227-0689-00L System Identification **28 September 2022**

Exercise 1: Data fit statistics & ML/MAP estimators

Background reading

The background material for this exercise is Section 7.4 of Ljung's book (*System Identification: Theory for the User*, 2nd Ed., Prentice-Hall, 1998).

Problem 1

1. Let the two random vectors X and Y be jointly Gaussian with

$$\mathcal{E}\{X\} = m_X,\ \mathcal{E}\{Y\} = m_Y,$$
$$\mathcal{E}\left\{(X - m_X)(X - m_Y)^\mathsf{T}\right\} = P_X,\ \mathcal{E}\left\{(Y - m_Y)(Y - m_Y)^\mathsf{T}\right\} = P_Y,$$
$$\mathcal{E}\left\{(X - m_X)(Y - m_Y)^\mathsf{T}\right\} = P_{XY},$$

   Show that the conditional distribution of X given Y is

$$X \mid Y \sim \mathcal{N}\left(m_X + P_{XY}P_Y^{-1}(Y - m_Y), P_X - P_{XY}P_Y^{-1}P_{XY}^\mathsf{T}\right). \tag{1.1}$$

   Hint: If D and $A - BD^{-1}C$ are invertible, the following equalities hold:

$$\det\left(\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right) = \det(D)\det(A - BD^{-1}C), \tag{1.2}$$
$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} (A - BD^{-1}C)^{-1} & -(A - BD^{-1}C)^{-1}BD^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & D^{-1} + D^{-1}C(A - BD^{-1}C)^{-1}BD^{-1} \end{bmatrix}. \tag{1.3}$$

2. Consider a static linear model:

$$y = \theta x + v.$$

   We have obtained the prior distribution of θ: $\theta \sim \mathcal{N}(\mu, \sigma_\theta^2)$, the noise distribution $v \sim \mathcal{N}(0, \sigma_v^2)$, and a pair of measurements (x_0, y_0). Please use the result in part 1 to calculate the maximum a posteriori (MAP) estimate of θ.

Problem 2

Consider the model structure

$$\mathbf{x} = F(\theta)\mathbf{w},$$
$$\mathbf{y} = H(\theta)\mathbf{x} + \mathbf{e},$$

where $\mathbf{w}$ and $\mathbf{e}$ are two independent Gaussian random vectors with zero mean values and unit covariance matrices.

1. Find the maximum likelihood estimate of the parameter θ given the measurement $\mathbf{y}$.
2. Assume that the prior distribution of θ is flat, i.e., $p(\theta)$ is independent of θ. Find the joint MAP estimate of the parameter θ and the intermediate variable $\mathbf{x}$ given the measurement $\mathbf{y}$.

You only need to derive the optimization problems for obtaining the estimates.

MATLAB **exercises:**

Consider the problem of estimating $\theta = [\theta_1\ \theta_2]^\top$ in the following function

$$y = \theta_1 x + \theta_2(2x^2 - 1) + v\,,$$

where noise v comes from the normal distribution $v \sim \mathcal{N}(\mu, \sigma^2)$ with $\mu = 0.2$ and $\sigma^2 = 0.1$.

In order to estimate θ, a set of x and y measurements are collected. The test data are provided as variables `x` and `y` in `SysID_Exercise_1.mat`.

1. Use the maximum likelihood (ML) method to estimate the value of θ.
2. The prior knowledge of θ is characterized by the following distribution,

$$\theta \sim \mathcal{N}(\mu_\theta, \sigma_\theta^2 \cdot I), \text{ with } \mu_\theta = [1\ 0.4]^\top \text{ and } \sigma_\theta^2 = 0.01.$$

   Calculate the maximum a posteriori (MAP) estimate of θ.
3. To assess the accuracy of the ML and MAP estimates, additional measurements are collected for validation, which are provided as variables `x_v` and `y_v` in `SysID_Exercise_1.mat`. Which one is more accurate judging from the validation data?