CIS 419/519: Homework 4

Jiatong Sun

02/29/2020

Although the solutions are entirely my own, I consulted with the following people and sources while working on this homework:
$https://oeis.org/wiki/List_of_LaTeX_mathematical_symbols$
$https://tex.stackexchange.com/questions/122778/left-brace-including-several-lines-in-eqnarray$

1 Fitting an SVM by Hand

a. As is given in the problem:

$$x1 = 0, \quad x2 = \sqrt{2} \tag{1}$$

and

$$\phi(x) = [1, \sqrt{2}, x^2]^T \tag{2}$$

We know that

$$\phi(x_1) = [1, 0, 0]^T \quad \phi(x_2) = [1, 2, 2]^T \tag{3}$$

Since the optimal vector $\boldsymbol{w}$ is orthogonal to the dicision to the decision boundary, it is parallel to the vector connecting $\phi(x_1)$ and $\phi(x_2)$.

Since

$$\phi(x_2) - \phi(x_1) = [1, 2, 2]^T - [1, 0, 0]^T = [0, 2, 2]^T \tag{4}$$

So $[0, 2, 2]^T$ is a vector that is parallel to the optimal vector $\boldsymbol{w}$.

b. The margin is the distance between the two points in the 3D space.

$$margin = ||\phi(x_2) - \phi(x_1)|| = \sqrt{(1-1)^2 + (2-0)^2 + (2-0)^2} = 2\sqrt{2} \tag{5}$$

c. From the result of a, we can assume that

$$\boldsymbol{w} = [0, 2t, 2t]^T \tag{6}$$

So

$$||\boldsymbol{w}|| = \sqrt{0^2 + (2t)^2 + (2t)^t} = \sqrt{8t^2} = 2\sqrt{2}t \tag{7}$$

According to the relationship between $||w||$ and the length of the margin, we know that

$$d = \frac{2}{||w||} = \frac{1}{\sqrt{2}t} = 2\sqrt{2} \tag{8}$$

or

$$t = \frac{1}{4} \tag{9}$$

So

$$\boldsymbol{w} = [0, \frac{1}{2}, \frac{1}{2}]^T \tag{10}$$

d. According to SVM requirement,

$$\begin{cases} y_1(\boldsymbol{w}^T\phi(x_1) + w_0) \geqslant 1 \\ y_2(\boldsymbol{w}^T\phi(x_2) + w_0) \geqslant 1 \end{cases} \tag{11}$$

or

$$\begin{cases} -1 \times (0 + w_0) \geqslant 1 \\ 1 \times (2 + w_0) \geqslant 1 \end{cases} \tag{12}$$

$$-1 \leqslant w_0 \leqslant -1 \tag{13}$$

So

$$w_0 = -1 \tag{14}$$

e.

$$h(x) = \boldsymbol{w}^T\phi(x) + w_0 = \frac{x^2}{2} + \frac{\sqrt{2}x}{2} - 1 \tag{15}$$

2 Support Vector

There are two possibilies:

1. Size of maximum margin increases, if a support vector determining the shortest margin is removed.

2. Size of maximum margin stays the same, if the removed vector is not the one determining the shortest margin.