H.W-9

RAKSHITHA PANDURANGA
7890-1614-34
rpandura@usc.edu

① $J(w) = \frac{1}{2}\|\underline{w}\|^2$

s.t $z_i(\underline{w}^T \underline{u_i} + w_0) - 1 \geq 0 \quad \forall i$

a) No, it will not be correctly classified due to the presence of certain datapoints b/w the boundary 0 & the margin 1

b) $L = f(x) - \lambda h(x)$

$$= \frac{1}{2}\|\underline{w}\|^2 - \sum_{i=1}^{N} \lambda_i \left[z_i(\underline{w}^T u_i + w_0) - 1\right]$$

Constraints → $\lambda \geq 0$

$z_i(\underline{w}^T u_i + w_0) - 1 \geq 0$

$\lambda[z_i(w^T \underline{u_i} + w_0) - 1] = 0$

c) $L(\underline{w}, w_0, \lambda) = \frac{1}{2}\|\underline{w}\|^2 - \sum_{i=1}^{N} \lambda_i [z_i(\underline{w}^T u_i + w_0) - 1]$

(i)

$$\nabla_w L(\underline{w}, w_0, \lambda) = \frac{1}{2}(2\underline{w}) - \sum_{i=1}^{N} \lambda_i z_i u_i = 0$$

$$\boxed{\underline{w}^* = \sum_{i=1}^{N} \lambda_i z_i u_i}$$

$$\frac{\partial}{\partial w_0} L(w, w_0, \lambda) = 0 - \sum_{i=1}^{N} \lambda_i z_i = 0$$

$$\boxed{\sum_{i=1}^{N} \lambda_i z_i = 0}$$

(ii) (3) Substituting for w^*,

$$L = \frac{1}{2}\left(\sum_{j=1}^{N} \lambda_j z_j \underline{u}_j^T\right)\left(\sum_{i=1}^{N} \lambda_i z_i \underline{u}_i\right)$$
$$- \sum_{i=1}^{N} \left(\lambda_i z_i \left(\sum_{j=1}^{N} \lambda_j z_j u_j^T\right) \underline{u}_i - \cancel{\sum_{i=1}^{N} \lambda_i z_i w_0} + \sum_{i=1}^{N} \lambda_i\right)$$

$$= \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j \underline{u}_i u_j^T \; z_i z_j$$
$$- \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j u_i u_j^T z_i z_j$$
$$- \cancel{\sum_{i=1}^{N} \lambda_i z_i w_0} + \sum_{i=1}^{N} \lambda_i$$

Subtracting, we get,

$$L_D(\lambda) = -\frac{1}{2} \sum_{j=1}^{N} \sum_{i=1}^{N} \lambda_i \lambda_j z_i z_j \underline{u}_j^T u_j + \sum_{i=1}^{N} \lambda_i$$

→ 2 KKT conditions

(1) $\lambda_i \geq 0 \quad \forall i$,

(2) $\lambda_i [z_i (w^T u_i + w_0) - 1] = 0 \quad \forall i$

Scanned by CamScanner

(3) $$J(\underline{w}) = \frac{1}{2}\|\underline{w}\|^2 + C\sum_{i=1}^{N}\xi_i$$

s.t $\quad z_i(\underline{w}^T \underline{x}_i + w_0) \geq 1 - \xi_i \quad \forall i$

$\xi_i \geq 0 \quad \forall i$

a) $$\boxed{L(\underline{w}, w_0, \underline{\xi}, \underline{\lambda}, \underline{\mu}) = \frac{1}{2}\|\underline{w}\|^2 + C\sum_{i=1}^{N}\xi_i - \sum_{i=1}^{N}\lambda_i\left[z_i(\underline{w}^T \underline{x}_i + w_0) - 1 + \xi_i\right] - \sum_{i=1}^{N}\mu_i\xi_i} \quad \text{— eqn (1)}$$

KKT conditions:

(1) $\lambda_i \geq 0 \quad \forall i$, (2) $\lambda_i\left[z_i(\underline{w}^T\underline{x}_i + w_0) - 1 + \xi_i\right] = 0 \quad \forall i$

(3) $\mu_i \geq 0 \quad \forall i$, (4) $\mu_i\xi_i = 0 \quad \forall i$.

(5) $\xi_i \geq 0 \quad \forall i$, (6) $z_i(\underline{w}^T\underline{x}_i + w_0) \geq 1 - \xi_i \quad \forall i$

b) Differentiating eqn (1)
(Taking gradient)

$$\nabla_{\underline{w}} L(\underline{w}, w_0, \underline{\xi}, \underline{\lambda}, \underline{\mu}) = (2\underline{w}) \times \frac{1}{2} + 0 - \sum_{i=1}^{N}\lambda_i z_i \underline{x}_i = 0$$

$$\therefore \boxed{\underline{w}^* = \sum_{i=1}^{N}\lambda_i z_i \underline{x}_i}$$

$$\nabla_{w_0} L(\underline{w}, w_0, \underline{\xi}, \underline{\lambda}, \underline{\mu}) = 0 + 0 - \sum_{i=1}^{N}\lambda_i z_i = 0$$

$$\boxed{\sum_{i=1}^{N}\lambda_i z_i = 0}$$

$$\nabla_{\underline{\xi}} L(\underline{w}, w_0, \underline{\xi}, \underline{\lambda}, \underline{\mu}) = CN - \sum_{i=1}^{N}\lambda_i - \sum_{i=1}^{N}\mu_i = 0$$

$$= C - \lambda_i - \mu_i = 0$$

$\therefore \boxed{c = \lambda_i + \mu_i} \quad 0 \le \lambda_i \le c$

$$L_D(\underline{\lambda}, \underline{\mu}) = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j z_i z_j u_i^T u_j + c \sum_{i=1}^{N} \xi_i - \sum_{i=1}^{N} \lambda_i \xi_i + \sum_{i=1}^{N} \lambda_i$$

w.r.t $\quad c\xi_1 - \lambda_1 \xi + c\xi_2 - \lambda_2 \xi_2$

$$\xi_1(c - \lambda_1) + \xi_2(c - \lambda_2) = \xi_1(\cancel{\lambda_i} + \mu_1 - \cancel{\lambda_i}) + \xi_2(\mu_2) = 0$$

$$\xi_1 \mu_1 + \xi_2 \mu_2 \cdots = 0. \quad [\because \mu_i \xi_i = 0 \ \forall i]$$

$$\therefore L_D(\underline{\lambda}, \underline{\mu}) = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} z_i z_j \lambda_i \lambda_j u_j^T u_j + \sum_{i=1}^{N} \lambda_i$$

KKT conditions $\rightarrow$

$0 \le \lambda_i \le c \ \forall i$

$\mu_i \ge 0 \quad \forall i$

$z_i [\underline{w}^T \underline{u_i} + w_0] - 1 + \xi \ge 0 \ \forall i$

$\lambda_i [z_i(\underline{w}^T \underline{u_i} + w_0) - 1 + \xi] = 0.$

$c = \lambda_i + \mu_i \quad 0 \le \lambda_i \le 0$

2)
a) $\underline{u_1} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$$L_D'(\underline{\lambda}, \mu) = \sum_{i=1}^{N} \lambda_i - \frac{1}{2}\left[\sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j z_i z_j \underline{u_i}^T \underline{u_j}^T\right] + \mu\left(\sum_{i=1}^{N} z_i \lambda_i\right)$$

$$\sum_{i=1}^{N} \lambda_i = \lambda_1 + \lambda_2$$

$$\sum_{i=1}^{N} \mu_i = \mu_1 + \mu_2 \qquad \text{Sub.}$$

$$L_D'(\underline{\lambda}, \mu) = \lambda_1 + \lambda_2 - \frac{1}{2}[\lambda_1^2 z_1^2 u_1^T u_1 + \lambda_2^2 z_2^2 u_2^T u_2 + \lambda_1 \lambda_2 z_1 z_2 u_1^T u_1 + \lambda_1 \lambda_2 z_1 z_2 u_2^T u_2]$$

$$= \lambda_1 + \lambda_2 - \frac{1}{2}[\lambda_1^2 + \lambda_2^2 + \lambda_1 \lambda_2 z_1 z_2 + \lambda_1 \lambda_2 z_1 z_2]$$

(~~Because~~

Scanned by CamScanner

$$\left[\because \quad u_1^T u_1 = \begin{bmatrix} -1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \end{bmatrix} = 1 \qquad u_2^T u_2 = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1 \right]$$

and sub. for $\sum_{i=1}^{N} \lambda_i z_i = 0$

$$\therefore \quad L_D'(\lambda, \mu) = \lambda_1 + \lambda_2 - \frac{\lambda_1^2}{2} - \frac{\lambda_2^2}{2}$$

By constraint equation we have,

~~[illegible]~~ $\sum_{i=1}^{N} \lambda_i z_i = 0$

$$\lambda_1 z_1 + \lambda_2 z_2 = 0$$

$$\lambda_1 - \lambda_2 = 0$$

$$\boxed{\lambda_1 = \lambda_2}$$

$$L_D'(\lambda, \mu) = 2\lambda_1 - \lambda_1^2$$

$$\nabla L_D'(\lambda, \mu) = 2 - 2\lambda_1 = 0$$

$$2\lambda_1 = 2$$

$$\boxed{\lambda_1 = 1} \quad \boxed{\lambda_2 = 1}$$

$$\underline{w}^* = \sum_{i=1}^{2} \lambda_i z_i \underline{u}_i$$

$$= 1(1) \begin{bmatrix} -1 \\ 0 \end{bmatrix} + (1)(-1) \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

$$\boxed{\underline{w}^* = \begin{bmatrix} -1 \\ -1 \end{bmatrix}}$$

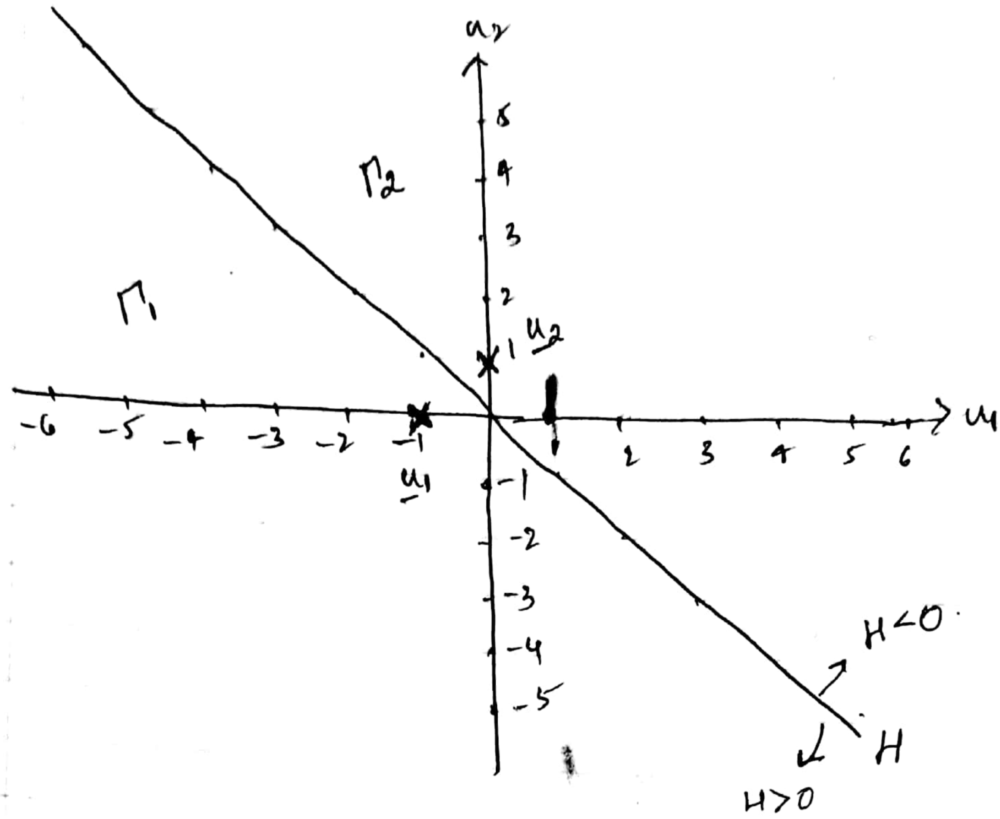
$$z_i(w^T u_i + w_0) - 1 \geq 0$$

$$1 \left[\begin{bmatrix} -1 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \end{bmatrix} + w_0 - 1 \geq 0 \right.$$

$$1 \left[1 \right] + w_0 - 1$$

$$\boxed{w_0 = 0}$$

Scanned by CamScanner



b) $H \Rightarrow \; -u_1 - u_2$

$$d.(H, u_1) = \frac{g(u_1)}{\|w\|} = \frac{1}{\sqrt{2}}$$

$$d(H, u_2) = \frac{g(u_2)}{\|w\|} = \frac{-1}{\sqrt{2}}$$

No, there is no possible linear boundary in $\underline{u}$ space that would give larger values for both distances (margins) than H gives.

Scanned by CamScanner