Задание 1

№1 1) Док-ть, что если $a \in \mathbb{R}^n$, $x \in \mathbb{R}^n$, то $\frac{\partial(a^T x)}{\partial x} = a$;

док-во: $a^T x = (a_1\, a_2 \dots a_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = a_1 x_1 + a_2 x_2 + \dots + a_n x_n$ (1×1)

$$\frac{\partial(a^T x)}{\partial x} = \begin{pmatrix} \frac{\partial(\sum_i a_i x_i)}{\partial x_1} \\ \frac{\partial(\sum_i a_i x_i)}{\partial x_n} \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = a$$

2) Док-ть, что если $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$, то $\frac{\partial(Ax)}{\partial x} = A$;

док-во:

$$A = \underbrace{\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}}_{m \times n}, \quad x = \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}}_{n \times 1}, \quad Ax = \begin{bmatrix} a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n \\ a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n \\ \vdots \\ a_{m1} x_1 + a_{m2} x_2 + \dots + a_{mn} x_n \end{bmatrix}$$

$$\frac{\partial(Ax)}{\partial x} = \begin{bmatrix} \frac{\partial(\sum_i a_{1i} x_i)}{\partial x_1} & \frac{\partial(\sum_i a_{1i} x_i)}{\partial x_2} & \dots & \frac{\partial(\sum_i a_{1i} x_i)}{\partial x_n} \\ \vdots & & & \vdots \\ \frac{\partial(\sum_i a_{mi} x_i)}{\partial x_1} & \frac{\partial(\sum_i a_{mi} x_i)}{\partial x_2} & \dots & \frac{\partial(\sum_i a_{mi} x_i)}{\partial x_n} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = A$$

3) Док-ть, что если $A \in \mathbb{R}^{n \times n}$, $x \in \mathbb{R}^n$, то $\frac{\partial(x^T A x)}{\partial x} = (A + A^T)x$, в частности, если $A^T = A$, то $\frac{\partial(x^T A x)}{\partial x} = 2Ax$

док-во: $x^T A = (x_1 \dots x_n)\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} = \left(\sum_{i=1}^n a_i x_i \quad \dots \quad \sum_{i=1}^n a_{in} x_n\right)$

$$x^T A x = \left(\sum_{i=1}^n a_{i1} x_i\right) x_1 + \dots + \left(\sum_{i=1}^n a_{in} x_i\right) x_n$$

$$\frac{\partial(x^T A x)}{\partial x} = \begin{pmatrix} 2a_{11} x_1 + \sum_{i=2}^n a_{i1} x_i + \dots + a_{1n} x_n \\ \vdots \\ a_{n1} x_n + \dots + 2a_{nn} x_n + \sum_{i=1}^{n-1} a_{in} x_i \end{pmatrix}$$

в общем виде $\frac{\partial(x^T A x)}{\partial x_k} = 2a_{kk} x_k + \sum_{j \neq k}^n x_j (a_{jk} + a_{kj})$

эл-ты $\begin{cases} x_j (a_{jk} + a_{kj}), & j \neq k \\ 2 a_{kk} x_j, & j = k \end{cases}$ $(*)$

$$(A + A^T)x = \begin{pmatrix} 2a_{11} & a_{21} + a_{12} & \dots & a_{n1} + a_{1n} \\ \vdots & & & \\ a_{1n} + a_{n1} & \dots & & 2a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

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

4) Док-ть, что если $x \in \mathbb{R}^n$, то $\frac{\partial \|x\|^2}{\partial x} = 2x$

док-во:

$$\|x\|^2 = x_1^2 + \ldots + x_n^2, \text{ тогда } \frac{\partial \|x\|^2}{\partial x} = \begin{pmatrix} 2x_1 \\ \vdots \\ 2x_n \end{pmatrix} = 2x$$

5) Док-ть, что если g - скалярная ф-я и под $g(x)$ понимается применение ф-ии g к каждой компоненте вектора $x \in \mathbb{R}^n$, то

$\frac{\partial g(x)}{\partial x} = diag(g'(x))$, где $diag(a)$ - диагональная матр. с диаг-ю a

док-во:

$$g(x) = g\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \frac{\partial g(x)}{\partial x} = \frac{\partial g(x)}{\partial x_1} + \ldots + \frac{\partial g(x)}{\partial x_n}$$

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

6) Док-ть, что если $h: \mathbb{R}^n \to \mathbb{R}^m$, $g: \mathbb{R}^m \to \mathbb{R}^p$, $x \in \mathbb{R}^n$, то

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

док-во:

$$\frac{\partial g(h(x))}{\partial h} \frac{\partial h(x)}{\partial x} = \begin{pmatrix} \frac{\partial g_1(h(x))}{\partial h_1} & \cdots & \frac{\partial g_1(h(x))}{\partial h_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_p(h(x))}{\partial h_1} & \cdots & \frac{\partial g_p(h(x))}{\partial h_m} \end{pmatrix} \begin{pmatrix} \frac{\partial h_1(x)}{\partial x_1} & \cdots & \frac{\partial h_1(x)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_m(x)}{\partial x_1} & \cdots & \frac{\partial h_m(x)}{\partial x_n} \end{pmatrix} =$$

$$= \begin{pmatrix} \sum_{i=1}^m \frac{\partial g_1(h(x))}{\partial h_i} \frac{\partial h_i(x)}{\partial x_1} & \cdots & \sum_{i=1}^m \frac{\partial g_1(h(x))}{\partial h_i} \frac{\partial h_i}{\partial x_n} \\ \vdots & \ddots & \\ \sum_{i=1}^m \frac{\partial g_p(h(x))}{\partial h_i} \frac{\partial h_i(x)}{\partial x_1} & \cdots & \sum_{i=1}^m \frac{\partial g_p(h(x))}{\partial h_i} \frac{\partial h_i}{\partial x_n} \end{pmatrix} = \frac{\partial g(h(x))}{\partial x}.$$

Задание 3

№3 Дана обучающая выборка

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

- методом наименьш. квадратов построить модель вида $f(x) = \beta_0 + \beta_1 x + \beta_2 x^2$
- построить модель того же вида методом ридж-регрессии с параметром регуляризации $\lambda = 1$.

Решение: 1) МНК-метод:

Составляем матрицу X и вектор y:

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

Решим систему нормальных ур-й $X^T X \beta = X^T y$, найдем β:

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

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

$$\begin{bmatrix} 5 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 3 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 16 \\ 2 \\ 14 \end{bmatrix} \Rightarrow \beta = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$$

Т.о., модель: $1 - x + 4x^2$

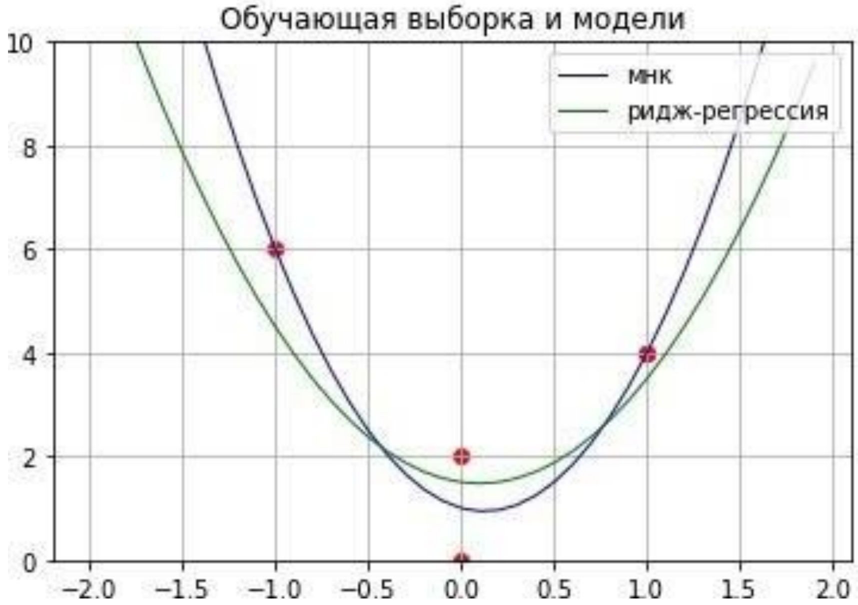
2) ридж-регрессия: для $\lambda = 1$ получаем

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

Решая регуляризованную систему норм. ур-й $(X^T X + \lambda I)\beta = X^T y$, находим:

$$\begin{bmatrix} 6 & 1 & 3 \\ 1 & 4 & 1 \\ 3 & 1 & 4 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 16 \\ 2 \\ 14 \end{bmatrix} \Rightarrow \beta = \begin{bmatrix} 3/2 \\ -1/2 \\ 5/2 \end{bmatrix}$$

Т.о., модель: $3/2 - 1/2\, x + 5/2\, x^2$


Обучающая выборка и модели
мнк
ридж-регрессия
10
8
6
4
2
0
−2.0
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0

Задание 9

№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) Методом линейного дискр. анализа для каждого класса построить дискр. ф-ю и записать ур-е разделяющей пов-ти

2) Методом квадр. дискр. анализа построить дискр. ф-и. Изобразить точки и разделяющие поверхности.

Решение:

1) $\hat{Pr}\{Y=0\}=\frac{5}{8}$, $\hat{Pr}\{Y=1\}=\frac{3}{8}$, $\hat{\mu}_0=\begin{pmatrix}1\\0\end{pmatrix}$, $\hat{\mu}_1=\begin{pmatrix}3\\1\end{pmatrix}$

$$\hat{\Sigma}_0=\frac{1}{4}\left(\begin{pmatrix}-1\\-1\end{pmatrix}(-1\ -1)+\begin{pmatrix}0\\0\end{pmatrix}(0\ 0)+\begin{pmatrix}-1\\0\end{pmatrix}(-1\ 0)+\begin{pmatrix}1\\0\end{pmatrix}(1\ 0)+\begin{pmatrix}1\\1\end{pmatrix}(1\ 1)\right)=$$

$$=\frac{1}{4}\left(\begin{pmatrix}1&1\\1&1\end{pmatrix}+\begin{pmatrix}0&0\\0&0\end{pmatrix}+\begin{pmatrix}1&0\\0&0\end{pmatrix}+\begin{pmatrix}1&0\\0&0\end{pmatrix}+\begin{pmatrix}1&1\\1&1\end{pmatrix}\right)=\frac{1}{4}\begin{pmatrix}4&2\\2&2\end{pmatrix}$$

$$\hat{\Sigma}_1=\frac{1}{2}\left(\begin{pmatrix}-1\\-1\end{pmatrix}(-1\ -1)+\begin{pmatrix}1\\0\end{pmatrix}(1\ 0)+\begin{pmatrix}0\\1\end{pmatrix}(0\ 1)\right)=\frac{1}{2}\left(\begin{pmatrix}1&1\\1&1\end{pmatrix}+\begin{pmatrix}1&0\\0&0\end{pmatrix}+\begin{pmatrix}0&0\\0&1\end{pmatrix}\right)=$$

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

$$\hat{\Sigma}=\frac{1}{6}\left(\begin{pmatrix}4&2\\2&2\end{pmatrix}+\begin{pmatrix}2&1\\1&2\end{pmatrix}\right)=\frac{1}{6}\begin{pmatrix}6&3\\3&4\end{pmatrix}=\begin{pmatrix}1&0{,}5\\0{,}5&2/3\end{pmatrix}$$

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

$$\left.\begin{aligned}\delta_0(x)&=\frac{8}{5}x_1-\frac{6}{5}x_2-\frac{4}{5}+\ln\frac{5}{8}\\ \delta_1(x)&=\frac{18}{5}x_1-\frac{6}{5}x_2-\frac{24}{5}+\ln\frac{3}{8}\end{aligned}\right\}\text{ линейные дискр. ф-ии}$$

$\delta_0(x)=\delta_1(x)$ – разделяющая поверхность

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

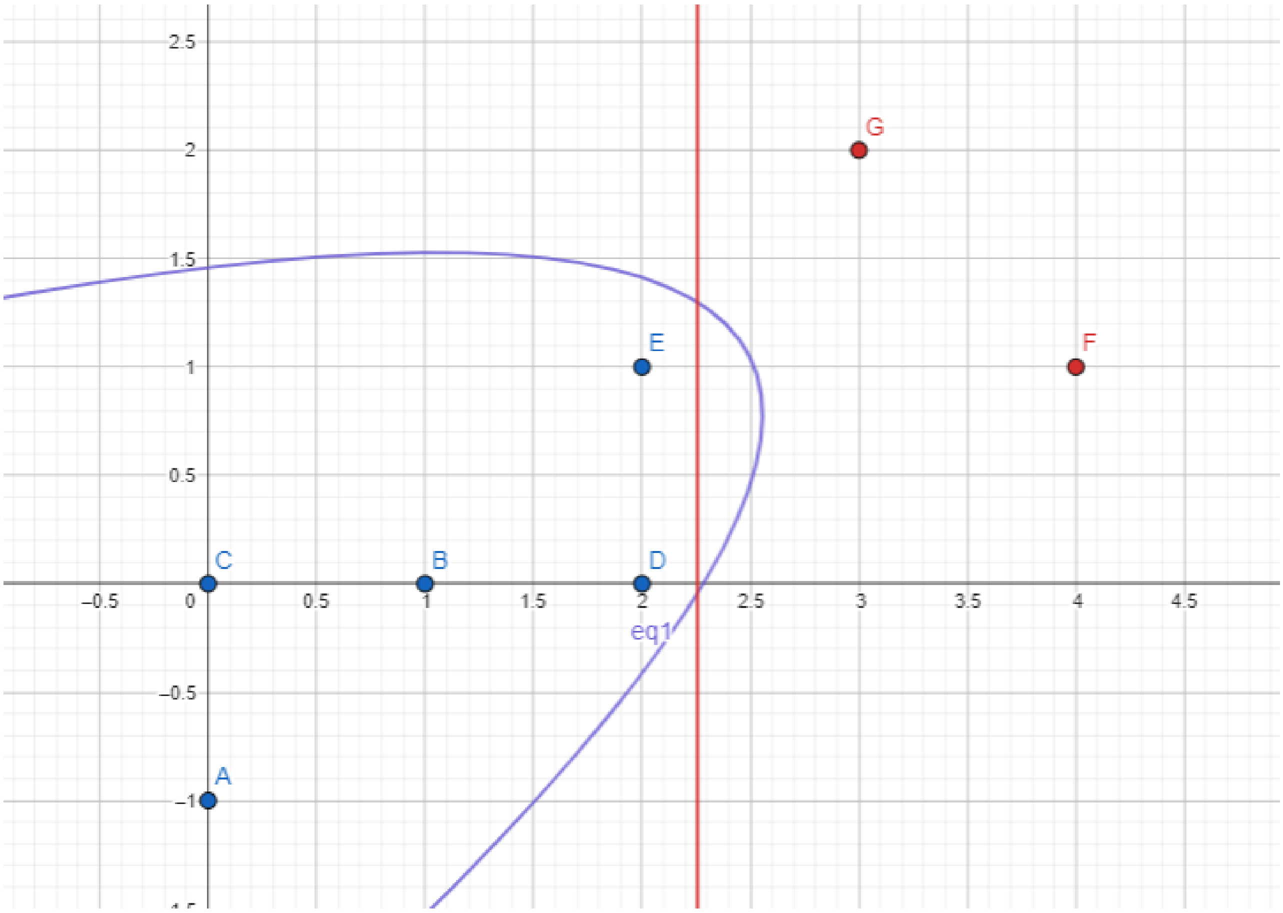
2) $$\delta_0(x)=-\frac{1}{2}\ln\frac{1}{4}\left(-\frac{1}{2}\right)\left((x_1-1\ \ x_2)\begin{pmatrix}2&-2\\-2&4\end{pmatrix}\begin{pmatrix}x_1-1\\x_2\end{pmatrix}\right)\ln\frac{5}{8}=$$

$$=\cancel{1}-x_1^2-2x_2^2+2x_1x_2+2x_1-2x_2-\cancel{1}+\ln\frac{5}{8}$$

$$\delta_1(x)=-\frac{1}{2}\ln\frac{3}{4}+\left(-\frac{2}{3}\right)x_1^2-\frac{2}{3}x_2^2+\frac{2}{3}x_1x_2+\frac{10}{3}x_1-\frac{2}{3}x_2-\frac{14}{3}+\ln\frac{3}{8}$$

$\delta_0(x)=\delta_1(x)$ – разделяющая поверхность

$$\boxed{\frac{1}{3}x_1^2+\frac{4}{3}x_2^2-\frac{4}{3}x_1x_2+\frac{4}{3}x_1+\frac{4}{3}x_2-\frac{14}{3}+\frac{1}{2}\ln\frac{12}{25}=0}$$



Задание 15

№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 $Pr(Y=0 \mid X_1=1, X_2=1)$, $Pr(Y=1 \mid X_1=1, X_2=1)$

Решение:

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

$\hat{Pr}\{X_1=1 \mid Y=0\} = \frac{2}{5}$, $\hat{Pr}\{X_1=1 \mid Y=1\} = \frac{3}{5}$, $\hat{Pr}\{X_2=1 \mid Y=0\} = \frac{3}{5}$

$\hat{Pr}\{X_2=1 \mid Y=1\} = 1$

$$Pr\{Y=0 \mid X_1=1, X_2=1\} = \frac{Pr\{X_1=1 \mid Y=0\}\, Pr\{X_2=1 \mid Y=0\}\, Pr\{Y=0\}}{Pr\{X_1=1, X_2=1\}} = \frac{\frac{2}{5}\cdot\frac{3}{5}\cdot\frac{1}{2}}{\frac{21}{50}} = \frac{2}{7}$$

$$Pr\{Y=1 \mid X_1=1, X_2=1\} = \frac{Pr\{X_1=1 \mid Y=1\}\, Pr\{X_2=1 \mid Y=1\}\, Pr\{Y=1\}}{Pr\{X_1=1, X_2=1\}} = \frac{\frac{3}{5}\cdot 1\cdot\frac{1}{2}}{\frac{21}{50}} = \frac{5}{7}.$$

Ответ: $\frac{2}{7}$, $\frac{5}{7}$.