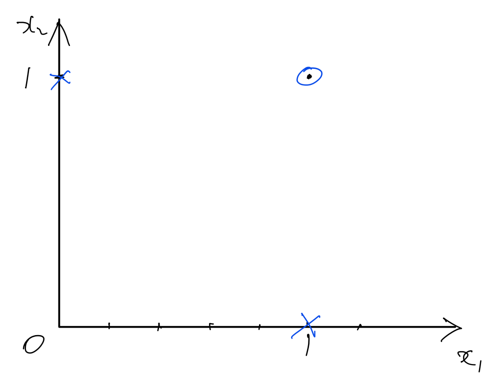
SVM 作业

Tuesday, November 9, 2021 3:41 PM

Training data set

$$D = \{([0,1]^T, -1), ([1,0]^T, -1), ([1,1]^T, 1)\}.$$



1). Assume the equation of the separation hyperplane is

$$w_1 x_1 + w_2 x_2 + b = 0 \qquad (1)$$

Hence, the SVM model can be formulated as the following:

$$\min \ \frac{1}{2}(w_1^2 + w_2^2)$$

$$\text{s.t.} \quad -(w_2 + b) \geq 1 \quad \text{(linearly separable)}$$
$$-(w_1 + b) \geq 1 \qquad (2)$$
$$w_1 + w_2 + b \geq 1$$
$$w_1 \in \mathbb{R}, \ w_2 \in \mathbb{R}, \ b \in \mathbb{R},$$

2) We first write (2) into the equivalent standard form as the following:

$$\min \ \frac{1}{2}(w_1^2 + w_2^2)$$

$$\text{s.t.} \quad 1 + w_2 + b \leq 0$$
$$1 + w_1 + b \leq 0 \qquad (3)$$
$$1 - w_1 - w_2 - b \leq 0$$
$$w_1 \in \mathbb{R}, \ w_2 \in \mathbb{R}, \ b \in \mathbb{R}.$$

The Lagrangian function of problem (3) is

$$L(w_1, w_2, b, \alpha_1, \alpha_2, \alpha_3)$$
$$= \frac{1}{2}(w_1^2 + w_2^2) + \alpha_1(1 + w_2 + b) + \alpha_2(1 + w_1 + b) + \alpha_3(1 - w_1 - w_2 - b) \qquad (4)$$

Where $\alpha_1, \alpha_2, \alpha_3$ are Lagrangian multipliers and $\alpha_1 \geq 0$, $\alpha_2 \geq 0$, $\alpha_3 \geq 0$.

Take the derivative of L with respect to each primal variable.

$$\begin{cases} \frac{\partial L}{\partial w_1} = w_1 + \alpha_2 - \alpha_3 \\ \frac{\partial L}{\partial w_2} = w_2 + \alpha_1 - \alpha_3 \\ \frac{\partial L}{\partial b} = \alpha_1 + \alpha_2 - \alpha_3 \end{cases} \xrightarrow[\text{equal to zeros}]{\text{let all derivatives}} \begin{cases} w_1 = -\alpha_2 + \alpha_3 \\ w_2 = -\alpha_1 + \alpha_3 \\ \alpha_1 + \alpha_2 - \alpha_3 = 0 \end{cases} \qquad (5)$$

Replace w_1, w_2 in (4) with $\alpha_1, \alpha_2, \alpha_3$, we have.

$$L(\alpha_1, \alpha_2, \alpha_3) = \frac{1}{2}(\alpha_1^2 + \alpha_2^2 + 2\alpha_3^2 - 2\alpha_2\alpha_3 - 2\alpha_1\alpha_3) + (-\alpha_1^2 - \alpha_2^2 - 2\alpha_3^2 + 2\alpha_1\alpha_3 + 2\alpha_2\alpha_3)$$
$$+ (\alpha_1 + \alpha_2 + \alpha_3)$$
$$= -\frac{1}{2}((\alpha_2 - \alpha_3)^2 + (\alpha_1 - \alpha_3)^2) + (\alpha_1 + \alpha_2 + \alpha_3)$$

$\Rightarrow$ The Lagrangian dual can be formulated as the following:

$$\max \ (\alpha_1 + \alpha_2 + \alpha_3) - \frac{1}{2}[(\alpha_2 - \alpha_3)^2 + (\alpha_1 - \alpha_3)^2]$$

$$\text{s.t.} \quad \alpha_1 + \alpha_2 - \alpha_3 = 0$$
$$\alpha_1 \geq 0, \ \alpha_2 \geq 0, \ \alpha_3 \geq 0$$

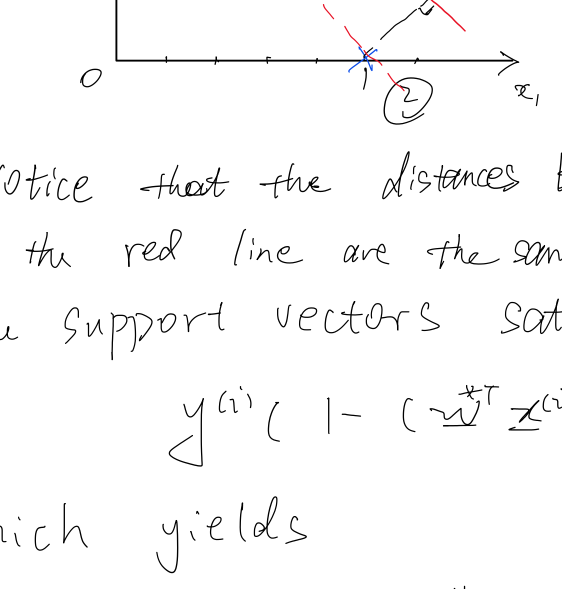
3) By (3), (4) and (5) in part 2), the KKT condition of problem (3) is

$$\begin{cases} 1 + w_2 + b \leq 0 \\ 1 + w_1 + b \leq 0 \\ 1 - w_1 - w_2 - b \leq 0 \\ \alpha_1(1 + w_2 + b) = 0 \\ \alpha_2(1 + w_1 + b) = 0 \\ \alpha_3(1 - w_1 - w_2 - b) = 0 \\ \alpha_1 \geq 0, \\ \alpha_2 \geq 0 \\ \alpha_3 \geq 0 \\ w_1 = -\alpha_2 + \alpha_3 \\ w_2 = -\alpha_1 + \alpha_3 \\ \alpha_1 + \alpha_2 - \alpha_3 = 0 \end{cases}$$

4)

There are a few methods.

Method 1:



Notice that the distances between each point to the red line are the same. Hence, they must be the support vectors satisfying

$$y^{(i)}(1 - (\underline{w}^{*T} \underline{x}^{(i)} + b^*)) = 1$$

which yields

$$\begin{cases} 1 + w_2^* + b^* = 0 \\ 1 + w_1^* + b^* = 0 \\ 1 - w_1^* - w_2^* - b^* = 0 \end{cases} \Rightarrow \begin{cases} w_1^* = 2 \\ w_2^* = 2 \\ b^* = -3 \end{cases}$$

By (5), we have

$$\begin{cases} 2 = -\alpha_2^* + \alpha_3^* \\ 2 = -\alpha_1^* + \alpha_3^* \\ 0 = \alpha_1^* + \alpha_2^* - \alpha_3^* \end{cases} \Rightarrow \begin{cases} \alpha_1^* = 2 \\ \alpha_2^* = 2 \\ \alpha_3^* = 4 \end{cases}$$