Problem 1

(a) Yes, all the training data will be correctly classified

(b)

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

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

(c)

(i) $\nabla_{\underline{w}} L(\underline{w}, w_0, \underline{\lambda}) = 0$

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

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

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

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

(ii) let $\underline{w}$ be $\sum_{i=1}^{N} \lambda_i z_i \underline{v}_i$

$$L_D(\underline{\lambda}) \Rightarrow \frac{1}{2} \sum_{i=1}^{N} \lambda_j z_i \underline{v}_i - \sum_{i=1}^{N} \lambda_i [z_i(\underline{w}^T \underline{v}_i + w_0) - 1]$$

$$L_D(\underline{\lambda}) \Rightarrow -\frac{1}{2}\left[\sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j z_i z_j \underline{v}_i^T \underline{v}_j\right] + \sum_{i=1}^{N} \lambda_i - w_0 \sum_{i=1}^{N} \lambda_i z_i$$

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

so, $L_D(\lambda) = -\frac{1}{2}\left[\sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i \lambda_j z_i z_j \underline{v_i}^T \underline{v_j}\right] + \sum_{i=1}^{N} \lambda_i$

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

$\lambda_i \geq 0$ , $\lambda_i [z_i(\underline{w}^{*T} \underline{v_i} + w_0 - 1] = 0 \quad \forall i$

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

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

Problem 2: (a)

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

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

so,

$-\frac{1}{2}(\lambda_1^2 + \lambda_2^2) + \lambda_1 + \lambda_2$

subject to $\lambda_1 - \lambda_2 = 0$

$\lambda_i \geq 0 \quad i = 1, 2$

let $\lambda_1 = \lambda_2$

so, we have $-\frac{1}{2}(2\lambda_1^2) + 2\lambda_1$

$$\frac{\partial f(\lambda)}{\lambda_1} = 0$$

$\Rightarrow$ we have $-2\lambda_1 + 2 = 0$

so, $\lambda_1 = \lambda_2 = 1$ and it satify $\lambda_i \geq 0$

so, we have $\lambda_1 = \lambda_2 = 1$

Because $w^* = \sum_{i=1}^{N} \lambda_i z_i \underline{v_i}$

so, $w^*_1 = w^*_2 = -1$

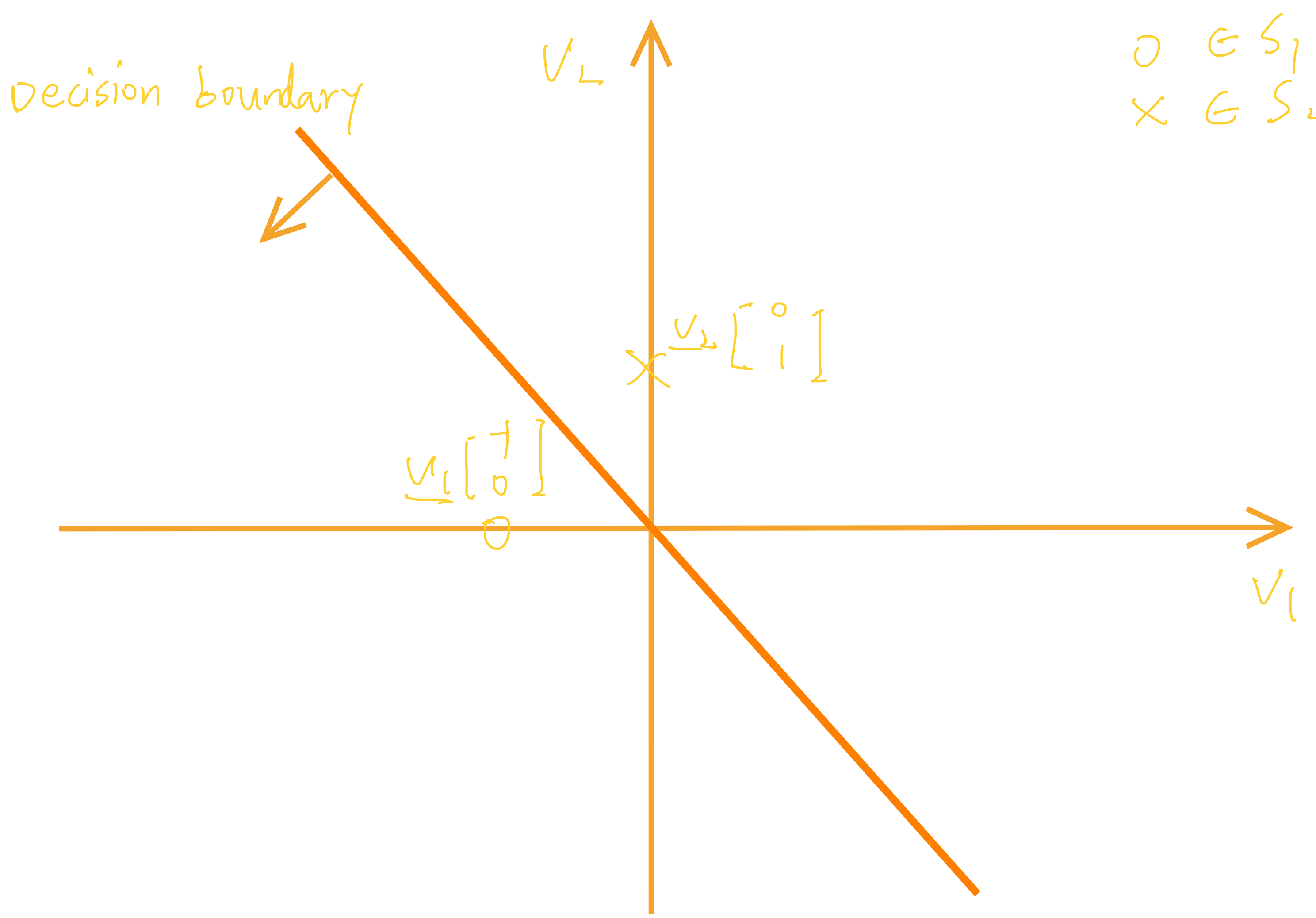
and because $\lambda_i [z_i(\underline{w}^{*T} \underline{v_i} + w_0) - 1] = 0$

and $\lambda_i > 0$

so, $z_i(\underline{w}^{*T} \underline{v_i} + w_0) - 1 = 0$

so $w_0 = 0$

so, the hyperplane will be $-u_1 - u_2 = 0$



(b)

The distance between $\underline{v_1}$ and $H$ is $\frac{g(\underline{u_1})}{\|w\|} = \frac{\sqrt{2}}{2}$

the distance between $\underline{v_2}$ and $H$ is $\frac{g(\underline{u_2})}{\|w\|} = \frac{\sqrt{2}}{2}$

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