Problem Set 3

1 Regularized Normal Equation for Linear Regression

Given a data set $\{x^{(i)}, y^{(i)}\}_{i=1,\cdots,m}$ with $x^{(i)} \in \mathbb{R}^n$ and $y^{(i)} \in \mathbb{R}$, the general form of regularized linear regression is as follows

$$\min_{\theta} \frac{1}{2m} \left[\sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2 + \lambda \sum_{j=1}^{n} \theta_j^2 \right] \tag{1}$$

Derive the normal equation.

2 Lagrange Duality

Formulate the Lagrange dual problem of the following linear programming problem

$$\begin{aligned} \min \quad & c^T x \\ s.t. \quad & Ax \preceq b \end{aligned}$$

where $x \in \mathbb{R}^n$ is variable, $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{k\times n}$, $b \in \mathbb{R}^k$.

3 SVM

3.1 Convex Functions

Prove $f(\omega) = \omega^T \omega$ (where $\omega \in \mathbb{R}^n$) is a convex function.

3.2 Soft-Margin for Separable Data

Consider training a soft-margin SVM with C set to some positive constant. Suppose the training data is linearly separable. Since increasing the ξ_i can only increase the objective of the primal problem (which we are trying to minimize), at the optimal solution to the primal problem, all the training examples will have functional margin at least 1 and all the ξ_i will be equal to zero. True or false? Explain! Given a linearly separable dataset, is it necessarily better to use a a hard margin SVM over a soft-margin SVM?

3.3 In-bound Support Vectors in Soft-Margin SVMs

Examples $x^{(i)}$ with $\alpha_i > 0$ are called *support vectors* (SVs). For soft-margin SVM we distinguish between *in-bound* SVs, for which $0 < \alpha_i < C$, and *bound* SVs for which $\alpha_i = C$. Show that in-bound SVs lie exactly on the margin. Argue that bound SVs can lie both on or in the margin, and that they will "usually" lie in the margin. Hint: use the KKT conditions.