Sistemas Inteligentes
Cuestiones y ejercicios del bloque 2, tema 2
Regresión Logística

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

18 de noviembre de 2024

Cuestiones

1 $\boxed{\text{C}}$ Sea un modelo de regresión logística en notación compacta (homogénea) para un problema de clasificaciónn en $C = 3$ clases y datos representados por vectores de dimensión $D = 2$.

$$p(\boldsymbol{y} \mid \boldsymbol{x}, \mathbf{W}) = \text{Cat}(\boldsymbol{y} \mid \mathcal{S}(\mathbf{W}^t\boldsymbol{x}) \text{ con } \mathbf{W}^t = \begin{pmatrix} 0 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{C \times D} .$$

Dado $\boldsymbol{x} = (0.5, 0.5)^t$, la probabilidad P de que $\boldsymbol{x}$ pertenezca a la clase 1 es:

A) $P < 0.25$
B) $0.25 \leq P < 0.5$
C) $0.5 \leq P < 0.75$
D) $0.75 \leq P$

2 $\boxed{\text{D}}$ Indica cuál de las siguientes afirmaciones sobre regresión logística es *incorrecta* (o escoge la última opción si las tres primeras son correctas):

A) Regresión logística es un modelo probabilístico de clasificación basado en la función softmax
B) Al tratarse de un modelo probabilístico de clasificación, regresión logística permite aplicar reglas de decisión más generales que la MAP (decidirse por la clase de máxima probabilidad a posteriori)
C) Al tratarse de un modelo probabilístico de clasificación, regresión logística permite plantear su aprendizaje probabilísticamente, con criterios estándar como máxima verosimilitud
D) Las tres afirmaciones anteriores son correctas

3 $\boxed{\text{D}}$ Indica cuál de las siguientes afirmaciones sobre regresión logística es *incorrecta* (o escoge la última opción si las tres primeras son correctas):

A) Regresión logística es un modelo probabilístico de clasificación basado en una función predictora de logits lineal con la entrada
B) Al tratarse de un modelo probabilístico de clasificación, regresión logística permite aplicar reglas de decisión más generales que la MAP (decidirse por la clase de máxima probabilidad a posteriori)
C) Regresión logística es un modelo probabilístico de clasificación basado en la distribución categórica
D) Las tres afirmaciones anteriores son correctas

Problemas

1. Sea un modelo de regresión logística en notación compacta (homogénea) para un problema de clasificación en $C = 3$ clases y datos representados por vectores de dimensión $D = 2$.

$$p(\boldsymbol{y} \mid \boldsymbol{x}, \mathbf{W}) = \text{Cat}(\boldsymbol{y} \mid \mathcal{S}(\mathbf{W}^t\boldsymbol{x})) \text{ con } \mathbf{W}^t = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \in \mathbb{R}^{C \times D}$$

Actualiza el valor de $\mathbf{W}$ con una iteración de descenso por gradiente con el conjunto de entrenamiento $\mathcal{D} = \{(\boldsymbol{x} = (1, 1, 1)^t, y = 1)\}$ y factor de aprendizaje $\eta = 0.1$.

$$\boldsymbol{a} = \mathbf{W}^t \boldsymbol{x} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

$$\boldsymbol{\mu} = \mathcal{S}(\boldsymbol{a}) = \begin{pmatrix} \frac{e^0}{e^0+e^1+e^1} \\ \frac{e^1}{e^0+e^1+e^1} \\ \frac{e^1}{e^0+e^1+e^1} \end{pmatrix} = \frac{1}{1+2e} \begin{pmatrix} 1 \\ e \\ e \end{pmatrix} = \begin{pmatrix} 0.1554 \\ 0.4223 \\ 0.4223 \end{pmatrix}$$

$$\mathbf{W} = \mathbf{W} - \eta \boldsymbol{x} (\mu - \boldsymbol{y})^t =$$

$$\begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} - 0.1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} (-0.8446, 0.4223, 0.4223) =$$

$$\begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} -0.0845 & 0.0422 & 0.0422 \\ -0.0845 & 0.0422 & 0.0422 \\ -0.0845 & 0.0422 & 0.0422 \end{pmatrix} =$$

$$\begin{pmatrix} 1.0845 & -0.0422 & 0.9578 \\ -0.9155 & 0.9578 & -1.0422 \\ 0.0845 & -0.0422 & 0.9578 \end{pmatrix}$$

2. La siguiente tabla presenta un conjunto de 2 muestras de entrenamiento de 2 dimensiones procedentes de 2 clases:

n	x_{n1}	x_{n2}	c_n
1	1	1	2
2	0	1	1

Adicionalmente, la siguiente tabla representa una matriz de pesos iniciales con los pesos de cada clase dispuestos por columnas:

$\mathbf{w}_1$	$\mathbf{w}_2$
0.	0.
-0.25	0.25
0.	0.

Se pide:

a) (0.5 puntos) Calcula el vector de logits asociado a cada muestra de entrenamiento.

b) (0.25 puntos) Aplica la función softmax al vector de logits de cada muestra de entrenamiento.

c) (0.25 puntos) Clasifica todas las muestras de entrenamiento. En caso de empate, elige cualquier clase.

d) (0.5 puntos) Calcula el gradiente de la función NLL en el punto de la matriz de pesos iniciales.

e) (0.5 puntos) Actualiza la matriz de pesos iniciales aplicando descenso por gradiente con factor de aprendizaje $\eta = 1.0$.

Solución:

a) Vector de logits para cada muestra de entrenamiento:

n	a_{n1}	a_{n2}
1	-0.25	0.25
2	0.	0.

b) Aplicación de la función softmax:

n	μ_{n1}	μ_{n2}
1	0.38	0.62
2	0.5	0.5

c) Clasificación de cada muestra:

n	$\hat{c}(x_n)$
1	2
2	1

d) Gradiente:

$\mathbf{g}_1$	$\mathbf{g}_2$
-0.06	0.06
0.19	-0.19
-0.06	0.06

e) Matriz de pesos actualizada:

$\mathbf{w}_1$	$\mathbf{w}_2$
0.06	-0.06
-0.44	0.44
0.06	-0.06

3. La siguiente tabla presenta un conjunto de 2 muestras de entrenamiento de 2 dimensiones procedentes de 2 clases:

n	x_{n1}	x_{n2}	c_n
1	1	0	2
2	1	1	1

Adicionalmente, la siguiente tabla representa una matriz de pesos iniciales con los pesos de cada clase dispuestos por columnas:

$\mathbf{w}_1$	$\mathbf{w}_2$
0.	0.
0.	0.
0.25	-0.25

Se pide:

a) (0.5 puntos) Calcula el vector de logits asociado a cada muestra de entrenamiento.

b) (0.25 puntos) Aplica la función softmax al vector de logits de cada muestra de entrenamiento.

c) (0.25 puntos) Clasifica todas las muestras de entrenamiento. En caso de empate, elige cualquier clase.

d) (0.5 puntos) Calcula el gradiente de la función NLL en el punto de la matriz de pesos iniciales.

e) (0.5 puntos) Actualiza la matriz de pesos iniciales aplicando descenso por gradiente con factor de aprendizaje $\eta = 1.0$.

Solución:

a) Vector de logits para cada muestra de entrenamiento:

n	a_{n1}	a_{n2}
1	0.	0.
2	0.25	-0.25

b) Aplicación de la función softmax:

n	μ_{n1}	μ_{n2}
1	0.5	0.5
2	0.62	0.38

c) Clasificación de cada muestra:

n	$\hat{c}(x_n)$
1	2
2	1

d) Gradiente:

$\mathbf{g}_1$	$\mathbf{g}_2$
0.06	-0.06
0.06	-0.06
-0.19	0.19

e) Matriz de pesos actualizada:

$\mathbf{w}_1$	$\mathbf{w}_2$
-0.06	0.06
-0.06	0.06
0.44	-0.44

4. La siguiente tabla presenta por filas un conjunto de 2 muestras de entrenamiento de 2 dimensiones procedentes de 2 clases:

n	x_{n1}	x_{n2}	c_n
1	0	1	1
2	0	0	2

Adicionalmente, la siguiente tabla representa una matriz de pesos iniciales con los pesos de cada clase dispuestos por columnas:

$\mathbf{w}_1$	$\mathbf{w}_2$
0.	0.
0.	0.
0.25	-0.25

Se pide:

a) (0.5 puntos) Calcula el vector de logits asociado a cada muestra de entrenamiento.

b) (0.25 puntos) Aplica la función softmax al vector de logits de cada muestra de entrenamiento.

c) (0.25 puntos) Clasifica todas las muestras de entrenamiento. En caso de empate, elige cualquier clase.

d) (0.5 puntos) Calcula el gradiente de la función NLL en el punto de la matriz de pesos iniciales.

e) (0.5 puntos) Actualiza la matriz de pesos iniciales aplicando descenso por gradiente con factor de aprendizaje $\eta = 1.0$.

Solución:

a) Vector de logits para cada muestra de entrenamiento:

n	a_{n1}	a_{n2}
1	0.25	-0.25
2	0.	0.

b) Aplicación de la función softmax:

n	μ_{n1}	μ_{n2}
1	0.62	0.38
2	0.5	0.5

c) Clasificación de cada muestra:

n	$\hat{c}(x_n)$
1	1
2	2

d) Gradiente:

$\mathbf{g}_1$	$\mathbf{g}_2$
0.06	-0.06
0.	0.
-0.19	0.19

e) Matriz de pesos actualizada:

$\mathbf{w}_1$	$\mathbf{w}_2$
-0.06	0.06
0.	0.
0.44	-0.44

5. La siguiente tabla presenta por filas un conjunto de 2 muestras de entrenamiento de 2 dimensiones procedentes de 2 clases:

n	x_{n1}	x_{n2}	c_n
1	0	0	2
2	1	1	1

Adicionalmente, la siguiente tabla representa una matriz de pesos iniciales con los pesos de cada clase dispuestos por columnas:

$\mathbf{w}_1$	$\mathbf{w}_2$
0.	0.
0.25	-0.25
0.25	-0.25

Se pide:

a) (0.5 puntos) Calcula el vector de logits asociado a cada muestra de entrenamiento.

b) (0.25 puntos) Aplica la función softmax al vector de logits de cada muestra de entrenamiento.

c) (0.25 puntos) Clasifica todas las muestras de entrenamiento. En caso de empate, elige cualquier clase.

d) (0.5 puntos) Calcula el gradiente de la función NLL en el punto de la matriz de pesos iniciales.

e) (0.5 puntos) Actualiza la matriz de pesos iniciales aplicando descenso por gradiente con factor de aprendizaje $\eta = 1.0$.

Solución:

a) Vector de logits para cada muestra de entrenamiento:

n	a_{n1}	a_{n2}
1	0.	0.
2	0.5	-0.5

b) Aplicación de la función softmax:

n	μ_{n1}	μ_{n2}
1	0.5	0.5
2	0.73	0.27

c) Clasificación de cada muestra:

n	$\hat{c}(x_n)$
1	2
2	1

d) Gradiente:

$\mathbf{g}_1$	$\mathbf{g}_2$
0.12	-0.12
-0.13	0.13
-0.13	0.13

e) Matriz de pesos actualizada:

$\mathbf{w}_1$	$\mathbf{w}_2$
-0.12	0.12
0.38	-0.38
0.38	-0.38