jxu72364@usc.edu

Problem 3

(a) $\underline{\underline{S}}_{\underline{w}} = \underline{\underline{S}}_1 + \underline{\underline{S}}_2$

$$\underline{\underline{S}}_{\underline{w}} = \begin{pmatrix} \sigma_1^2+\rho_1^2 & & & 0 \\ & \sigma_2^2+\rho_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_D^2+\rho_D^2 \end{pmatrix}$$

$$\underline{w} = \underline{\underline{S}}_w^{-1}[\underline{m}_1 - \underline{m}_2]$$

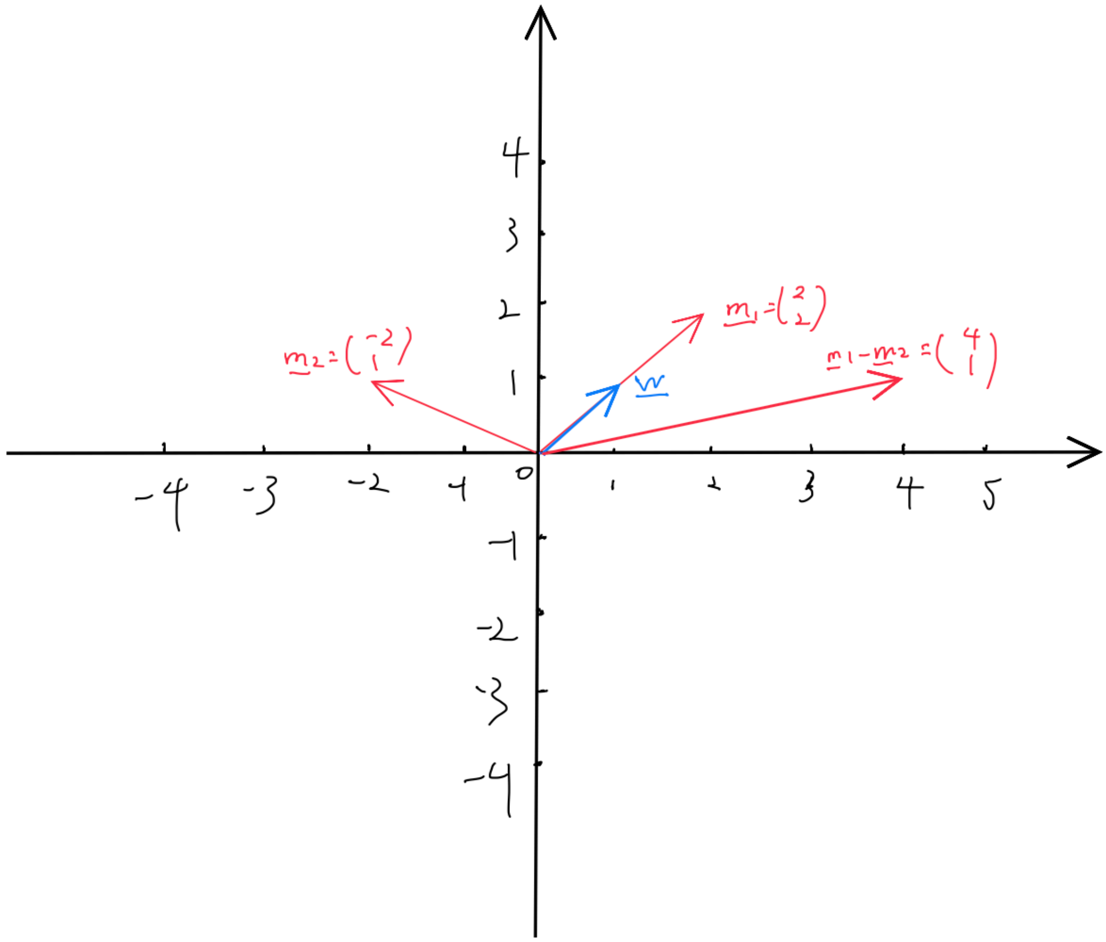
$$\underline{m}_1 - \underline{m}_2 = \begin{bmatrix} m_1^{(1)} - m_1^{(2)} \\ m_2^{(1)} - m_2^{(2)} \\ \vdots \\ m_D^{(1)} - m_D^{(2)} \end{bmatrix}, \text{ and } \underline{\underline{S}}_w^{-1} = \begin{pmatrix} \frac{1}{\sigma_1^2+\rho_1^2} & & & 0 \\ & \frac{1}{\sigma_2^2+\rho_2^2} & & \\ & & \ddots & \\ 0 & & & \frac{1}{\sigma_D^2+\rho_D^2} \end{pmatrix}$$

$$\text{so, } \underline{w} = \begin{pmatrix} \frac{m_1^{(1)} - m_1^{(2)}}{\sigma_1^2+\rho_1^2} \\ \frac{m_2^{(1)} - m_2^{(2)}}{\sigma_2^2+\rho_2^2} \\ \vdots \\ \frac{m_D^{(1)} - m_D^{(2)}}{\sigma_D^2+\rho_D^2} \end{pmatrix}$$

(b) $$\underline{m}_1 - \underline{m}_2 = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$$

$$\underline{w} = \begin{pmatrix} \frac{4}{4\sigma_2^2+4\rho_2^2} \\ \frac{1}{\sigma_2^2+\rho_2^2} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sigma_2^2+\rho_2^2} \\ \frac{1}{\sigma_2^2+\rho_2^2} \end{pmatrix} = \frac{1}{\sigma_2^2+\rho_2^2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

So, we can plot:



(C)

Although for this problem the direction for (m1 – m2) and W is same, I think W makes more sense for a 1D feature direction, because m1 – m2 only consider about the distance between projected class means, but we should consider another target, which is to make data points in each class as close as possible.

Therefore, W makes more sense for a 1D feature space direction.