Фильчев Виктор,
3823М1ПМвм

Домашняя работа №4

№45

Распишем $E_D(f(x,D)-f^*(x))^2$:

$E_D[(f(x,D)-f^*(x))^2] = E_D[(f(x,D)+E_D[f(x,D)]-E_D[f(x,D)]-$
$-f^*(x))^2] = E_D[(f(x,D)-E_D[f(x,D)])^2+(E_D[f(x,D)]-f^*(x))^2+$
$+2(f(x,D)-E_D[f(x,D)])(E_D[f(x,D)]-f^*(x))]$

Вычислим мат. ожидание отдельных слаг-х:

① $E_D[(f(x,D)-E_D[f(x,D)])^2] = D_D[f(x,D)]$ - дисперсия $f(x,D)$

② $E_D[(E_D[f(x,D)]-f^*(x))^2] = (E_D[f(x,D)-f^*(x)])^2$ - квадрат матожид-х смещ-я.

③ $2E_D[(f(x,D)-E_D[f(x,D)])(E_D[f(x,D)]-f^*(x))] = E_D[f(x,D)E_D[f(x,D)]-$
$-E_D[f(x,D)]^2-f(x,D)f^*(x)+f^*(x)E_D[f(x,D)]] = E_D[f(x,D)]^2-$
$-E_D[f(x,D)]^2-E_D[f(x,D)]E_D[f^*(x)]+E_D[f^*(x)]E_D[f(x,D)]=0$

Получаем таким образом:

$E_D[(f(x,D)-f^*(x))^2] = D_D[f(x,D)]+(E_D[f(x,D)-f^*(x)])^2$.

№46

① $Corr(X_1,X_2) = \frac{cov(X_1,X_2)}{\sigma X_1 \cdot \sigma X_2}$

$\sigma X = \sqrt{D[X]}$, $cov(X_1,X_2) = E((X_1-EX_1)(X_2-EX_2))$

№ 46

Два класса, пр-во признаков двумерное, $P_1 = P_2 = \frac{1}{2}$.

① $\Sigma_1 = \Sigma_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

$P_1 = P_2 = \frac{1}{2} \Rightarrow E(x) = \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 = \frac{1}{2}\begin{pmatrix} -1 \\ -1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$\Sigma = \sum_{k=1}^{2} P_k(\Sigma_k + (\mu_k - E(x))(\mu_k - E^T(x))$

$\Sigma = \frac{1}{2}\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right) + \frac{1}{2}\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right) =$

$= \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$

Ковар. матрица диагональна ⇒ ков. между x_1 и x_2 равна нулю.

$\rho_{x_1,x_2} = \frac{cov(x_1, x_2)}{\sigma_{x_1}\sigma_{x_2}} = 0$

② Байесов классификатор:

$C(x_1) = \begin{cases} c_1, & \text{если } x_1 < 0 \\ c_2, & \text{если } x_2 \geq 0 \end{cases}$

Вычисление байесовой ошибки:

Для C_1: Ошибка при $x_1 \geq 0$

$\mathcal{E}_1 = \frac{1}{2} P(x_1 \geq 0 / C_1) = \frac{1}{2}\left(1 - \Phi\left(\frac{0-(-1)}{1}\right)\right) = \frac{1}{2}(1 - \Phi(1)) = 0,0793.$

Для $C2$: Ошибка при $x_1 < 0$

$\mathcal{E}_2 = \frac{1}{2} P(x_1 < 0 / C_2) = \frac{1}{2}\Phi\left(\frac{0-1}{1}\right) = \frac{1}{2}\Phi(-1) = 0,0793.$

Общая ошибка:

$\mathcal{E} = \mathcal{E}_1 + \mathcal{E}_2 = 2 \cdot 0,0793 = 0,1587.$

Знач-я ф-ии расп-я:

$\Phi(1) = 0,8413, \; \Phi(-1) = 1 - \Phi(1) = 0,1587.$

③ Ковариац. матрицы и априорные вер-ти одинаковы для обоих классов.

$\delta(x) = (\mu_2 - \mu_1)^T x + \frac{1}{2}(\|\mu_2\|^2 - \|\mu_1\|^2)$

Разности.

$\mu_2 - \mu_1 = \binom{1}{1} - \binom{-1}{1} = \binom{2}{0}$

$\|\mu_2\|^2 - \|\mu_1\|^2 = (1^2 + 1^2) - ((-1)^2 + 1^2) = 2 - 2 = 0$

$\delta(x) = 2x_1$

Классификатор.

$C(x) = \begin{cases} c_1, \text{ если } \delta(x) < 0 \text{ (т.е. } x_1 < 0) \\ c_2, \text{ если } \delta(x) \geq 0 \text{ (т.е. } x_1 \geq 0) \end{cases}$

Порог классификации и распр. x_1 останутся те же, => байесова ошибка та же, 0,1587.
Для обоих пр-в байесова ошибка не уменьшается, остаётся 0,1587.

④ $x_2 | C_1 \sim N(1,1)$

$x_2 | C_2 \sim N(1,1)$

$\mu_{C_2} - \mu_{C_1} = \binom{1}{1} - \binom{-1}{1} = \binom{2}{0}$

$\delta(x) = (\mu_{C_2} - \mu_{C_1})^T x = 2x_1$

$x_1 > 0$ при C_2, $x_1 < 0$ при C_1.

x_2 не влияет на реш-е классиф-ии, та же байесова ошибка.

Исп-е второй пер-й не приводит к уменьш. ошибки. (x_2 не содержит информации для различ. классов).

№48

$x^{(0)} = 0$, $x^{(1)} = 1$; шумовой пр-к на отр $[-1;1]$;
$x = (0{,}32, 0)$.

$d(x, x^{(0)}) = \sqrt{0{,}32^2 + (0 - n^{(0)})^2} = \sqrt{0{,}1024 + (n^{(0)})^2}$

$d(x, x^{(1)}) = \sqrt{(-0{,}68)^2 + (-n^{(1)})^2} = \sqrt{0{,}4624 + (n^{(1)})^2}$

$P(d(x, x^{(1)}) < d(x, x^{(0)}))$ - найдем:

$\sqrt{0{,}4624 + (n^{(1)})^2} < \sqrt{0{,}1024 + (n^{(0)})^2}$

$0{,}4624 + (n^{(1)})^2 < 0{,}1024 + (n^{(0)})^2$

$(n^{(1)})^2 - (n^{(0)})^2 < -0{,}36$

$|n^{(0)}| > 0{,}6$

$n^{(1)} < \sqrt{0{,}36 - (n^{(0)})^2}$

$P(|n^{(0)}| > 0{,}6) = 2 \cdot \frac{0{,}4}{2} = 0{,}4$

$P(|n^{(1)}| < \sqrt{0{,}36 - (n^{(0)})^2}) = 2\int_0^1 \sqrt{0{,}36 - (n^{(0)})^2}\, d\left(\frac{n^{(0)}}{2}\right) =$
$= 0{,}21.$

№49

$\forall x \in D$, $\exists \{y_1, \dots y_n\}$

$d(x, y_i) = \|x - y_i\|_2$

$\|x - y_j\|_2 \leq \|x - y_i\|_2$ $\forall i, j \in$ набору из k соседей

$\|x - y_i\|_2^2 = \|x\|_2^2 - 2x^T y_i + \|y_i\|_2^2$

($\|x\|_2^2$ - одинаково для $\forall$ рассм-й)

$\|x - y\|_2^2 \leq \|x - y_j\|_2^2$

$$-2x^T y_i + \|y_i\|_2^2 \leq -2x^T y_j + \|y_j\|_2^2$$

$$x^T \underbrace{(y_j - y_i)}_{a_{ij}} \leq \underbrace{\frac{1}{2}(\|y_j\|_2^2 - \|y_i\|_2^2)}_{b_{ij}}$$

$$a_{ij}^T x = b_{ij}$$

$i, j \in [1, k]$ => имеем сист-у из мн. усл-й, пересечение которых – выпуклый многогранник. Пар-в конечное число => область является полиэдром.