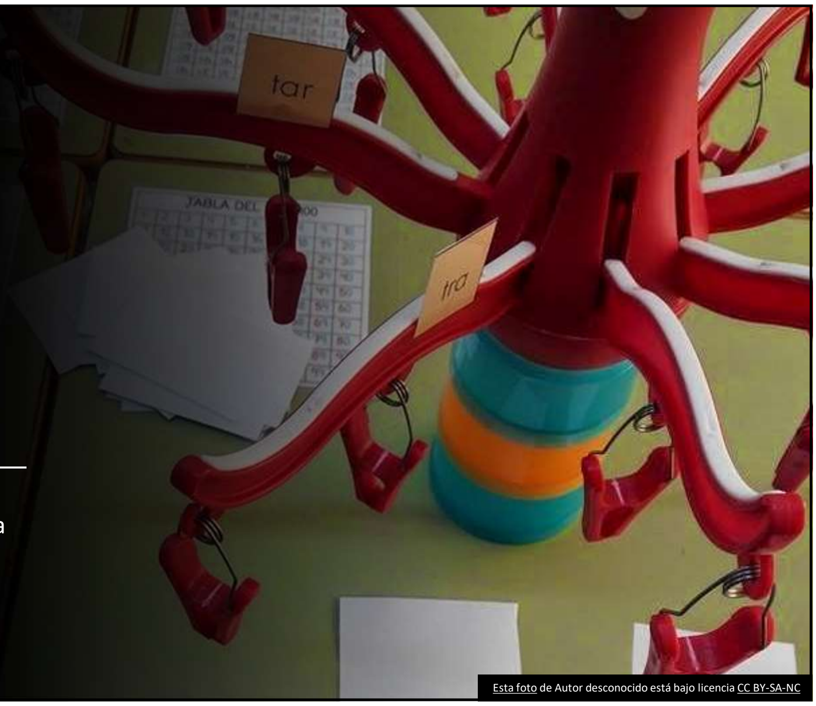
ESCOM-IPN
Pattern Recognition

Ej. Clasificador paramétrico Bayesiano

(Diseño de un clasificador, para una FD para clases con Distribución Gaussiana o ***dM***)

M. En C. María Elena Cruz Meza
mcruzm@ipn.mx



1

Algoritmo:

1.*Se elige una muestra de patrones clasificada de antemano con n clases* {c_1, c_2,..., c_n} *y la métrica* d_M (distancia de Mahalanobis).

2.*Con base en la muestra y para cada clase* C_i *estimar los parámetros: Media para cada clase* {m_i , m_j }*, así como las matrices de covarianza de cada clase* {C_i , C_j}

Media: $m_i = \frac{1}{p}\sum_{k=1}^{p}(x_{ij})$ Matriz de covarianzas: $C_k = \frac{1}{p-1}\sum_{k=1}^{p}(x_i - m_i)(x_i - m_i)^t$

Siempre i es distinto de j, y m_i *es el vector medio o representante de la clase*

3.*Generar funciones discriminantes* $d_{M(i,j)}(x)$ *para cada par de clases* c_i *y* c_j *de forma que:*

$$d_M^2(x_k, m_i) = (x_k - m_i)C_i^{-1}(x_k - m_i)^t$$

4.*En el momento de clasificar (recuperación), el patrón x será clasificado en la clase* c_i *si cumple lo siguiente:*

$$x \in c_i \; siid_M^2(x_k, m_i) < d_M^2(x_k, m_j), \; \forall i \neq j, \; i,j = 1,2,\dots c$$

2

Ej. Suponiendo los siguientes datos:

1. Sea un sistema biclase {c1, c2}, *k* patrones, distribuidos en un espacio bidimensional, tal como se muestra en el siguiente esquema:

C_1	$x_{1,1}$	1	3	1	2	3
	$x_{1,2}$	2	3	5	2	3

C_2	$x_{2,1}$	6	6	7	8	8
	$x_{2,2}$	4	3	4	4	5

a) Calcular la separabilidad entre las clases
b) Diseñar un clasificador basado en la información a priori del problema.
b) Dado el patrón desconocido x=(4,5), probar su pertenencia a una de las dos clases, conforme la fd diseñada en a).

3

Resolviendo...

Fase de Aprendizaje

Dado de que se tiene el conocimiento previo de ambas clases la cardinalidad de los patrones y sus rasgos x1 y x2). Se pueden estimar los vectores media y las matrices de covarianza correspondientes a cada una de las clases como sigue:

Para cada c_k:, siendo *k* el número de patrones pertenecientes a cada clase c_k, desde *k=1* hasta *p* patrones

- Media: $m_i = \frac{1}{p} \sum_{k=1}^{p} (x_{ij})$

- Matriz de Covarianza: $C_k = \frac{1}{p-1} \sum_{k=1}^{p} (x_i - m_i)(x_i - m_i)^t$

4

Solución:

2. Estimando los vectores media y las matrices de covarianza correspondientes a cada una de las clases como sigue:

Para C_1:

$$m_1 = \frac{1}{5}\sum_{k=1}^{5}(x_{ij}) = \frac{1}{5}\begin{bmatrix}10\\15\end{bmatrix} = \begin{bmatrix}2\\3\end{bmatrix}$$

$$C_1 = \frac{1}{5-1}(x+a)^n = \frac{1}{4}\left[\begin{pmatrix}-1\\-1\end{pmatrix}\begin{pmatrix}-1 & -1\end{pmatrix} + \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1 & 0\end{pmatrix} + \begin{pmatrix}-1\\2\end{pmatrix}\begin{pmatrix}-1 & 2\end{pmatrix} + \begin{pmatrix}0\\-1\end{pmatrix}\begin{pmatrix}0 & -1\end{pmatrix} + \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1 & 0\end{pmatrix}\right]$$

$$C_1 = \frac{1}{4}\left[\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix} + \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix} + \begin{pmatrix}1 & -2\\-2 & 4\end{pmatrix} + \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix} + \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}\right]$$

$$C_1 = \frac{1}{4}\begin{pmatrix}4 & -1\\-1 & 6\end{pmatrix}$$

$$C_1 = \begin{pmatrix}1.0 & -0.25\\-0.25 & 1.5\end{pmatrix}$$

5

2. Estimando los vectores media y las matrices de covarianza correspondiente para la segunda de las clases, como sigue:

Para C_2:

Media:

$$m_1 = \frac{1}{5}\sum_{k=1}^{5}(x_{ij}) = \frac{1}{5}\begin{bmatrix}35\\20\end{bmatrix} = \begin{bmatrix}7\\4\end{bmatrix}$$

Matriz de Covarianza

$$C_2 = \frac{1}{5-1}(x+a)^n = \frac{1}{4}\left[\begin{pmatrix}-1\\0\end{pmatrix}\begin{pmatrix}-1 & 0\end{pmatrix} + \begin{pmatrix}-1\\-1\end{pmatrix}\begin{pmatrix}-1 & -1\end{pmatrix} + \begin{pmatrix}0\\0\end{pmatrix}\begin{pmatrix}0 & 0\end{pmatrix} + \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1 & 0\end{pmatrix} + \begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}1 & 1\end{pmatrix}\right]$$

$$C_2 = \frac{1}{4}\left[\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix} + \begin{pmatrix}1 & 1\\1 & 1\end{pmatrix} + \begin{pmatrix}0 & 0\\0 & 0\end{pmatrix} + \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix} + \begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}\right]$$

$$C_2 = \frac{1}{4}\begin{pmatrix}4 & 2\\2 & 2\end{pmatrix}$$

$$C_2 = \begin{pmatrix}1.0 & 0.5\\0.5 & 0.5\end{pmatrix}$$

6

3. Dado lo anterior, podemos diseñar un clasificador bayesiano sabiendo que un patrón desconocido x_k, puede clasificarse de acuerdo con la forma:

$$p(y=c_i/x) = \frac{p(x\ /\ y=c_j)P(y=c_j)}{p(x)}$$

Donde:

$$p(x) = \sum_{j=1}^{c} p\,(x\ /y\ =\ c_j)P(y\ =\ c_j)$$

Y por la regla de decisión establecida como:

$$x\ \in C_i\ sii\ \ p(y\ = c_j\ /\ x) > \ \ p(y\ = c_j\ /x), \forall\ i\ \neq j, i, j = 1, 2, \dots, c$$

Notemos que p(x) no aporta nada a la decisión, por lo que optaremos por una forma alternativa para clasificar al vector:

$$x\ \in C_i\ sii\ \ p(\,x\ /\ y\ = c_i)p(y\ = c_i) > p(\,x\ /\ y\ = c_j)p(y\ = c_j), \forall\ i\ \neq j, i, j = 1, 2, \dots, c$$

7

3. Diseñando el clasificador...

Normalmente las distribuciones de densidad de probabilidad se eligen Normales o Gaussianas, así, debemos calcular la distancias a cada clase, en este caso, dado que las matrices de covarianzas son distintas, podemos diseñarlo usando la distancia de Mahalanobis, de la forma:

$$d_M^2\,(x_k\ -\ m_i) = (x_k\ -\ m_i)\,C_2^{-1}\,(x_k\ -\ m_i)^t$$

$$x\ \in C_i\ sii\ \ d_M^2\,(x_k\ -\ m_i) < d_M^2\,(x_k\ -\ m_j)\,, \qquad \forall\ i\ \neq j, i, j = 1, 2, \dots, c$$

Con *{m_i, m_j}* y *{C_i, C_j}* que corresponden a la Media y la Matriz de Covarianza estimada para cada clase, p*ara toda i distinta de j*

8

Fase de Aprendizaje

a) La separabilidad entre las dos clases, la Podemos obtener a partir de (10.28), (10.29) y (10.31) [concultar el capitulo 10 d ePajares y que fue copartido], obtenemos la separabilidad entre ambas clases:

Diver$_{1,2}$:

$$Diverg_{12} = 27.978 \qquad \cos\alpha_{12} = 0 \qquad J_{12} = 1.986$$

Dado que solo son dos clases, la separabilidad entre ellas no aporta gran información, para conoce esto es necesaria una relación entre más de dos clases.

9

Fase de Aprendizaje

4. Probando el clasificador para *xk*=(4,5), debemos obtener las distancias a las dos clases:

Para C1:

$$d_M^2\,(x_k - m_1) = (x_k - m_1)\,C_1^{-1}\,(x_k - m_1)^t$$

$$d_M^2\,(x_k - m_1) = (2\,,2)\begin{pmatrix}1.0435 & 0.1739\\ 0.139 & 0.6957\end{pmatrix}\begin{pmatrix}2\\2\end{pmatrix} = 8.35$$

Para C2:

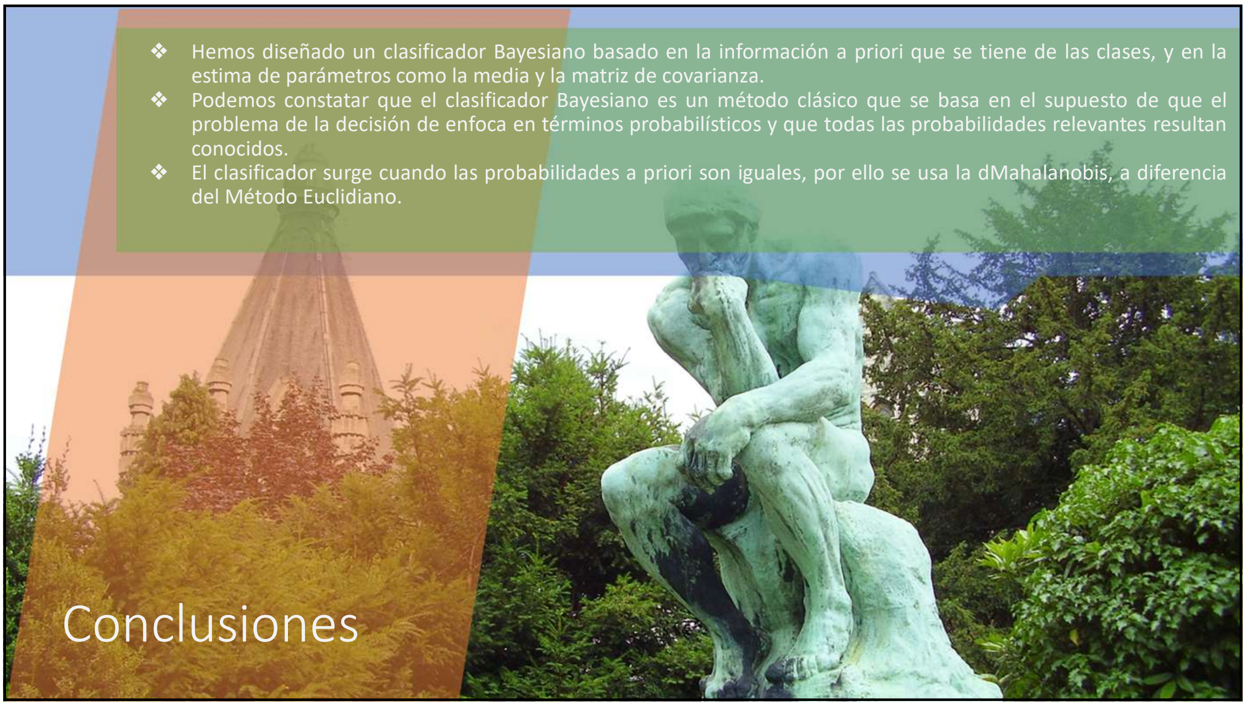
$$d_M^2\,(x_k - m_2) = (x_k - m_2)\,C_2^{-1}\,(x_k - m_2)^t$$

$$d_M^2\,(x_k - m_1) = (-3,1)\begin{pmatrix}2.0 & -2.0\\ 2.0 & 4.0\end{pmatrix}\begin{pmatrix}-3\\1\end{pmatrix} = 34.00$$

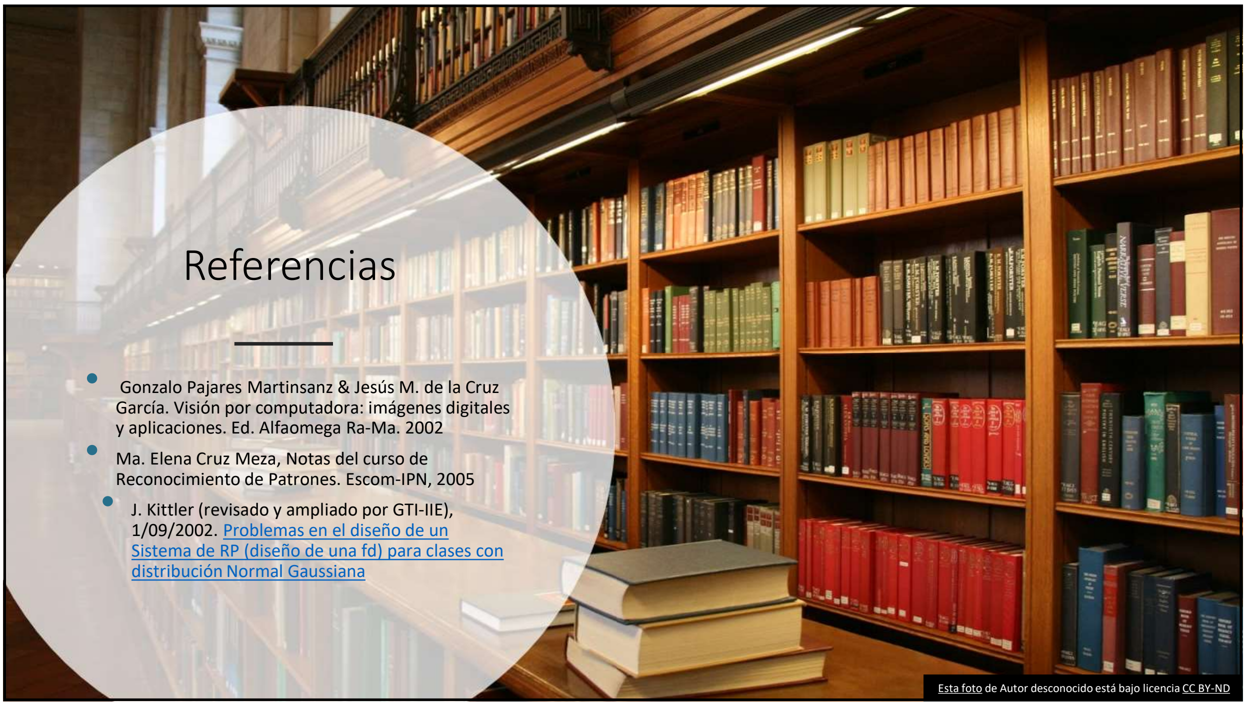
Concluimos que:

$$x \in C_1 dado\ que\ d_M^2\,(x_k - m_1) < d_M^2\,(x_k - m_1)$$

10


❖ Hemos diseñado un clasificador Bayesiano basado en la información a priori que se tiene de las clases, y en la estima de parámetros como la media y la matriz de covarianza.
❖ Podemos constatar que el clasificador Bayesiano es un método clásico que se basa en el supuesto de que el problema de la decisión de enfoca en términos probabilísticos y que todas las probabilidades relevantes resultan conocidos.
❖ El clasificador surge cuando las probabilidades a priori son iguales, por ello se usa la dMahalanobis, a diferencia del Método Euclidiano.
Conclusiones

11


Referencias
• Gonzalo Pajares Martinsanz & Jesús M. de la Cruz García. Visión por computadora: imágenes digitales y aplicaciones. Ed. Alfaomega Ra-Ma. 2002
• Ma. Elena Cruz Meza, Notas del curso de Reconocimiento de Patrones. Escom-IPN, 2005
• J. Kittler (revisado y ampliado por GTI-IIE), 1/09/2002. Problemas en el diseño de un Sistema de RP (diseño de una fd) para clases con distribución Normal Gaussiana
Esta foto de Autor desconocido está bajo licencia CC BY-ND

12