Sistemas Inteligentes
Cuestiones y ejercicios del bloque 2, tema 2
Aprendizaje de funciones discriminantes: Perceptrón

Escola Tècnica Superior d'Informàtica
Dep. de Sistemes Informàtics i Computació
Universitat Politècnica de València

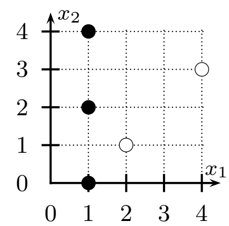
10 de noviembre de 2014

1. Cuestiones

1 ☐ El algoritmo Perceptrón es una técnica de aprendizaje...

A) supervisado de clasificadores lineales.
B) supervisado de clasificadores no-lineales.
C) no-supervisado de clasificadores lineales.
D) no-supervisado de clasificadores no-lineales.

2 ☐ En la figura de la derecha se representan cinco muestras de aprendizaje bidimensionales de 2 clases: $\circ$ y $\bullet$. Se desea contruir un clasificador lineal para cualquier $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, como sigue:



$$c(\mathbf{x}) = \begin{cases} \circ & \text{si } \mathbf{w}^t\,\mathbf{x} > 0 \\ \bullet & \text{si } \mathbf{w}^t\,\mathbf{x} \leq 0 \end{cases} \quad \text{donde} \quad \mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \text{ es un vector de pesos a escoger}$$

Si nuestro criterio de aprendizaje es la minimización del número de errores de clasificación (sobre las muestras de aprendizaje), elegiremos...

A) $\mathbf{w} = (1, 0)^t$
B) $\mathbf{w} = (1, 1)^t$
C) $\mathbf{w} = (1, -1)^t$
D) Ninguna de las anteriores, ya que hay otros vectores de pesos que producen menos errors sobre las muestras dadas.

3 ☐ En el algoritmo Perceptrón:

A) principalmente existen dos parámetros: número de clases y número de prototipos.
B) la constante de aprendizaje tiene que ser lo más alta posible, para que aprenda más.
C) el margen tiene que ser cero cuando las clases no son linealmente separables.
D) principalmente existen dos parámetros: la constante de aprendizaje α y el margen b.

4 ☐ En el algoritmo Perceptrón:

A) principalmente existen dos parámetros: número de clases y número de prototipos.
B) el uso del margen b permite encontrar soluciones adecuadas cuando el problema no es linealmente separable.
C) el valor del margen b depende del valor de la constante de aprendizaje α empleado.
D) principalmente existen dos parámetros: la constante de aprendizaje α y el número de iteraciones.

5 ☐ El parámetro del algoritmo Perceptrón que denominamos *margen*, b, es un valor real que, suponiendo que sea positivo (como suele ser), restringe el conjunto de soluciones a las que puede converger el algoritmo. Concretamente, dadas N muestras de entrenamiento $(\mathbf{x}_1, c_1), \ldots, (\mathbf{x}_N, c_N)$ de C clases, el algoritmo Perceptrón tratará de hallar funciones discriminantes lineales $g_1(\cdot), \ldots, g_C(\cdot)$ tales que, para todo $n = 1, \ldots, N$:

A) $g_{c_n}(\mathbf{x}_n) > g_c(\mathbf{x}_n)$ para toda clase $c \neq c_n$
B) $g_{c_n}(\mathbf{x}_n) > g_c(\mathbf{x}_n) + b$ para toda clase $c \neq c_n$
C) $g_{c_n}(\mathbf{x}_n) > g_c(\mathbf{x}_n) - b$ para toda clase $c \neq c_n$
D) Ninguna de las anteriores

6 Sea un problema de clasificación en C clases $C \in \{1, 2, \ldots, C\}$, en el que los objetos están representados mediante puntos en un espacio vectorial de D dimensiones, $\mathbf{x} \in \mathbb{R}^D$. Suponiendo que un $\mathbf{x}$ dado pertenece a la clase 1, el algoritmo Perceptrón:

A) Modifica el discriminante lineal $g_1(\mathbf{x})$ en todo caso.
B) Modifica el discriminante lineal $g_1(\mathbf{x})$ si existe $c \neq 1, g_c(\mathbf{x}) > g_1(\mathbf{x})$.
C) Modifica el discriminante lineal $g_c(\mathbf{x})$ si $g_c(\mathbf{x}) < g_1(\mathbf{x})$ con $c \neq 1$.
D) Modifica el discriminante lineal $g_1(\mathbf{x})$ sólo si $g_c(\mathbf{x}) > g_1(\mathbf{x})$ para todo $c \neq 1$.

7 En un problema de clasificación en dos clases se tienen los siguientes puntos en dos dimensiones: $\mathbf{x}_1 = (1,1)^t, \mathbf{x}_2 = (2,2)^t, \mathbf{x}_3 = (2,0)^t$; $\mathbf{x}_1$ y $\mathbf{x}_2$ pertenencen a la clase A y $\mathbf{x}_3$ a la clase B. Teniendo en cuenta que se emplea un clasificador basado en funciones discriminantes lineales con vectores de pesos $\mathbf{w}_A$ y $\mathbf{w}_B$ asociados a las clases A y B respectivamente, indica cuál de las siguientes afirmaciones es *falsa*:

A) Se puede encontrar una función discriminante lineal que clasifique $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$ con error=2/3.
B) Los pesos $\mathbf{w}_A = (1, -1, 1)^t$ y $\mathbf{w}_B = (1, 2, -4)^t$ clasifican $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$ sin error.
C) Los pesos $\mathbf{w}_A = (1, -1, 1)^t$ y $\mathbf{w}_B = (1, 2, -4)^t$ clasifican $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$ con error=1/3.
D) Se puede encontrar una función discriminante lineal que clasifique $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$ con error=1/3.

8 Sea un problema de clasificación en tres clases $\{A, B, C\}$. El espacio de representación de los objetos es bidimensional, $\mathbb{R}^2$. Se propone emplear un clasificador basado en funciones discriminantes lineales con los siguientes vectores de pesos para cada clase, $\mathbf{w}_A = (1, 1, 0)^t$, $\mathbf{w}_B = (-1, 1, -1)^t$ y $\mathbf{w}_C = (1, -2, 2)^t$. ¿Cuál es la clasificación de los objetos $\mathbf{x}_1 = (1, 1)^t$ y $\mathbf{x}_2 = (0, -1)^t$?

A) $c(\mathbf{x}_1) = B \quad c(\mathbf{x}_2) = C$
B) $c(\mathbf{x}_1) = A \quad c(\mathbf{x}_2) = B$
C) $c(\mathbf{x}_1) = B \quad c(\mathbf{x}_2) = A$
D) $c(\mathbf{x}_1) = A \quad c(\mathbf{x}_2) = A$

9 (Examen de SIN del 18 de enero de 2013)
Sean $g_1(\mathbf{y}) = y_1^2 + 2\,y_2^2$ y $g_2(\mathbf{y}) = 2\,y_1^2 + y_2^2$ dos funciones discriminantes para las clases 1 y 2, respectivamente. La frontera de decisiónentre estas clases es:

A) Una parábola.
B) Hiperesférica.
C) Viene dada por la ecuación $y_1^2 + y_2^2 = 0$.
D) Una recta.

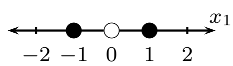
10 (Examen de SIN del 30 de enero de 2013)
Para un problema de clasificación de dos clases en $\Re^2$ se han construido tres clasificadores distintos. Uno está formado por las dos funciones discriminantes lineales siguientes: $g_1(y) = 3 + 4\,y_1 - 2\,y_2$ y $g_2(y) = -3 + 1.5\,y_1 + 5\,y_2$. El segundo clasificador por $g'_1(y) = 6 + 8\,y_1 - 4\,y_2$ y $g'_2(y) = -6 + 3\,y_1 + 10\,y_2$. El tercero por $g''_1(y) = -6 - 8\,y_1 + 4\,y_2$ y $g''_2(y) = 6 - 3\,y_1 - 10\,y_2$. ¿Los tres clasicadores son equivalentes? es decir ¿definen las mismas fronteras de decisión?

A) (g_1, g_2) y (g'_1, g'_2) son equivalentes.
B) Los tres son equivalentes.
C) (g_1, g_2) y $(g"_1, g"_2)$ son equivalentes.
D) (g'_1, g'_2) y $(g"_1, g"_2)$ son equivalentes.

11 (Examen de SIN del 30 de enero de 2013)
El algoritmo Perceptrón es una técnica de aprendizaje...

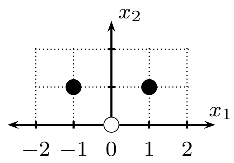
A) supervisado de clasificadores lineales.
B) supervisado de clasificadores cuadráticos.
C) no-supervisado de clasificadores lineales.
D) no-supervisado de clasificadores cuadráticos.

12 (Examen de SIN del 15 de enero de 2014; examen del bloque 2; cuestión 4)
En la figura de la derecha se representan tres muestras de aprendizaje unidimensionales de 2 clases: ○ y ●. ¿Cuál será el número de errores de clasificación cometidos por un clasificador lineal de mínimo error?



A) 0
B) 1
C) 2
D) 3

13 $\square$ (Examen de SIN del 15 de enero de 2014; examen del bloque 2; cuestión 5)
Supongamos que en el ejercicio anterior añadimos un nueva característica x_2 que se define como $x_2 = x_1^2$. De esta forma las tres muestras de aprendizaje pasan a ser bidimensionales como se observa en la figura de la derecha. En este caso, ¿cuál será el número de errores de clasificación cometidos por un clasificador lineal de mínimo error?



A) 0
B) 1
C) 2
D) 3

14 $\square$ (Examen de SIN del 15 de enero de 2014; examen del bloque 2; cuestión 6)
Sea un problema de clasificación en 2 clases, $c = 1, 2$, para objetos representados mediante vectores de características bidimensionales. Se tienen 2 muestras de entrenamiento: $\mathbf{x}_1 = (0, 0)^t$ de la clase $c_1 = 1$, y $\mathbf{x}_2 = (1, 1)^t$ de $c_2 = 2$. Asimismo, se tiene un clasificador lineal definido por los vectores de pesos: $\mathbf{w}_1 = (w_{10}, w_{11}, w_{12}) = (1, -1, -1)^t$ y $\mathbf{w}_2 = (w_{20}, w_{21}, w_{22}) = (-1, 1, 1)^t$. Si aplicamos una iteración del algoritmo Perceptrón a partir de estos vectores de pesos, con factor de aprendizaje $\alpha = 1$ y margen $b = 0.1$, entonces:

A) No se modificará ningún vector de pesos.
B) Se modificará el vector de pesos de la clase 1.
C) Se modificará el vector de pesos de la clase 2.
D) Se modificarán los vectores de pesos de ambas clases.

15 $\square$ (Examen de SIN del 15 de enero de 2014; examen del bloque 2; cuestión 7)
El algoritmo Perceptrón está gobernado por dos parámetros que denominamos *velocidad de aprendizaje*, α, y *margen*, b, siendo ambos valores reales. En caso de que no supiéramos si las muestras de aprendizaje son linealmente separables, ¿qué valores de los parámetros α y b proporcionan mayores garantías de obtener fronteras de decisión de mejor calidad?

A) $\alpha = 0.1$ y $b = 0.0$.
B) $\alpha = 0.0$ y $b = 0.0$.
C) $\alpha = 0.1$ y $b = 1.0$.
D) $\alpha = 0.0$ y $b = 1.0$.

16 $\square$ (Examen de SIN del 28 de enero de 2014; examen final; cuestión 1)
En un experimento de clasificación con 300 datos de *test* se han observado 15 errores. Con una confianza del 95 %, podemos afirmar que la verdadera probabilidad de error es:

A) $P(\text{error}) = 5\,\% \pm 0.3\,\%$
B) $P(\text{error}) = 0.05 \pm 0.3$
C) $P(\text{error}) = 0.05$, exactamente
D) $P(\text{error}) = 0.05 \pm 0.03$

17 $\square$ (Examen de SIN del 28 de enero de 2014; examen final; cuestión 3)
Sea un problema de clasificación en 2 clases, $c = A, B$, para objetos representados mediante vectores de características bidimensionales. Como resultado de la aplicación del algoritmo Perceptrón sobre un conjunto de entrenamiento, se han obtenido los vectores de pesos $\mathbf{w}_A = (1, 1, 0)^t$ y $\mathbf{w}_B = (-1, 0, 1)^t$. ¿En qué clases se clasifican $\mathbf{x}_1 = (-1, 0)^t$ y $\mathbf{x}_2 = (0, 3)^t$?

A) $\hat{c}(\mathbf{x}_1) = A$ y $\hat{c}(\mathbf{x}_2) = A$.
B) $\hat{c}(\mathbf{x}_1) = A$ y $\hat{c}(\mathbf{x}_2) = B$.
C) $\hat{c}(\mathbf{x}_1) = B$ y $\hat{c}(\mathbf{x}_2) = A$.
D) $\hat{c}(\mathbf{x}_1) = B$ y $\hat{c}(\mathbf{x}_2) = B$.

2. Problemas

1. Sea un problema de clasificación en tres clases, $c = \{A, B, C\}$, para objetos representados mediante vectores de características tridimensionales. Se tiene un clasificador lineal basado en funciones discriminantes lineales de la forma

$$g_c(\mathbf{x}) = \mathbf{w}_c \cdot \mathbf{x} \qquad \text{para toda clase } c$$

donde $\mathbf{w}_c$ y $\mathbf{x}$ se hallan en notación compacta; es decir: $\mathbf{w} = (w_0, w_1, w_2, w_3)^t \in \mathbb{R}^4$ y $\mathbf{x} = (x_0, x_1, x_2, x_3)^t \in \mathbb{R}^4$, con $x_0 = 1$. Teniendo en cuenta que:

$$\mathbf{w}_A = (1, 1, 1, 1)^t \qquad \mathbf{w}_B = (-1, 0, -1, -2)^t \qquad \text{y} \qquad \mathbf{w}_C = (-2, 2, -1, 0)^t$$

Se pide:

a) Clasifica el punto $\mathbf{x}' = (2, 1, 2)^t$.

b) Sabemos que el punto $\mathbf{x}' = (-1, 0, -1)^t$ pertenece a la clase A. ¿Qué valores tendrían $\mathbf{w}_A$,$\mathbf{w}_B$ y $\mathbf{w}_C$ tras aplicar el algoritmo Perceptrón para dicho punto, usando una constante de aprendizaje $\alpha = 0.1$?

c) Dado el punto $\mathbf{x}' = (1, -1, 2)^t$ que pertenece a la clase A ***(corrección: antes C)***, obtén valores posibles de las discriminantes que lo clasifiquen correctamente.

2. Sea un problema de clasificación en 3 clases, $c = \{1, 2, 3\}$, para objetos representados mediante vectores de características bidimensionales. Se tienen 3 muestras de entrenamiento: $\mathbf{x}_1 = (0, 0)^t$ de la clase $c_1 = 1$, $\mathbf{x}_2 = (0, 1)^t$ de $c_2 = 2$, y $\mathbf{x}_3 = (2, 2)^t$ de $c_3 = 3$. Encuentra un clasificador lineal de mínimo error mediante el algoritmo Perceptrón, con vectores de pesos iniciales de las clases nulos, factor de aprendizaje $\alpha = 1$ y margen $b = 0.1$. Presenta una traza de ejecución que incluya las sucesivas actualizaciones de los vectores de pesos de las clases.

3. Sea un problema de clasificación en 2 clases, $c = \{1, 2\}$, para objetos representados mediante vectores de características bidimensionales. Se tienen 2 muestras de entrenamiento: $\mathbf{x}_1 = (0, 0)^t$ de la clase $c_1 = 1$, y $\mathbf{x}_2 = (1, 1)^t$ de $c_2 = 2$. Encuentra un clasificador lineal de mínimo error mediante el algoritmo Perceptrón, con vectores de pesos iniciales de las clases nulos, factor de aprendizaje $\alpha = 1$ y margen $b = 0.1$. Presenta una traza de ejecución que incluya las sucesivas actualizaciones de los vectores de pesos de las clases.

4. (Examen de SIN del 18 de enero de 2013; tiempo estimado: 30 minutos)
En un problema de clasificación en 2 clases, A, B, para objetos representados mediante vectores de características bidimensionales, se tienen dos muestras de entrenamiento:

$$\mathbf{y}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \in A, \quad \mathbf{y}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \in B,$$

a) Inicializando a 0 todas las componentes de los vectores de pesos iniciales, mostrar una traza de ejecución del algoritmo Perceptrón, con factor de aprendizaje $\alpha = 1.0$ y margen $b = 0.1$. La traza debe incluir las sucesivas actualizaciones de los vectores de pesos de las clases e indicar los vectores de pesos obtenidos como solución final.

b) Obtener la ecuación de la frontera de separación entre clases correspondiente a la solución obtenida. Representar gráficamente esta frontera, junto con los datos de entrenamiento. ¿Es satisfactoria esta solución?