№1

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

$$a^T x = \sum_{i=1}^{n} a_i x_i$$

$$\frac{\partial a^T x}{\partial x_k} = \frac{\partial}{\partial x_k} \sum_{i=1}^{n} a_i x_i = a_k$$

Отсюда следует, что $\frac{\partial a^T x}{\partial x} = a$

2) Если $A \in R^{mn}$, $x \in R^n$, то $\frac{\partial Ax}{\partial x} = A$

$(Ax)_i = \sum_{k=1}^{m} a_m x_k$, отсюда следует ч

$$\frac{\partial (Ax)_i}{\partial x_j} = a_{ij}, \quad i = 1, \dots, m \quad j = 1, \dots, n$$

$$\frac{\partial Ax}{\partial x} = a$$

3) Если $A \in R^{m \cdot m}$ и $x \in R^n$, то

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

Если $A^T = A$, то $\frac{\partial x^T A x}{\partial x} = 2Ax$

$$\frac{\partial}{\partial x_k} x^T A x = \frac{\partial}{\partial x_k} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j =$$

$$= \sum_{j=1}^{n} a_{kj} x_j + \sum_{i=1}^{n} a_{ik} x_i =$$

$$= (Ax)_k + (A^T x)_k$$

Отсюда следует:

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

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

$$\frac{\partial \|x\|^2}{\partial x_k} = \frac{\partial}{\partial x_k} \left(\sum_{i=1}^{n} x_i^2 \right) = 2 x_k$$

Отсюда следует:

$$\frac{\partial \|x\|^2}{\partial x} = 2x$$ чтд

5) Если g - скалярная функция и под $g'(x)$ понимается прим-е функции g к каждой компоненте вектора $x \in R^n$, то

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

$$\frac{\partial g(x)}{\partial x} = \left(\frac{\partial g(x)}{\partial x_1}, \frac{\partial g(x)}{\partial x_2}, \dots, \frac{\partial g(x)}{\partial x_n}\right)$$

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

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

$\frac{\partial g(h(x))}{\partial h}$ — матрица $p \cdot m$

$\frac{\partial h(x)}{\partial x}$ — матрица $m \cdot n$

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

$$\frac{\partial h_k(x)}{\partial x_j}$$

№3

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

Методом наим. квадратов построить модель вида

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

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

$$X^T X = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix}^T \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 3 \end{pmatrix}$$

$$X^T Y = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & -1 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 4 \\ 0 \\ 2 \\ 6 \end{pmatrix} = \begin{pmatrix} 16 \\ 2 \\ 14 \end{pmatrix}$$

Найдем β из $X^T X \beta = X^T Y$

$$\begin{pmatrix} 5 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 3 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} 16 \\ 2 \\ 14 \end{pmatrix} \Rightarrow \left(\begin{array}{ccc|c} 5 & 1 & 3 & 16 \\ 1 & 3 & 1 & 2 \\ 3 & 1 & 3 & 14 \end{array}\right) \begin{array}{l} (1)-5(2) \\ \\ (3)-3(2) \end{array}$$

$$= \left(\begin{array}{ccc|c} 0 & -14 & -2 & 6 \\ 1 & 3 & 1 & 2 \\ 0 & -8 & 0 & 8 \end{array}\right) = \left(\begin{array}{ccc|c} 0 & 1 & 1 & -3 \\ 1 & 3 & 1 & 2 \\ 0 & 1 & 0 & -1 \end{array}\right) \begin{array}{l} (1)-1(3) \\ \\ (2)-3(3) \end{array}$$

$$= \left(\begin{array}{ccc|c} 0 & 0 & 1 & 4 \\ 1 & 0 & 1 & 5 \\ 0 & 1 & 0 & 1 \end{array}\right) (2)-(1) \quad = \left(\begin{array}{ccc|c} 0 & 0 & 1 & 4 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \end{array}\right)$$

$$\beta = \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix}$$

Найденная модель $f(x) = 1 - x + 4x^2$

Модель ридж-регрессии с параметром регуляризации $\lambda = 1$

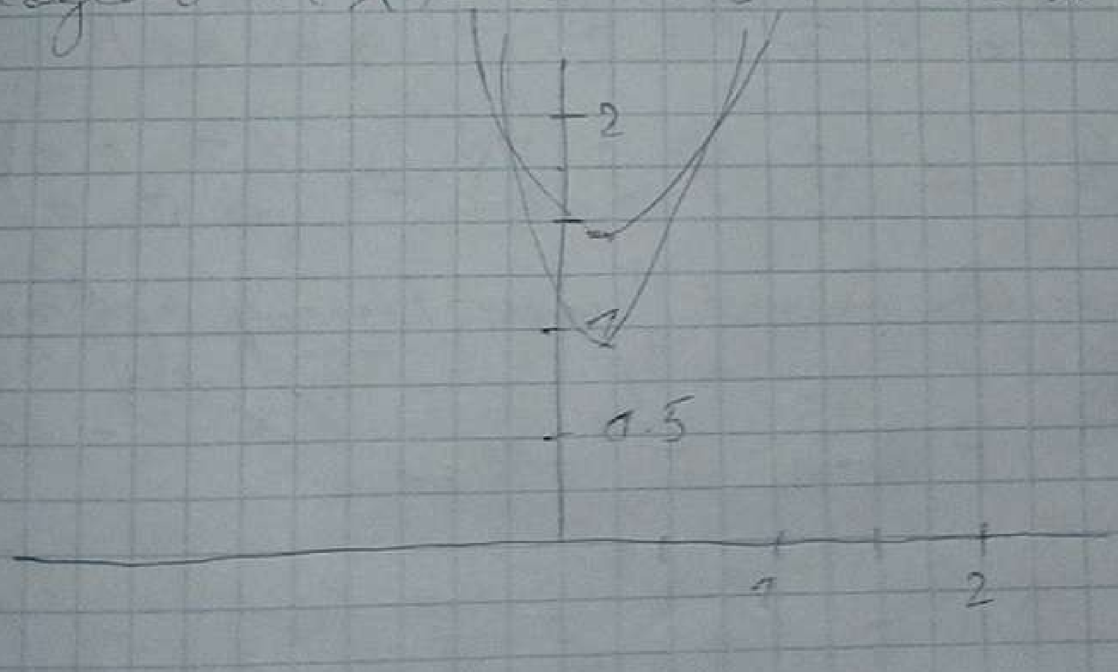
$$(X^T X + \lambda I) = \begin{pmatrix} 5 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} =$$

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

Из $(X^T X + \lambda I)\beta = X^T Y$, найдем β

$$\begin{pmatrix} 6 & 1 & 3 \\ 1 & 4 & 1 \\ 3 & 1 & 4 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} 16 \\ 2 \\ 14 \end{pmatrix} \Rightarrow \beta = \begin{pmatrix} 1.5 \\ -0.5 \\ 2.5 \end{pmatrix}$$

Модель $f(x) = 1.5 - 0.5x + 2.5x^2$



№9

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

Вероятности классов

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

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

Выборочные матрица ковариации для каждого класса

$$\hat{\Sigma}_j = \frac{1}{N_j - 1} \sum_{y^{(i)}=j} (x^{(i)} - \hat{\mu}_j) \cdot (x^{(i)} - \hat{\mu}_j)^T$$

$$\hat{\Sigma}_0 = \frac{1}{4} \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$$

$$\hat{\Sigma}_1 = \frac{1}{2} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$$

Оцениваем матрицу ковариации

$$\hat{\Sigma} = \frac{1}{N-K} \sum_k \sum_{y^{(i)}=k} (x^{(i)} - \hat{\mu}_k) \cdot (x^{(i)} - \hat{\mu}_k)^T = \frac{1}{6} \begin{pmatrix} 6 & 3 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0.5 \\ 0.5 & \frac{2}{3} \end{pmatrix}$$

Дискриминантные ф-ии

$\delta_j(x) = x^T \hat{\Sigma}^{-1}\hat{\mu}_j - \frac{1}{2}\hat{\mu}_j^T \hat{\Sigma}^{-1}\hat{\mu}_j +$
$+ \ln \hat{P}_r\{y=j\}$

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

$\hat{\Sigma}^{-1} = \begin{pmatrix} \frac{8}{5} & -\frac{6}{5} \\ -\frac{6}{5} & \frac{12}{5} \end{pmatrix}$

$\delta_0(x) = (x_1\ x_2) \begin{pmatrix} 8/5 & -6/5 \\ -\frac{6}{5} & \frac{12}{5} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \frac{1}{2}(1\ 0) \cdot$

$\cdot \begin{pmatrix} \frac{8}{5} & -\frac{6}{5} \\ -\frac{6}{5} & \frac{12}{5} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \ln\left(\frac{5}{8}\right) =$

$= \frac{8}{5}x_1 - \frac{6}{5}x_2 - \frac{4}{5} + \ln\left(\frac{5}{8}\right)$

$\delta_1(x) = \frac{18}{5}x_1 - \frac{6}{5}x_2 - \frac{24}{5} + \ln\frac{3}{8}$

Прямая $\delta_0 = \delta_1$ — разделяющая поверхность

$\frac{8}{5}x_1 - \frac{6}{5}x_2 - \frac{4}{5} + \ln\left(\frac{5}{8}\right) =$

$= \frac{18}{5} x_1 - \frac{6}{5} x_2 - \frac{24}{5} + \ln \frac{3}{8}$

$2x_1 - 4 + \ln \frac{3}{5} = 0$

$g_j^2(x) = -\frac{1}{2} \ln(\det \hat{\Sigma}_j) - \frac{1}{2} (x - \hat{\mu}_j)^T$

$\cdot \hat{\Sigma}_j^{-1} (x - \hat{\mu}_j)^T + \ln(\hat{P}\{y = j\})$

$g_0^2(x) = -x_1^2 - 2x_2^2 + 2x_1x_2 + 2x_1 -$

$- 2x_2 - 1 + \frac{\ln 4}{2} + \ln \frac{5}{8}$

$g_1^2(x) = -\frac{3}{2} \ln(\frac{3}{4} - 2x_1^2 - 2x_2^2 +$

$+ 2x_1x_2 + 10x_1 - 2x_2 - 14 + 3\ln \frac{3}{8}$

Разделяющая поверхность порядка

$g_0^2(x) = g_1^2(x)$

$x_1^2 + 4x_2 - 4x_1x_2 + 4x_1 + 4x_2 - 11 +$

$+ \frac{3}{2} \ln 3 - 3 \ln 5 = 0$

№ 15

x	0	0	1	1	1	0	0	1	1	1	0
x_2	0	1	0	0	1	1	1	1	1	1	1
y	0	0	0	0	0	0	1	1	1	1	1

$Pr(Y=0 \mid x_1=1, x_2=1) =$

$$= \frac{Pr(Y=0)\, Pr(x_1=1 \mid Y=0)\, Pr(x_2=1 \mid Y=0)}{Pr(x_1=1, x_2=1)}$$

$$Pr(Y=1 \mid x_1=1, x_2=1) = \frac{Pr(Y=1)\, Pr(x_1=1 \mid Y=1)}{Pr(x_1=1, x_2=1)}$$

$\cdot\ Pr(X_2=1 \mid Y=1)$

Оценивание априорные вер-ти

$\hat{Pr}(Y=0) = \frac{1}{2}$ $\quad \hat{Pr}(Y=1) = \frac{1}{2}$

Оценивание условные вероятност.

$Pr(X_1=0 \mid Y=0) = \frac{3}{5}$ $\quad Pr(X_1=1 \mid Y=0) = \frac{2}{3}$

$Pr(X_1=0 \mid Y=1) = \frac{2}{5}$ $\quad Pr(X_1=1 \mid Y=1) = \frac{3}{5}$

$Pr(X_2=0 \mid Y=0) = \frac{2}{3}$ $\quad Pr(X_2=1 \mid Y=0) = \frac{3}{5}$

$Pr(X_2=0 \mid Y=1) = 0$ $\quad Pr(X_2=1 \mid Y=1) = 1$

$Pr(X_1=1, X_2=1) = Pr\{Y=0\} \cdot Pr(X_1=1 \mid Y=0)$

$\cdot\ Pr(X_2=1 \mid Y=0) + Pr\{Y=1\} \cdot Pr\{Y=1\} \cdot$

$\cdot\ Pr(X_1=1 \mid Y=1) \cdot Pr(X_2=1 \mid Y=1) = \frac{27}{50}$

$$Pr(Y=0 \mid X_1=1, X_2=1) = \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \cdot \frac{8}{27} = \frac{2}{27}$$

$$Pr(Y=1 \mid X_1=1, X_2=1) = \frac{1}{2} \cdot \frac{3}{4} \cdot 1 \cdot \frac{20}{27} = \frac{5}{18}$$