HW-8

① Find the solution (x^*, y^*) to the following problem:
optimize xy; subject to $x+y=10$

$\rightarrow$ $x+y-10=0$

Lagragian $L(x, y, \beta) = xy + \beta(x+y-10)$

Taking partial derivatives;

$\nabla_x L(x, y, \beta) = y + \beta = 0 \longrightarrow ①$

$\nabla_y L(x, y, \beta) = x + \beta = 0 \longrightarrow ②$

$\nabla_\beta L(x, y, \beta) = x + y - 10 = 0 \longrightarrow ③$

Solving above equation for $x+y$:-

$y = -\beta$ (from ①)

$x = -\beta$ (from ②)

Substituting in ③,

$-\beta - \beta - 10 = 0$

$-2\beta - 10 = 0$

$-2\beta = 10$

$\therefore \boxed{\beta = -5}$ i.e. $x^* = 5$, $y^* = 5$

② The SVM optimization can be defined by the primal form:

$$\min_w \frac{1}{2}||w||^2 \text{ subject to } y_i(w^T x_i + b) \geq 1, \; i = 1, \dots, N$$

i) Write the primal problem in standard form.

$$g_i(w) = -y_i(w^T x_i + b) + 1 \leq 0$$

ii) Form the Lagragian function $L(w, b, \alpha)$

$$L(w, \alpha, \beta) = f(w) + \sum_{i=1}^{m} \alpha_i g_i(w) + \sum_{i=1}^{l} \beta_i h_i(w)$$

here, $L(w, b, \alpha) = \frac{1}{2}||w||^2 + \sum_{i=1}^{m} \alpha_i \left[-y_i(w^T x_i + b) + 1\right]$

as there are no equality constraints $h_i(w)$, there is no β_i term

iii) Find w and b that minimize $L(w, b, \alpha)$
$$= \frac{1}{2} \|w\|^2 + \sum_{i=1}^{N} \alpha_i \left[-y_i (w^T x_i + b) + 1 \right]$$

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

iv) Plug the results back into $L(w, b, \alpha)$
$$\|z\|^2 = z^T z$$
$$\therefore L(w, b, \alpha) = \frac{1}{2} w^T w + \sum_{i=1}^{N} \alpha_i \left[-y_i (w_{x_i}^T + b) + 1 \right]$$
$$\therefore L(w, b, \alpha) = \frac{1}{2} \sum_{i=1}^{N} \alpha_i y_i x_i^T \cdot \sum_{j=1}^{N} \alpha_j y_j x_j + \sum_{i=1}^{N} \alpha_i \left[-y \left[\left(\sum_{j=1}^{N} \alpha_j y_j x_j^T \right) x_i + b \right) + 1 \right]$$
$$= \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_i x_j) - \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_j y_i (x_j^T x_i) - \sum_{i=1}^{N} \alpha_i y_i b + \sum_{i=1}^{N} \alpha_i$$

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

also, from subequation (3), $\sum_{i=1}^{N} \alpha_i y_i b = 0$

$\therefore$ The objective function $J(\alpha) = L(w, b, \alpha)$
$$= \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j \cdot (x_i^T y_j)$$