Y|\theta}(\hat{\theta}_{LMMSE} - E_{Y|\theta}\hat{\theta}_{LMMSE})(\hat{\theta}_{LMMSE} - E_{Y|\theta}\hat{\theta}_{LMMSE})^T$

  $R_{\tilde{\theta}\tilde{\theta}}^{LMMSE} = E_\theta E_{Y|\theta}\tilde{\theta}_{LMMSE}\tilde{\theta}_{LMMSE}^T = C_{\theta\theta} - C_{\theta Y}C_{YY}^{-1}C_{Y\theta}$

  From point h: $E_\theta[b_{LMMSE}(\theta)b_{LMMSE}^T(\theta)] = C_{\theta\theta} - C_{\theta Y}C_{YY}^{-1}C_{Y\theta}$

  $E_\theta E_{Y|\theta}(\hat{\theta}_{LMMSE} - E_{Y|\theta}\hat{\theta}_{LMMSE})(\hat{\theta}_{LMMSE} - E_{Y|\theta}\hat{\theta}_{LMMSE})^T = C_{\theta\theta} - C_{\theta Y}C_{YY}^{-1}C_{Y\theta} - (C_{\theta\theta} - C_{\theta Y}C_{YY}^{-1}C_{Y\theta}) = 0$

2 Deterministic Parameter Estimation

- **Point a:**

$$\int_{-\infty}^{\infty} f(y \mid \lambda, \alpha, \beta)\, dy = 1$$

$$\int_{\alpha}^{\beta} \gamma e^{-\lambda y}\, dy = \frac{\gamma}{\lambda}\left(e^{-\alpha\lambda} - e^{-\beta\lambda}\right) = 1$$

$$\to \gamma = \frac{\lambda}{e^{-\alpha\lambda} - e^{-\beta\lambda}}, \quad f(y \mid \lambda, \alpha, \beta) = \frac{\lambda e^{-\lambda y}}{e^{-\alpha\lambda} - e^{-\beta\lambda}} \mathbf{1}_{[\alpha,\beta)}(y)$$

- **Point b:**

$$l(\alpha, \beta | Y, \lambda) = \left(\frac{\lambda}{e^{-\alpha\lambda} - e^{-\beta\lambda}}\right)^n e^{-\lambda \sum_{i=1}^n y_i} \mathbf{1}_{[\alpha,\infty)}(y_{\min}) \mathbf{1}_{(-\infty,\beta)}(y_{\max})$$

- **Point c:** Log-likelihood $\alpha \le y \le \beta \to y_{\min} \ge \alpha, y_{\max} \le \beta$

$$\begin{aligned} L(\alpha, \beta \mid y, \lambda) &= \ln[l(\alpha, \beta \mid Y, \lambda)] \\ &= n \ln \lambda - n \ln\left(e^{-\alpha\lambda} - e^{-\beta\lambda}\right) - \lambda \sum_{i=1}^n y_i \end{aligned}$$

$$\begin{aligned} \frac{\partial L(\alpha, \beta \mid Y, \lambda)}{\partial \alpha} &= \frac{n\lambda}{1 - e^{(\alpha-\beta)\lambda}} > 0, \quad \hat{\alpha} = y_{\min} \\ \frac{\partial L(\alpha, \beta \mid Y, \lambda)}{\partial \beta} &= \frac{n\lambda}{1 - e^{(\beta-\alpha)\lambda}} < 0, \quad \hat{\beta} = y_{\max} \end{aligned}$$

- **Point d:**

$$\frac{\partial L(\alpha, \beta\lambda \mid Y)}{\partial \lambda} = \frac{n}{\lambda} - \frac{n(\alpha e^{-\alpha\lambda} - \beta e^{-\beta\lambda})}{e^{-\alpha\lambda} - e^{-\beta\lambda}} - \sum_{i=1}^n y_i$$

From now on, $\alpha = 0, \beta = \infty \to f(y \mid \lambda) = \lambda e^{-\lambda y}$

- **Point e:**

$$\begin{aligned} m_y = E_y &= \int_0^\infty y \lambda e^{-\lambda y}\, dy \\ &= -\int_0^\infty -y e^{-\lambda y}\, dy \\ &= -\left(y e^{-\lambda y}\Big|_0^\infty - \int_0^\infty e^{-\lambda y}\, dy \right) \\ &= -\left[-\frac{e^{-\lambda y}}{-\lambda}\Big|_0^\infty \right] \\ &= -\left[\frac{1}{\lambda}(0-1)\right] = \frac{1}{\lambda} \end{aligned}$$

$$\begin{aligned} Ey^2 &= \int_0^\infty y^2 \lambda e^{-\lambda y}\, dy \\ &= -\left[y^2 e^{-\lambda y}\Big|_0^\infty - \int_0^\infty 2y e^{-\lambda y}\, dy \quad \text{Integrate again by parts} \right] \\ &= 2\left(-\frac{y e^{-\lambda y}}{\lambda}\Big|_0^\infty - \frac{e^{-\lambda y}}{\lambda^2}\Big|_0^\infty \right) \\ &= 2\left[0 - \left(0 - \frac{1}{\lambda^2}\right)\right] = \frac{2}{\lambda^2} \\ \sigma_y^2 = E_{y^2} - (Ey)^2 &= \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2} \end{aligned}$$

- **Point f:**

$$L(\lambda \mid Y) = \ln f(Y \mid \lambda) = \sum_{i=1}^n \ln(\lambda e^{-\lambda y_i}) = \sum_{i=1}^n (\ln \lambda - \lambda y_i) = n \ln \lambda - \lambda \sum_{i=1}^n y_i$$

- **Point g:**

$$\frac{\partial L(\lambda \mid Y)}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} y_i = 0 \rightarrow \hat{\lambda}_{ML} = \frac{n}{\sum_{i=1}^{n} y_i} = \frac{1}{\bar{y}}$$

- **Point h:** In the method of moments, we set sample moments equal to population moments. Here, the first sample moment is the sample mean, and from what we obtained from point e:

$$m_y = E[y] = g(\lambda) = \frac{1}{\lambda} \rightarrow \hat{\lambda} = \frac{1}{m_y} = \frac{1}{\bar{y}}$$

- **Point i:** We want to find the asymptotic properties of $\hat{\lambda}_{ML}$ as $n \rightarrow \infty$. If we put $y = \hat{\lambda}_{ML}$ and $m_y = E(y)$, the first-order expansion is given by $y - m_y = g'(\theta_0)(\hat{\theta} - \theta_0)$, where $g(\theta)$ is a function of the parameter and θ_0 is the true value. In this case, $g(\lambda) = \lambda$, and $\theta_0 = \lambda$

$$\rightarrow y - \lambda = g'(\lambda)(\hat{\lambda}_{ML} - \lambda)$$

  Taking the expectations on both sides: $E(y) - \lambda = \text{var}(\hat{\lambda}_{ML})$

  Asymptotic mean: $m_{b\lambda_{ML}} = E(\hat{\lambda}_{ML}) \sim \lambda$

  Asymptotic variance $\sigma^2_{e\lambda_{ML}} \sim \text{var}(\hat{\lambda}_{ML})$

  For the MLE to be asymptotically unbiased, $m_{b\lambda_{ML}}$ should be equal to λ, which is the case here, so $\hat{\lambda}_{ML}$ is asymptotically unbiased, and its asymptotic mean and variance can be expressed in terms of λ.

- **Point j:** Fisher information matrix for deterministic parameters is

$$J(\lambda) = -E_{Y|\lambda} \frac{\partial}{\partial \lambda} \left(\frac{\partial \ln f(Y \mid \lambda)}{\partial \lambda} \right)^T$$

  From point f, we have $L(\lambda|Y) = n \ln \lambda - \lambda \sum_{i=1}^{n} y_i$

  Computing its second derivative and evaluating it negatively, we obtain $J(\lambda) = \frac{n}{\lambda^2}$

  Cramer-Rao bound: $var(\hat{\lambda}) \geq J^{-1}(\lambda) = \frac{\lambda^2}{n}$

  The estimator for λ, in this case, is asymptotically efficient because the variance of the estimator approaches the Cramer-Rao Bound as the sample size becomes very large ($\lim_{n\rightarrow\infty} CRB = 0$).