7. a)

$$\frac{\partial}{\partial \alpha}\left(\|\vec{w}\|^2 - \sum_{i=1}^{n} \lambda_i\left(y_i(x_i \cdot \vec{w} + \alpha) - 1\right)\right) = 0$$

$$-\sum_{i=1}^{n} \frac{\partial}{\partial \alpha}\left(\lambda_i y_i x_i \cdot \vec{w} + \lambda_i y_i \alpha\right) = 0$$

$$\sum_{i=1}^{n} \lambda_i y_i = 0 \longleftarrow$$ condition for α^* (α that minimizes the objective fn.)

$$\frac{\partial}{\partial w}\left(\|\vec{w}\|^2 - \sum_{i=1}^{n} \lambda_i\left(y_i(x_i \cdot \vec{w} + \alpha) - 1\right)\right) = 0$$

$$2w^* - \sum_{i=1}^{n} \frac{\partial}{\partial w}\left(\lambda_i y_i X_i \cdot w + \lambda_i y_i \alpha - \lambda_i\right) = 0$$

$$w^* = \frac{1}{2} \sum_{i=1}^{n} \lambda_i y_i X_i$$

7. a.) continued...

$$\therefore \max_{\lambda_i \geq 0} \min_{w, \alpha} \|\vec{w}\|^2 - \sum_{i=1}^{n} \lambda_i \left(y_i (X_i \cdot \vec{w} + \alpha) - 1 \right)$$

$$= \max_{\lambda_i \geq 0} \left(\frac{1}{2} \sum_{i=1}^{n} \lambda_i y_i X_i \right)^2 - \sum_{i=1}^{n} \lambda_i y_i X_i \left(\frac{1}{2} \sum_{i=1}^{n} \lambda_i y_i X_i \right) - \alpha^* \underbrace{\sum_{i=1}^{n} \lambda_i y_i}_{0} + \sum_{i=1}^{n} \lambda_i$$

$$= \max_{\lambda_i \geq 0} -\frac{1}{4} \left(\lambda_1^2 y_1^2 X_1^2 + \lambda_1 \lambda_2 y_1 y_2 X_1 X_2 + \ldots + \lambda_i \lambda_j y_i y_j X_i X_j + \ldots + \lambda_n^2 y_n^2 X_n^2 \right) + \sum_{i=1}^{n} \lambda_i$$

$$= \max_{\lambda_i \geq 0} \sum_{i=1}^{n} \lambda_i - \frac{1}{4} \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j y_i y_j X_i X_j \quad , \quad \underline{\sum_{i=1}^{n} \lambda_i y_i = 0} \quad \square$$

* The new constraint comes from minimizing over α.

7. b.) From 7. a.) $W^* = \frac{1}{2}\sum_{i=1}^{n}\lambda_i y_i X_i$, & $\sum_{i=1}^{n}\lambda_i y_i = 0 \Rightarrow \alpha^*$

$$\therefore\quad r(x) = \begin{cases} +1 & \text{if } \omega\cdot x + \alpha \geq 0 \\ -1 & \text{otherwise} \end{cases}$$

$$= \begin{cases} +1 & \text{if } \left(\frac{1}{2}\sum_{i=1}^{n}\lambda_i^* y_i X_i\right)\cdot x + \alpha^* \geq 0 \\ -1 & \text{otherwise} \end{cases}$$

$$= \begin{cases} +1 & \text{if } \alpha^* + \frac{1}{2}\sum_{i=1}^{n}\lambda_i^* y_i X_i \cdot x \geq 0 \\ -1 & \text{otherwise} \end{cases}$$

$\square$

- Here, having λ_i^* implies ω^*

7.

c.) for $\lambda_i^*(y_i(\vec{x}_i \cdot \vec{w}^* + \alpha^*) - 1) = 0$

if $\lambda_i^* > 0$ it must be that

$$y_i(\vec{x}_i \cdot \vec{w}^* + \alpha^*) - 1 = 0$$

$$\text{or } y_i(\vec{x}_i \cdot \vec{w}^* + \alpha^*) = 1$$

this means when the class $y_i = -1$

then $\vec{x}_i \cdot \vec{w}^* + \alpha = -1$

which puts $\vec{x}_i$ on the margin for the -1 class

7 c.) continued...

when $y_i = +1$ then $\vec{x}_i \cdot \vec{w}^* + \alpha = 1$

putting $\vec{x}_i$ on the margin for the +1 class.

In other words, for $\lambda_i^* > 0$, the corresponding $\vec{x}_i$ are support vectors

7. d.) The decision rule is determined by the decision boundary given by

$$\vec{X}_i \cdot \vec{w} + \alpha = 0$$

& we find the margin, $\frac{1}{\|\vec{w}\|}$ (which yields $\vec{w}$) & α by first finding the $\vec{X}_i$ which solve the equality constraint

$$y_i(\vec{X}_i \cdot \vec{w} + \alpha) = 1$$

& thus allowing us to solve for $\vec{w}$ & α

7. d.) for all other $\vec{X}_i$ where $y_i(\vec{X}_i \cdot \vec{w} + \alpha) > 1$

No single solution $(\vec{w}, \alpha)$ can be found. Thus only the support vectors determine the location of the decision boundary.

7. e.) Suppose that w' & α' solve

$$\min_{w,\alpha} \|w\|^2 , \quad y_i(x_i \cdot w + \alpha) \geq 1 \quad \forall i \in \{1, \dots, n\}$$

but suppose at least one class has no support vectors x_i. This implies the margin width $\frac{2}{\|w'\|}$ could be infinite & implying $\|w'\| \to 0$. Suppose then there is some other w where $\|w'\| < \|w\|$

7e. cont. & $\frac{2}{\|w'\|} > \frac{2}{\|w\|}$; but say there is some $\varepsilon > 0$ where $\frac{2}{\|w'\|} = \frac{\varepsilon + 2}{\|w\|}$.

Then $\varepsilon = 2\frac{\|w\|}{\|w'\|} - 2 = 2\left(\frac{\|w\|}{\|w'\|} - 1\right)$