TTT4275 Summary for February 1th
Spring 2019

Lecturer: Magne Hallstein Johnsen,

IES, NTNU

Maximum Likelihood Estimator (MLE)

- The LSE estimator (linear model approximation) and BLUE did not need knowledge of $p(x, \theta)$. Thus CRLB can not be found, and the estimator quality is generally unknown.

- The MLE requires knowledge of $p(x, \theta)$, thus CRLB can be found

- The term likelihood means $L(\theta/x) = p(x, \theta)$ where x is known and θ is unknown/variable

- MLE is generally not efficient, but is always asymptoticallt efficient, i.e.

$$\begin{aligned} lim_{N\to\infty} E\{\widehat{\theta}\} &= \theta \quad (1) \\ lim_{N\to\infty} var(\widehat{\theta}) &= CRLB \end{aligned}$$

- As the name MLE indicates the estimator is found by

$$\widehat{\theta} = argmax_{\theta} L(\theta/x) \quad (2)$$

Bayesian estimation -1

- Classical estimation (LSE, BLUE, MLE) : θ is unknown but deterministic, i.e.
$$p(x,\theta) = p(x/\theta)p(\theta) = p(x/\theta)$$

- Bayesian estimation (BMSE, MAP): θ is unknown and a stochastic variable with prior density $p(\theta)$

- Bayesian MSE is given by
$$BMSE(\acute{\theta}) = E\{(\theta - \acute{\theta})^2\} = \int\int (\theta - \acute{\theta})^2 p(x,\theta) d\theta dx \quad (3)$$

- Utilizing $p(x,\theta) = p(\theta/x)p(x)$
$$\begin{aligned} BMSE(\acute{\theta}) &= \int F(\acute{\theta}, x) p(x) dx \text{ where} \\ F(\acute{\theta}, x) &= \int (\theta - \acute{\theta})^2 p(\theta/x) d\theta \end{aligned}$$

Bayesian estimation -2

- Minimizing $BMSE(\acute{\theta})$ is eqivalent to minimizing the integrand $F(\acute{\theta}, x)$ as all variables inside the integrals are positive

- Minimizing $F(\acute{\theta}, x)$ by setting the derivative wrt. $\acute{\theta}$ equal to zero results in
$$\hat{\theta} = E\{\theta/x\} = \int \theta p(\theta/x) dx \tag{4}$$

- The strategy then is first to find the posterior density $p(\theta/x)$ from
$$p(\theta/x) = \frac{p(x/\theta)p(\theta)}{p(x)}$$
and then calculate the integral in eq. 4.

- In practice it is seldom easy to calculate the integral. Thus a (suboptimal) strategy is to use the maximum value of the posterior (MAP)
$$\hat{\theta} = argmax_{\theta}\ p(\theta/x) = argmax_{\theta}\ \frac{p(x/\theta)p(\theta)}{p(x)} = argmax_{\theta}\ p(x/\theta)p(\theta) \tag{5}$$

- If the posterior is symmetric MAP and minimum BMSE are identical