5 Д/З

№1 $\exists\ g: R^n \to R^m,\ x \in R^n$

Матрица Якоби

$$\frac{\partial g}{\partial x} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1} & & \frac{\partial g_m}{\partial x_n} \end{pmatrix}$$

Док-ть:

1) если $a \in R^n$, $x \in R^n$, то $\frac{\partial (a^T x)}{\partial x} = a$

$$\frac{\partial f}{\partial x} = \left(\frac{\partial f}{\partial x_1} \cdots \frac{\partial f}{\partial x_n}\right) = (a_1 \dots a_n) = a$$

2) если $A \in R^{m \times n}$, $x \in R^n$, то $\frac{\partial (Ax)}{\partial x} = A$

$$Ax = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \cdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix} = g$$

$$\frac{\partial g_i}{\partial x_j} = a_{ij} \Rightarrow \frac{\partial g}{\partial x} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

$$\begin{cases} \frac{\partial (Ax)}{\partial x} = \frac{\partial g}{\partial x} \\ \frac{\partial g}{\partial x} = A \end{cases} \Rightarrow \frac{\partial (Ax)}{x} = A$$

4) если $x \in R^n$, то $\frac{\partial \|x\|^2}{\partial x} = 2x$

$\|x\|^2 = x_1^2 + \dots + x_n^2$

$$\frac{\partial \|x\|^2}{\partial x} = \frac{\partial(x_1^2 + \dots + x_n^2)}{\partial x} = 2(x_1 + \dots + x_n) = 2x$$

5) если g - скаляр. ф-я и под $g(x)$ понимается применение ф-ии g к каждой компоненте вектора $x \in R^n$, то $\frac{\partial g(x)}{\partial x} = \mathrm{diag}(g'(x))$, где $\mathrm{diag}(a)$ - диагональная матрица с диагональю a.

$$\frac{\partial g(x)}{\partial x} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_n}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \dots & \frac{\partial g_n}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \frac{\partial g(x_1)}{\partial x_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{\partial g(x_n)}{\partial x_n} \end{pmatrix} \Rightarrow$$

$$\Rightarrow \frac{\partial g(x)}{\partial x} = \mathrm{diag}(g'(x))$$

6) если $h: R^n \to R^m$, $g: R^m \to R^p$, $x \in R^n$, то

$$\frac{\partial g(h(x))}{\partial x} = \frac{\partial g(h(x))}{\partial h} \frac{\partial h(x)}{\partial x}$$

$$g(h) = \begin{pmatrix} g_1(h_1 \dots h_m) \\ \vdots \\ g_p(h_1 \dots h_m) \end{pmatrix} \qquad h(x) = \begin{pmatrix} h_1(x_1 \dots x_n) \\ \vdots \\ h_m(x_1 \dots x_n) \end{pmatrix}$$

3) если $A \in R^{n\times n}$, $x \in R^n$, то $\frac{\partial(x^T A x)}{\partial x} = (A + A^T)x$;

в частности, если $A^T = A$, то $\frac{\partial (x^T A x)}{\partial x} = 2Ax$;

$$x^T A x = (x_1 \dots x_n)\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn}\end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n\end{pmatrix} = x_1 \sum_i a_{1i} x_i + \dots + x_n \sum_i a_{ni} x_i$$

$$\Rightarrow (x^T A x) = \sum_{j=1}^{n} \left(x_j \sum_{i=1}^{n} a_{ji} x_i\right) = g$$

$$\begin{cases} \frac{\partial g}{\partial x_1} = \sum_{i=1}^{n} a_{1i} x_i + \sum_{i=1}^{n} a_{i1} x_i \\ \vdots \\ \frac{\partial g}{\partial x_n} = \sum_{i=1}^{n} a_{ni} x_i + \sum_{i=1}^{n} a_{in} x_i \end{cases}$$

$$\Downarrow$$

$$\frac{\partial g}{\partial x} = \begin{pmatrix} \frac{\partial g}{\partial x_1} \\ \vdots \end{pmatrix} = \underbrace{\begin{pmatrix} \sum_{i=1}^{n} a_{1i} x_i \\ \vdots \\ \sum_{i=1}^{n} a_{ni} x_i \end{pmatrix}}_{= Ax} + \underbrace{\begin{pmatrix} \sum_{i=1}^{n} a_{i1} x_i \\ \vdots \\ \sum_{i=1}^{n} a_{in} x_i \end{pmatrix}}_{= A^T x} \Rightarrow$$

$$\Rightarrow \frac{\partial (x^T A x)}{\partial x} = (A + A^T)x$$

$$\frac{\partial g_i(h(x))}{\partial x_i} = \sum_{k=1}^{m} \frac{\partial g_i}{\partial h_k}\frac{\partial h_k}{\partial x_i}$$

$$\frac{\partial g(h(x))}{\partial x} = \begin{pmatrix} \sum_{k=1}^{m} \frac{\partial g_1}{\partial h_k}\frac{\partial h_k}{\partial x_1} & \dots & \sum_{k=1}^{m} \frac{\partial g_1}{\partial h_k}\frac{\partial h_k}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{m} \frac{\partial g_p}{\partial h_k}\frac{\partial h_k}{\partial x_1} & \dots & \sum_{k=1}^{m} \frac{\partial g_p}{\partial h_k}\frac{\partial h_k}{\partial x_n} \end{pmatrix}$$

С другой стороны

$$\frac{\partial g}{\partial h} = \begin{pmatrix} \frac{\partial g_1}{\partial h_1} & \dots & \frac{\partial g_1}{\partial h_m} \\ \vdots & & \\ \frac{\partial g_p}{\partial h_1} & \dots & \frac{\partial g_p}{\partial h_m} \end{pmatrix} = (1) \quad \frac{\partial h}{\partial x} = \begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \dots & \frac{\partial h_1}{\partial x_n} \\ \vdots & & \\ \dots & & \frac{\partial h_m}{\partial x_n} \end{pmatrix} = (2)$$

$$\frac{\partial g(h)}{\partial h}\frac{\partial h}{\partial x} = (1)(2) = \begin{pmatrix} \sum_{k=1}^{m} \frac{\partial g_1}{\partial h_k}\frac{\partial h_k}{\partial x_1} & \dots & \sum \frac{\partial g_1}{\partial h_k}\frac{\partial h_k}{\partial x_n} \\ \vdots & \ddots & \\ \vdots & \dots & \sum \frac{\partial g_p}{\partial h_k}\frac{\partial h_k}{\partial x_n} \end{pmatrix}$$

$\Rightarrow$ из (1) и (2) $\Rightarrow$, что $\frac{\partial g(h)}{\partial h}\frac{\partial h}{\partial x} = \frac{\partial(g(h))}{\partial x}$

№3 Дана обучающая выборка

x	1	1	0	0	-1
y	4	4	0	2	6

2) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix}$ $y = \begin{pmatrix} 4 \\ 4 \\ 0 \\ 2 \\ 6 \end{pmatrix}$ $X^T X = \begin{pmatrix} 5 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 3 \end{pmatrix}$

$X^T y = \begin{pmatrix} 16 \\ 2 \\ 14 \end{pmatrix}$

$$\hat{\mu}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \hat{\mu}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \Rightarrow \begin{cases} X_0 = \begin{pmatrix} 0 & 1 & 0 & 2 & 2 \\ -1 & 0 & 0 & 0 & 1 \end{pmatrix} \\ \tilde{X}_0 = \begin{pmatrix} -1 & 0 & -1 & 1 & 1 \\ -1 & 0 & 0 & 0 & 1 \end{pmatrix} \\ X_1 = \begin{pmatrix} 2 & 4 & 3 \\ 0 & 1 & 2 \end{pmatrix} \Rightarrow \tilde{X}_1 = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \end{cases}$$

Выборочные матрицы ковариации для каждого класса

$$\hat{\Sigma}_0 = \frac{1}{N_0 - 1} \sum_{y^{(i)}=0} (x^{(i)} - \hat{\mu}_0)(x^{(i)} - \hat{\mu}_0)^T = \{N_0 = 5\} =$$

$$= \frac{1}{4} \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}$$

$$\hat{\Sigma}_1 = \frac{1}{2} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Оценка матрицы ковариации

$$\hat{\Sigma} = \frac{1}{N-K} \sum_k \sum_{y^{(i)}=k} (x^{(i)} - \hat{\mu}_k)(x^{(i)} - \hat{\mu}_k)^T = \frac{1}{6} \begin{pmatrix} 6 & 3 \\ 3 & 4 \end{pmatrix}$$

Обратные матрицы:

$$\hat{\Sigma}_0^{-1} = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix} \qquad \hat{\Sigma}_1^{-1} = \begin{pmatrix} 4/3 & -2/3 \\ -2/3 & 4/3 \end{pmatrix} \qquad \hat{\Sigma}^{-1} = \begin{pmatrix} 8/5 & -6/5 \\ -6/5 & 12/5 \end{pmatrix}$$

Лин диск. ф-и

$$S_0(x) = \frac{8}{5} x_1 - \frac{6}{5} x_2 - \frac{4}{5} + \ln 5 - 3 \ln 2 \quad (1)$$

$$S_1(x) = \frac{18}{5} x_1 - \frac{6}{5} x_1 - \frac{6}{5} x_2 - \frac{24}{5} + \ln 3 - 3 \ln 2 \quad (2)$$

Найдем β из сист. норм. ур $X^T X\beta = X^T Y$:

$$\begin{cases} 5\beta_0 + \beta_1 + 3\beta_2 = 16 \\ \beta_0 + 3\beta_1 + \beta_2 = 2 \\ 3\beta_0 + \beta_1 + 3\beta_2 = 14 \end{cases} \Rightarrow \beta = \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix} \Rightarrow$$

$$\Rightarrow f(x) = 1 - x + 4x^2$$

3) $X^T X + \lambda I = \begin{pmatrix} 6 & 1 & 3 \\ 1 & 4 & 1 \\ 3 & 1 & 4 \end{pmatrix}$

$(X^T X + \lambda I)\beta = X^T Y$

$$\begin{cases} 6\beta_0 + \beta_1 + 3\beta_2 = 16 \\ \beta_0 + 4\beta_1 + \beta_2 = 2 \\ 3\beta_0 + \beta_1 + 4\beta_2 = 14 \end{cases} \Rightarrow \beta = \begin{pmatrix} 3/2 \\ -1/2 \\ 5/2 \end{pmatrix}$$

$$f_1(x) = {}^3/_2 - {}^1/_2 x + {}^5/_2 x^2$$

№9 Дана обучающая выборка

x_1	0	1	0	2	2	2	4	3
x_2	-1	0	0	0	1	0	1	2
Y	0	0	0	0	0	1	1	1

1) Оценим вероятность классов

$\hat{P}_0(\varphi = 0) = 5/8$

$\hat{P}_1(\varphi = 1) = 3/8$

Оценивем среднее для классов

Раздел. поверхность прямая с уравнением, которое получается при прир. $S_0(x) = S_1(x)$

$$x_1 = 2 + \frac{1}{2}\ln\left(\frac{5}{3}\right)$$

Квадрат. дискр. ф-ия:

$$S_0(x) = -\frac{1}{2}\ln\det\hat{\Sigma}_0 - \frac{1}{2}(x-\hat{\mu}_0)^T\hat{\Sigma}_0^{-1}(x-\hat{\mu}_0) +$$
$$+ \ln\hat{P}_r\{\varphi=0\}$$

$$S_0(x) = \ln 5 - 2\ln 2 - (x_1^2 - 2x_1x_2 + 2x_2^2 + 2x_2 - 2x_1 + 1)$$

$$S_1(x) = -\frac{1}{3}(2x_1^2 - 2x_1x_2 + 2x_2^2 - 10x_1 + 2x_2 + 14) + \frac{1}{2}\ln 3 - 2\ln 2$$

Раздел. пов-ть - парабола, прир. S_0 и S_1:

$$(x_1 - 2x_2)^2 + 4x_1 + 4x_2 - 11 - 3\ln\frac{5}{\sqrt{3}} = 0$$

№15 Дана обучающая выборка

X_1	0	0	1	1	0	0	1	1	1	0
X_2	0	1	0	1	1	1	1	1	1	1
Y	0	0	0	0	0	1	1	1	1	1

С помощью наивного байесова классификатора оценить вероятности $P_r(Y=0 \mid X_1=1, X_2=1)$; $P_r(Y=1 \mid X_1=1, X_2=1)$

$\hat{P}_r\{\varphi=0\} = \frac{1}{2}$ $\hat{P}_r\{\varphi=1\} = \frac{1}{2}$

$\hat{P}_r\{x_1=1 \mid \varphi=0\} = \frac{2}{5}$ $\hat{P}_r\{x_2=1 \mid \varphi=0\} = 3/5$

$\hat{P}_r\{x_1=1 \mid \varphi=1\} = 3/5$ $\hat{P}_r\{x_2=1 \mid \varphi=1\} = 1$

Основное предположение байесова классификатора

I $\hat{P}_r\{\varphi=0 \mid X_1=1;\ X_2=1\} = \dfrac{P_r\{X_1=1 \mid \varphi=0\}\, P_r\{X_2=1 \mid \varphi=0\}\, P_r(\varphi=0)}{P_r\{X_1=1,\ X_2=1\}}$

II $\hat{P}_r\{\varphi=1 \mid X_1=1;\ X_2=1\} = \dfrac{P_r\{X_1=1 \mid \varphi=1\}\, P_r\{X_2=1 \mid \varphi=1\}\, P_r\{\varphi=1\}}{P_r\{X_1=1,\ X_2=1\}}$

I $= \frac{3}{25} \cdot \frac{1}{P_r\{\ldots\}}$ II $= \frac{3}{10} \cdot \frac{1}{P_r\{\ldots\}}$

$P_r\{\ldots\} = P_r\{X_1=1,\ X_2=1\} = \frac{3}{25} + \frac{3}{10} = \frac{21}{50}$

I $= \frac{2}{7}$ II $= \frac{5}{7}$

4.1

Док-во:

$W_k = v_0 + L(v_1 \dots v_k)$ - k-мерное мин. многообр

$v_0 = \underset{\alpha_0}{\arg\min}\left(\sum_{i=1}^{N} dist^2(x_i, \alpha_0)\right) = \underset{a_0 \in R^n}{\arg\min}\left(\sum_{i=1}^{N} \|x^{(i)} - a_0\|^2\right) =$

$= \frac{1}{N} \sum_{i=1}^{N} x^{(i)} = \overline{x}$

График для задачи № 3

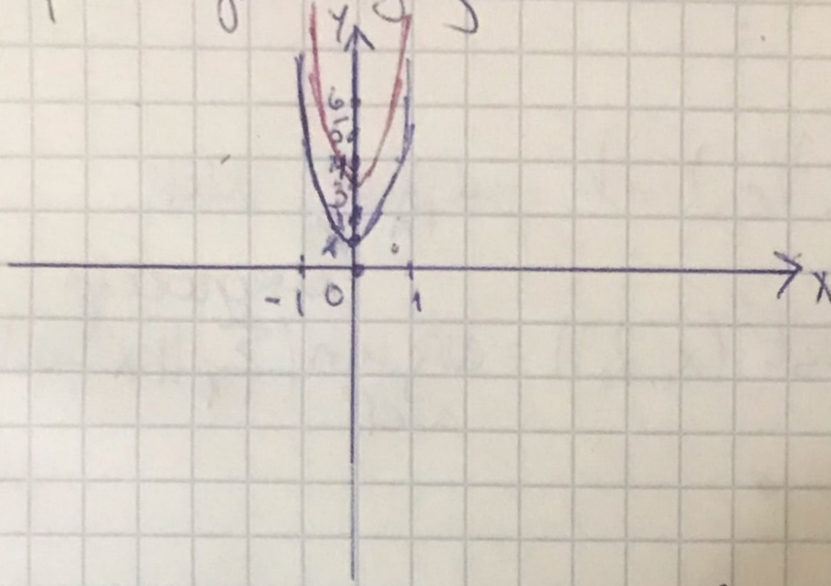

График для задачи № 9