Machine Learning Worksheet 09

Gaussian Process Regression

Let $\boldsymbol{X} = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix} = \begin{pmatrix} 0 & 1 \\ 0.5 & 2 \end{pmatrix}$ and $\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$ be observations and $\boldsymbol{x}_* = \begin{pmatrix} 0.5 & 1 \end{pmatrix}$ be a datapoint with unknown function value $\boldsymbol{f}_* = f(\boldsymbol{x}_*)$.

A distribution over the function f is given by a Gaussian process $f \sim GP(m, K)$ with mean function $m(\boldsymbol{x}) = 0$ and covariance function $K(\boldsymbol{x}, \boldsymbol{x}') = \exp(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}')^T(\boldsymbol{x} - \boldsymbol{x}'))$.

Problem 1: Assuming a noise-free scenario, write down the joint distribution for $\boldsymbol{f}_{\text{jt}} = \begin{pmatrix} y_1 \\ y_2 \\ f(\boldsymbol{x}_*) \end{pmatrix}$.

$$\boldsymbol{f}_{\text{jt}} \sim \mathcal{N}(\boldsymbol{m}_{\text{jt}}, \boldsymbol{K}_{\text{jt}}) = \mathcal{N}\left(\begin{bmatrix} \boldsymbol{m}(\boldsymbol{X}) \\ m(\boldsymbol{x}_*) \end{bmatrix}, \begin{bmatrix} \boldsymbol{K}(\boldsymbol{X}, \boldsymbol{X}) & \boldsymbol{K}(\boldsymbol{X}, \boldsymbol{x}_*) \\ \boldsymbol{K}(\boldsymbol{x}_*, \boldsymbol{X}) & \boldsymbol{K}(\boldsymbol{x}_*, \boldsymbol{x}_*) \end{bmatrix} \right) \sim \mathcal{N}\left(\boldsymbol{0}, \begin{pmatrix} 1. & 0.54 & 0.88 \\ 0.54 & 1. & 0.61 \\ 0.88 & 0.61 & 1. \end{pmatrix} \right)$$

Problem 2: Write down the conditional distribution $p(f_* | \boldsymbol{y}, \boldsymbol{X})$ using the rules for conditionals of an MVN (MVN/GP lecture slides).

Setting $\boldsymbol{y}_2 = f_*$, $\boldsymbol{y}_1 = \boldsymbol{y}$, $\boldsymbol{\Sigma}_{22} = \boldsymbol{K}(\boldsymbol{x}_*, \boldsymbol{x}_*)$, $\boldsymbol{\Sigma}_{21} = \boldsymbol{K}(x_*, X) = \boldsymbol{\Sigma}_{12}^T$, $\boldsymbol{\Sigma}_{11} = \boldsymbol{K}(\boldsymbol{X}, \boldsymbol{X})$, $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \boldsymbol{0}$:

$$f_* | \boldsymbol{y} = \boldsymbol{y}_2 | \boldsymbol{y}_1 \sim \mathcal{N}(\boldsymbol{\mu}_{2|1}, \boldsymbol{\Sigma}_{2|1})$$

$$\boldsymbol{\mu}_{2|1} = \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} (\boldsymbol{y}_1 - \boldsymbol{\mu}_1) \sim \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0.88 & 0.61 \end{pmatrix} \begin{pmatrix} 1 & 0.54 \\ 0.54 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 2 \\ 5 \end{pmatrix} \sim 2.50$$

$$\boldsymbol{\Sigma}_{2|1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12} \sim 1 - \begin{pmatrix} 0.88 & 0.61 \end{pmatrix} \begin{pmatrix} 1 & 0.54 \\ 0.54 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0.88 \\ 0.61 \end{pmatrix} \sim 0.20$$

$$f_* | \boldsymbol{y} \sim \mathcal{N}(2.50, 0.20)$$

Problem 3: Now, assume instead that the observations are disturbed by Gaussian noise with variance σ_n^2. Write down the joint distribution $p(\boldsymbol{y}, f(\boldsymbol{x}_1))$.

$$\begin{pmatrix} y_1 \\ y_2 \\ f(\boldsymbol{x}_1) \end{pmatrix} \sim \mathcal{N}\left(\boldsymbol{0}, \begin{pmatrix} 1 + \sigma_n^2 & 0.54 & 0.88 \\ 0.54 & 1 + \sigma_n^2 & 0.61 \\ 0.88 & 0.61 & 1 \end{pmatrix} \right)$$
