EE559 Week 9 HW Pin-Hsuan Lee 1139093874

1、(a) If $z_i(\underline{w}^T\underline{u}_i+w_0)-1\geq 0$ is satisfied, which means $z_i(\underline{w}^T\underline{u}_i+w_0)\geq 0$ for all i and all of the training data is correctly classified.

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

$\lambda_i\geq 0\ ,\ \forall i$

$z_i(\underline{w}^T\underline{u}_i+w_0)-1\geq 0\ ,\ \forall i$

$\lambda_i[z_i(w^Tu_i+w_0)-1]=0\ ,\ \forall i$

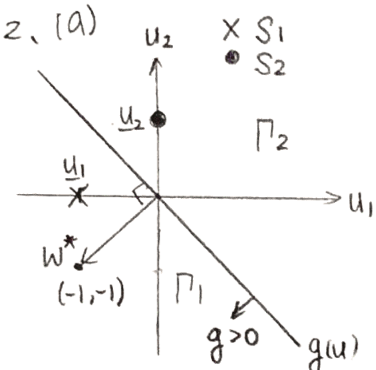
(c)(i) $\nabla_{\underline{w}}L=\ w-\sum_{j=1}^{N}\lambda_j z_j\underline{u}_j=0\ \Rightarrow\ w=\sum_{j=1}^{N}\lambda_j z_j\underline{u}_j$

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

(ii)
$$\begin{aligned}\mathcal{L}(\underline{w},w_0,\lambda)&=\frac{1}{2}\|w\|^2-\sum_{i=1}^{N}\lambda_i z_i w^Tu_i-\sum_{i=1}^{N}\lambda_i z_i w_0+\sum_{i=1}^{N}\lambda_i\\&=\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_j z_i z_j\underline{u}_i^T\underline{u}_j-\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_j z_i z_j\underline{u}_i^T\underline{u}_j-\underbrace{\sum_{i=1}^{N}\lambda_i z_i w_0}_{0}+\sum_{i=1}^{N}\lambda_i\\&=-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_j z_i z_j\underline{u}_i^T\underline{u}_j+\sum_{i=1}^{N}\lambda_i\end{aligned}$$

KKT $\begin{cases}\lambda_i\geq 0,\ \forall i\\ \lambda_i[z_i(w^Tu_i+w_0)-1]=0,\ \forall i\end{cases}$

2、(a)



$\underline{u}_1=\begin{bmatrix}-1\\0\end{bmatrix}\ ,\ \underline{u}_2=\begin{bmatrix}0\\1\end{bmatrix}$

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

$N=2\Rightarrow$
$$\begin{aligned}L_D'(\lambda,\mu)&=\lambda_1+\lambda_2-\frac{1}{2}(\lambda_1\lambda_1 z_1 z_1\underline{u}_1^T u_1+\lambda_2\lambda_1 z_2 z_1\underline{u}_2^T\underline{u}_1+\lambda_1\lambda_2 z_1 z_2\underline{u}_1^T\underline{u}_2\\&\quad+\lambda_2\lambda_2 z_2 z_2\underline{u}_2^T\underline{u}_2]+\mu z_1\lambda_1+\mu z_2\lambda_2\\&=\lambda_1+\lambda_2-\frac{1}{2}(\lambda_1\lambda_1 z_1 z_1\cdot 1+\cancel{\lambda_2\lambda_1 z_2 z_1\cdot 0}+\cancel{\lambda_1\lambda_2 z_1 z_2\cdot 0}+\lambda_2\lambda_2 z_2 z_2\cdot 1)\\&\quad+\mu z_1\lambda_1+\mu z_2\lambda_2\\&=\lambda_1+\lambda_2-\frac{1}{2}(\lambda_1^2 z_1^2+\lambda_2^2 z_2^2)+\mu z_1\lambda_1+\mu z_2\lambda_2\end{aligned}$$

$$\nabla_\lambda L_D'(\lambda,\mu)=\begin{bmatrix}1\\1\end{bmatrix}-\begin{bmatrix}\lambda_1 z_1^2\\\lambda_2 z_2^2\end{bmatrix}+\begin{bmatrix}\mu z_1\\\mu z_2\end{bmatrix}=\begin{bmatrix}1-\lambda_1+\mu z_1\\1-\lambda_2+\mu z_2\end{bmatrix}$$
$$=\begin{bmatrix}1-\lambda_1+\mu\\1-\lambda_2-\mu\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

Γ_1 is in positive side of the boundary

$\therefore z_1=1,\ z_2=-1$

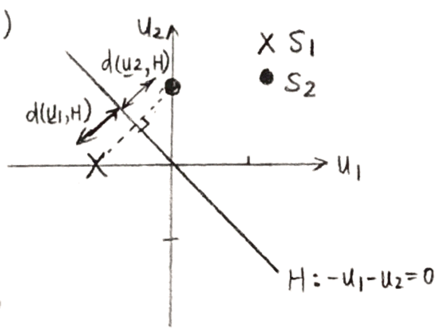
$\nabla_\mu L_D'(\lambda,\mu)=z_1\lambda_1+z_2\lambda_2=\lambda_1-\lambda_2=0$

$\begin{cases}1-\lambda_1+\mu=0\\1-\lambda_2-\mu=0\\\lambda_1-\lambda_2=0\end{cases}\Rightarrow\ \begin{matrix}\lambda_1=\lambda_2=1\\\mu=0\end{matrix}\ \Rightarrow\ w^*=\sum_{i=1}^{2}\lambda_j z_j\underline{u}_j=\lambda_1 z_1\underline{u}_1+\lambda_2 z_2\underline{u}_2=\begin{bmatrix}-1\\0\end{bmatrix}+\begin{bmatrix}0\\-1\end{bmatrix}=\begin{bmatrix}-1\\-1\end{bmatrix}$

satisfy the KKT condition that $\lambda_i\geq 0,\ \forall i$

$\lambda_i[z_i(w^Tu_i+w_0)-1]=0\Rightarrow\begin{matrix}\lambda_1[z_1(1+w_0^*)-1]=0\\\lambda_2[z_2(-1+w_0^*)-1]=0\end{matrix}\Big\}\ w_0^*=0$

(b)



$d(\underline{u}_1,H)=\dfrac{1\cdot([-1\ -1]\begin{bmatrix}-1\\0\end{bmatrix}+0)}{\sqrt{2}}=\dfrac{1}{\sqrt{2}}$

$d(\underline{u}_2,H)=\dfrac{-1\cdot([-1\ -1]\begin{bmatrix}0\\1\end{bmatrix}+0)}{\sqrt{2}}=\dfrac{1}{\sqrt{2}}$

$\because d(u_i,H)=\dfrac{z_i(\underline{w}^T\underline{u}_i+w_0)}{\|w\|}$

There is no other possible linear boundary in $\underline{u}$-space that would give larger values for both distances than H gives.

3. (a)

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

$$\left.\begin{aligned}
&\lambda_i \geq 0 \ , \ \forall i \\
&z_i(\underline{w}^T\underline{u}_i + w_0) - 1 + \xi_i \geq 0 \ , \ \forall i \\
&\lambda_i\left[z_i(\underline{w}^T\underline{u}_i + w_0) - 1 + \xi_i\right] = 0 \ , \ \forall i \\
&\mu_i \geq 0 \ , \ \forall i \\
&\xi_i \geq 0 \ , \ \forall i \\
&\mu_i\xi_i = 0 \ , \ \forall i
\end{aligned}\right\} \text{ KKT conditions}$$

(b)

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

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

$$\nabla_{\xi}L = \begin{bmatrix} c \\ c \\ \vdots \\ c \end{bmatrix}_{N\times 1} - \sum_{i=1}^{N}\lambda_i - \sum_{i=1}^{N}\mu_i = 0 \Rightarrow \begin{bmatrix} c \\ c \\ \vdots \\ c \end{bmatrix}_{N\times 1} = \sum_{i=1}^{N}\lambda_i + \sum_{i=1}^{N}\mu_i \Rightarrow C = \lambda_i + \mu_i \ , \ \forall i$$

$$L(\underline{w}, w_0, \xi, \underline{\lambda}, \underline{\mu}) = \frac{1}{2}\|w^*\|^2 + \cancel{C\sum_{i=1}^{N}\xi_i} - \sum_{i=1}^{N}\lambda_i z_i \underline{w}^{*T}\underline{u}_i - \underbrace{\sum_{i=1}^{N}\lambda_i z_i w_0}_{0} + \sum_{i=1}^{N}\lambda_i - \cancel{\sum_{i=1}^{N}\lambda_i\xi_i} - \cancel{\sum_{i=1}^{N}\mu_i\xi_i}$$

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

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

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

$$0 \leq \lambda_i \leq C \ , \ \forall i$$

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

$$\mu_i\xi_i = 0 \ , \ \forall i$$

$$\mu_i \geq 0 \ , \ \forall i$$