Notiztitel 17.10.2008

14.05.2014

Eigenvalue problem written in matrix-vector notation:

$$U = \underbrace{[\underline{e}_1, \underline{e}_2, \ldots \underline{e}_N]}_{\text{Eigenvectors of matrix } \underline{G}}^T$$

$$\Lambda = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{bmatrix}$$

Eigenvalues

Reformulate Eigenvector equation:

$$C_x \cdot U^T = U^T \cdot \Lambda \quad \Big| \; U$$

$$U C_x \cdot U^T = \underbrace{U \cdot U^T}_{= I} \cdot \Lambda$$

$$U \cdot C_x \cdot U^T = \Lambda$$

Transformation formula:

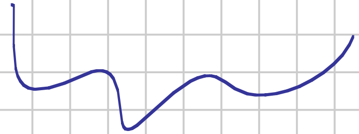
$$\underline{y} = U \cdot (\underline{x} - \underline{m}_x)$$

$$\underline{y} = \begin{array}{l} \\ \\ M \\ \\ \end{array} \begin{bmatrix} \underline{e}_1^T \\ \underline{e}_2^T \\ \vdots \\ \underline{e}_N^T \end{bmatrix} \cdot (\underline{x} - \underline{m}_x)$$

Statistical properties of new vectors $\underline{y}$

$$\underline{m}_y = E\{\underline{y}\} = U \cdot E\{\underline{x}\} - U \cdot \underline{m}_x$$

$= \underline{0}$



m_x

New covariance matrix:

$$C_y = E\{\underline{y} \cdot \underline{y}^T\} - \underline{m}_y \cdot \underline{m}_y^T$$

$$= U \cdot \underbrace{E\{(\underline{x} - \underline{m}_x) \cdot (\underline{x} - \underline{m}_x)^T\}}_{C_x} U^T - \underline{0}$$

$$= U \cdot C_x \cdot U^T = \Lambda$$

$\Rightarrow$ Vectors $\underline{y}$ are uncorrelated in their components

⇒ Major applications of PCA :

1) Dimensionality reduction

Reverse transformation :

PCA $\underline{y} = U[\underline{x} - \underline{m}_x] \;\Big|\; U^T$

$$U^T \cdot \underline{y} = \underbrace{U^T \cdot U}_{\underline{I}}[\underline{x} - \underline{m}_x]$$

$$\Rightarrow \underline{x} = U^T \cdot \underline{y} + \underline{m}_x$$

$$= \left[\underline{e}_1, \underline{e}_2, \cdots \underline{e}_M \right] \cdot \underline{y} + \underline{m}_x$$

$$= \sum_{i=1}^{N} \underline{e}_i \cdot y_i + \underline{m}_x$$

Now, consider a reduced matrix

U, consisting only of M (instead of $N \geq M$) Eigenvectors

$\Rightarrow$ the dimension of vectors y will reduced from N to M

$\Rightarrow$ the reconstructed vector $\underline{\tilde{x}}$ will still have the dimension $N > M$

$\Rightarrow$ Error resulting, between the reconstructed vectors $\underline{\tilde{x}}$ and the original vectors $\underline{x}$

mean square error can be expressed as

$$\varepsilon = \sum_{n=1}^{N} \lambda_n - \sum_{m=1}^{M} \lambda_m$$

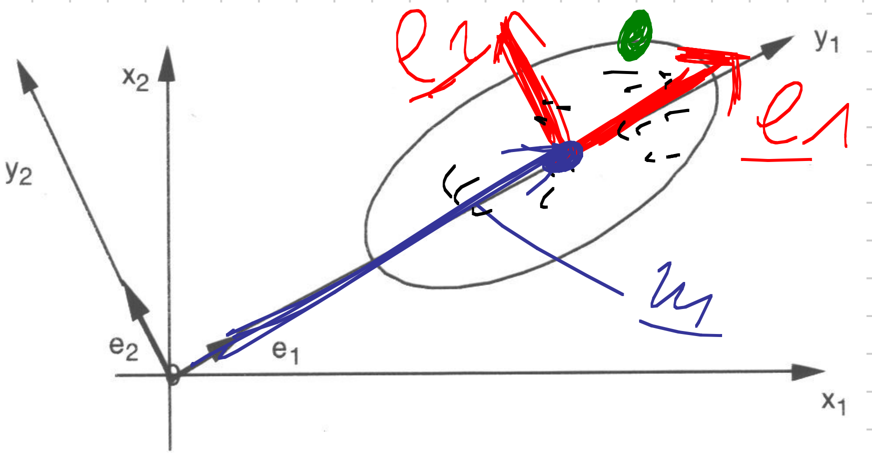
$$= \sum_{n=M+1}^{N} \lambda_n$$

if $M = N \longrightarrow \epsilon = 0$

else: create a minimal error by sorting the Eigenvectors in descending size of their Eigenvalues

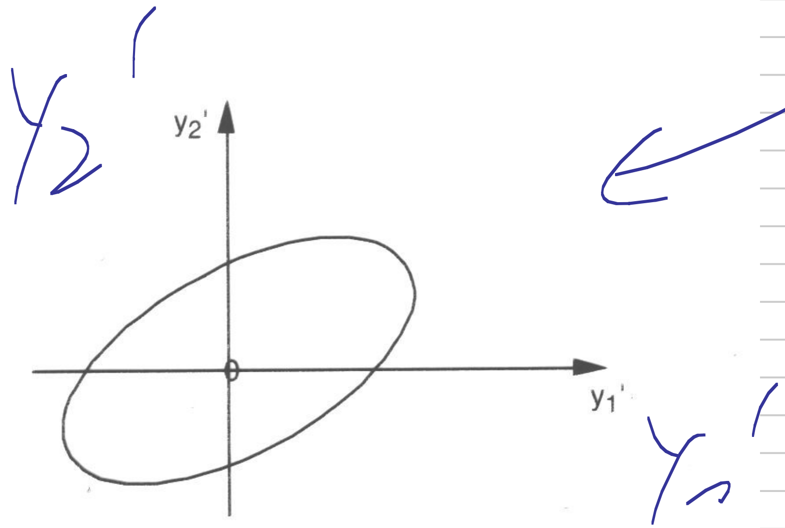
2. Application:

Pattern normalization

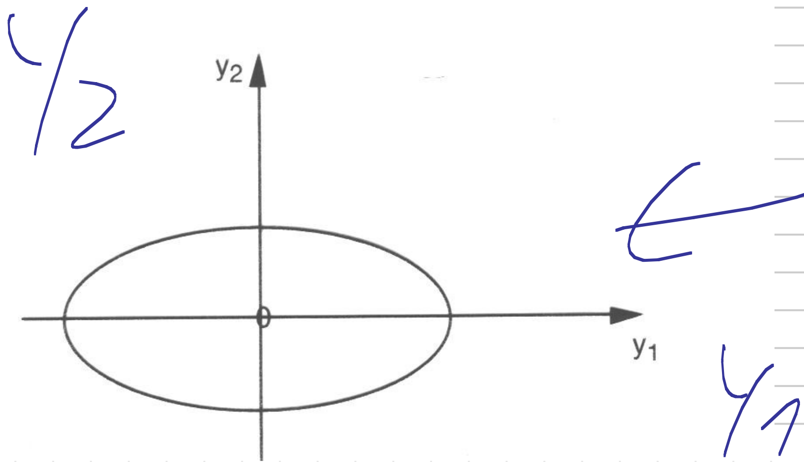


$$\underline{y} = U \cdot \underbrace{[\underline{x} - m_x]}_{\underline{y}'}$$

$$\underline{y}' = \underline{x} - \underline{u}_1$$



$$\underline{y} = U \cdot \underline{y}'$$



Reconsider reconstruction formula:

$$\underline{x} = \sum_{i=1}^{N} \underline{e}_i \cdot y_i + \underline{m}$$

for $N = 2$

$$\underline{x} = \underline{m} + y_1 \cdot \underline{e}_1 + y_2 \cdot \underline{e}_2$$

2.3.2 Linear Discriminant Analysis (LDA)

Consider again a collection of vectors $\underline{x}(k)$, $k = 1, \dots, K$

where the class information of these vectors is known

Consider N different classes

$\underline{x}(0) \quad \underline{x}(1) \quad \underline{x}(2) \quad \cdots \quad \underline{x}(K)$

$\begin{bmatrix} \vdots \end{bmatrix} \quad \begin{bmatrix} \vdots \end{bmatrix} \quad \begin{bmatrix} \vdots \end{bmatrix} \quad \cdots \quad \begin{bmatrix} \vdots \end{bmatrix}$

$K_7^* \quad K_1^* \quad K_3^* \quad \cdots \quad K_8^*$ (memberships)

Compute the following variables:

each class k contain K_k members

$$\sum_{k=1}^{N} K_k = K$$

Class probabilities:

$$p_h = \frac{K_h}{K}$$

Center of gravity of each class K_h

$$\underline{m}_h = \frac{1}{K_h} \sum_{\substack{k=1 \\ \underline{x}(k) \in K_h^*}}^{K_h} \underline{x}(k)$$

- Class covariance:

$$C_n = \frac{1}{K_n} \sum_{\substack{k=1 \\ x(k) \in K_n^*}}^{K_N} \underline{x}(k)\,\underline{x}(k)^T - \underline{m}_n \cdot \underline{m}_n^T$$

- Overall center of all patterns:

$$\underline{m} = E\{\underline{m}_n\} = \sum_{n=1}^{N} p_n \cdot \underline{m}_n$$

- intra class scattering matrix =
(average covariance matrix)

$$C_a = E\{C_n\} = \sum_{n=1}^{N} p_n \cdot C_n$$

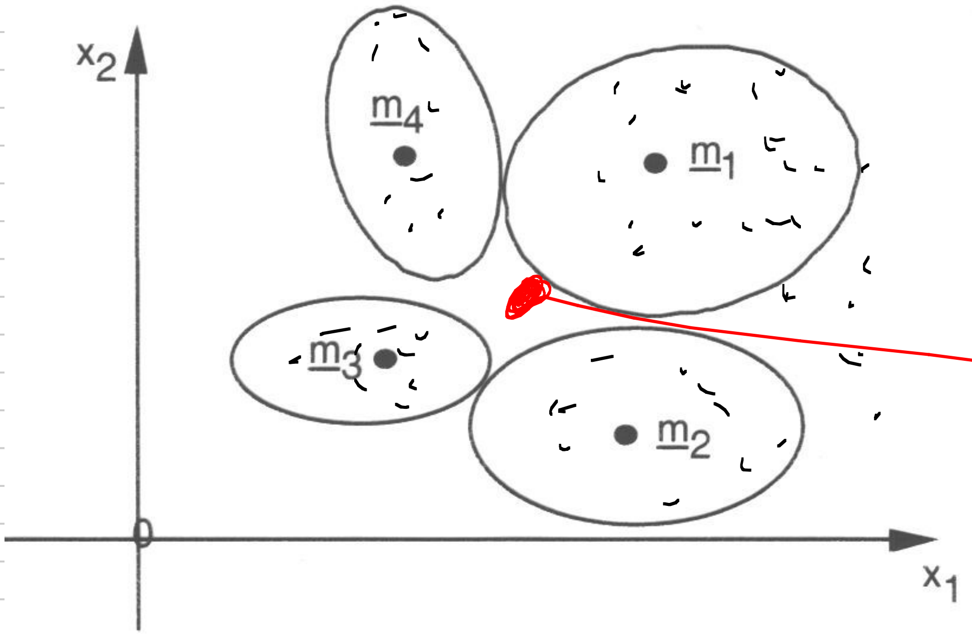
- interclass scattering matrix :

$$C_b = E\left\{(\underline{m}_n - \underline{m})(\underline{m}_n - \underline{m})^T\right\}$$

$$= \sum_{k=1}^{N} p_k \left(\underline{m}_k - \underline{m}\right)\left(\underline{m}_k - \underline{m}\right)^T$$

- Overall covariance matrix of entire data

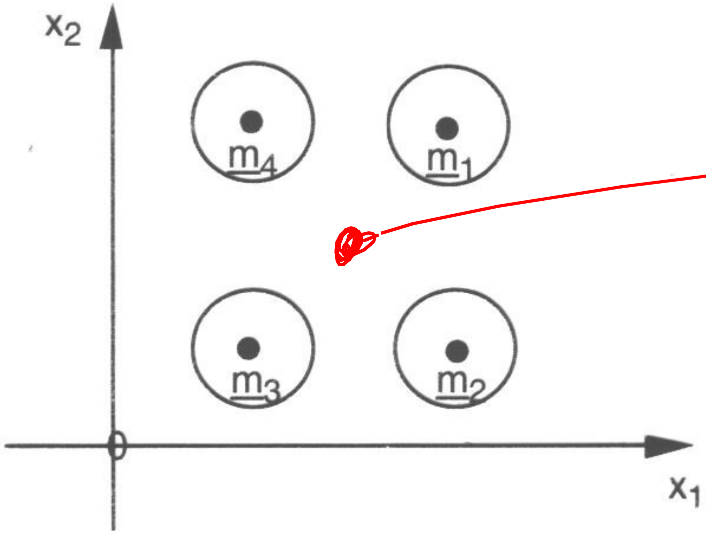
$$C = C_a + C_b$$



$\underline{m}$

Original situation, before LDA

- relatively large C_a
- relatively small C_b



$\underline{m}$

After LDA:

desired:

- small C_a (compact classes

- and additionally large C_b
(spread-out classes)

$\Rightarrow$ no overlap between classes