4. Bayesian Estimation

4.1 Variance of a Gaussian Random Variable

Consider a scalar Gaussian random variable $\mathsf{x}|\theta \sim \mathcal{N}(\mu, \sigma^2 = \theta^{-1})$ with known mean μ and unknown inverse of the variance θ. Prior knowledge on θ is given by prior PDF $f_\theta(\theta)$. In contrast to Maximum-Likelihood estimation, Bayesian estimation considers prior statistical knowledge on the parameter θ to be estimated.

a) What are advantages of modeling the unknown parameter θ as a random variable? Are there potential disadvantages as well?

A prior PDF $f_\theta(\theta)$ is called a conjugate prior with respect to $f_{\mathsf{x}|\theta}(x|\theta)$ if the posterior PDF $f_{\theta|\mathsf{x}}(\theta|x)$ has the same structure as the prior PDF $f_\theta(\theta)$. For conjugate priors, the posterior PDF can be obtained in closed form. If θ parametrizes the inverse of the variance, the conjugate prior with respect to a Gaussian likelihood is given by the Gamma distribution.

In the following, we assume the prior PDF $\theta \sim \Gamma(\alpha, \beta)$ with $\alpha, \beta > 0$. In addition, we assume an estimation of θ based on N conditionally i.i.d. statistics $\mathsf{x}_1, \ldots, \mathsf{x}_N|\theta$ with the conditional joint PDF $f_{\mathsf{x}_1,\ldots,\mathsf{x}_N|\theta}(x_1, \ldots, x_N|\theta) = \prod_{i=1}^{N} f_{\mathsf{x}|\theta}(x_i|\theta)$.

b) State Bayes' rule for the posterior PDF $f_{\theta|\mathsf{x}_1,\ldots,\mathsf{x}_N}(\theta|x_1, \ldots, x_N)$.

c) Determine the posterior PDF $f_{\theta|\mathsf{x}_1,\ldots,\mathsf{x}_N}(\theta|x_1, \ldots, x_N)$.

Hint: The PDF $f_b(b)$ of a Gamma distribution $b \sim \Gamma(\alpha, \beta)$ is given as

$$f_b(b; \alpha, \beta) = \beta^\alpha \frac{1}{\Gamma(\alpha)} b^{\alpha-1} e^{-\beta b} \quad \text{for } b \geq 0 \text{ and } \alpha, \beta > 0.$$

d) Determine the Bayes estimator $\hat{\theta}_{\mathrm{CM}} = \mathrm{E}[\theta|x_1, \ldots, x_N]$. Compare the result to the Maximum-Likelihood estimator $\hat{\sigma}^2_{\mathrm{ML}}$ for the variance σ^2 of a Gaussian distribution with known mean μ as derived in problem 2.1f). Discuss the behavior of the estimated variance/precision for N approaching zero or infinity.

Hint: For $b \sim \Gamma(\alpha, \beta)$, it follows that $\mathrm{E}[b] = \alpha/\beta$.

4.2 Mean of a Gaussian Random Variable

Consider an estimator for the expected value of the random variable x given $\boldsymbol{\theta}$ with $\mathsf{x}|\boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\theta}, \boldsymbol{C}_{\mathsf{x}|\boldsymbol{\theta}})$. The parameter to be estimated is therefore $\mathrm{E}[\mathsf{x}|\boldsymbol{\theta}] = \boldsymbol{\theta}$. The statistical knowledge on $\boldsymbol{\theta}$ is given by the prior distribution $\boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{m}_{\boldsymbol{\theta}}, \boldsymbol{C}_{\boldsymbol{\theta}})$. We assume N i.i.d. statistics $\mathsf{x}_1, \ldots, \mathsf{x}_N$ with the conditional joint PDF $f_{\mathsf{x}_1,\ldots,\mathsf{x}_N|\boldsymbol{\theta}}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N|\boldsymbol{\theta}) = \prod_{i=1}^{N} f_{\mathsf{x}|\boldsymbol{\theta}}(\boldsymbol{x}_i|\boldsymbol{\theta})$.

a) State Bayes' rule for the posterior PDF $f_{\boldsymbol{\theta}|\mathsf{x}_1,\ldots,\mathsf{x}_N}(\boldsymbol{\theta}|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N)$.

b) Determine the posterior PDF $f_{\boldsymbol{\theta}|\mathsf{x}_1,\ldots,\mathsf{x}_N}(\boldsymbol{\theta}|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N)$.

Hint: Compare the result to the structure of a Gaussian PDF.

c) Determine the Bayes estimator $\hat{\boldsymbol{\theta}}_{\mathrm{CM}} = \mathrm{E}[\boldsymbol{\theta}|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N]$ using the result from b).

Hint: $(\boldsymbol{A}^{-1} + \boldsymbol{B}^{-1})^{-1} = \boldsymbol{A}(\boldsymbol{A} + \boldsymbol{B})^{-1}\boldsymbol{B} = \boldsymbol{B}(\boldsymbol{A} + \boldsymbol{B})^{-1}\boldsymbol{A}$.

d) Determine the Bayes estimator $\hat{\theta}_{\mathrm{CM}} = \mathrm{E}[\theta|x_1, \ldots, x_N]$ using your result for the posterior PDF $f_{\theta|\mathsf{x}_1,\ldots,\mathsf{x}_N}(\theta|x_1, \ldots, x_N)$ for scalar random variables. Give an interpretation on how the prior knowledge on θ is combined with the empirical information obtained from the observations $x_1, \ldots, x_N$.

4.3 Gaussian Transmission Model

Consider the transmission model given as

$$\boldsymbol{x} = \boldsymbol{H}\boldsymbol{\theta} + \boldsymbol{n},$$

where $\boldsymbol{x} \in \mathbb{R}^M$ is the receive symbol, $\boldsymbol{H} \in \mathbb{R}^{M\times L}$ is the known channel matrix, $\boldsymbol{\theta} \in \mathbb{R}^L$ is the random transmit symbol, and $\boldsymbol{n} \in \mathbb{R}^M$ is additive Gaussian noise with $\boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{C}_{\boldsymbol{n}})$.

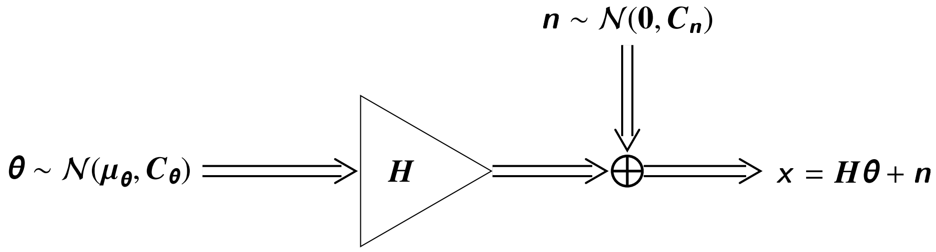

Fig. 4.1: Illustration of the linear transmission model with the channel matrix $\boldsymbol{H}$, the transmitted signal and parameter of interest $\boldsymbol{\theta}$, and the additive Gaussian noise $\boldsymbol{n}$.

The covariance matrix $\boldsymbol{C}_{\boldsymbol{n}}$ of the noise is assumed to be known. Prior knowledge on the random transmit symbol $\boldsymbol{\theta}$ is given by the prior PDF $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$. It is assumed that $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ is Gaussian, i.e., $\boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{\theta}}, \boldsymbol{C}_{\boldsymbol{\theta}})$. The random variables $\boldsymbol{\theta}$ and $\boldsymbol{n}$ are assumed to be independent.

a) Determine the posterior PDF $f_{\boldsymbol{\theta}|\boldsymbol{x}}(\boldsymbol{\theta}|\boldsymbol{x})$.

Hint: If two random variables $\boldsymbol{a} \sim \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{a}}, \boldsymbol{C}_{\boldsymbol{a}})$ and $\boldsymbol{b} \sim \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{b}}, \boldsymbol{C}_{\boldsymbol{b}})$ are jointly Gaussian with expected value $\boldsymbol{\mu}_{[\boldsymbol{a}^{\mathrm{T}}\ \boldsymbol{b}^{\mathrm{T}}]^{\mathrm{T}}} = [\boldsymbol{\mu}_{\boldsymbol{a}}^{\mathrm{T}}\ \boldsymbol{\mu}_{\boldsymbol{a}}^{\mathrm{T}}]^{\mathrm{T}}$ and covariance matrix

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}_{\boldsymbol{a}} & \boldsymbol{C}_{\boldsymbol{ab}} \\ \boldsymbol{C}_{\boldsymbol{ba}} & \boldsymbol{C}_{\boldsymbol{b}} \end{bmatrix},$$

the conditional PDF $f_{\boldsymbol{b}|\boldsymbol{a}}(\boldsymbol{b}|\boldsymbol{a})$ of $\boldsymbol{b}$ given $\boldsymbol{a}$ is again Gaussian with the conditional mean

$$\mathrm{E}_{\boldsymbol{b}|\boldsymbol{a}}[\boldsymbol{b}|\boldsymbol{a}] = \mathrm{E}\left[\boldsymbol{b}\right] + \boldsymbol{C}_{\boldsymbol{ba}}\boldsymbol{C}_{\boldsymbol{a}}^{-1}(\boldsymbol{a} - \mathrm{E}\left[\boldsymbol{a}\right])$$

and the conditional covariance matrix

$$\boldsymbol{C}_{\boldsymbol{b}|\boldsymbol{a}} = \boldsymbol{C}_{\boldsymbol{b}} - \boldsymbol{C}_{\boldsymbol{ba}}\boldsymbol{C}_{\boldsymbol{a}}^{-1}\boldsymbol{C}_{\boldsymbol{ab}}.$$

b) Determine the Bayes estimator $\hat{\boldsymbol{\theta}}_{\mathrm{CM}} = \mathrm{E}[\boldsymbol{\theta}|\boldsymbol{x}]$.

c) Let $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{\theta}$ be the random mean of $\boldsymbol{x}$ with realization $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{\theta}$. Show that the Bayes estimator $\hat{\boldsymbol{y}}_{\mathrm{CM}} = \mathrm{E}[\boldsymbol{y}|\boldsymbol{x}]$ of the mean $\boldsymbol{y}$ is given as $\hat{\boldsymbol{y}}_{\mathrm{CM}} = \boldsymbol{H}\hat{\boldsymbol{\theta}}_{\mathrm{CM}}$.