

# Homework 02: The Bias-variance Tradeoff, SVMs and Kernel Methods

CS 6316 Machine Learning

Due on Oct. 24, 2023 11:59 PM

1. **Bias-Variance Tradeoff** (11 points)

Please refer to the attached iPython notebook file.

2. **KKT Conditions** (5 points) The Lagrangian form of SVMs with slack variable  $\xi$  is formulated as

$$\begin{aligned} L(\mathbf{w}, b, \xi, \alpha, \beta) = & \frac{1}{2} \|\mathbf{w}\|_2^2 + C \sum_{i=1}^m \xi_i \\ & - \sum_{i=1}^m \alpha_i (y_i (\mathbf{w}^\top \mathbf{x}_i + b) - 1 + \xi_i) \\ & - \sum_{i=1}^m \beta_i \xi_i \end{aligned} \quad (1)$$

Similar to SVMs in separable cases (slide 19), to find the KKT conditions (as in slide 28, we need to compute the derivative with respect to all parameters  $\{\mathbf{w}, b, \xi, \alpha, \beta\}$ . Overall, please show that the KKT conditions can be represented to the following equations

$$\mathbf{w} = \sum_{i=1}^m \alpha_i y_i \mathbf{x}_i \quad (2)$$

$$\sum_{i=1}^m \alpha_i y_i = 0 \quad (3)$$

$$\alpha_i + \beta_i = C \quad (4)$$

$$\alpha_i = 0 \quad \text{or} \quad y_i (\mathbf{w}^\top \mathbf{x}_i + b) = 1 - \xi_i \quad (5)$$

$$\beta_i = 0 \quad \text{or} \quad \xi_i = 0 \quad (6)$$

3. **Kernel Methods** (4 points) In our lecture, we show that a special case of the polynomial kernels

$$K(\mathbf{x}, \mathbf{x}') = (\langle \mathbf{x}, \mathbf{x}' \rangle + c)^d \quad (7)$$

with  $d = 2$  and  $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^2$ . On page 54 of the slides, we should show how this special case can be decomposed as a dot product with a nonlinear mapping  $\Phi(\cdot)$

$$K(\mathbf{x}, \mathbf{x}') = \langle \Phi(\mathbf{x}), \Phi(\mathbf{x}') \rangle. \quad (8)$$

In this problem, consider  $d = 3$  with  $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^2$  and show how the  $\Phi(\mathbf{x})$  is defined in this case.