ELE 535 - Machine Learning and Pattern Recognition

Fall 2018

HOMEWORK 8: Theory

Q1 Let $\{(x_i, y_i)\}_{i=1}^m$ with $x_i \in \mathbb{R}^n$ and $y_i \in \{\pm 1\}$, $i \in [1:m]$, be a linearly separable set of training data. Show that if C is sufficiently large, the solution of the primal SVM problem will give the unique maximum margin separating hyperplane. How large does C need to be?

Q2 Let $\{(x_i, y_i\}_{i=1}^m$ with $x_i \in \mathbb{R}^n$ and $y_i \in \{\pm 1\}$, $i \in [1\!:\!m]$, be a training dataset. For a fixed value of C, let the corresponding SVM classifier have parameters $w^\star, b^\star$.

(a) Let $h \in \mathbb{R}^n$ and $Q \in \mathcal{O}_n$, and form the second training set: $\{Q(x_i - h), y_i)\}_{i=1}^m$. Show that the SVM classifier for this second dataset using the same value of C has parameters $Qw^\star, {w^\star}^T h + b^\star$.

(b) If we first center the training examples, how does this change the SVM classifier?

Q3 Give a clear and concise derivation of the dual of the primal linear SVM problem shown below and explain the origin of each of the constraints in the dual problem.

$$\begin{aligned} \min_{w\in\mathbb{R}^n, b\in\mathbb{R}, s\in\mathbb{R}^m} \quad & 1/2\, w^T w + C\mathbf{1}^T s \\ \text{s.t.} \quad & Z^T w + by + s - \mathbf{1} \geq \mathbf{0} \\ & s \geq \mathbf{0}. \end{aligned}$$

Q4 Suppose that instead of using $C\sum_{i=1}^m s_i$ as the penalty term in the objective of the primal SVM problem we use the quadratic penalty $1/2 C \sum_{i=1}^m s_i^2$, while maintaining the constraint $s_i \geq 0$.

(a) Formulate the new primal problem in vector form. When is the primal problem feasible?

(b) Does strong duality hold for this problem? Justify your answer.

(c) Write down the KKT conditions.

(d) Find the dual problem.

Q5 You are provided with $m > 1$ data points $\{x_j \in \mathbb{R}^n\}_{j=1}^m$ of which at least d, with $1 < d \leq m$ are distinct. Let $X = [x_1, \ldots, x_m]$ and consider the one class SVM problem:

$$\begin{aligned} \min_{R\in\mathbb{R}, a\in\mathbb{R}^n, s\in\mathbb{R}^m} \quad & R^2 + C\mathbf{1}^T s \\ \text{s.t.} \quad & \|x_i - a\|_2^2 \leq R^2 + s_i, \quad i = 1, \ldots, m, \\ & s \geq \mathbf{0}. \end{aligned}$$

(a) Show that this is a feasible convex program and that strong duality holds. [Hint: let $r = R^2$]

(b) Write down the KKT conditions.

(c) Show that $\alpha^\star \neq \mathbf{0}$ and that if $C > 1/(d-1)$ then $(R^2)^\star > 0$ (harder).

(d) What are the support vectors for this problem?

(e) Derive the dual problem.

(f) Assume $C > 1/(d-1)$. Given the dual solution, how should a and R^2 be selected?