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:

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

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:

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

Align to more common used format:

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

To conveniently writing, following derivation would omit $\top$ and $\mathcal{L}$ subscription.

Based on the KKT first condition, stationarity, we respectively derivate w and b:

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

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

Based on equation (7) and (8), we substitute **w** and **b**:

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

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

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