№ 1

ML №2

Док-ть:

1) $$\frac{\partial (a^T x)}{\partial x} = a$$

$$\rightarrow \frac{\partial (a^T x)}{\partial x} = \frac{\partial (a^T)}{\partial x} \cdot x + \frac{a^T \partial x}{\partial x} =$$

$$= 0 \cdot x + \frac{\partial x}{\partial x} \cdot a^T = a$$

2) $$\frac{\partial (Ax)}{\partial x} = A$$

$$\rightarrow \frac{\partial (Ax)}{\partial (x)} = \frac{\partial A}{\partial x} \cdot x + A \cdot \frac{\partial x}{\partial x} = 0 \cdot x + A \cdot 1 = A$$

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

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

$$= \Big(\frac{\partial x^T}{\partial x} A + x^T \underbrace{\frac{\partial A}{\partial x}}_{=0}\Big) x + x^T A = \Big(\frac{\partial x}{\partial x}\Big)^T A x + x^T A =$$

$$= A \cdot x + A^T \cdot x = (A + A^T) x$$

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

$$(A + A^T) x = (A + A) x = 2Ax$$

4) $$\frac{\partial \|x\|^2}{\partial x} = 2x$$

$$\rightarrow \frac{\partial \|x\|^2}{\partial x} = \frac{\partial (x, x)}{\partial x} = \frac{\partial x}{\partial x} \cdot x + x \frac{\partial x}{\partial x} = x + x = 2x$$

5) g - ск. ф. , $g(x)$ прим. к каждой компоненте

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

$$\longrightarrow \frac{\partial g(x)}{\partial x} = \left(\frac{\partial g}{\partial x_1} \quad \frac{\partial g}{\partial x_2} \quad \frac{\partial g}{\partial x_3} \quad \dots \quad \frac{\partial g}{\partial x_n}\right)$$ - градиент

$$\frac{\partial g(x)}{\partial x} = \left(\frac{\partial g}{\partial x_1} \quad \frac{\partial g}{\partial x_2} \quad \frac{\partial g}{\partial x_3} \quad \dots \quad \frac{\partial g}{\partial x_n}\right) \cdot E = \begin{pmatrix} \frac{\partial g}{x_1} & 0 & \dots & 0 \\ 0 & \frac{\partial g}{x_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \frac{\partial g}{x_n} \end{pmatrix}$$

$$= diag(g(x))$$

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

по правилу дифф. сложной функции: $(f(g(x)))' = \frac{\partial f}{\partial g} \cdot \frac{\partial g}{\partial x}$

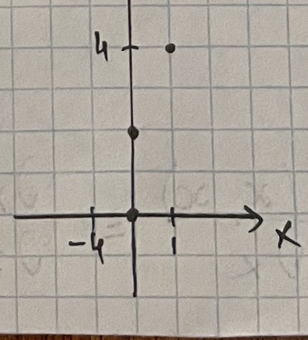
$$f(g(h(x)))' = \frac{\partial g(h(x))}{\partial h} \cdot \frac{\partial h(x)}{\partial x}$$

№ 3

Выборка:

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

1) изображ. точки:



2) Построить модель методом наименьш. квадратов и ~~ногр~~ построить график этой функции.

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

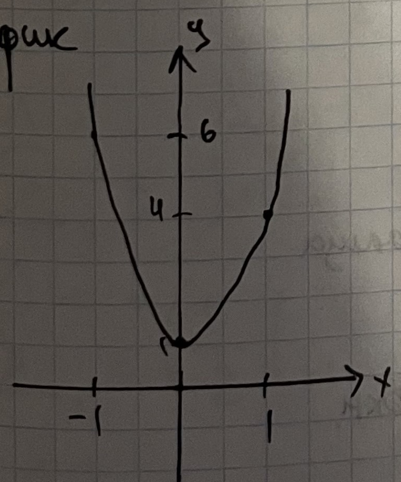
$$X^T X = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & -1 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 3 \end{pmatrix}$$

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

Решаем систему:

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

График



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

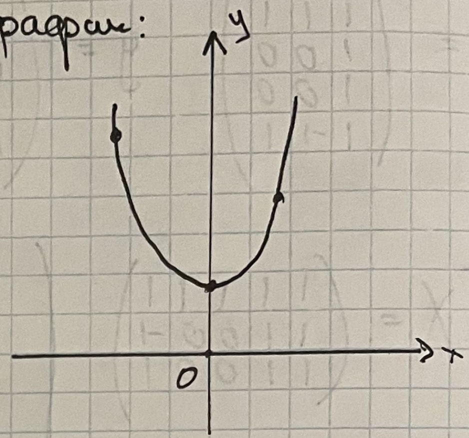
3) метод ридж-регрессии, с параметром $\lambda = 1$ + график

~~это~~ система: $(X^T X + \lambda I)\beta = X^T y$

из пункта 2:

$$\begin{pmatrix} 5+1 & 1 & 3 \\ 1 & 3+1 & 1 \\ 3 & 1 & 3+1 \end{pmatrix} \cdot \beta = \begin{pmatrix} 16 \\ 2 \\ 14 \end{pmatrix}$$

Графи:



$$\begin{pmatrix} 6 & 1 & 3 \\ 1 & 4 & 1 \\ 3 & 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 16 \\ 2 \\ 14 \end{pmatrix}$$

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

Тогда: $f(x) = \frac{3}{2} - \frac{1}{2}x + \frac{5}{2}x^2$

№9 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 дискриминантного анализа

а) построить дискр. функцию

б) записать уравнение раздел. поверхн.

$\hat{P}_F \{Y = 0\} = \frac{5}{8} \Rightarrow \{Y = 1\} = \frac{3}{8}$

$$\hat{\mu}_0 = \sum_{i=0} \frac{x_i}{N} = \frac{1+2+2}{5} = 1$$
$$= \frac{-1+1}{5} = 0 \Rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\hat{\mu}_1 = \frac{2+4+3}{\cancel{8}\,3} = 3$$
$$= \frac{1+2}{3} = 1 \Rightarrow \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

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

$$= \frac{1}{4} \begin{pmatrix} -1 & 0 & -1 & 1 & 1 \\ -1 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & -1 \\ 0 & 0 \\ -1 & 0 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}$$

$$\hat{\Sigma}_1 = \text{аналогично} \ldots = \frac{1}{2} \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Тогда:

$$\hat{\Sigma} = \frac{1}{N-K} \sum_K \sum_{y(i)=K} (x^{(i)} - \hat{\mu}_K)(x^{(i)} - \hat{\mu}_K)^T = \frac{1}{6} \left(\begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \right) =$$

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

$$\hat{\Sigma}_0^{-1} = \cancel{\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}^{-1}} = \cancel{\begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix}} \left(\frac{1}{4} \cdot \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} \right)^{-1} = 4 \cdot \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix} = \begin{pmatrix} 2 & - \\ -2 & \end{pmatrix}$$

$$\hat{\Sigma}_1^{-1} = \left(\frac{1}{2} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \right)^{-1} = 2 \cdot \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} \frac{4}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{4}{3} \end{pmatrix}$$

$$\hat{\Sigma}^{-1} = \left(\frac{1}{6} \begin{pmatrix} 6 & 3 \\ 3 & 4 \end{pmatrix} \right)^{-1} = 6 \cdot \begin{pmatrix} \frac{4}{15} & -\frac{1}{5} \\ -\frac{1}{5} & \frac{2}{5} \end{pmatrix} = \begin{pmatrix} \frac{8}{5} & -\frac{6}{5} \\ -\frac{6}{5} & \frac{-12}{5} \end{pmatrix}$$

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

$$\begin{pmatrix}x_1\\x_2\end{pmatrix}\begin{pmatrix}1.6&-1.2\\-1.2&2.4\end{pmatrix}\cdot\begin{pmatrix}1\\0\end{pmatrix}-\frac{1}{2}(1\ \ 0)\begin{pmatrix}1.6&-1.2\\-1.2&2.4\end{pmatrix}\cdot\begin{pmatrix}1\\0\end{pmatrix}+\ln\frac{5}{8}=$$

$$=1.6x_1-1.2x_2-0{,}8+\ln\frac{5}{8}$$

$\delta_1(x)=$ аналогично $\ldots = \begin{pmatrix}x_1\\x_2\end{pmatrix}(\ldots)\begin{pmatrix}3\\1\end{pmatrix}-\frac{1}{2}(3\ \ 1)(\ldots)\begin{pmatrix}3\\1\end{pmatrix}+\ln\frac{3}{8}=$

$$=3.6x_1-1.2x_2-4.8+\ln\frac{3}{8}$$

б) Поверхность разделяющая:

$\cancel{4x_1-10x_2+15=0}$

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

$\Downarrow$

$$1.6x_1-\cancel{1.2x_2}-0{,}8+\ln\frac{5}{8}=3.6x_1-\cancel{1.2x_2}-4.8+\ln\frac{3}{8}$$

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

2) методом квадратичного дискр. анализа

а) построить дискр. функции

б) изобраз точки и разд. поверх

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

$$=\underbrace{-\frac{1}{2}\ln(1)}_{0}-\frac{1}{2}(x_1-1\ \ \ x_2-0)\begin{pmatrix}x_1-1\\x_2-0\end{pmatrix}\hat{\Sigma}_0^{-1}+\ln\frac{5}{8}=$$

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

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

$$= (x_1 - 1 \quad x_2) \cdot \begin{pmatrix} -1 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 - 1 \\ x_2 \end{pmatrix} + \ln \frac{5}{8} =$$

$$= (-x_1 + x_2 + 1 \quad x_1 + 2x_2 - 1) \begin{pmatrix} x_1 - 1 \\ x_2 \end{pmatrix} + \ln \frac{5}{8} =$$

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

$$\delta_1(x) = \dots = -\frac{1}{2} \ln (\cancel{0{,}75}\ 1{,}5) - \frac{1}{2}(x_1 - 3 \quad x_2 - 1) \begin{pmatrix} \frac{4}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{4}{3} \end{pmatrix} \cdot$$

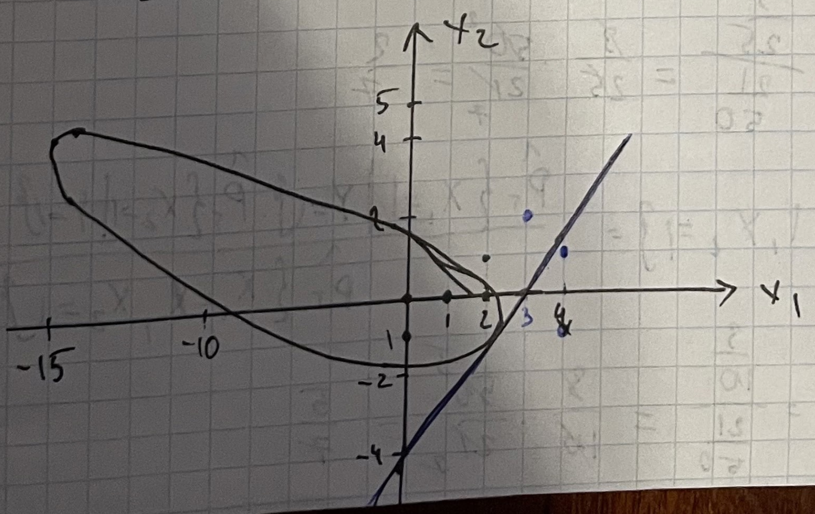
$$\cdot \begin{pmatrix} x_1 - 3 \\ x_2 - 1 \end{pmatrix} + \ln \frac{3}{8} = -\frac{1}{2} \ln(1{,}5) + \ln \frac{3}{8} + (x_1 - 3 \quad x_2 - 1) \begin{pmatrix} -\frac{4}{6} & \frac{2}{6} \\ \frac{2}{6} & -\frac{4}{6} \end{pmatrix} \cdot$$

$$\cdot \begin{pmatrix} x_1 - 3 \\ x_2 - 1 \end{pmatrix} = -\frac{1}{2} \ln(1{,}5) + \ln \frac{3}{8} + \left(-\frac{4}{6}x_1 + \frac{2}{6}x_2 - \frac{10}{6} \quad \frac{2}{6}x_1 - \frac{4}{6}x_2 - \frac{2}{6}\right)$$

$$\cdot \begin{pmatrix} x_1 - 3 \\ x_2 - 1 \end{pmatrix} = -\frac{1}{2} \ln(1{,}5) + \ln \frac{3}{8} - \frac{4}{6}x_1^2 - \frac{4}{6}x_2^2 + \frac{4}{6}x_1x_2 - \frac{10}{6}x_2 + \frac{32}{6}$$

б) разделяющая поверхность: – эллипс

$$-\frac{2}{6}x_1^2 + \frac{16}{6}x_2^2 + \frac{8}{6}x_1x_2 + 2x_1 - \frac{2}{6}x_2 - \frac{38}{6} + \ln \frac{5}{8} - \ln \frac{3}{8} + \frac{1}{2}\ln(1{,}5)$$



???

№ 15

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

$Pr(Y=0 \mid X_1=1, X_2=1)$ – ?

$Pr(Y=1 \mid X_1=1, X_2=1)$ – ?

1) Априорные вероятности:

$$\hat{P}_r\{Y=0\} = \frac{5}{10} = \frac{1}{2} \qquad \hat{P}_r\{Y=1\} = \frac{1}{2}$$

2) Условные вероятности:

$$\hat{P}_r\{X_1=1 \mid Y=0\} = \frac{2}{5} \Rightarrow \hat{P}_r\{X_2=1 / Y=0\} = \frac{3}{5}$$

$$\hat{P}_r\{X_1=1 \mid Y=1\} = \frac{3}{5} \Rightarrow \hat{P}_r\{X_2=1 / Y=1\} = 1$$

3) По основному предположению наивного байесова классификатора:

$$\overset{*}{P}_r\{Y=0 \mid X_1=1, X_2=1\} = \frac{\hat{P}_r\{X_1=1 \mid Y=0\} \cdot \hat{P}_r\{X_2=1 \mid Y=0\} \, P_r\{Y=0\}}{\hat{P}_r\{X_1=1, X_2=1\}} =$$

$$\approx \frac{\frac{2}{5}\cdot\frac{3}{5}\cdot\frac{1}{2}}{\frac{3}{25}+\frac{3}{10}} = \frac{\frac{3}{25}}{\frac{21}{50}} = \frac{3}{25}\cdot\frac{50}{21} = \frac{2}{7}$$

$$P_r\{Y=1 \mid X_1=1, X_2=1\} = \frac{\hat{P}_r\{X_1=1 \mid Y=1\} \cdot \hat{P}_r\{X_2=1 \mid Y=1\} \cdot P_r\{Y=1\}}{\hat{P}_r\{X_1=1, X_2=1\}} =$$

$$\approx \frac{\frac{3}{5}\cdot 1\cdot\frac{1}{2}}{\frac{21}{50}} = \frac{\frac{3}{10}}{\frac{21}{50}} = \frac{3}{10}\cdot\frac{50}{21} = \frac{5}{7}$$