Problem 1 (5 points)

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

$$\text{optimize } xy$$

$$\text{subject to } 2x + 2y = 20$$

Make sure the constraint is in standard form.

Standard form

$$\text{optimize } xy$$

$$\text{subject to } 2x + 2y - 20 = 0$$

Lagrangian: $L(x, y, \beta) = xy + \beta(2x + 2y - 20)$

Partial Derivatives:

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

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

$$\nabla_\beta L(x, y, \beta) = 2x + 2y - 20 = 0$$

Solve for x, y

$$\boldsymbol{x^* = 5, y^* = 5}$$

Problem 2 (5 points)

Define the Lagrangian function for the following optimization problem.

$$\text{minimize } 7x_1^3 + 8x_2^2 + 9x_1$$

$$\text{subject to } 6x_1 + 11x_2^2 = 13x_2$$

$$5x_2^2 \leq 4x_1 - 7$$

$$20x_1 + 5 \geq x_2$$

Do not solve for the variables, i.e., do not try to compute the partial derivatives.

Make sure the constraints are in standard form.

Standard form

$$\text{minimize } 7x_1^3 + 8x_2^2 + 9x_1$$

$$\text{subject to } 6x_1 + 11x_2^2 - 13x_2 = 0$$

$$5x_2^2 - 4x_1 + 7 \leq 0$$

$$-20x_1 - 5 + x_2 \leq 0$$

$$\begin{aligned}\boldsymbol{L(x_1, x_2, \alpha_1, \alpha_2, \beta_1) = 7x_1^3 + 8x_2^2 + 9x_1} +& \\ \boldsymbol{\alpha_1\left(5x_2^2 - 4x_1 + 7\right)} +& \\ \boldsymbol{\alpha_2(-20x_1 - 5 + x_2)} +& \\ \boldsymbol{\beta_1\left(6x_1 + 11x_2^2 - 13x_2\right)}&\end{aligned}$$

Bonus Problem (5 points)

The SVM optimization can be defined by the primal form:

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

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

Or by its the dual form:

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

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

Show the objective function $J(\alpha)$ is the Lagrangian function $L(\boldsymbol{w}, b, \alpha)$ evaluated at $\boldsymbol{w}$ that minimizes that function.

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 $L\ (\boldsymbol{w}, b, \alpha)$

Note $\|\boldsymbol{z}\|^2$ is the inner product $< \boldsymbol{z}_i, \boldsymbol{z}_j >$ meaning $\boldsymbol{z}^T\boldsymbol{z}$ between points in the vector space. If $\boldsymbol{z} = \sum_i \boldsymbol{x}_i$ then $\boldsymbol{z}^T\boldsymbol{z} = \sum_i \boldsymbol{x}_i \sum_j \boldsymbol{x}_j$

(1). Write the primal problem in standard form

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

(2). Form the Lagrangian function

$$L(\boldsymbol{w}, b, \alpha) = \frac{1}{2}\|\boldsymbol{w}\|^2 + \sum_{i=1}^{N} \alpha_i[-y_i(\boldsymbol{w}^T\boldsymbol{x}_i + b) + 1] \rightarrow (\text{Eq. } 1)$$

- Note there is no β_i because there are no equality constraints

(3). Find $\boldsymbol{w}$ and b that minimize $L\ (\boldsymbol{w}, b, \alpha)$

$$\nabla_{\boldsymbol{w}}\ L\ (\boldsymbol{w}, b, \alpha) = \boldsymbol{w} - \sum_{i=1}^{N} \alpha_i y_i \boldsymbol{x}_i = 0 \rightarrow \boldsymbol{w} = \sum_{i=1}^{N} \alpha_i y_i \boldsymbol{x}_i \rightarrow (\text{Eq. } 2)$$

$$\nabla_b\ L\ (\boldsymbol{w}, b, \alpha) = \sum_{i=1}^{N} \alpha_i y_i = 0 \rightarrow (\text{Eq. } 3)$$

(4). Plug (Eq. 2) and (Eq. 3) back into (Eq. 1). Notice index i in each equation is only relevant for that equation.

$$L(\boldsymbol{w}, b, \alpha) = \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + \sum_{i=1}^{N} \alpha_i[-y_i(\boldsymbol{w}^T\boldsymbol{x}_i + b) + 1]$$

$$L(\boldsymbol{w}, b, \alpha) = \frac{1}{2}\sum_{i=1}^{N} \alpha_i y_i \boldsymbol{x}_i^T \cdot \sum_{j=1}^{N} \alpha_j y_j \boldsymbol{x}_j + \sum_{i=1}^{N} \alpha_i \left[-y_i\left(\left(\sum_{j=1}^{N} \alpha_j y_j \boldsymbol{x}_j^T\right)\boldsymbol{x}_i + b\right) + 1\right]$$

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

But

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

And from (Eq. 3)

$$\sum_{i=1}^{N} \alpha_i\, y_i b = 0$$

Thus,

$$J(\boldsymbol{\alpha}) = L(\boldsymbol{w}, \boldsymbol{b}, \boldsymbol{\alpha}) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j \left(x_i^T x_j \right)$$