

# Homework #5: Machine Learning

Kyler Little

April 14, 2018

## Problem #1

Consider two finite-dimensional feature transforms  $\Phi_1$  and  $\Phi_2$  and their corresponding kernels  $K_1$  and  $K_2$ .

(a) Define  $\Phi(x) = (\Phi_1(x), \Phi_2(x))$ . Express the corresponding kernel of  $\Phi$  in terms of  $K_1$  and  $K_2$ .

(b) Consider the matrix  $\Phi_1(x)\Phi_2(x)^T$  and let  $\Phi(x)$  be the vector representation of the matrix (say, by concatenating all the rows). Express the corresponding kernel of  $\Phi$  in terms of  $K_1$  and  $K_2$ .

(c) Hence, show that if  $K_1$  and  $K_2$  are kernels, then so are  $K_1 + K_2$  and  $K_1 K_2$ .

## Problem #2

Exercise 8.16 (e-Chap:8-42) in LFD. Please ignore part (c) and only do (a), (b), (d).

Exercise 8.16 Show that the optimization problem in (8.30) is a QP-problem.

(a) Show that the optimization variable is  $u = \begin{bmatrix} b \\ w \\ \xi \end{bmatrix}$ , where  $\xi = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$ .

(b) Show that  $u^* \leftarrow \text{QP}(Q, p, A, c)$ , where

$$Q = \begin{bmatrix} 0 & 0_d^T & 0_N^T \\ 0_d & I_d & 0_{d \times N} \\ 0_N & 0_{N \times d} & 0_{N \times N} \end{bmatrix}, p = \begin{bmatrix} 0_{d+1} \\ C \cdot 1_N \end{bmatrix}, A = \begin{bmatrix} YX & I_N \\ 0_{N \times d+1} & I_N \end{bmatrix}, \text{ and } c = \begin{bmatrix} 1_N \\ 0_N \end{bmatrix},$$

and  $YX$  is the signed data matrix from Exercise 8.4.

(d) How do you determine which data points violate the margin, which data points are on the edge of the margin and which data points are correctly separated and outside the margin?

## Problem #3

(a) Describe what are hinge loss, logistic regression loss, and 0-1 loss mathematically. Describe their similarities and differences using the unified picture we developed in class.

(b) By relying on the result in the above question, consider a point that is correctly classified and distant from the decision boundary. Why would SVMs decision boundary be unaffected by this point, but the one learned by logistic regression be affected?

## Problem #4

Summarize your observations from the coding portion into a short report. In your report, please report the accuracy result and total support vector number of each model. A briefly analysis based on the results is also needed.