HW 5.

↳ Hardik Prajapati (2678294168)

1.

$E[x] = 0$, $R = E[xx^T]$

$$\hat{x} = \sum_{i=1}^{M} a_i e_i = a^T e$$

$$J = E[\|x - \hat{x}\|^2]$$

$$= E[(x - \hat{x})^T (x - \hat{x})]$$

$$= E[(x - a^T e)^T (x - a^T e)]$$

$$= E[x^T x - x^T a^T e - e^T a x + e^T a a^T e]$$

2.

$p(x|w_i) \sim$ Mean μ_i & covariance matrices Σ_i

$y = w^T x$

$p(y|w_i) \sim$ Mean $\tilde{\mu}_i$ & variances σ_i^2

$$J_1(w) = \frac{(\tilde{\mu}_1 - \tilde{\mu}_2)^2}{\sigma_1^2 + \sigma_2^2}$$

$$\tilde{\mu}_i = \frac{1}{\# D_i} \sum_{x \in y} \cancel{y} = \frac{1}{\# D_i} \sum_{x \in D_i} w^T x$$

$$= \frac{1}{\# D_i} w^T \sum_{x \in D_i} x$$

$$= w^T \mu_i$$

Numerator :- $(\tilde{\mu}_1 - \tilde{\mu}_2)^2 = (w^T \mu_1 - w^T \mu_2)^2$

$$= [w^T \mu_1 - w^T \mu_2][w^T \mu_1 - w^T \mu_2]^T$$

$$= w^T (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T w$$

$$= w^T S_B w$$

where, $S_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T$

Denominator:- $\sigma_1^2 + \sigma_2^2$, $\sigma_i^2 = \sum_{y \in w_i} (y - \tilde{\mu}_i)^2$

$$= \sum_{x \in w_i} (w^T x - w^T \mu_i)^2$$

$$= \sum_{x \in w_i} \left[w^T (x-\mu_i)\right]\left[w^T (x-\mu_i)\right]^T$$

$$= \sum_{x \in w_i} w^T (x-\mu_i)(x-\mu_i)^T w$$

$$= w^T \left[\sum_{x \in w_i} (x-\mu_i)(x-\mu_i)^T\right] w$$

$$= w^T S_1 w + w^T S_2 w$$

$$= w^T S_w w \quad , \quad S_w = S_1 + S_2$$

$$J(w) = \frac{w^T S_B w}{w^T S_w w}$$

Comparing to Rayleigh Quotient,

$$S_w = C^T C \qquad C = \sqrt{S_w}$$

$$J(w) = \frac{w^T S_B w}{w^T C^T C w}$$

$$Cw = z \qquad w = C^{-1} z$$

$$J(z) = \frac{z^T C^{-T} S_B C^{-1} z}{z^T z}$$

Rayleigh Quotient

$\Rightarrow$ z that maximizes $J(z)$ is eigenvalue associated with $\lambda_{max}(C^{-T} S_B C^{-1})$

$$C^{-1} S_B \underbrace{C^{-1} z}_{w} = \lambda_{max} \underbrace{z}_{Cw}$$

$$\Rightarrow C^{-T} S_B w = \lambda_{max} C w$$

$$\Rightarrow S_B w = \lambda_{max} \underbrace{C^T C}_{S_w} w$$

$$\Rightarrow S_B w = \lambda_{max} S_w w$$

$$\Rightarrow \boxed{S_w^{-1} S_B w = \lambda_{max} w}$$

$$S_w^{-1} (\mu_1 - \mu_2) \underbrace{(\mu_1 - \mu_2)^T w}_{\alpha} = \lambda_{max} w$$

$$w = \frac{\alpha}{\lambda_{max}} S_w^{-1} (\mu_1 - \mu_2)$$

$$\boxed{w = S_w^{-1} (\mu_1 - \mu_2)}$$

$$\Rightarrow \boxed{w = (\Sigma_1 + \Sigma_2)^{-1} (\mu_1 - \mu_2)}$$

//

Q2. b) $$J_2(w) = \frac{(\mu_1 - \mu_2)^2}{P(\omega_1)\sigma_1^2 + P(\omega_2)\sigma_2^2}$$

Denominator:-

$$P(\omega_1)\sigma_1^2 + P(\omega_2)\sigma_2^2, \quad = P(\omega_1)\{w^T S_1 w\} + P(\omega_2)\{w^T S_2 w\}$$

$$= w^T \{P(\omega_1)S_1 + P(\omega_2)S_2\} w$$

$$\Rightarrow J_2(w) = \frac{w^T S_B w}{w^T \{P(\omega_1)S_1 + P(\omega_2)S_2\} w}$$

$$\{P(\omega_1)S_1 + P(\omega_2)S_2\} = C^T C$$

$$\Rightarrow C = \sqrt{P(\omega_1)S_1 + P(\omega_2)S_2}$$

$$\Rightarrow J(w) = \frac{w^T S_B w}{w^T C^T C w}$$

$Cw = z$ $\qquad w = C^{-1} z$

$$J(z) = \frac{z^T C^{-T} S_B C^{-1} z}{z^T z}$$

Rayleigh Quotient

$\Rightarrow$ The $\underline{z}$ that maximizes $J(z)$ is eigen vector that is associated with $\lambda_{max}(C^{-T} S_B C^{-1})$:

$$C^{-T} S_B \underbrace{C^{-1} z}_{w} = \lambda_{max} \underbrace{z}_{Cw}$$

$$\Rightarrow C^{-T} S_B w = \lambda_{max} C w$$

$$\Rightarrow S_B w = \lambda_{max} C^T C w$$

$$\Rightarrow S_B w = \lambda_{max} \{P(w_1) S_1 + P(w_2) S_2\} w.$$

$$\Rightarrow \{P(w_1) S_1 + P(w_2) S_2\}^{-1} S_B w = \lambda_{max} w$$

$$\Rightarrow \{P(w_1) S_1 + P(w_2) S_2\}^{-1} (\mu_1 - \mu_2) \underbrace{(\mu_1 - \mu_2)^T w}_{\alpha} = \lambda_{max} w$$

$$\Rightarrow w = \frac{\alpha}{\lambda_{max}} \{P(w_1) S_1 + P(w_2) S_2\}^{-1} (\mu_1 - \mu_2)$$

$$\Rightarrow \boxed{w = (P(w_1) \Sigma_1 + P(w_2) \Sigma_2)^{-1} (\mu_1 - \mu_2)}$$

Q2. c) $J(w_1)$ is "closer" to criterion that is used by Fisher's LDA.