Homework 4, CPSC 8420, Fall 2020

Sadeghi Tabas, Sadegh

Nov 30, 2020

Solution to Problem 1

$$J = a^T Zb + \lambda_1(a^T Xa - 1) + \lambda_2(b^T Yb - 1) \tag{1}$$

$$\frac{\delta J}{\delta a} = 0 \longrightarrow Zb + \lambda_1 Xa = 0 \longrightarrow Zb = -\lambda_1 Xa \tag{2}$$

$$\frac{\delta J}{\delta b} = 0 \longrightarrow Za + \lambda_2 Yb = 0 \longrightarrow Za = -\lambda_2 Yb \tag{3}$$

$$(2) * a^T \longrightarrow \lambda_1 a^T Xa = -a^T Zb \tag{4}$$

$$\frac{\delta J}{\delta \lambda_1} = 0 \longrightarrow a^T Xa = 1 \longrightarrow \lambda_1 = -a^T Zb \longrightarrow \lambda_1 = -b^T Za \tag{5}$$

$$(3) * b^T \longrightarrow \lambda_2 b^T Yb = -b^T Za \tag{6}$$

$$\frac{\delta J}{\delta \lambda_2} = 0 \longrightarrow b^T Yb = 1 \longrightarrow \lambda_2 = -b^T Za \tag{7}$$

$$(5),(7) \longrightarrow \lambda_1 = \lambda_2 = \lambda = -b^T Za = -a^T Zb \tag{8}$$

substituting λ_1 and λ_2 in (2) and (3):

$$Zb = -(-b^T Za)Xa = -(-a^T Zb)Xa \tag{9}$$

$$Za = -(-a^T Zb)Yb = -(-b^T Za)Yb \tag{10}$$

Solution to Problem 2

$$L = \frac{1}{2}||w||_2^2 - \sum_i \alpha_i[y_i(w^T x_i + b) - 1] \tag{11}$$

$$\frac{\delta L}{\delta w} = 0 \longrightarrow w = \sum_i \alpha_i y_i x_i = 0 \tag{12}$$

$$\frac{\delta L}{\delta b} = 0 \longrightarrow \sum_i \alpha_i y_i = 0 \tag{13}$$

substituting (12) in (11):

$$MaxL = \sum_i \alpha_i - \frac{1}{2}\sum_{i,j} \alpha_i\alpha_j y_i y_j (x_i^T x_j) \quad Subeject \quad to : \alpha_i \geq 0, \quad \alpha_i[y_i(w^T x_i + b) - 1] = 0 \tag{14}$$

Considering the constraints:

$$if \quad \alpha_i > 0 \longrightarrow y_i(w^T x_i + b) = 1 \tag{15}$$

it says this point drops on the margin boundary so it is support vector.

$$if \quad y_i(w^T x_i + b) > 1 \longrightarrow \alpha_i = 0 \tag{16}$$

based on the equation (12), hyperplane only depends on the support vectors, if (x_i, y_i) is a support vector, then $w^T x_i + b = y_i$. because $w = \sum_j \alpha_j y_j x_j \longrightarrow b = y_i - \sum_j \alpha_j y_j x_j^T x_i$.

$$*\alpha_i y_i \longrightarrow \sum_i \alpha_i y_i b = \sum_i \alpha_i y_i^2 - \sum_{i,j} \alpha_i \alpha_j y_i y_j (x_i^T x_j) \tag{17}$$

As $y_i^2 = 1, w = \sum_i \alpha_i y_i x_i$ and (13), we have:

$$\sum_i \alpha_i - ||w||^2 = 0 \tag{18}$$

$$margin \quad \gamma = \frac{1}{||w||} \longrightarrow \gamma^2 = \frac{1}{||w||^2} = \frac{1}{\sum_i \alpha_i} = \frac{1}{||\alpha||_1} \tag{19}$$