Дз №3 №15

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

$Pr(y=0 \mid X_1=1,\ X_2=1)$

$Pr(y=1 \mid X_1=1,\ X_2=1)$

$\widehat{Pr}\{\cancel{y_0=1}\}$ $\widehat{Pr}\{y=0\} = \frac{1}{2}$ $\widehat{Pr}\{y=1\} = \frac{1}{2}$

$\widehat{Pr}\{X_1=1 \mid y=0\} = \frac{2}{5}$ $\widehat{Pr}\{X_2=1 \mid y=0\} = \frac{3}{5}$

$\widehat{Pr}\{X_1=1 \mid y=1\} = \frac{3}{5}$ $\widehat{Pr}\{X_2=1 \mid y=1\} = 1$

$Pr\{y=0 \mid X_1=1\ X_2=1\} = \widehat{Pr}\{X_1=1 \mid y=0\}$

$\widehat{Pr}\{X_2=1 \mid y=0\} \ / \ (Pr\ X_1=1, X_2=1\}$

$\widehat{Pr}\{y=0\} = \left(\frac{2}{5} \cdot \frac{3}{5} \cdot \frac{1}{2}\right) : \left(\frac{3}{25} + \frac{3}{10}\right)$

$= \frac{3}{25} \cdot \frac{50}{21} = \frac{2}{7}$

$Pr\{y=1 \mid X_1=1, X_2=2\} = Pr\{X=1 \mid y=1\}\ \widehat{Pr}\{X_2=1 \mid y=1\}$

$\cdot\ \widehat{Pr}\{y=1\} = \left(\frac{3}{5} \cdot 1 \cdot \frac{1}{2}\right) : \left(\frac{3}{10} + \frac{3}{25}\right) = \frac{5}{7}$ $Pr\{X_1=1, X_2=1\}$

Ответ: $\frac{2}{7}$; $\frac{5}{7}$

№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

$$\hat{P}_r\{y=0\} = \frac{5}{8} \qquad \hat{P}_r = \{y=1\} = \frac{3}{8}.$$

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

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

$$= \frac{1}{4}\left[\begin{pmatrix} -1 \\ -1 \end{pmatrix}\begin{pmatrix} -1 & -1 \end{pmatrix} + \begin{pmatrix} -1 \\ 0 \end{pmatrix}\begin{pmatrix} -1 & 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \end{pmatrix} + \right.$$

$$\left. + \begin{pmatrix} 1 \\ 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \end{pmatrix}\right] = \frac{1}{4}\left[\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \right.$$

$$\left. + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\right] = \frac{1}{4} \cdot \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}$$

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

$$= \frac{1}{2}\left[\begin{pmatrix}-1\\-1\end{pmatrix}(-1\ -1) + \begin{pmatrix}1\\0\end{pmatrix}(1\ 0) + \begin{pmatrix}0\\1\end{pmatrix}(0\ 1)\right] =$$

$$= \frac{1}{2}\left[\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix} + \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix} + \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}\right] = \frac{1}{2}\cdot\begin{pmatrix}2 & 1\\1 & 2\end{pmatrix}$$

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

$$\hat{\Sigma} = \frac{1}{N-K}\sum_{k}\sum_{y^{i}=k}(x^{(i)} - \hat{\mu}_k)(x^{(i)} - \hat{\mu}_k)^T$$

$$= \frac{1}{6}\left(\begin{pmatrix}4 & 2\\2 & 2\end{pmatrix} + \begin{pmatrix}2 & 1\\1 & 2\end{pmatrix}\right) = \frac{1}{6}\cdot\begin{pmatrix}6 & 3\\3 & 4\end{pmatrix} = \begin{pmatrix}1 & 1/2\\1/2 & 2/3\end{pmatrix}$$

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

$$\hat{\Sigma}^{-1} = \begin{pmatrix}1/6 & -1/2\\-1/2 & 2.4\end{pmatrix}$$

$$\delta_0(x) = x^T + \hat{\Sigma}^{-1}\hat{\mu}_0 \frac{1}{2}\mu_0^T \Sigma^{-1}\hat{\mu}_0 +$$

$$+ \ln \hat{P}_r\{y=0\} = (x_1\ x_2)^\lambda \cdot \begin{pmatrix}1{,}6 & -1{,}2\\-1{,}2 & 2{,}4\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}$$

$$-\left(\frac{1}{2}\right)(1\ 0)\begin{pmatrix}1,6 & -1,2\\ -1,2 & 2,4\end{pmatrix}\cdot\begin{pmatrix}1\\0\end{pmatrix}+\ln\frac{5}{8}=$$

$$=1,6x_1-1,2x_2-\frac{1}{2}\cdot 1,6+\ln\frac{5}{8}=$$

$$=1,6x_1-1,2x_2-0,8+\ln\frac{5}{8}$$ дискриминантная ф-я, для 0

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

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

$$+\ln\frac{3}{8}=3,6x_1-1,2x_2-4,8+\ln\frac{3}{8}$$

дискрим. ф-ия для класса 1

$$\delta_0(x)=\delta_1(x)\to l_1:\ 2x_1-4+\ln\frac{3}{5}$$

ур-е разделяющей поверхности

Квадратичные дискримин. ф-ии

$$\delta_0(x)=-\frac{1}{2}\ln\det\hat{\Sigma}_0-\frac{1}{2}(x-\hat{\mu}_0)^T\Sigma_0^{-1}(x-\hat{\mu}_0)$$

$$+\ln\hat{P}_r\{Y=0\}=-\frac{1}{2}\ln\frac{1}{4}-\frac{1}{2}(x_1-1\ \ x_2)\begin{pmatrix}2 & -2\\ -2 & 4\end{pmatrix}$$

$\begin{pmatrix} x_1-1 \\ x_2 \end{pmatrix} + \ln\frac{5}{8} = \ln\frac{\sqrt{3}}{4} - \frac{1}{2}\left\{(x_1-1\ x_2)\right.$

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

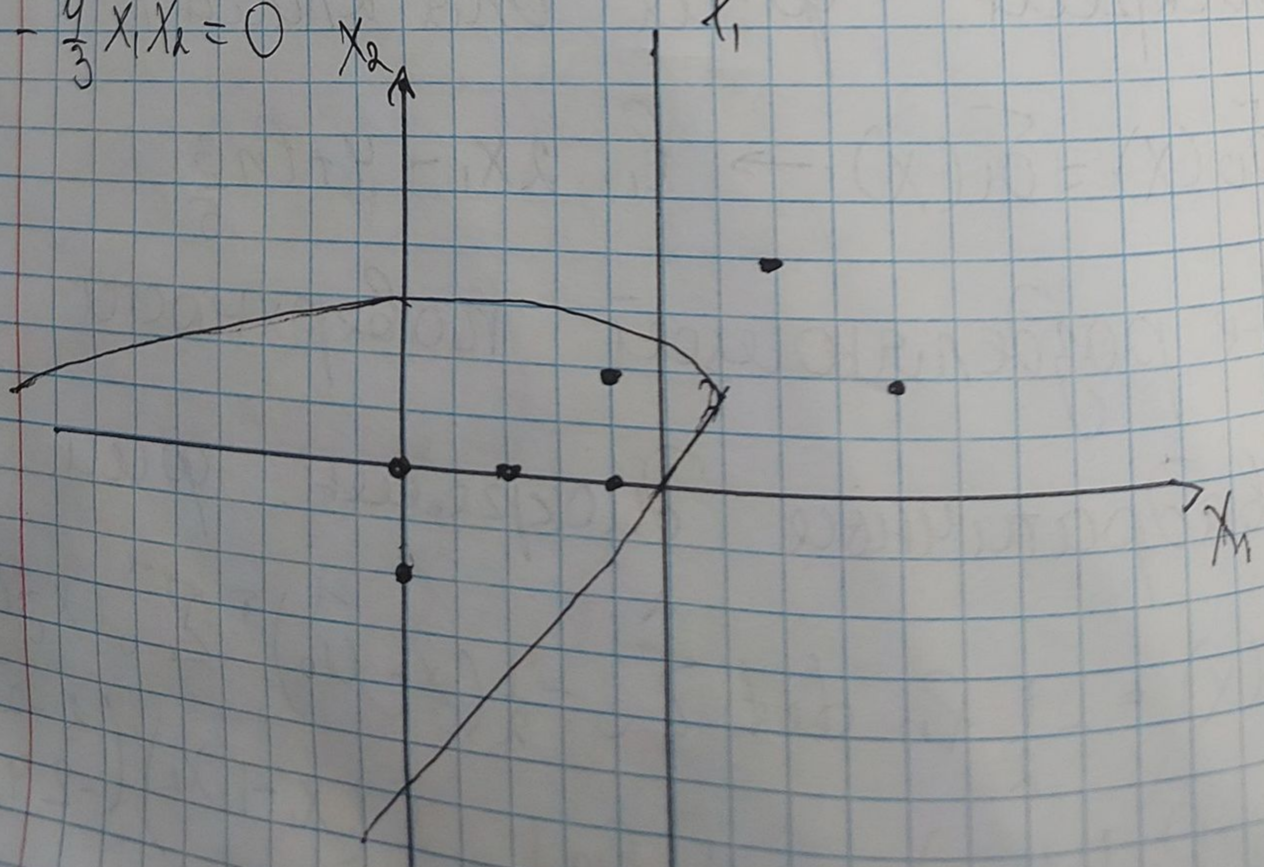
$\delta_1(x) = -\frac{2}{3}x_1^2 + \frac{2}{3}x_1x_2 + \frac{10}{3}x_1 - \frac{2}{3}x_2 - \frac{14}{3} -$

$-\frac{2}{8}x^2 + \ln\frac{2}{5\sqrt{3}})$

Разделяющая поверхность

$\ell_1: \frac{1}{8}x_1^2 + \frac{4}{3}x_1 + \frac{4}{3}x_2 + \frac{4}{3}x_2^2 - \frac{11}{3} + \ln\left(\frac{3}{5\sqrt{3}}\right) -$

$-\frac{4}{3}x_1x_2 = 0$



№19

$$\frac{\partial g_k}{\partial S_1} = g_k(I(k=1) - g_1) \qquad g_k \neq S_1, S_2, \dots, S_k = \frac{e^{S_k}}{\sum e^{S(\,)}}$$

$$\frac{\partial g_k}{\partial S_e} = \frac{h(s)\frac{\partial f(s)}{\partial S_e} - f(s)\cdot\frac{\partial h(s)}{\partial S_e}}{(h(s))^2} \qquad f(s) = e^{S_k} \quad h(s) = \sum e^{S}$$

$$\frac{\partial f(s)}{\partial S_k} = e^{S_k} \qquad \frac{\partial f(s)}{\partial S_e} = 0 \text{ при } l \neq k \qquad \frac{\partial h(s)}{\partial S_e} = e^{S}$$

а) $k = l$: $$\frac{\partial g_k}{\partial S_e} = \sum_{l=1}^{k} \frac{e^{S_l} e^{S_k} - e^{S_k} e^{S_k}}{(\sum e^{S_l})^2} = \frac{e^{S_k}}{\sum_{l=1}^{k} e^{S_l}} \cdot$$

$$\cdot \frac{\sum^{k} e^{S_l} - e^{S_k}}{\sum_{i=1} e^{S_i}} = g_k(1 - g_k)$$

б) $k \neq l$: $$\frac{\partial g_l}{\partial S_e} = \sum \frac{e^{S_l}\cdot 0 - e^{S_k}\cdot e^{S_l}}{(\sum e^{S_l})^2} = -\frac{e^{S_k} e^{S_l}}{\sum e^{S_l}} = -g_k(\,$$

(а, б)

$$\Rightarrow \frac{\partial g_k}{\partial S_e} = g_k(I(k=l) - g_l)$$

2) $\frac{\partial R^{(i)}}{\partial g_k} = -\frac{I(g^{(k)}=k)}{g_k}$ $R^{(i)} = -\sum I(y^{(i)}=k)$

$\frac{\partial R^{(i)}}{\partial g_k} = -\sum_{j=1}^{K} I(y^{(i)}=j\} \frac{\partial}{\partial g_k} \ln g_i \cdot \frac{\partial}{\partial g_2} \cdot \frac{\partial}{\partial g_k}$

$\ln g_j \cdot = \frac{1}{g} \cdot \frac{\partial g_j}{\partial g_k}$

$k=j$; $\frac{\partial g_k}{\partial g_n} = 1$ $k \neq j$ $\frac{\partial g_i}{\partial g_k} = 0 \Rightarrow \frac{\partial R^{(i)}}{\partial g_k} = -\frac{I(y^{(i)}=k)}{g_k}$

3) $\frac{\partial R^{(i)}}{\partial S_e} = g_e - I(e = y^{(i)})$

$\frac{\partial R^{(i)}}{\partial S_e} = -\sum I(y^{(i)}=k) \frac{\partial}{\partial S_1} \ln g_k(S_1, S_2, ..., S_n)$

$\frac{\partial \ln g_k}{\partial S_e} = \frac{1}{g_k} \frac{\partial g_k}{\partial S_e} \Rightarrow \frac{\partial R^{(i)}}{\partial S_e} = -\sum_{k=1}^{K} I(y^{(i)}=k)\frac{1}{g_k}$

$\frac{\partial g_k}{\partial S_e}$

№20 $\frac{\partial R^{(i)}}{\partial S_k} = g_k - y_k^{(i)}$

$g_k(S_k) = \frac{1}{1+e^{S_k}}$ $R^{(i)} = -\sum(\cdot y_k^{(i)} \ln g_k + (1 - g_k$

$\ln(1-g_k)$

$\frac{\partial g_k}{\partial S_k} = g_k(1-g_k)$ $\frac{\partial}{\partial S_k}\Big(-\big(y_k^{(i)} \ln g_k + (1-y_k^{(i)})$

$\ln(1-g_k)\big)\Big) = -\frac{\partial}{\partial S_k}(y_k^{(i)} \ln g_k) - \frac{\partial}{\partial S_k}\big((1-y_k^{(i)})\ln(1-g_k$

$\frac{\partial}{\partial S_k}\big((1-y_k^{(i)})\ln(1-g_k)\big) = -(1-y_k^{(i)})\frac{1}{1-g_k}$

$\Big(\cdot\frac{\partial g_k}{\partial S_k}\Big) = (1-y_k^{(i)})\frac{g_k(1-g_k)}{1-g_k} = (1-y_k^{(i)})g_k$

$\frac{\partial R^{(i)}}{\partial S_k} = \big(-y_k^{(i)}(1-g_k) + (1-y_k^{(i)})g_k\big) = \big(-y_k^{(i)} +$

$+ y_k^{(i)} g_k + g_k - y_k^{(i)} g_k\big) = g_k - y_k^{(i)}$

k) $g_k = \dfrac{e^{S_k}}{h(S)}$ $h(S) = \sum\limits_{e=1} e^{Se}$

$$\frac{\partial g_k}{\partial S_e} = \frac{\partial}{\partial S_e}\left(\frac{e^{S_k}}{h(S)}\right) = \frac{e^{S_k}\delta_{ke}\, h(s) - e^{S_k}\frac{\partial h(s)}{\partial Se}}{h^2(S)}$$

$$\delta_{ke} = \begin{cases} 1, & \text{если } k = e \\ 0, & \text{если } k \neq e \end{cases}$$

$$\frac{\partial h(S)}{\partial Se} = e^{Se} \qquad \frac{\partial g_k}{\partial Se} = \frac{e^{S_k}\delta_{ke} h(S) - e^{S_k}e^{Se}}{h^2(S)} = \frac{e^{S}}{2}$$

$$\frac{2\delta_{ke} - e^{S_k}}{2} = g_k(\delta_{ke} - g_e) \Rightarrow \frac{\partial R^{(i)}}{\partial S_e} =$$

$$= -\sum I(y^{(i)} = k)(\delta_{ke} - g.)$$

при $k = e$: $-\sum\limits_k I(y^{(i)} = k)\,\delta_{ke} = -I(y^{(i)} = e)$

$k \neq e$: $\sum I(y^{(i)} = k) g_e \Rightarrow \dfrac{\partial R^{(i)}}{\partial S_e} = g_e I(y^{(i)} =$