UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Funciones discriminantes

Jorge Civera
Alfons Juan
Albert Sanchis

DSIC
Departamento de Sistemas
Informáticos y Computación

Objetivos formativos

- Aplicar funciones discriminantes
- Calcular la frontera de decisión entre dos clases
- Identificar el tipo de frontera de decisión
- Calcular las regiones de decisión de un clasificador
- Obtener e identificar clasificadores equivalentes

Índice

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

1. Introducción

Un ***clasificador*** es una función definida como:

$$c(x) = \arg\max_{c} \; g_c(x)$$

donde para cada clase c se define su función discriminante g_c.

El grado de pertenencia del objeto $\boldsymbol{x}$ a la clase c es $g_c(\boldsymbol{x})$.

$c(\boldsymbol{x})$ es la clase a la que el objeto $\boldsymbol{x}$ pertenece en mayor grado.

Introducción

Ejemplo: Un clasificador en 3 clases para $\boldsymbol{x} = (x_1, x_2)^t \in \{0, 1\}^2$:

x_1	x_2	$g_1(\boldsymbol{x})$	$g_2(\boldsymbol{x})$	$g_3(\boldsymbol{x})$	$c(\boldsymbol{x})$
0	0	1.0	0.0	0.0	1
0	1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1
1	0	0.25	0.5	0.25	2
1	1	0.01	0.01	0.98	3

El clasificador de Bayes se obtiene como $g_c(x) = p(c \mid x)$:

$$c(x) = \arg\max_c \; p(c \mid x)$$

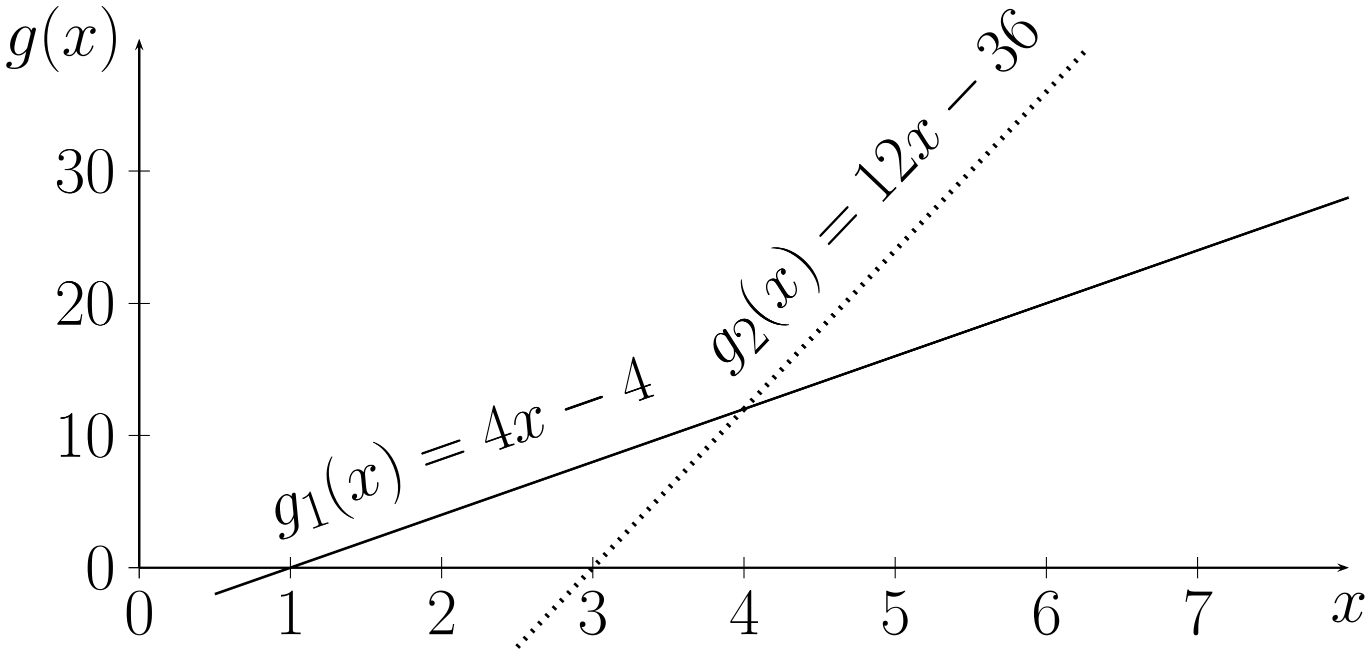
2. Clasificadores lineales

Un ***clasificador lineal*** se define en términos de f.d. lineales:

$$g_c(\boldsymbol{x}) = \sum_d w_{cd}\, x_d + w_{c0} = \boldsymbol{w}_c^t \boldsymbol{x} + w_{c0}$$

donde $\boldsymbol{w}_c$ es el vector de pesos de la clase c y w_{c0}, el peso umbral.

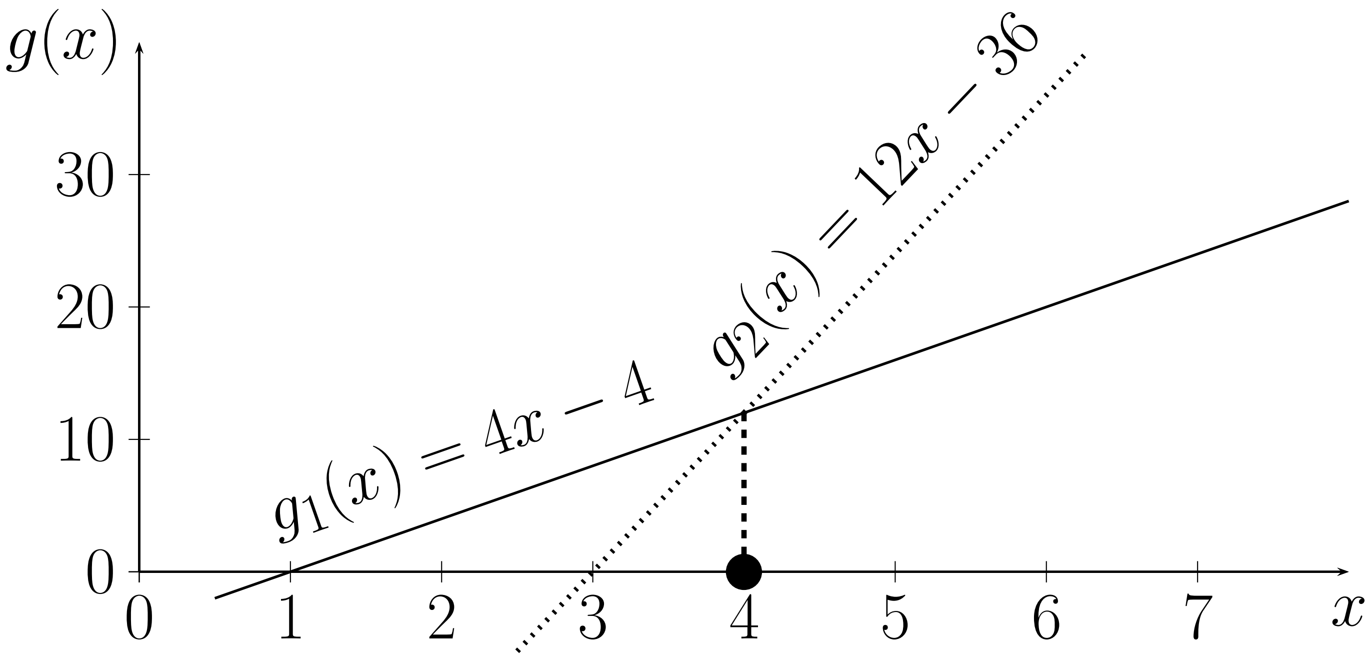
Ejemplo: Clasificador lineal en 2 clases para x unidimensional



3. Fronteras de decisión

La ***frontera de decisión*** entre dos clases i, j es el lugar geométrico de los puntos $\mathbf{x} \in E$ donde se cumple:

$$g_i(\mathbf{x}) = g_j(\mathbf{x}) \qquad i \neq j, \;\; 1 \leq i, j \leq C$$



$$g_1(x) = g_2(x) \rightarrow 4x - 4 = 12x - 36 \rightarrow x = \frac{32}{8} = 4$$

Fronteras de decisión

La ***frontera de decisión*** entre dos clases i, j con $\mathbf{x} \in E$ es:

- Un punto, si $E \equiv \mathbb{R}$
- Una línea (ej. *rectas*), si $E \equiv \mathbb{R}^2$
- Una superficie (ej. *planos*), si $E \equiv \mathbb{R}^3$

En general son *hipersuperficies* definidas por las ecuaciones:

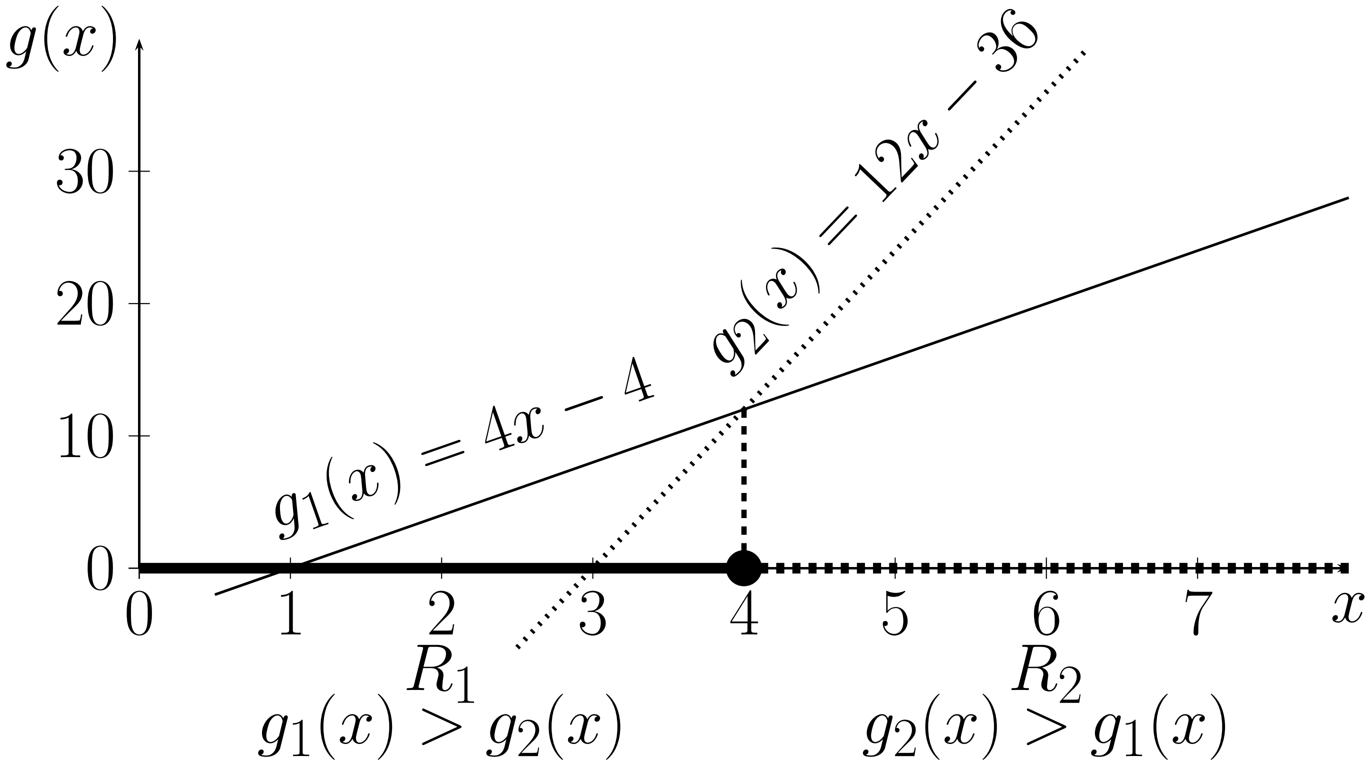
$$g_i(\mathbf{x}) - g_j(\mathbf{x}) = 0 \qquad i \neq j, \;\; 1 \leq i, j \leq C$$

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

4. Regiones de decisión

Un clasificador en C clases divide el espacio de representación de x en C ***regiones de decisión***, $R_1, \dots, R_C$:

$$R_j = \{\mathbf{x} \in E : \quad g_j(\mathbf{x}) > g_i(\mathbf{x}) \quad i \neq j,\ 1 \leq i \leq C\}$$



5. Clasificadores equivalentes

Dos ***clasificadores*** $(g_1, \ldots, g_C)$ y $(g'_1, \ldots, g'_C)$ son ***equivalentes*** si definen las mismas fronteras y regiones de decisión, es decir:

$$g_i(\mathbf{x}) > g_j(\mathbf{x}) \Leftrightarrow g'_i(\mathbf{x}) > g'_j(\mathbf{x}) \qquad \forall j \neq i, \;\; \forall \mathbf{x} \in E$$

¿Cómo obtener clasificadores equivalentes?

$g'_i(\mathbf{x}) = a \cdot g_i(\mathbf{x}) + b$	con $a > 0$	$1 \leq i \leq C$
$g'_i(\mathbf{x}) = ln\, g_i(\mathbf{x})$	con $g_i(\mathbf{x}) > 0$	$1 \leq i \leq C$

Dado (g_1, g_2) de la traspa anterior, un clasificador equivalente sería:

$$g'_1(x) = x - 1 \;\text{ y }\; g'_2(x) = 3x - 9 \rightarrow x = \frac{8}{2} = 4$$

donde $a = \frac{1}{4}$ y $b = 0$.

6. Conclusiones

Hemos visto:

- La aplicación de funciones discriminantes
- El cálculo de sus fronteras y regiones de decisión asociadas
- La obtención de clasificadores equivalentes