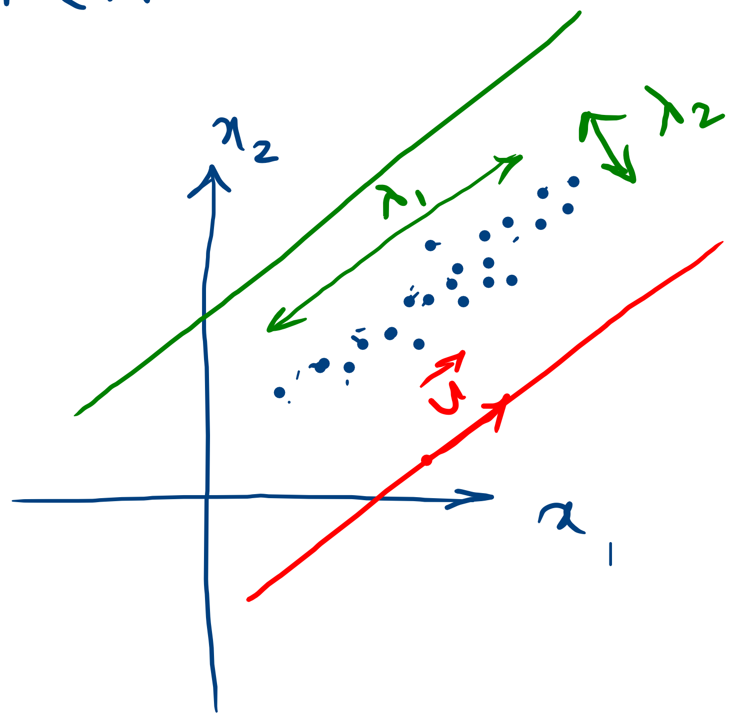
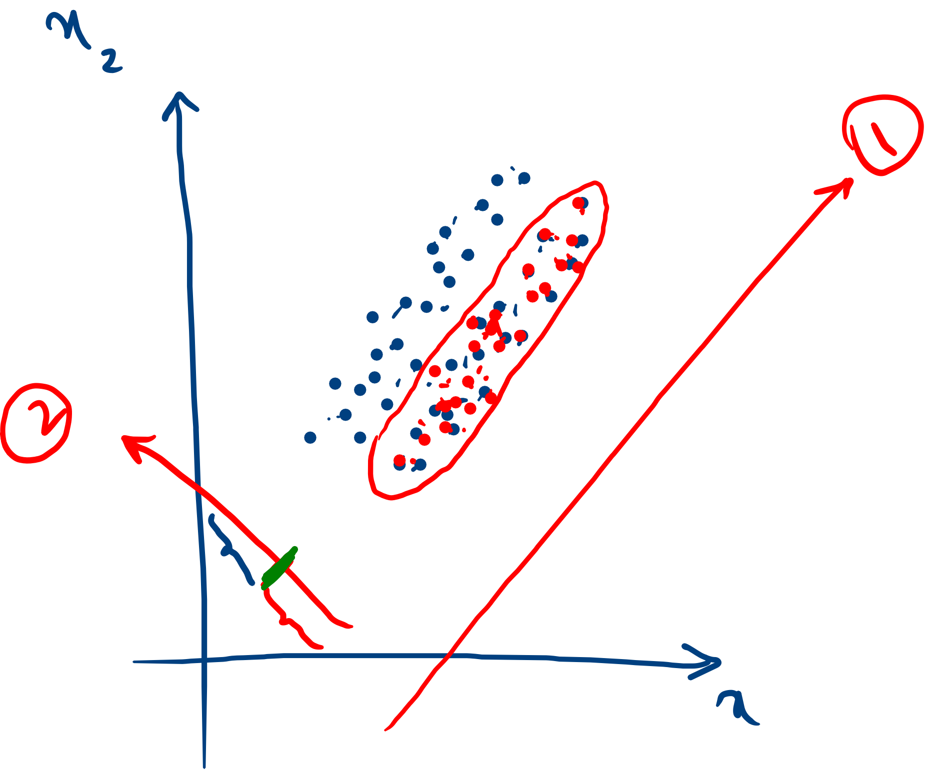
Dimension Reduction

PCA

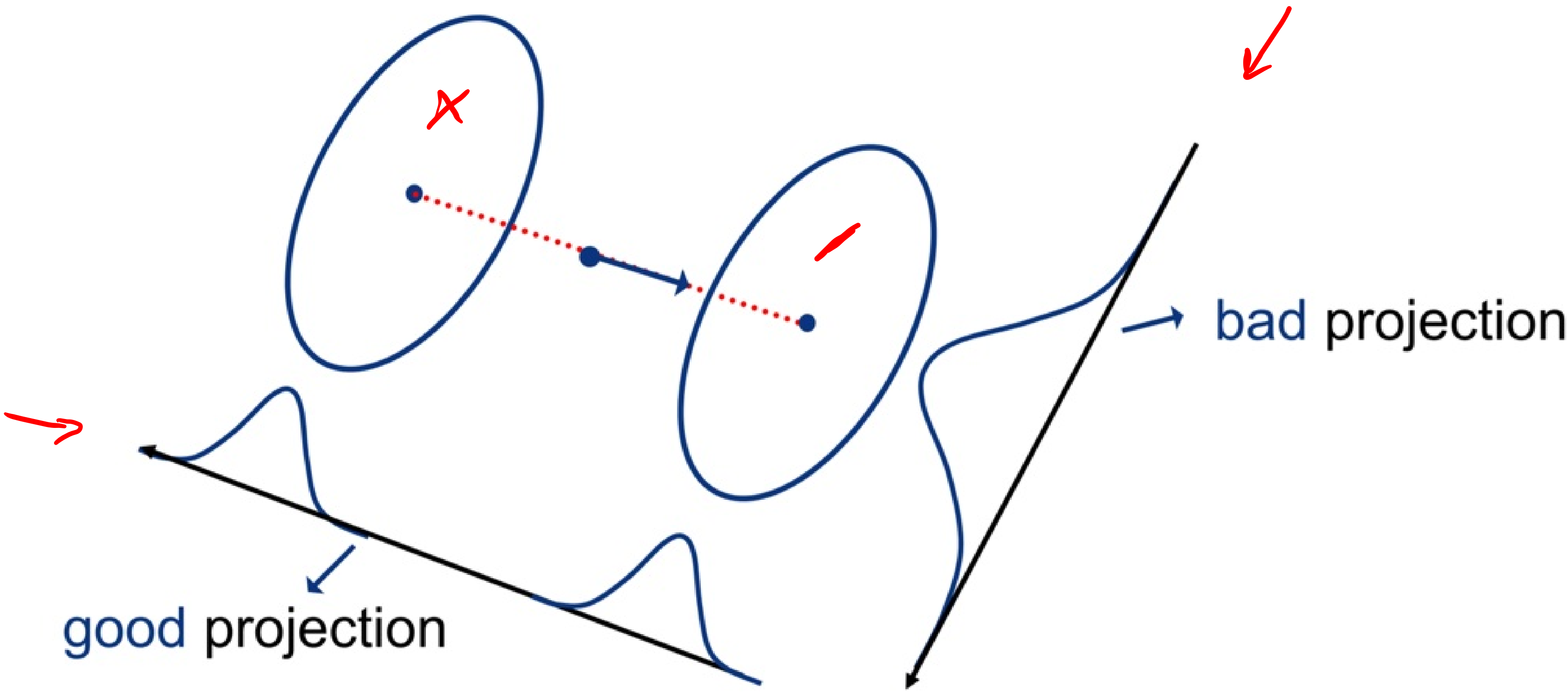


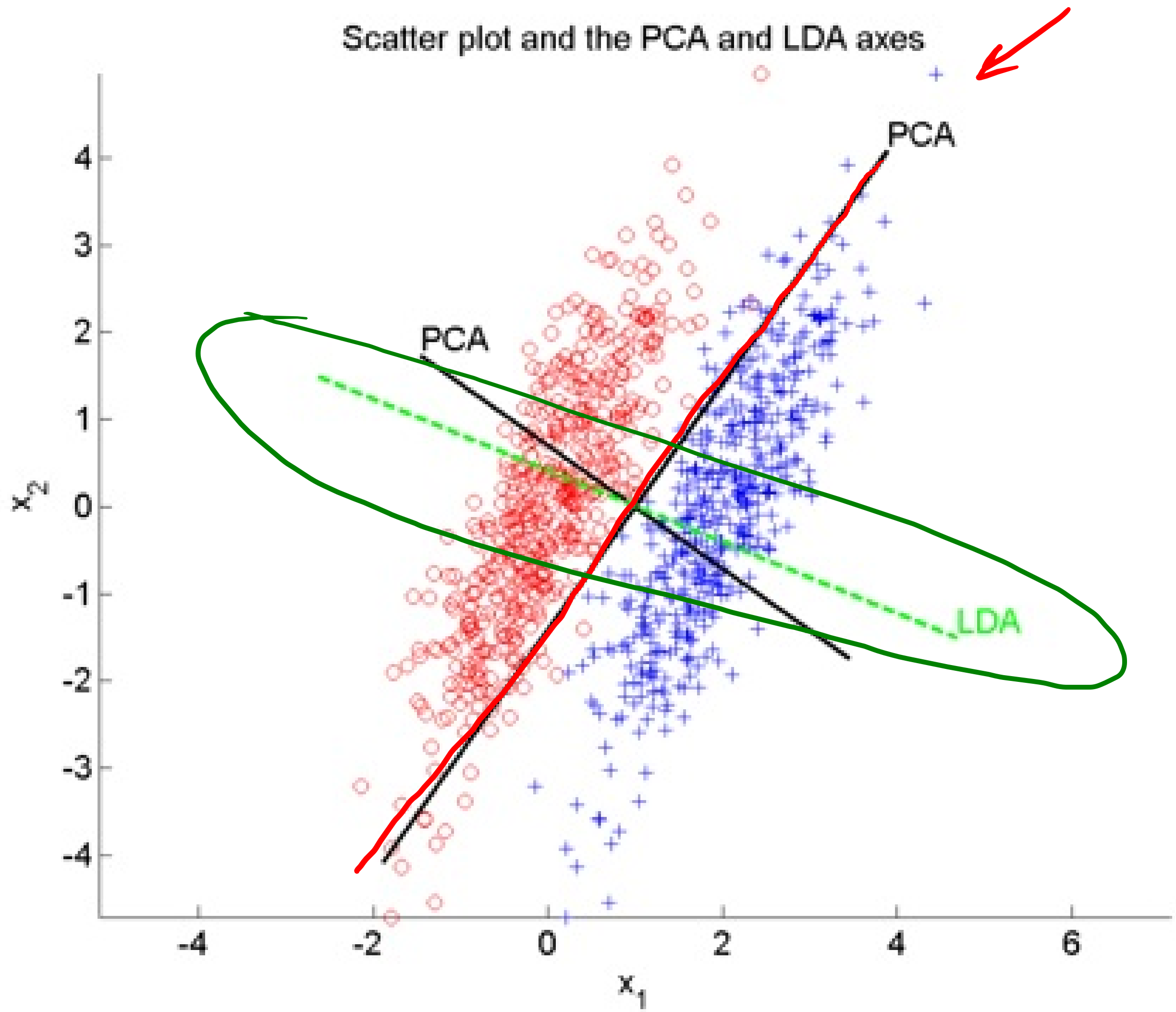
$$\boxed{\max_{u} \ \mathrm{var}(z) \quad \text{s.t.} \quad \|u\|_2^2 = 1}$$

$$\boxed{S = \frac{1}{n}\sum_{i=1}^{n} x_i x_i^T}$$



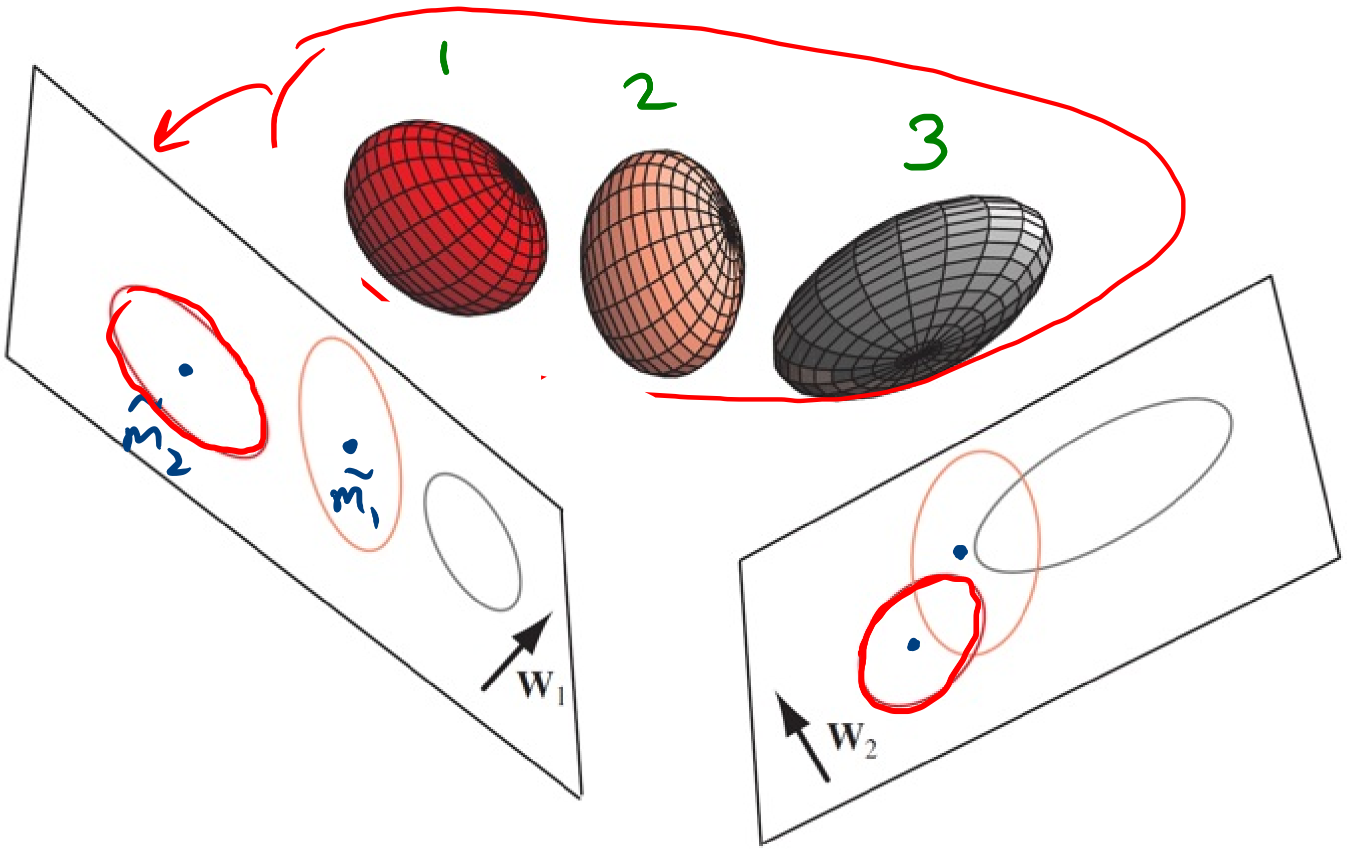
x_2
x_1
1
2

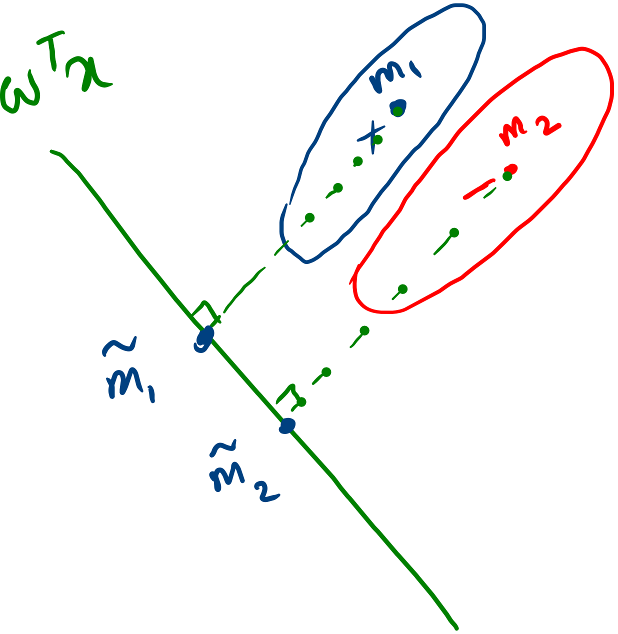

bad projection
good projection



Scatter plot and the PCA and LDA axes
PCA
PCA
LDA
4
3
2
1
0
-1
-2
-3
-4
-4
-2
0
2
4
6
x_2
x_1

Three-dimensional distributions are projected onto two-dimensional subspaces described by a normal vectors $\mathbf{w}_1$ and $\mathbf{w}_2$





$$\boxed{J(w) = \frac{(\tilde{m}_1 - \tilde{m}_2)^2}{\tilde{S}_1^2 + \tilde{S}_2^2}}$$

Scatter

$$(\tilde{m}_1 - \tilde{m}_2)^2 - \alpha\left(\tilde{S}_1^2 + \tilde{S}_2^2\right)$$

LDA: linear Discriminant Analysis

$$\underset{d \times d}{S} = \frac{1}{n} \sum_{i=1}^{n} x_i x_i^T \qquad x_i \in \mathbb{R}^d$$

$$\max_{\omega} \frac{(\tilde{m}_1 - \tilde{m}_2)^2}{\tilde{S}_1^2 + \tilde{S}_2^2}$$

$$\boxed{g(x) = \underline{\underline{\omega^T}} x \atop ?}$$

$$\tilde{m}_1 = \frac{1}{n}\sum_{i=1}^{n} z_i = \frac{1}{n}\sum_{i=1}^{n} \omega^T x_i = \omega^T \underbrace{\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)}_{m_1} = \omega^T m_1$$

$$\tilde{m}_2 = \omega^T m_2$$

$$\tilde{S}_1^2 = \sum_{i=1}^{n} z_i z_i^T = \sum_{i=1}^{n} (\omega^T x_i)(\omega^T x_i)^T = \sum_{i} \omega^T x_i x_i^T \omega$$

$$= \omega^T \underbrace{\left(\sum_{i=1}^{n} x_i x_i^T\right)}_{S_1^2} \omega = \omega^T S_1^2 \omega$$

$$\tilde{S}_2^2 = \omega^T S_2^2 \omega$$

$$J(w) = \frac{(w^T m_1 - w^T m_2)^2}{w^T s_1^2 w + w^T s_2^2 w} = \frac{w^T \overbrace{(m_1 - m_2)(m_1 - m_2)^T}^{S_B} w}{w^T \underbrace{(s_1^2 + s_2^2)}_{} w}$$

S_W : within cluster scatter

S_B : Between cluster scatter

$$\boxed{\max_{w} J(w) = \frac{w^T S_B w}{w^T S_W w}}$$

$$S_B = \underset{d \times 1}{(m_1 - m_2)} \underset{1 \times d}{(m_1 - m_2)^T}$$

Rayleigh Quotient

$$\boxed{\begin{aligned} &\max_{w} \ w^T S_B w \\ &\text{s.t.} \quad w^T S_w w = K \end{aligned}}$$

$$\max_{w} L(w) = \underbrace{w^T S_B w} - \lambda \left(\underbrace{w^T S_w w} - K \right)$$

$$\lambda \geq 0$$

$$2 S_B w - 2\lambda S_w w = 0 \implies \boxed{S_B w = \lambda S_w w}$$

$$\boxed{\underbrace{S_w^{-1} S_B}_{A} w = \lambda w}$$

$$\max w^T S_B w = \max w^T \lambda S_w w = \max \lambda \underbrace{w^T S_w w}_{K} = \max \lambda K$$

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

$$S_B = \underbrace{\overbrace{(m_1 - m_2)}\,(m_1 - m_2)^T}$$

$$\lambda w = S_w^{-1} (m_1 - m_2) \underbrace{(m_1 - m_2)^T w}_{\alpha} \Rightarrow w = \frac{\alpha}{\lambda} S_w^{-1} (m_1 - m_2)$$

$$\Rightarrow \boxed{w^* = S_w^{-1} (m_1 - m_2)}$$

$$\boxed{S_B w} = a \underbrace{a^T w}_{\alpha} = \alpha a$$

$$a = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$S = aa^T = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$w = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$Sw = \begin{bmatrix} 6 \\ 12 \\ 18 \end{bmatrix} = 6 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 6a$$

$$2 \longrightarrow 1$$

$$c \longrightarrow c-1$$