Dimension reduction

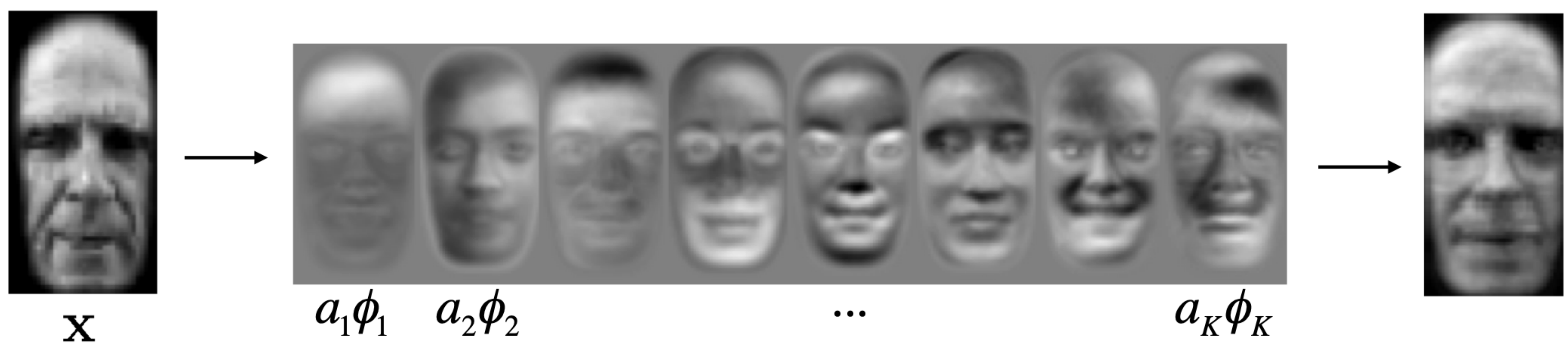
PCA examples with eigenfaces

Compute each training image x_i 's projections as

$$x_i \rightarrow \left(x_i^c \cdot \phi_1, x_i^c \cdot \phi_2, \ldots, \; x_i^c \cdot \phi_K \right) \equiv \left(a_1, a_2, \ldots, a_K \right) \qquad x_i^c \text{ : mean-centered } x_i$$

Visualize the estimated training face x_i

$$x_i \approx \mu + a_1\phi_1 + a_2\phi_2 + \ldots + a_K\phi_K$$



For a new image t

$$t \rightarrow \left((t-\mu)\cdot\phi_1, (t-\mu)\cdot\phi_2, \ldots, \; (t-\mu)\cdot\phi_K \right) \equiv \left(w_1, w_2, \ldots, w_K \right)$$

Compare the projection w with all known projections to find the closest one

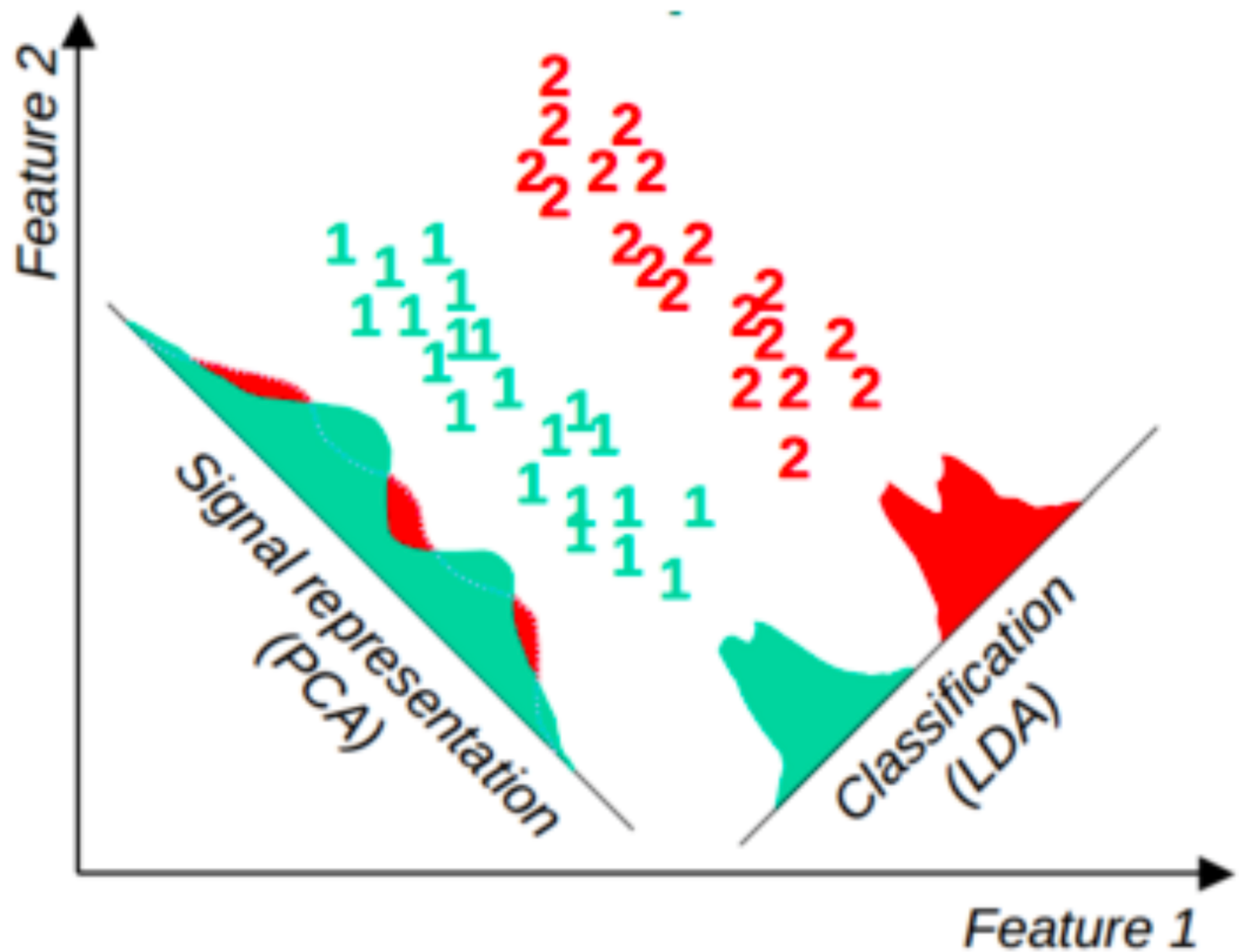
PCA examples with eigenfaces

K = 4

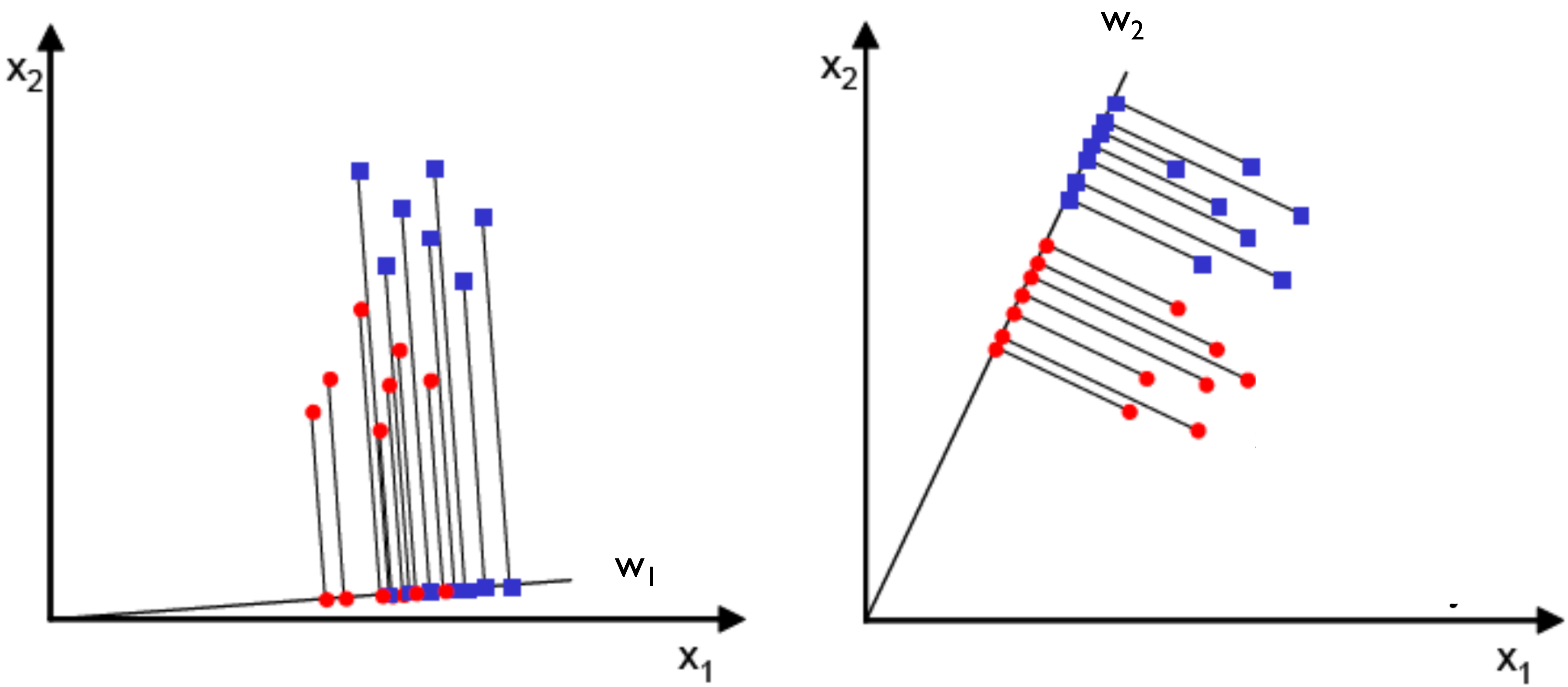
K = 200

K = 400

PCA for the data with labels?



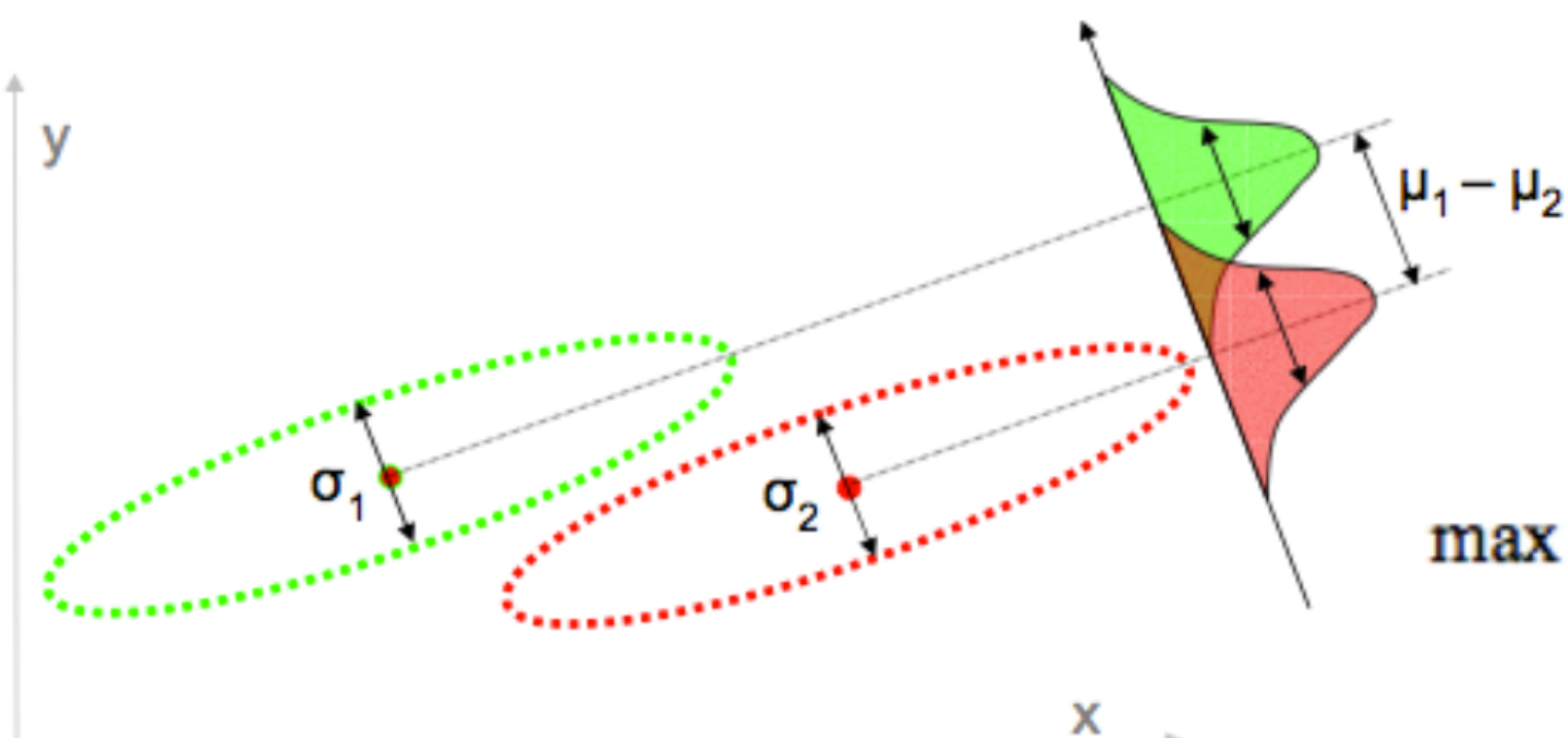
PCA for the data with labels?



Linear Discriminant Analysis

- LDA focuses on maximizing the seperability among known categories
- Supervised feature extraction
- Pick a new dimension that represents
  - max seperation between means of projected classes
  - min variance within each projected class

→ Eigenvectors based on between-class and within-class covariance matrices



$$\max \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} \quad \begin{matrix} S_B \\ S_w \end{matrix}$$

$S_w = S_1 + S_2$

Linear Discriminant Analysis

$$J(w) = \frac{\left|\tilde{\mu}_1 - \tilde{\mu}_2\right|^2}{\tilde{s}_1^2 + \tilde{s}_2^2}$$

$$\mu_i = \frac{1}{N_i} \sum_{x \in \omega_i} x \quad \textit{and} \quad \tilde{\mu}_i = \frac{1}{N_i} \sum_{y \in \omega_i} y = \frac{1}{N_i} \sum_{x \in \omega_i} w^T x$$

$$= w^T \frac{1}{N_i} \sum_{x \in \omega_i} x = w^T \mu_i$$

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

$$= w^T \underbrace{\left(\mu_1 - \mu_2\right)\left(\mu_1 - \mu_2\right)^T}_{S_B} w$$

$$= w^T S_B w = \tilde{S}_B$$

Linear Discriminant Analysis

$$J(w) = \frac{\left|\tilde{\mu}_1 - \tilde{\mu}_2\right|^2}{\tilde{s}_1^2 + \tilde{s}_2^2}$$

$$\tilde{s}_i^2 = \sum_{y \in \omega_i} \left(y - \tilde{\mu}_i\right)^2 = \sum_{x \in \omega_i} \left(w^T x - w^T \mu_i\right)^2$$

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

$$= w^T \left(\sum_{x \in \omega_i} \left(x - \mu_i\right)\left(x - \mu_i\right)^T \right) w = w^T S_i w$$

$$\tilde{s}_1^2 + \tilde{s}_2^2 = w^T S_1 w + w^T S_2 w = w^T \left(S_1 + S_2\right) w = w^T S_W w = \tilde{S}_W$$

$$J(w) = \frac{\left|\tilde{\mu}_1 - \tilde{\mu}_2\right|^2}{\tilde{s}_1^2 + \tilde{s}_2^2} = \frac{w^T S_B w}{w^T S_W w}$$

Linear Discriminant Analysis

To find the maximum J(w)

$$\frac{d\frac{f(x)}{g(x)}}{dx} = \frac{f'(x)g(x) - g'(x)f(x)}{g(x)^2}$$

$$\frac{d}{dw}J(w) = \frac{d}{dw}\left(\frac{w^T S_B w}{w^T S_W w}\right) = 0$$

$$\Rightarrow \left(w^T S_W w\right)\frac{d}{dw}\left(w^T S_B w\right) - \left(w^T S_B w\right)\frac{d}{dw}\left(w^T S_W w\right) = 0$$

$$\Rightarrow \left(w^T S_W w\right) 2 S_B w - \left(w^T S_B w\right) 2 S_W w = 0$$

$$Dividing \;\; by \;\; 2w^T S_W w:$$

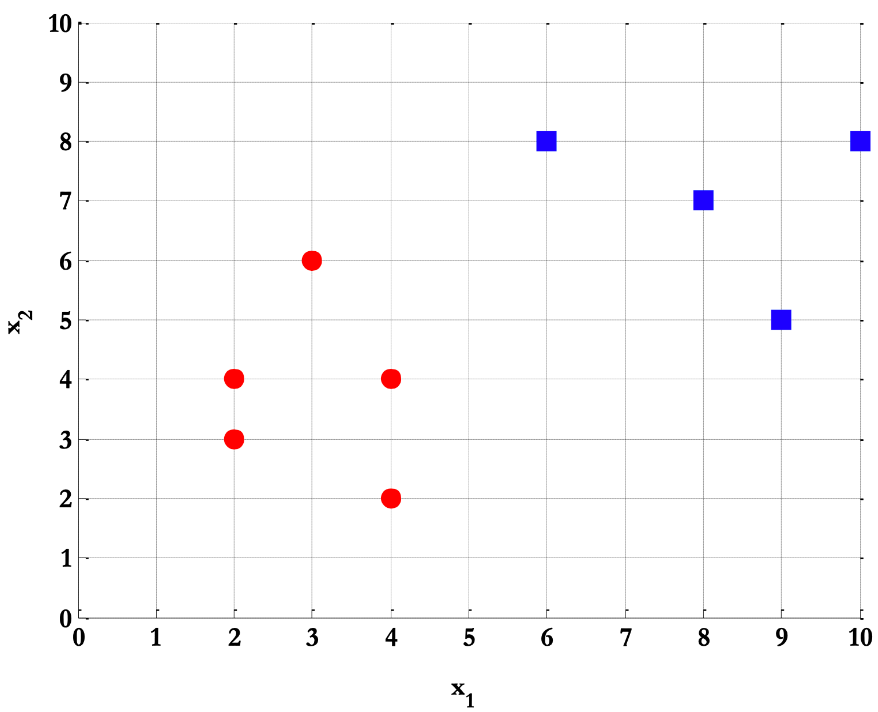
$$\Rightarrow \left(\frac{w^T S_W w}{w^T S_W w}\right) S_B w - \left(\frac{w^T S_B w}{w^T S_W w}\right) S_W w = 0$$

$$\Rightarrow S_B w - J(w) S_W w = 0$$

$$\Rightarrow S_W^{-1} S_B w - J(w) w = 0$$

$$w^* = \arg\max_w J(w) = S_W^{-1}(\mu_1 - \mu_2)$$

Linear Discriminant Analysis



- Samples for class $\boldsymbol{\omega_1}$: $\mathbf{X_1}$=(x_1,x_2)={(4,2),(2,4),(2,3),(3,6),(4,4)}
- Sample for class $\boldsymbol{\omega_2}$: $\mathbf{X_2}$=(x_1,x_2)={(9,10),(6,8),(9,5),(8,7),(10,8)}

$$w^* = S_W^{-1}(\mu_1 - \mu_2) = \begin{pmatrix} 3.3 & -0.3 \\ -0.3 & 5.5 \end{pmatrix}^{-1} \left[\begin{pmatrix} 3 \\ 3.8 \end{pmatrix} - \begin{pmatrix} 8.4 \\ 7.6 \end{pmatrix} \right]$$

$$= \begin{pmatrix} 0.3045 & 0.0166 \\ 0.0166 & 0.1827 \end{pmatrix} \begin{pmatrix} -5.4 \\ -3.8 \end{pmatrix}$$

$$= \begin{pmatrix} 0.9088 \\ 0.4173 \end{pmatrix}$$