Errata and Extra Tips for Lab
Assignment 3
Machine Learning 1, 17/18

This document contains errata and extra tips for lab assignment 3.

Part 1: Gaussian Processes, 2. Predictive distribution

Extension of equations 6.66 and 6.67 for an arbitrary number of test points:

- Consider a data set $\mathbf{x}_{\text{train}} = (x_{\text{train},1}, ..., x_{\text{train},N_{\text{train}}})$, $\mathbf{t} = (t_1, ..., t_{N_{\text{train}}})$, $\mathbf{x}_{\text{test}} = (x_{\text{test},1}, ..., x_{\text{test},N_{\text{test}}})$ and a kernel function $k(x_n, x_m)$
- $\mathbf{K}$: $K_{n,m} = k(x_{\text{train},n}, x_{\text{train},m})$; $\mathbf{K}$ is a $N_{\text{train}} \times N_{\text{train}}$ matrix
- $\mathbf{C}$: $\mathbf{C} = \mathbf{K} + \beta^{-1}\mathbf{I}$; $\mathbf{C}$ is a $N_{\text{train}} \times N_{\text{train}}$ matrix
- $\mathbf{k}$: $k_{n,m} = k(x_{\text{train},n}, x_{\text{test},m})$; $\mathbf{k}$ is a $N_{\text{train}} \times N_{\text{test}}$ matrix
- $\mathbf{c}$: $c_{n,m} = k(x_{\text{test},n}, x_{\text{test},m}) + \beta^{-1}$; $\mathbf{k}$ is a $N_{\text{test}} \times N_{\text{test}}$ matrix
- Equation 6.66 becomes: $\mu = \mathbf{k}^T\mathbf{C}^{-1}\mathbf{t}$; μ is a vector with length N_{test}
- Equation 6.67 becomes: $\mathbf{\Sigma} = \mathbf{c} - \mathbf{k}^T\mathbf{C}^{-1}\mathbf{k}$; $\mathbf{\Sigma}$ is a $N_{\text{test}} \times N_{\text{test}}$ matrix

The model + target noise for every data point in the test vector $\mathbf{x}_{\text{test}}$ is given by

$$\sigma = \sqrt{diag(\mathbf{\Sigma})},$$

where σ is a vector with length N_{test}.