№ 1

Зиннин И.

$$\frac{\partial y}{\partial x} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \dots & \frac{\partial y_1}{\partial x_n} \\ \dots & & & \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \dots & \frac{\partial y_m}{\partial x_n} \end{pmatrix}$$

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

<u>Д-во</u>: $(a_1 \dots, a_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (a_1 x_1, \dots, a_n x_n) = \sum_i x_i a_i$

$$\frac{\partial \sum_i x_i a_i}{\partial x_k} = \begin{pmatrix} \frac{\partial \sum_i x_i a_i}{\partial x_1} \\ \vdots \\ \frac{\partial \sum_i x_i a_i}{\partial x_k} \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = a$$

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

<u>Д-во</u>: Аналогично пункту 1)

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

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

<u>Д-во</u>: $x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$, $x^T = (x_1, \dots, x_m)$, $A = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \\ a_{m1} & \dots & a_{mm} \end{pmatrix}$

$$x^T A = (x_1 a_{11} + \dots + x_m a_{m1}) + \dots + (x_1 a_{1m} + \dots + x_m a_{mm})$$

$$x^T A x = (x_1 a_{11} + \dots + x_m a_{m1}) x_1 + \dots + (x_1 a_{1m} + \dots + x_m a_{mm}) x_m$$

$$\frac{\partial (x^T A x)}{\partial x_k} = \frac{\partial}{\partial x_k}\left(x_1^2 a_{11} + x_1 (x_2 a_{21} + \dots) + x_1 a_{12} x_2 + \dots\right) =$$

$$= \frac{\partial}{\partial x_k}\left(x_k^2 a_{kk} + x_k \sum_i x_i a_{ik} + x_k \sum_i a_{ki} x_i + \dots\right) =$$

$$= 2 x_k a_{kk} + \sum_i x_i a_{ik} + \sum_i a_{ki} x_i = 2 x_k a_{kk} + \sum_i x_i (a_{ik} + a_{ki}) =$$

$$= \sum_{i=1}^{m} x_i (a_{ik} + a_{ki}) = (A + A^T)\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = (A + A^T)\dot{x}$$

4) $x \in R^n$; $\frac{\partial \|x\|^2}{\partial x} = 2 \cdot x$

Д-во: $\sum_i \|x_i\|^2 \to \frac{\partial \sum_i \|x_i\|^2}{\partial x_k} = 2x$

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

Д-во: $\frac{\partial g}{\partial x} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \dots & \frac{\partial g_1}{\partial x_m} \\ \dots & & & \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \dots & \frac{\partial g_m}{\partial x_n} \end{pmatrix}$

Т.к. g - скалярная функция, то производные $\frac{\partial g_i}{\partial x_j}$ $i \neq j$ $= 0 \Rightarrow \frac{\partial g(x)}{\partial x} = diag(g'(x))$

№3

Обучающая выборка

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

2) $f(x) = \beta_0 + \beta_1 x + \beta_2 x^2$

$$X = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix}, \quad y = \begin{pmatrix} 4 \\ 4 \\ 0 \\ 2 \\ 6 \end{pmatrix}$$

$$X^T = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & -1 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

$$XX^T = \begin{pmatrix} 5 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 3 \end{pmatrix}, \quad X^T y = \begin{pmatrix} 16 \\ 2 \\ 14 \end{pmatrix}$$

Решаем систему: $X^T X \beta = X^T y$

$$\begin{pmatrix} 5 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 3 \end{pmatrix} \begin{pmatrix} \beta_x \\ \beta_y \\ \beta_z \end{pmatrix} = \begin{pmatrix} 16 \\ 2 \\ 14 \end{pmatrix} \Rightarrow \beta = \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix}$$

модель: $1 - x + 4x^2$: Синий цвет на рисунке

3) $\lambda = 1$

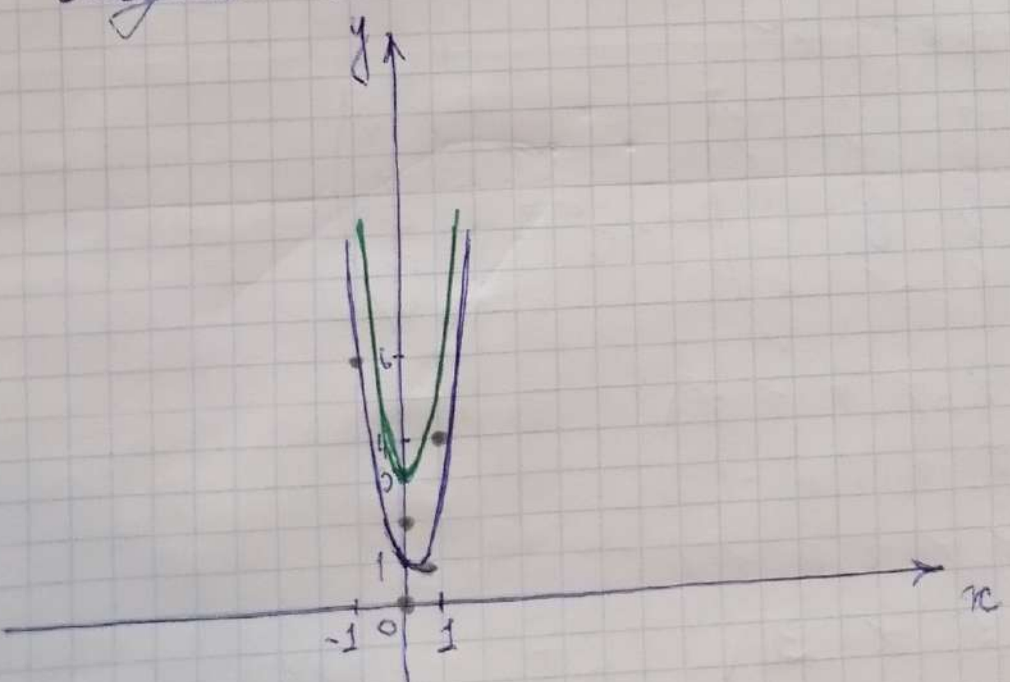
$$(X^T X + \lambda I)\beta = X^T y$$

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

Решаем систему: $(X^T X + \lambda I)\beta = X^T y$

$$\beta = \begin{pmatrix} 3 \\ -1 \\ 5 \end{pmatrix}$$

модель: $3 - x + 5x^2$: Зеленый цвет на рисунке



№9

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) Вероятности классов:

$\hat{P}_r\{Y=0\} = 5/8, \quad \hat{P}_r\{Y=1\} = \frac{3}{8}$

Средние для классов:

$$\hat{\mu}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \hat{\mu}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

Выборочные матрицы ковариации:

$$\hat{\Sigma}_0 = \frac{1}{N_0 - 1} \sum_{y^{(i)}=0} (x^{(i)} - \hat{\mu}_0)(x^{(i)} - \hat{\mu}_0)^T = \frac{1}{2}\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

$$\hat{\Sigma}_1 = \frac{1}{N_1 - 1} \sum_{y^{(i)}=1} (x^{(i)} - \hat{\mu}_1)(x^{(i)} - \hat{\mu}_1)^T = \frac{1}{2}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Матрица ковариации:

$$\hat{\Sigma} = \frac{1}{N-K} \sum_{k} \sum_{y^{(i)}=k} (x^{(i)} - \hat{\mu}_k)(x^{(i)} - \hat{\mu}_k)^T = \frac{1}{6}\left(\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\right) =$$

$$= \frac{1}{6}\begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 2/3 & 1/3 \\ 1/3 & 1/2 \end{pmatrix}$$

$$\hat{\Sigma}_0^{-1} = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}, \quad \hat{\Sigma}_1^{-1} = \frac{1}{3}\begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix}, \quad \hat{\Sigma}^{-1} = \begin{pmatrix} 2{,}25 & -1{,}5 \\ -1{,}5 & 3 \end{pmatrix}$$

Линейные дискриминантные функции:

$$\delta_0(x) = x^T \hat{\Sigma}^{-1} \hat{\mu}_0 - \frac{1}{2} \hat{\mu}_0^T \hat{\Sigma}^{-1} \hat{\mu}_0 + \ln \hat{P}_r\{Y=0\} =$$

$$= \frac{9}{4} x_1 - \frac{6}{4} x_2 - \frac{9}{8} + \ln \frac{5}{8}$$

$$\delta_1(x) = x^T \hat{\Sigma}^{-1} \hat{\mu}_1 - \frac{1}{2} \hat{\mu}_1^T \hat{\Sigma}^{-1} \hat{\mu}_1 + \ln \hat{P}_r\{Y=1\} =$$

$$= \frac{21}{4} x_1 - \frac{6}{4} x_2 - \frac{57}{8} + \ln \frac{3}{8}$$

$$\delta_0(x) = \delta_1(x)$$

$$7{,}5x_1 + 6 + \ln\frac{5}{3} = 0$$ – разделяющая поверхность (Изображена синим цветом)

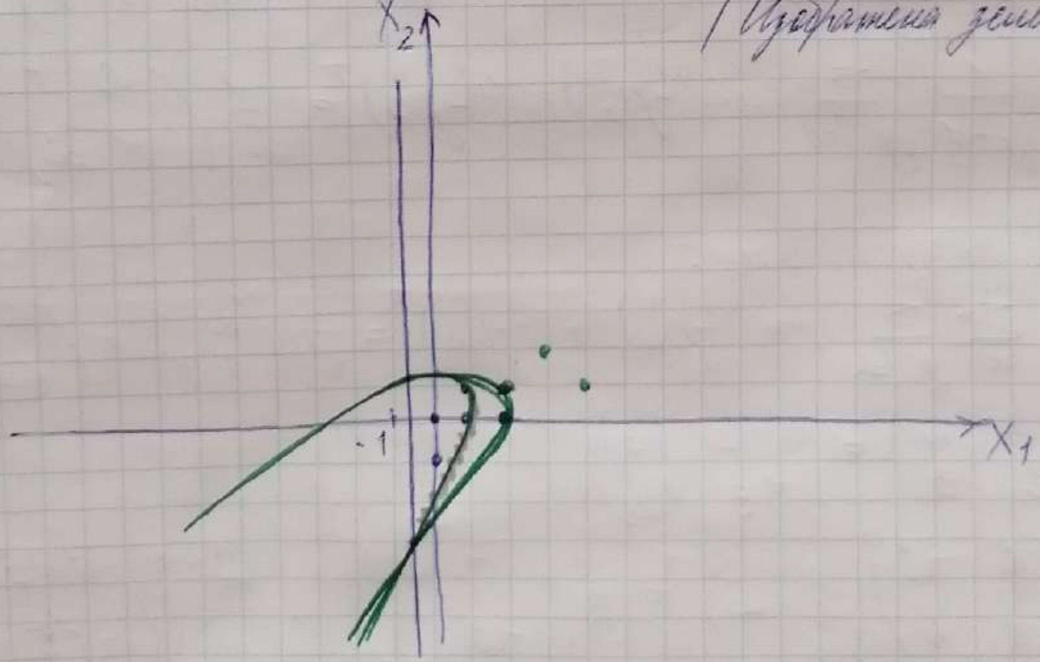
2) Квадратичные дискриминантные функции

$$\delta_0(x) = -\frac{1}{2}\ln\det\hat{\Sigma}_0 - \frac{1}{2}(x-\hat{\mu}_0)^T\hat{\Sigma}_0^{-1}(x-\hat{\mu}_0) + \ln\hat{P}_c\{Y=0\} =$$
$$= \frac{1}{2}\ln 2 - x_1^2 - 2x_2^2 + 2x_1x_2 + 2x_1 - 2x_2 - 1 + \ln\frac{5}{8}$$

$$\delta_1(x) = -\frac{1}{2}\ln\det\hat{\Sigma}_1 - \frac{1}{2}(x-\hat{\mu}_1)^T\hat{\Sigma}_1^{-1}(x-\hat{\mu}_1) + \ln\hat{P}_c\{Y=1\} =$$
$$= -\frac{1}{2}\ln\frac{3}{2} - \frac{1}{3}(2x_1^2 + 2x_2^2 - 2x_1x_2 - 10x_1 + 2x_2 + 14) + \ln\frac{3}{8}$$

$$\delta_0(x) = \delta_1(x)$$
$$\frac{3}{2}\ln 3 + 3\ln\frac{5}{8} - x_1^2 - 4x_2^2 + 4x_1x_2 - 4x_1 - 4x_2 + 11 = 0$$

(Изображена зелёным цветом)



№ 15.

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

Оценить: $P_r(Y=0 / X_1=1, X_2=1)$; $P_r(Y=1 / X_1=1, X_2=1)$

Априорные вероятности:

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

Условные вероятности:

$\hat{P}_r\{x_1=0 / Y=0\} = \frac{3}{5}$; $\hat{P}_r\{X_1=1 / Y=0\} = \frac{2}{5}$

$\hat{P}_r\{X_2=0 / Y=0\} = \frac{2}{5}$; $\hat{P}_r\{X_2=1 / Y=0\} = \frac{3}{5}$

$\hat{P}_r\{X_1=0 / Y=1\} = \frac{2}{5}$; $\hat{P}_r\{X_1=1 / Y=1\} = \frac{3}{5}$

$\hat{P}_r\{X_2=0 / Y=1\} = 0$; $\hat{P}_r\{X_2=1 / Y=1\} = 1$

Получаем:

$$P_r\{Y=0 / X_1=1, X_2=1\} = \frac{P_r\{X_1=1/Y=0\} \cdot P_r\{X_2=1/Y=0\} \cdot P_r(Y=0)}{P_r\{X_1=1, X_2=1\}}$$

$$= \frac{2/5 \cdot 3/5 \cdot 1/2}{3/25 + 3/10} = \frac{6}{21}$$

$$P_r\{Y=1 / X_1=1, X_2=1\} = \frac{P_r\{X_1=1/Y=1\} \cdot P_r\{X_2=1/Y=1\} \cdot P_r\{Y=1\}}{P_r\{X_1=1, X_2=1\}} =$$

$$= \frac{3/5 \cdot 1 \cdot 1/2}{3/25 \cdot 3/10} = \frac{5}{7}$$