Machine Learning Worksheet 08

Kernels and Constrained Optimisation

1 Kernelised k-nearest neighbours

To classify the point $\boldsymbol{x}$ the k-nearest neighbours finds the k training samples $\mathcal{N} = \{\boldsymbol{x}^{(s_1)}, \boldsymbol{x}^{(s_2)}, \ldots, \boldsymbol{x}^{(s_k)}\}$ that have the shortest distance $|\boldsymbol{x} - \boldsymbol{x}^{(s_i)}|_2$ to $\boldsymbol{x}$. Then the label that is mostly represented in the neighbour set $\mathcal{N}$ is assigned to $\boldsymbol{x}$.

Problem 1: Formulate the k-nearest neighbours algorithm in feature space by introducing the feature map $\boldsymbol{\phi}(\boldsymbol{x})$. Then rewrite the k-nearest neighbours algorithm so that it only depends on the scalar product in feature space $K(\boldsymbol{\phi}(\boldsymbol{x}), \boldsymbol{\phi}(\boldsymbol{y})) = \boldsymbol{\phi}(\boldsymbol{x})^T \boldsymbol{\phi}(\boldsymbol{y})$.

2 Convex functions

Problem 2: Given two convex functions $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ show that the sum $h(\boldsymbol{x}) = f(\boldsymbol{x}) + g(\boldsymbol{x})$ and the scaled function $u(\boldsymbol{x}) = cf(\boldsymbol{x})$ with $c \geq 0$ are convex.

Problem 3: Consider the family of convex functions $f_\lambda(\boldsymbol{x})$, that is for every $\lambda \in \mathbb{R}$ the function $f_\lambda(\boldsymbol{x})$ is convex. Prove that the point-wise maximum $g(\boldsymbol{x}) = \max_\lambda f_\lambda(\boldsymbol{x})$ is convex.

Problem 4: Show that the Lagrange dual function $g(\boldsymbol{\alpha}) = \min_{\boldsymbol{x}} L(\boldsymbol{x}, \boldsymbol{\alpha})$ is concave. (A function $f(x)$ is concave if and only if $-f(x)$ is convex.)

3 A simple constrained optimization problem

Consider the simple optimization problem

$$\begin{aligned} &\text{minimize } f_0(x) = x^2 + 1 \\ &\text{subject to } f_1(x) = (x-2)(x-4) \leq 0\,. \end{aligned}$$

Problem 5: Plot the objective $f_0(x)$ and the constraint $f_1(x)$ versus x in one plot. Show the feasible points. Use this plot to directly give the solution of the optimisation problem.

Problem 6: Derive the Lagrangian $L(x, \alpha)$ and use a computer program to plot it for $\alpha \in \{0, 0.5, 1, 1.5, 2, 3, 4, 5, 8\}$. For which regions is the value of the Lagrangian larger than the objective function? For which regions is the value of the Lagrangian smaller than the objective function? Which points are unaffected? What is the upper bound of $\min_x L(x, \alpha)$ for all $\alpha \geq 0$?

Problem 7: Derive and plot the Lagrange dual function $g(\alpha)$. State the dual problem.

Problem 8: Find the dual optimal value and the dual optimal solution α^*.

Problem 9: Is the dual optimal value also the minimum of the original optimisation problem?

Problem 10: Is the constraint f_1 active or inactive? Can you also see this from the plot of the primal problem? What does it mean when a constraint is active, i.e. what is the effect of an active constraint on the solution?