Notiztitel 21. 5. 2014 17.10.2008

Transformations of LDA

Generally: $\underline{x}' = \Phi \underline{x}$

What is the impact of such a transformation to matrices C_a and C_b?

- class centers:

$$\underline{m}_n' = \Phi \cdot \underline{m}_n$$

- overall center: $\underline{m}' = \Phi \cdot \underline{m}$

- class - covariances:

$$C_n' = \Phi \cdot C_n \cdot \Phi^T$$

Therefore, we can now easily derive the formulas for the matrices C_a'

and C_b':

$$C_a' = E\{C_n'\} = E\{\Phi \cdot C_n \cdot \Phi^T\}$$

$$= \Phi \cdot \underbrace{E\{C_n\}}_{C_a} \cdot \Phi^T = \Phi \cdot C_a \cdot \Phi^T$$

$$C_b' = E\{(\underline{m}_n' - \underline{m}') \cdot (\underline{m}_n' - \underline{m}')^T\}$$

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

$$= \Phi\,\underbrace{E\left\{\left(\underline{m}_n - \underline{m}\right)\cdot\left(\underline{m}_n - \underline{m}\right)^T\right\}}_{C_b}\,\Phi^T$$

$$\Rightarrow \quad C_b' = \Phi\cdot C_b\cdot\Phi^T$$

For the first transformation, we want to achieve the fact that the new matrix C_a' should be transformed into the identity matrix I.

Because for orthogonal transformations, we have:

$$\Phi \cdot I \cdot \Phi^T = \Phi \cdot \Phi^T$$
$$= \Phi \cdot \Phi^{-1} = I$$

Such a transformation can be realized with the following transformation matrix:

$$\Phi_a = \Lambda_a^{-1/2} \cdot U_a$$

Contains the Eigenvalues of Ca in diagonal form

Contains the Eigenvectors of Ca

$\Rightarrow$ Now, the new matrix C_a' will be:

$$C_a' = \Phi_a \, C_a \cdot \Phi_a^T$$

$$= \left(\Lambda_a^{-1/2} \cdot U_a\right) \cdot C_a \left(\Lambda_a^{-1/2} \cdot U_a\right)^T$$

$$= \Lambda_a^{-1/2} \cdot U_a \cdot C_a \cdot U_a^T \cdot \left(\Lambda_a^{-1/2}\right)^T$$

$$= \Lambda_a^{-1/2} \cdot \Lambda_a \cdot \left(\Lambda_a^{-1/2}\right)^T = I$$

Now, we look for the 2nd transformation:

1) $\underline{x}' = \Phi_a \cdot \underline{x}$

2) $\underline{x}'' = \Phi_b \cdot \underline{x}'$

With $\Phi_b = U_b$

With U_b as matrix of the Eigen-vectors derived from matrix C_b'.

For the matrix C_b'' we get then:

$$C_b'' = \Phi_b \cdot C_b' \cdot \Phi_b^T$$

$$= U_b \cdot C_b' \cdot U_b^T = \Lambda_b'$$

matrix of Eigenvalues
of matrix C_b!

Summarization of LDA procedure:

1) Calculate matrices C_a and C_b from pattern vector $\underline{x}(k)$, $k=1, \dots, K$

2) Calculate the Eigenvalues and

Eigenvectors from matrix C_a.
Calculate the transformation matrix $\Phi_a = \Lambda_a^{-1/2} \cdot U_a$

3) Calculate matrix C_b' according

$$C_b' = \Phi_a \cdot C_b \cdot \Phi_a^T$$

4) Calculate Eigenvectors of C_b',

build matrix U_b' from these Eigenvectors, and set $\Phi_b = U_b'$

5) Calculate the overall transformation matrix:

$$\Phi = \Phi_b \cdot \Phi_a$$

$\Rightarrow$ Now, all vectors can be

transformed according to

$$\underline{x}''(k) = \Phi \cdot \underline{x}(k)$$

resulting into a new vector space, with "better" vectors $\underline{x}''$, having the following features:

- average covariance matrix: $C_a'' = \underline{I}$

- Scattering matrix $\underline{\underline{C}}_b' = \underline{\underline{\Lambda}}_b'$

3. Pattern Classification

major notations :

$\underline{X}(k)$: original vectors

k : running index

$\underline{y}(k)$: pre-processed vectors (optimized, dimension N)

$$\Omega = \{\Omega_1, \Omega_2, \dots, \Omega_M\}$$

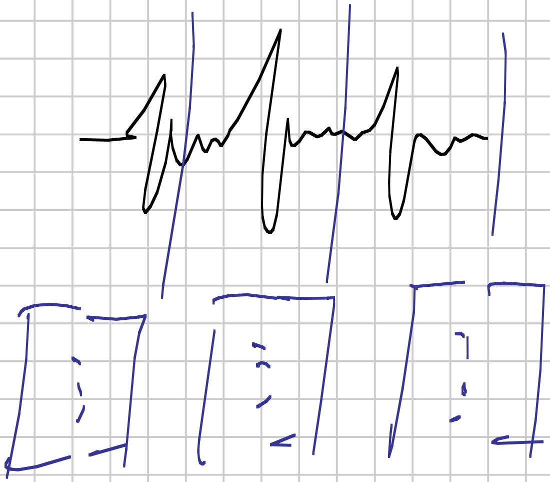
$\Rightarrow$ M different pattern classes

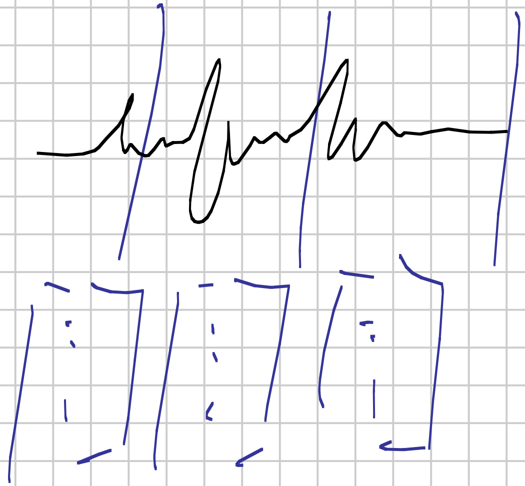
Distinction between numerical and syntactical classification:

Example: Speech recognition

Word 1

Word 2





numerical

Comparison

/a/s/o/

syntactical

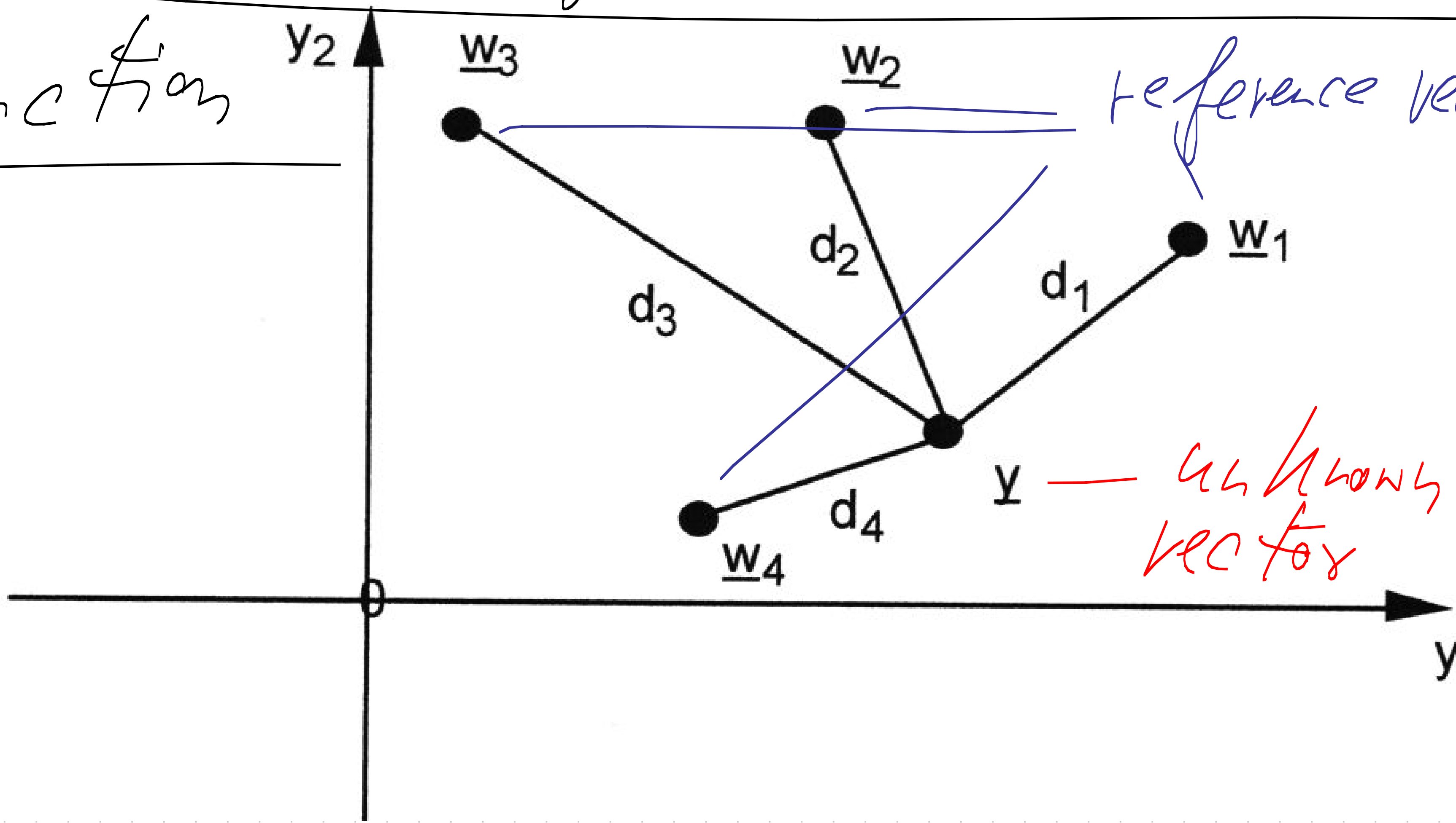
(symbolic)

Comparison

/a/f/a/

3.2 Numerical classification methods

3.2.1 Classification with distance function



$\Rightarrow$ find similarity measures between $\underline{y}$ and reference vectors $\underline{w}_m$

3.2.1.1 Distance measures

$\Rightarrow$ most insightful approach: Euclidean distance between $\underline{y}$ and $\underline{w}_m$

$$d(\underline{x}, \underline{y}) = \sqrt{\sum_{i=1}^{N} (x_i - y_i)^2}$$

or an equivalent distance measure

$$d(\underline{x}, \underline{y}) = \sum_{i=1}^{N} (x_i - y_i)^2$$

$$= (\underline{x} - \underline{y})^T \cdot (\underline{x} - \underline{y})$$

possible alternatives:

$$d(\underline{x},\underline{y}) = \sum_{i=1}^{N} |x_i - y_i|$$

(Manhattan distance)
(city block distance)

another possibility:

$$d(\underline{x},\underline{y}) = \sum_{i=1}^{N} c_i (x_i - y_i)^2$$