Guangrui Shen

28012207 3

1. (a) Yes.

If $z_i(\underline{w}^T \underline{u}_i + w_0) \geq 1$, then $z_i(\underline{w}^T \underline{u}_i + w_0) \geq 0$, which guarantees the correct classification in linearly separable case.

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

KKT conditions:

$$\begin{cases} \lambda_i \geq 0 & \forall i, \\ \lambda_i [z_i(\underline{w}^{*T} \underline{u}_i + w_0^*) - 1] = 0 & \forall i \\ z_i(\underline{w}^T \underline{u}_i + w_0) - 1 \geq 0 & \forall i \end{cases}$$

(c) (i) $\nabla_{\underline{w}} L = \underline{w} - \sum_{i=1}^{N} \lambda_i z_i \underline{u}_i = 0$

$$\underline{W}^* = \sum_{i=1}^{N} \lambda_i z_i \underline{u}_i$$

$$\frac{\partial L}{\partial w_0} = -\sum_{i=1}^{N} \lambda_i z_i = 0$$

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

(ii) $$L_D(\lambda) = \frac{1}{2}\left(\sum_{i=1}^{N} \lambda_i z_i \underline{u}_i\right)^T \left(\sum_{i=1}^{N} \lambda_i z_i \underline{u}_i\right) - \sum_{i=1}^{N} \lambda_i \left[z_i\left(\left(\sum_{i=1}^{N} \lambda_i z_i \underline{u}_i\right)^T \underline{u}_i + w_0\right) - 1\right]$$

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

$$= -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j z_i z_j \underline{u}_i^T \underline{u}_j + \sum_{i=1}^{N} \lambda_i$$

Constraints:

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

$$\lambda_i \geq 0 \quad \forall i$$

$$\lambda_i [z_i(\underline{w}^{*T} \underline{u}_i + w_0) - 1] = 0 \quad \forall i$$

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

$$z_i = \begin{cases} 1 & \text{if } \underline{u}_i \in S_1 \\ -1 & \text{if } \underline{u}_i \in S_2 \end{cases}$$

Then,

$$L_D'(\underline{\lambda}, \mu) = \lambda_1 + \lambda_2 - \frac{1}{2}\left(\lambda_1^2 \underline{u}_1^T \underline{u}_1 - \lambda_1 \lambda_2 \underline{u}_1^T \underline{u}_2 - \lambda_1 \lambda_2 \underline{u}_2^T \underline{u}_1 + \lambda_2^2 \underline{u}_2^T \underline{u}_2\right) + \mu(\lambda_1 - \lambda_2)$$

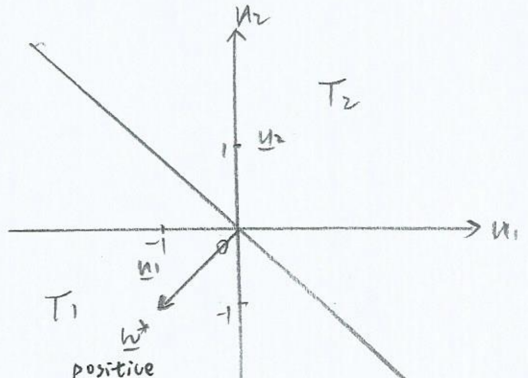
$$= \lambda_1 + \lambda_2 - \frac{1}{2}\left(\lambda_1^2 + \lambda_2^2\right) + \mu(\lambda_1 - \lambda_2)$$

$$\left.\begin{aligned} \frac{\partial L_D'(\underline{\lambda}, \mu)}{\partial \lambda_1} &= 1 - \lambda_1 + \mu = 0 \\ \frac{\partial L_D'(\underline{\lambda}, \mu)}{\partial \lambda_2} &= 1 - \lambda_2 - \mu = 0 \\ \frac{\partial L_D'(\underline{\lambda}, \mu)}{\partial \mu} &= \lambda_1 - \lambda_2 = 0 \end{aligned}\right\} \Rightarrow \begin{cases} \lambda_1 = \lambda_2 = 1 \\ \mu = 0 \end{cases}$$

So, $\underline{w}^* = \sum_{i=1}^{N} \lambda_i z_i \underline{u}_i = \begin{bmatrix} -1 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$

Because $\lambda_i\left[z_i(\underline{w}^{*T} \underline{u}_i + w_0^*) - 1\right] = 0 \quad \forall i$, we get $w_0^* = 0$

So the discriminant function is $g(\underline{u}) = \underline{w}^{*T} \underline{u} = 0 \Rightarrow \begin{bmatrix} -1 & -1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = -u_1 - u_2 = 0$ (say $\underline{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$)



(b) $d(\underline{u}_1, H) = \left|\frac{g(\underline{u}_1)}{\|\underline{w}^*\|}\right| = \left|\frac{\begin{bmatrix} -1 & -1 \end{bmatrix}\begin{bmatrix} -1 \\ 0 \end{bmatrix}}{\sqrt{(-1)^2 + (-1)^2}}\right| = \frac{1}{\sqrt{2}}$

$d(\underline{u}_2, H) = \left|\frac{g(\underline{u}_2)}{\|\underline{w}^*\|}\right| = \left|\frac{\begin{bmatrix} -1 & -1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix}}{\sqrt{(-1)^2 + (-1)^2}}\right| = \frac{1}{\sqrt{2}}$

No.