# Machine Learnig Homework 10

## Problem 1

$$\begin{pmatrix} y \\ f(x_*) \end{pmatrix} \sim \mathcal{N}\begin{pmatrix} \mu & K & K_* \\ \mu_* ' & K_*^T & K_{**} \end{pmatrix} \tag{1}$$

Since the mean function $m(X) = 0$ for every $X$, both $\mu$ and $\mu_*$ are equal to zero and:

$$\begin{aligned} K_{ij} &= K(X_i, X_j) \\ K_{*ij} &= K(X_i, X_{*j}) \\ K_{**ij} &= K(X_{*i}, X_{*j}) \end{aligned} \tag{2}$$

Therefore

$$\begin{pmatrix} y \\ f(x_*) \end{pmatrix} \sim \mathcal{N}(\mathbf{0}, \Sigma) \tag{3}$$

where

$$\begin{aligned} \Sigma = \begin{pmatrix} K & K_* \\ K_*^T & K_{**} \end{pmatrix} &= \begin{pmatrix} K(x_1, x_1) & K(x_1, x_2) & K(x_1, x_*) \\ K(x_2, x_1) & K(x_2, x_1) & K(x_2, x_*) \\ K(x_*, x_1) & K(x_*, x_2) & K(x_*, x_*) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0.54 & 0.88 \\ 0.54 & 1.2 & 0.61 \\ 0.88 & 0.61 & 1 \end{pmatrix} \end{aligned} \tag{4}$$

## Problem 2

$$p(f_* | \mathbf{y}, \mathbf{X}) \sim \mathcal{N}\left(\mu_{f_*|\mathbf{y},\mathbf{X}}, \Sigma_{f_*|\mathbf{y},\mathbf{X}}\right) \tag{5}$$

where

$$\begin{aligned} \mu_{f_*|\mathbf{y},\mathbf{X}} &= K_*^T K^{-1} y = 2.5039 \\ \Sigma_{f_*|\mathbf{y},\mathbf{X}} &= K_{**} - K_*^T K^{-1} K_* = 0.196 \end{aligned} \tag{6}$$

## Problem 3

With noisy observations assumed, the covariance matrix takes the form of $\widetilde{K} = K + \sigma_y^2 I$. Consequently, the covariance matrix of the joint distribution is

$$\widetilde{\Sigma} = \begin{pmatrix} \widetilde{K} & K_* \\ K_*^T & K_{**} \end{pmatrix} = \begin{pmatrix} K + \sigma_y^2 I & K_* \\ K_*^T & K_{**} \end{pmatrix} \tag{7}$$

Finally, the predictive distribution is given by

$$\begin{aligned}
&\mathcal{N}(\widetilde{\mu_*}, \widetilde{\Sigma_*}) \\
&\widetilde{\mu_*} = m(f(x_*)) + K_*^T (K + \sigma_y^2 I)^{-1} (f - m(y)) = K_*^T (K + \sigma_y^2 I)^{-1} f \\
&\widetilde{\Sigma_*} = K_{**} - K_*^T (K + \sigma_y^2 I)^{-1} K_* = 1 - K_*^T (K + \sigma_y^2 I)^{-1} K_*
\end{aligned} \tag{8}$$