Name: Inchara Raveendra

SCU ID: 00001653600

Assignment - 8

Problem 1 (5 points)

Find the solution (x^*, y^*) to the following problem.

$$\text{optimize } xy$$
$$\text{subject to } x + y = 10$$

Standard form

$$\text{optimize } \quad xy$$
$$\text{subject to } \quad x + y - 10 = 0$$

Lagrangian $\quad L(x, y, \beta) = xy + \beta(x + y - 10) = 0$

Partial derivatives:

$$\nabla_x L(x, y, \beta) = y + \beta = 0 \rightarrow (1)$$
$$\nabla_y L(x, y, \beta) = x + \beta = 0 \rightarrow (2)$$
$$\nabla_\beta L(x, y, \beta) = x + y - 10 = 0 \rightarrow (3)$$

$$x = y = -\beta$$

$(3) \rightarrow \quad \therefore \quad 2x - 10 = 0 \qquad x^* = y^* = 5$

Problem 2 (5 points)

The SVM optimization can be defined by the primal form:

$$\min_w \frac{1}{2}\|\boldsymbol{w}\|^2$$

$$\text{subject to } y_i(\boldsymbol{w}^T\boldsymbol{x}_i + b) \geq 1, \qquad i = 1, \ldots, N$$

Or by its the dual form:

$$\max_\alpha J(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i\alpha_j y_i y_j \left(\boldsymbol{x}_i^T\boldsymbol{x}_j\right)$$

$$\text{subject to } \alpha_i \geq 0, i = 1, \ldots N \text{ and } \sum_{i=1}^{N} \alpha_i y_i = 0$$

What is the Lagrangian function $L(\boldsymbol{w}, b, \alpha)$ evaluated at $\boldsymbol{w}$ that minimizes that function?

Note this is the objective function $J(\alpha)$.

Hints:

1. Write the primal problem in standard form
2. Form the Lagrangian function $L(\boldsymbol{w}, b, \alpha)$
3. Find $\boldsymbol{w}$ and b that minimize $L(\boldsymbol{w}, b, \alpha)$
4. Plug the results back into $L(\boldsymbol{w}, b, \alpha)$

① $\min_w \frac{1}{2}\|W\|^2$ subject to $g_i(W) = -y_i(w^T x_i + b) + 1 \leq 0$

② $L(W, b, \alpha) = \frac{1}{2}\|W\|^2 + \sum_{i=1}^{N} \alpha_i[-y_i(w^T x_i + b) + 1]$ → Eqn (A)

③ $\nabla_w L(w, b, \alpha) = w - \sum_{i=1}^{N} \alpha_i y_i x_i = 0$

$\Rightarrow w = \sum_{i=1}^{N} \alpha_i y_i x_i$ → Eqn (B)

$\nabla_b L(W, b, \alpha) = \sum_{i=1}^{N} \alpha_i y_i = 0$ → Eqn (C)

④ Substitute Ⓑ and Ⓒ in Ⓐ

$$L(w, b, \alpha) = \frac{1}{\alpha} w^T w + \sum_{i=1}^{N} \alpha_i [-y_i (w^T x_i + b) + 1]$$

$$L(w, b, \alpha) = \frac{1}{\alpha} \sum_{i=1}^{N} \alpha_i y_i x_i^T \cdot \sum_{j=1}^{N} \alpha_j y_j x_j$$

$$+ \sum_{i=1}^{N} \alpha_i \left[-y_i \left(\left(\sum_{j=1}^{N} \alpha_j y_j x_j \right) x_i + b \right) + 1 \right]$$

$$L(w, b, \alpha) = \frac{1}{\alpha} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_i^T x_j) - \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_j^T x_i)$$

$$- \sum_{i=1}^{N} \alpha_i y_i b + \sum_{i=1}^{N} \alpha_i$$

From, $\sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_j^T x_i) = \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_i^T x_j)$

and $\sum_{i=1}^{N} \alpha_i y_i = 0 \Rightarrow \sum_{i=1}^{N} \alpha_i y_i b = 0$ from Ⓒ

Therefore, $J(\alpha) = L(w, b, \alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{\alpha} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_i^T x_j)$