Reconocimiento de patrones

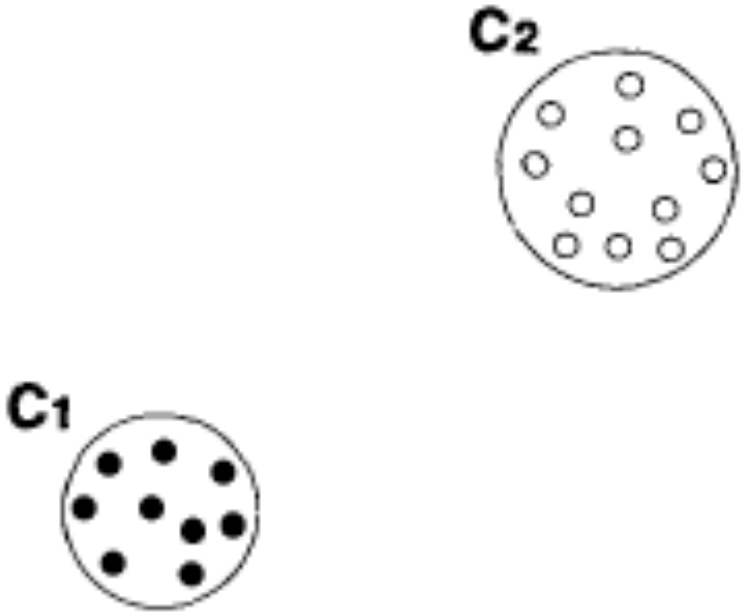
Clase 7: Clasificación por medio de funciones de distancia

Para el día de hoy…

- Introducción
- Clasificación de distancia mínima
    - Prototipo único
    - Múltiples prototipos

Introducción

- Si un patrón es representado por un vector en $\mathbb{R}^n$ el decir que x e y son similares significa que los dos vectores están "cerca"
- La distancia entre ellos es "pequeña"
- Si las clases pueden ser representadas por un prototipo o varios, los métodos basados en distancia pueden ser eficientes



Clasificación de distancia mínima

1

Es uno de los primeros conceptos en reconocimiento de patrones

2

Es una técnica efectiva cuando los patrones tienen limitada variabilidad

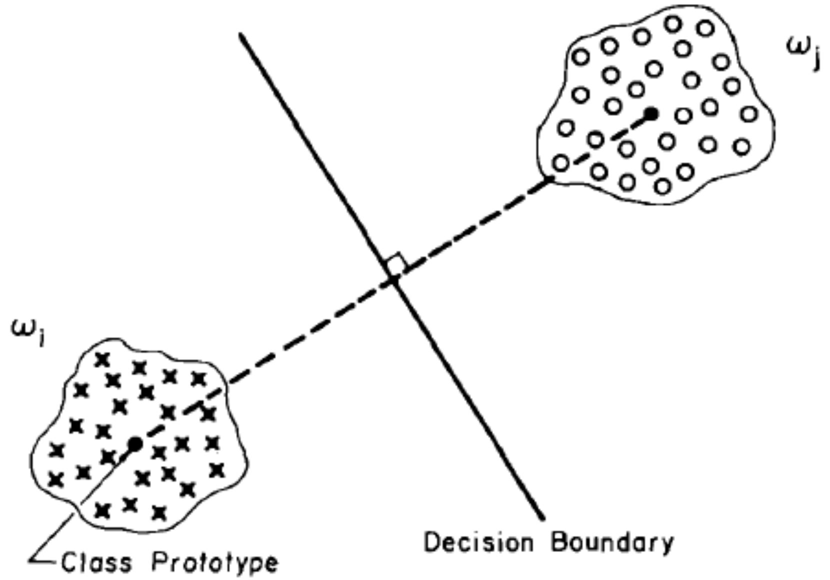
Prototipo único

- Sean $C_1, \dots, C_m$ m clases en $\mathbb{R}^n$ representadas por un prototipo único $y_1, \dots, y_m$ respectivamente.
- Las distancias entre un nuevo patrón y un prototipo es

$$D_i = ||x - y_i|| = \left((x - y_i)^T (x - y_i)\right)^{\frac{1}{2}}, 1 \leq i \leq m$$

- Un clasificador de distancia mínima clasificará a x en C_j tal que

$$D_j = \min ||x - y_j||$$

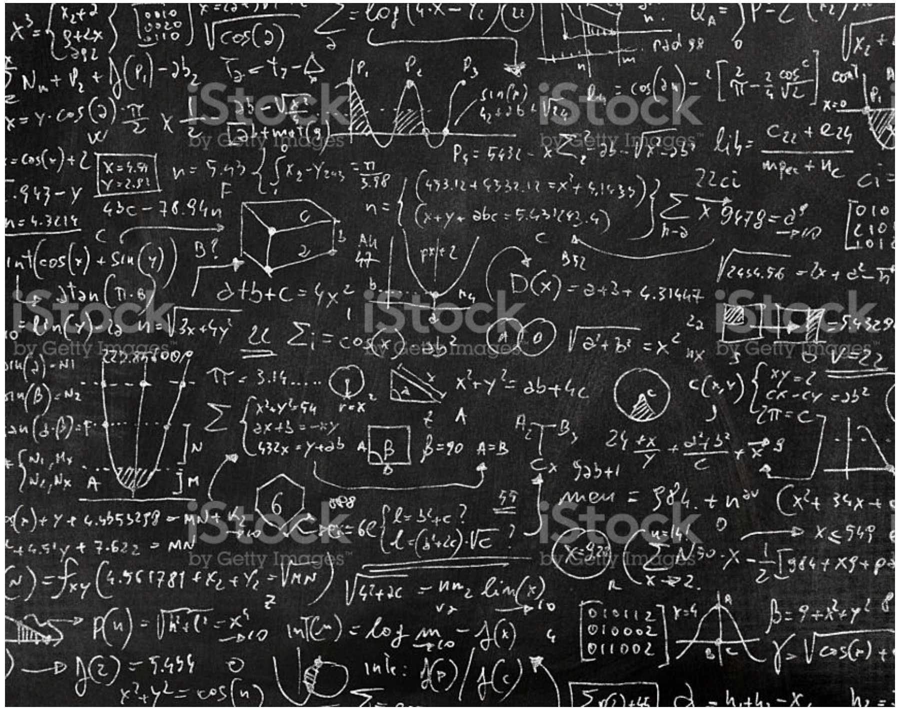


Un poco de manipulación

- Es preferible minimizar D_i^2

  $$D_i^2 = (x - y_i)^T (x - y_i) = x^T x - 2x^T y_i + y_i^T y_i$$

- ¿Por qué?



Un poco de manipulación

- Es preferible minimizar D_i^2

$$D_i^2 = (x - y_i)^T(x - y_i) = x^T x - 2x^T y_i + y_i^T y_i$$

Un poco de manipulación

- Es preferible minimizar D_i^2
$$D_i^2 = (x - y_i)^T (x - y_i) = x^T x - 2x^T y_i + y_i^T y_i$$
- $x^T x$ es constante y lo podemos eliminar. Entonces podemos utilizar
$$2x^T y_i - y_i^T y_i, 1 \leq i \leq m$$

La función de decisión

- Podemos definir la función de decisión como

$$d_i(x) = x^T y_i - \frac{1}{2} y_i^T y_i, 1 \leq i \leq m$$

- $x \in C_i$ si y solo si $d_i(x) > d_j(x), j \neq i$
- La función de decisión lineal

$$d_i(x) = w_i^T x, 1 \leq i \leq m$$

- Donde $x = (1, x_1, \dots x_n)^T, w_i = (w_{i0}, w_{i1}, w_{i2}, \dots, w_{in})^T$ son determinados por $y_i = (y_{i1}, \dots, y_{in})^T, 1 \leq i \leq m$ como

$$w_{ij} = y_{ij}, 1 \leq i \leq m, 1 \leq j \leq n$$

$$w_{i0} = -\frac{1}{2} y_i^T y_i, 1 \leq i \leq m$$

Ejercicio

- Calcule la función de decisión del clasificador de distancia mínima para el caso de dos patrones en dos dimensiones donde solo hay dos clases.

Solución

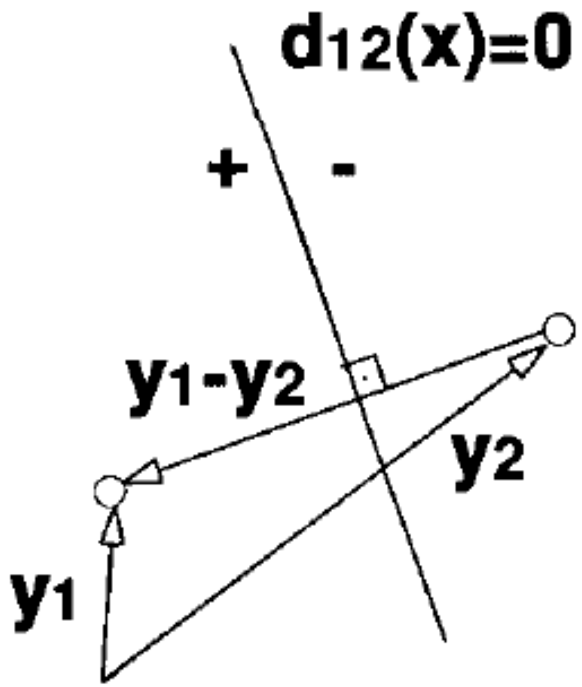
- Calcule la función de decisión del clasificador de distancia mínima para el caso de dos patrones

$$d_{12}(x) = d_1(x) - d_2(x)$$

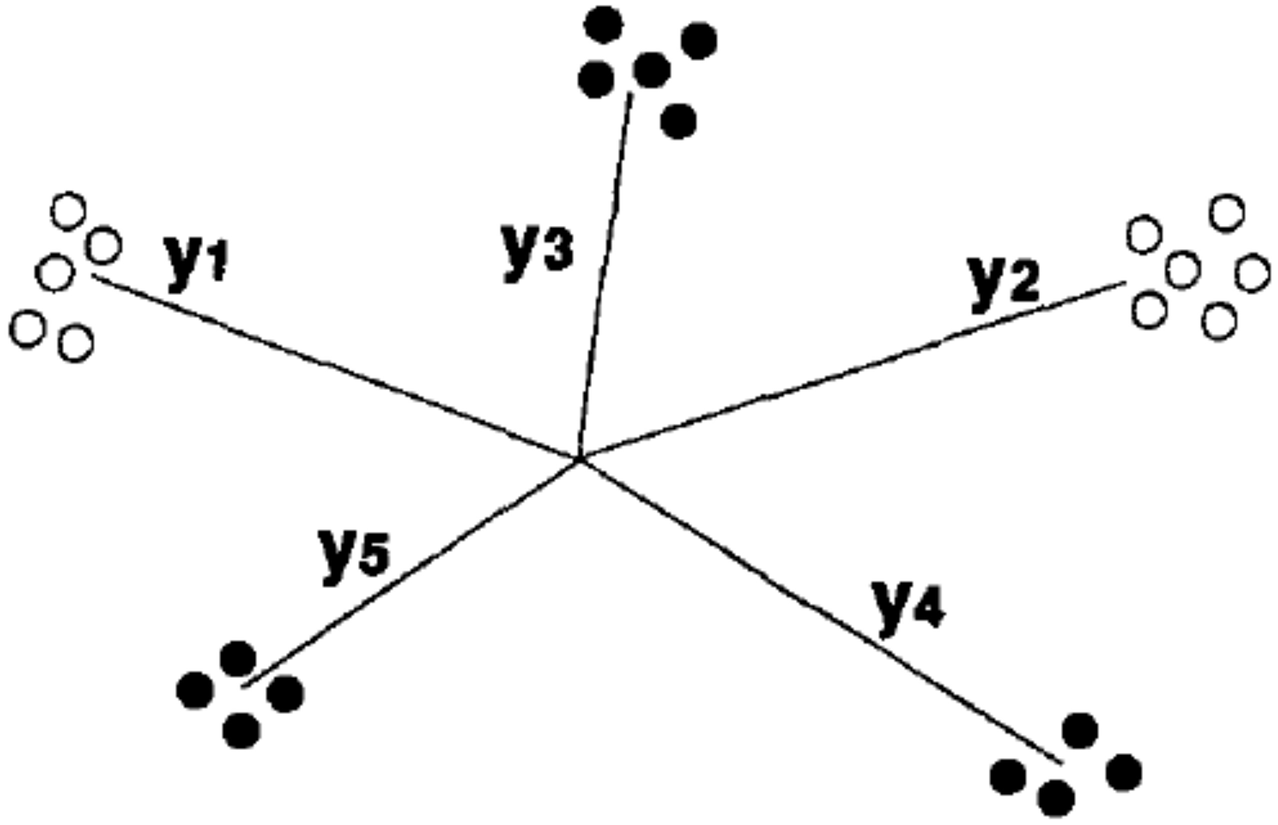
$$= x^T(y_1 - y_2) - \frac{1}{2}y_1^T y_1 + \frac{1}{2}y_2^T y_2 = 0$$

- Sustituyendo $x = \frac{y_1+y_2}{2}$

$$d_{12}(x) = \frac{1}{2}(y_1 + y_2)^T(y_1 - y_2) - \frac{1}{2}y_1^T y_1 + \frac{1}{2}y_2^T y_2 = 0$$



Múltiples prototipos



- Ahora, cada clase tiene varios grupos
- Cada grupo es representado por un prototipo
- Entonces, cada clase puede ser caracterizada por un conjunto finito de prototipos
- Ejemplo
  - $C_1 = \{y_1, y_2\}$
  - $C_2 = \{y_3, y_4, y_5\}$

Clasificador de distancia mínima

- Sean $C_1, \dots, C_m$ varias clases con múltiples prototipos y C_i incluye los prototipos $y_i^{(i)}, \dots, y_i^{(n_i)}$ para $1 \leq i \leq m$
- La distancia de un patrón z a C_i se define como

$$D_i = \min_{1 \leq j \leq n_i} \left\| z - y_i^{(j)} \right\|$$

- El máximo ocurre en $j = j(i, z)$. La función de decisión para z es

$$d_i(z) = z^T \left(y_i^{(j(i,z))} \right) - \frac{1}{2} \left(y_i^{(j(i,z))} \right)^T y_i^{(j(i,z))}, 1 \leq i \leq m$$

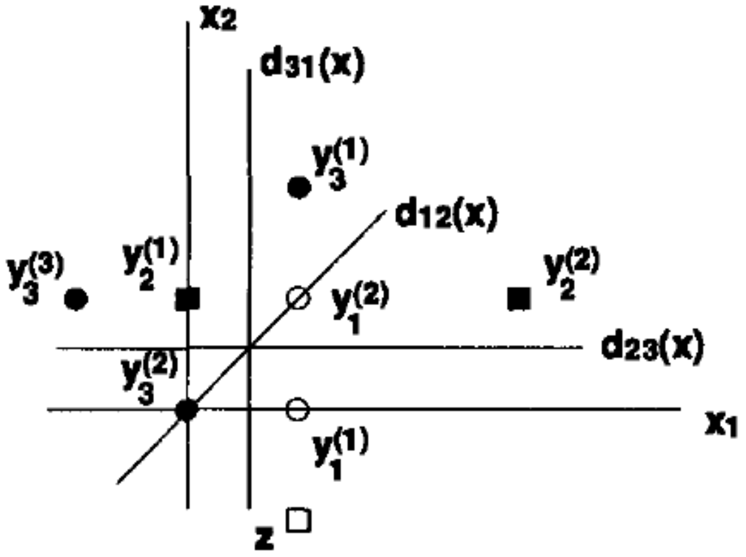
- z es clasificado en C_i si y solo si $d_i(z) > d_j(z)$, para todo $j \neq i$

Ejercicio

- Considere el problema de tres clases en $\mathbb{R}^2$ donde las clases están representadas por:
  - C_1: $(1,0), (1,1)$
  - C_2: $(0,1), (3,1)$
  - C_3: $(1,2), (0,0), (-1, 1)$
- Clasifique el patrón $z = (1, -1)$

Solución

$$j(1,z) = 1 \quad , \quad y_1^{(j(1,z))} = (1,0)^T$$
$$j(2,z) = 1 \quad , \quad y_2^{(j(2,z))} = (0,1)^T$$
$$j(3,z) = 2 \quad , \quad y_3^{(j(3,z))} = (0,0)^T$$



$$d_1(\boldsymbol{x}) = (x_1, x_2)(1,0)^T - (1,0)(1,0)^T = x_1 - \tfrac{1}{2}$$
$$d_2(\boldsymbol{x}) = (x_1, x_2)(0,1)^T - (0,1)(0,1)^T = x_2 - \tfrac{1}{2}$$
$$d_3(\boldsymbol{x}) = (x_1, x_2)(0,0)^T - (0,0)(0,0)^T = 0$$

$$d_{12}(\boldsymbol{x}) = d_1(\boldsymbol{x}) - d_2(\boldsymbol{x}) = x_1 - x_2 = 0$$
$$d_{23}(\boldsymbol{x}) = d_2(\boldsymbol{x}) - d_3(\boldsymbol{x}) = x_2 - \frac{1}{2} = 0$$
$$d_{31}(\boldsymbol{x}) = d_3(\boldsymbol{x}) - d_1(\boldsymbol{x}) = \frac{1}{2} - x_1 = 0$$

$z \in C_1$

Algoritmo MC-MP

- Entrada: n el número de dimensiones, m el número de clases, n_i el número de prototipos, $\{y_i^{(j)}\}$ los prototipos, x una nueva observación
- Salida: k la clase en la cual se clasifica a x
- Paso 1:
  - Para $i = 1, \ldots, m$ encontrar $j(i, x)$ tal que
  - $x^T y_i^{(j(i,x))} - \frac{1}{2}\left(y_i^{(j(i,x))}\right)^T y_i^{(j(i,x))} = \max_{1 \le j \le n_i} [x^T y_i^{(j)} - \frac{1}{2}(y_i^{(j)})^T y_i^{(j)}]$
- Paso 2:
  - Encontrar k que satisfaga
  - $x^T y_k^{(j_k)} - \frac{1}{2}\left(y_k^{(j_k)}\right)^T y_k^{(j_k)} = \max_{1 \le i \le m} [x^T y_i^{(j_i)} - \frac{1}{2}\left(y_i^{(j_i)}\right)^T y_i^{(j_i)}]$ donde $j_i = j(i, x)$

Notas

- ¿Es correcto?
- ¿Complejidad en tiempo?
- ¿Complejidad en espacio?
- ¿Existen regiones indeterminadas?

Clasificador multiprototipo

- Dado un conjunto de clases C_i con un conjunto de prototipos y_j
- Dado un nuevo ejemplo x
  - Calculamos el prototipo más cercano de cada clase
  - Asignamos x a la clase que tenga el prototipo más cercano a x
- ¿Cómo elegimos los prototipos de cada clase?

Para la otra vez…

- Clasificador de vecino más cercano
- Ref
  - Julius T. Tou y Rafael C. Gonzalez
  - Pág. 75-83