

Question 1

Primal Hard Margin:

$$\begin{aligned} & \underset{\mathbf{w}, b}{\text{minimize}} \quad \frac{1}{2} \|\mathbf{w}\|^2 \\ & \text{subject to} \quad y_i(\mathbf{w}^\top \mathbf{x}_i + b) \geq 1, \quad \forall i \end{aligned} \tag{1}$$

General **convex optimization problem** format:

$$\min_x f(x); \quad s.t. \quad g_i(x) \leq 0 \tag{2}$$

Transform the primal hard margin into such format:

$$\begin{aligned} & \underset{\mathbf{w}, b}{\text{minimize}} \quad f(x) = \frac{1}{2} \|\mathbf{w}\|^2 \\ & \text{subject to} \quad g_i(x) = 1 - y_i(\mathbf{w}^\top \mathbf{x}_i + b) \leq 0, \quad \forall i \end{aligned} \tag{3}$$

General convex primal corresponded **Lagrangian function** is:

$$\mathcal{L}(x, \{\alpha_i\}) = f(x) + \sum_i \alpha_i g_i(x) \tag{4}$$

Based on equation (1) and (3), we get Lagrangian dual equation:

$$\mathcal{L}(x, \{\alpha_i\}) = \frac{1}{2} \|\mathbf{w}\|^2 + \sum_i \alpha_i [1 - y_i(\mathbf{w}^\top \mathbf{x}_i + b)] \tag{5}$$

Align to more common used format:

$$\mathcal{L}(x, \{\alpha_i\}) = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_i \alpha_i [y_i(\mathbf{w}^\top \mathbf{x}_i + b) - 1] \tag{6}$$

To conveniently writing, following derivation would omit \top and \mathcal{L} sub-
script.

Partially derivate \mathbf{w} first:

$$\begin{aligned} \frac{\partial L}{\partial w} &= w - \left[w \cdot \sum_i \alpha_i y_i x_i + \sum_i \alpha_i y_i b \right]' + \left(\sum_i \alpha_i \right)' \\ &= w - \sum_i \alpha_i y_i x_i - 0 + 0 = 0 \end{aligned} \tag{7}$$

Partially derivate \mathbf{b} :

$$\begin{aligned}\frac{\partial L}{\partial b} &= \left(\frac{1}{2}||w||^2\right)' - \left[\sum_i \alpha_i y_i w x_i + b \cdot \sum_i \alpha_i y_i\right]' + \left(\sum_i \alpha_i\right)' \\ &= 0 - 0 - \sum_i \alpha_i y_i + 0 = 0\end{aligned}\tag{8}$$

Based on equation (7) and (8), we substitute \mathbf{w} and \mathbf{b} :

$$\begin{aligned}\mathcal{L} &= \frac{1}{2} \left(\sum_i \alpha_i y_i x_i\right) \left(\sum_j \alpha_j y_j x_j\right) - \sum_i \alpha_i y_i x_i \cdot \mathbf{w} - \sum_i \alpha_i y_i \cdot \mathbf{b} + \sum_i \alpha_i \\ &= \frac{1}{2} \left(\sum_i \alpha_i y_i x_i\right) \left(\sum_j \alpha_j y_j x_j\right) - \left(\sum_i \alpha_i y_i x_i\right) \cdot \left(\sum_j \alpha_j y_j x_j\right) \\ &\quad - \sum_i \alpha_i y_i \cdot \mathbf{b} + \sum_i \alpha_i \\ &= -\frac{1}{2} \left(\sum_i \alpha_i y_i x_i\right) \cdot \left(\sum_j \alpha_j y_j x_j\right) - 0 + \sum_i \alpha_i\end{aligned}\tag{9}$$

Finally, we get the **Dual Hard Margin** equation:

$$\mathcal{L} = \sum_i \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i x_j\tag{10}$$