PCA:

$$\underset{s\times d}{U^T} \underset{d\times n}{X} \rightarrow \underset{s\times n}{X'} \qquad X \rightarrow n\times d.$$

$X \rightarrow d\times n.$ $\qquad$ ω

$U \rightarrow d\times s$ $\qquad$ $s < d$

Main objective. $\qquad$ **Constraint:**

$$\max\left[\mathrm{Var}(U^T X)\right] \; ; \; U^T U = I$$

$$\mathrm{Var}(U^T X) \rightarrow E\left[\left(U^T X - E[U^T X]\right)\left(U^T X - E[U^T X]\right)^T\right]$$

$$E[U^T X] \Rightarrow U^T E[X]$$

$$= E\left[U^T(X - E[X])\left(U^T(X - E[X])\right)^T\right]$$

$$= E\left[U^T \underbrace{(X - E[X])(X - E[X])^T}_{S} U\right]$$

$$= E\left[U^T S U\right] = U^T S U$$

$$\rightarrow \max\left[U^T S U\right] \quad \text{Constraint } U^T U = I$$

$$\max_U \left[U^T S U - \lambda(U^T U - 1)\right]$$

$$= \frac{\partial}{\partial U}\left[U^T S U - \lambda(U^T U - 1)\right]$$

$$= 2SU - \lambda 2U = 0$$

$$SU = \lambda U$$

$\max(U^T S U)$

$\max(U^T \lambda U)$

$\max(\lambda) \rightarrow$ Largest eigen vectors

Linear assumption

→ $\boxed{\text{Sensitive to Scaling}}$

→ Sensitive to outliers.

LDA

LDA

two classes C_1, C_2
$\downarrow \quad \downarrow$
$n_1 \quad n_2$

$\mu_1, \mu_2 \rightarrow$ class means before projection.

$$\cancel{\tilde{\mu}_1 = \frac{1}{n_1} \sum_{x_i \in C_1}^{n_1} v^T x_i} = v^T \left(\frac{1}{n} \sum_{x_i \in C_1}^{n} x_i \right)$$

$$\tilde{\mu}_1 = v^T \mu_1$$

$$\tilde{\mu}_2 = v^T \mu_2$$

objective:

1. Maximize Seperation between classes
2. Minimize within class variance

$q_i \rightarrow \omega^T x_i$

$$\tilde{\mu}_1 = \frac{1}{n_1} \sum_{x_i \in C_1} \omega^T x_i = \omega^T \frac{1}{n_1} \sum_{x_i \in C_1} x_i$$

$$\tilde{\mu}_1 = \omega^T \mu_1$$

$$\tilde{\mu}_2 = \omega^T \mu_2$$

1st objective $\max\limits_{v} (\mu_1 - \mu_2)$

$\downarrow$

$\max\limits_{v} (\mu_1 - \mu_2)^2$

$$\min \quad s_1^2 = \sum_{x_i \in C_1} (a_i - \tilde{\mu}_1)^2, \quad s_2^2 = \sum_{x_i \in C_2} (a_i - \tilde{\mu}_2)^2$$

2nd obs: ↗

$$\max_{v} \frac{(\tilde{\mu}_1 - \tilde{\mu}_2)^2}{s_1^2 + s_2^2}$$

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

$$= v^T \underbrace{(\mu_1 - \mu_2)(\mu_1 - \mu_2)^T}_{S_b} v$$

$$\boxed{(\tilde{\mu}_1 - \tilde{\mu}_2)^2 = \omega^T S_b \omega}$$

$$s_j^2 = \sum_{x_i \in C_j} (a_i - \tilde{\mu}_j)^2 = \sum_{x_i \in C_j} (\omega^T x_i - \omega^T \mu_j)^2$$

$$= \sum_{x_i \in C_j} \omega^T (x_i - \mu_j)(x_i - \mu_j)^T v$$

$$= \omega^T \left[\underbrace{\sum_{x_i \in C_j} (x_i - \mu_j)(x_i - \mu_j)^T}_{S_j} \right] \omega$$

$$s_j^2 = \omega^T S_j \omega$$

$$s_1^2 + s_2^2 = \omega^T S_1 \omega + \omega^T S_2 \omega$$

$$= \omega^T \underbrace{(S_1 + S_2)}_{S_\omega} \omega = \omega^T S_\omega \omega$$

$$\boxed{s_1^2 + s_2^2 = \omega^T S_\omega \omega}$$

$$\max_{w} \frac{w^T S_b w}{w^T S_w w}$$

$$\frac{d}{dw}\left(\frac{w^T S_b w}{w^T S_w w}\right) = \frac{S_b w \, w^T S_w w - w^T S_b w \, S_w w}{(w^T S_w w)^2}$$

$$S_b w \, w^T S_w w = w^T S_b w \, S_w w$$

$$\cancel{S_b w \frac{w^T S_w w}{w^T S_b w} = S_w w}$$

$$\cancel{S_b}$$

$$S_b w = \underbrace{\frac{w^T S_b w}{w^T S_w w}}_{\lambda} S_w w$$

$$S_b w = S_w \lambda w$$

$$\underbrace{S_w^{-1} S_b}_{m} w = \lambda w$$

$$m w = \lambda w$$

→ Linearity assumption

→ Sensitive to outliers

→ Sensitive to class imbalance (bias towards the majority class)