№13 $\exists$ K классов

$$g_k(s_1, s_2 \dots s_K) = \frac{e^{s_k}}{\sum_{l=1}^{K} e^{s_l}}$$

В качестве штрафа:

$$R^{(i)} = -\sum_{k=1}^{K} I(y^{(i)} = k) \ln g_k(s_1, s_2, \dots s_K)$$

где $g_k(s_1, s_2 \dots s_K)$ - softmax - ф-я.

Доказать, что:

1) $\frac{\partial g_k}{\partial s_l} = g_k \cdot (I(k=l) - g_l)$

$$\frac{\partial g_k}{\partial s_l} = \frac{\frac{\partial}{\partial s_l}(e^{s_k}) \sum_{l=1}^{K} e^{s_l} - e^{s_k} \cdot \frac{\partial}{\partial s_l}\left(\sum_{l=1}^{K} e^{s_l}\right)}{\left(\sum_{l=1}^{K} e^{s_l}\right)^2} =$$

$$= \frac{I(k=l) e^{s_k} \sum_l e^{s_l} - e^{s_k} e^{s_l}}{\left(\sum_l e^{s_l}\right)^2} = \frac{e^{s_k}}{\sum_l e^{s_l}}$$

$$\left(\left(\frac{I(k=l)\sum e^{s_l}}{\sum e^{s_l}}\right) - \frac{e^{s_l}}{\sum_l e^{s_l}}\right) =$$

$$= g_k \cdot (I(k=l) - g_l) \qquad \square$$

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

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

$\square$

3) $$\frac{\partial R^{(i)}}{\partial s_l} = g_l - I(l = y^{(i)})$$

$$\frac{\partial R^{(i)}}{\partial s_l} = -\sum_k I(y^{(i)}=k) \cdot \frac{\frac{\partial g_k}{\partial s_l}}{g_k} =$$

$$= -\sum_k I(y^{(i)}=k) \cdot (I(k=l) - g_l) =$$

$$= g_l - I(l = y^{(i)})$$

$\square$

№20 ∀ $\mathbb{K}$ задач двуклассовой классификации последний слой нейронной сети вычисляет $\mathbb{K}$ сигмоидальных ф-й

$$g_k(s_k) = \frac{1}{1+e^{-s_k}} \quad (k = 1, 2, \dots \mathbb{K})$$

В качестве штрафа используется:

$$R^{(i)} = -\sum_{k=1}^{\mathbb{K}} \left(y_k^{(i)} \ln g_k + (1 - y_k^{(i)}) \ln(1 - g_k) \right)$$

Доказать, что $\dfrac{\partial R^{(i)}}{\partial s_k} = g_k - y_k^{(i)}$

$$\frac{\partial R^{(i)}}{\partial s_k} = -\left[y_k^{(i)} (1 - g_k) + (1 - y_k^{(i)}) \cdot \frac{-g_k(1-g_k)}{(1-g_k)} \right]$$

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

$$= g_k - y_k^{(i)}$$

$\square$

№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 $P(Y=0 / X_1=1, X_2=1)$ и $P(Y=1 / X_1=1, X_2=1)$

$$P(Y=0 / X_1=1, X_2=1) = \frac{P(Y=0)P(X_1=1 / Y=0)P(X_2=1 / Y=0)}{P(X_1=1, X_2=1)} = \frac{\frac{1}{2}\cdot\frac{2}{5}\cdot\frac{3}{5}}{\frac{4}{10}} = \frac{\frac{6}{50}}{\frac{4}{10}} =$$

$$= \frac{3}{10}$$

$$P(Y=1 / X_1=1, X_2=1) = 0{,}7$$

№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 дискриминантного анализа построить дискриминантную ф-ю и записать уравнение разделяющей поверхности:

Вероятности классов: $P\{Y=0\} = \frac{5}{8}$

$P\{Y=1\} = \frac{3}{8}$

Средние для классов: $\mu_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $\mu_1 = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$

Выборочные матрицы ковариации для каждого класса

$$M_0 = \frac{1}{N_0 - 1} \sum (x^{(i)} - \mu_0)(x^{(i)} - \mu_0)^T = \frac{1}{4} \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}$$

$$M_1 = \frac{1}{2} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Матрица ковариации

$$M = \frac{1}{N-K} \sum_{K} \sum_{y^{(i)}} (x^{(i)} - \mu_K)(x^{(i)} - \mu_K)^T =$$

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

Обратные матрицы

$$M_0^{-1} = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix} \qquad M_1^{-1} = \begin{pmatrix} \frac{8}{5} & -\frac{6}{5} \\ -\frac{6}{5} & \frac{12}{5} \end{pmatrix}$$

Лин. дискриминантные ф-и:

$$\delta_0(x) = x^T M^{-1} \mu_0 - \frac{1}{2} \mu_0^T \cancel{M^{-1}} \mu_0 +$$

$$+ \ln P\{Y=0\} = (x_1 \ x_2) \begin{pmatrix} \frac{8}{5} & -\frac{6}{5} \\ -\frac{6}{5} & \frac{12}{5} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$- \frac{1}{2} (1 \ 0) \begin{pmatrix} \frac{8}{5} & -\frac{6}{5} \\ -\frac{6}{5} & \frac{12}{5} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \ln \frac{5}{8} =$$

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

$$\delta_1(x) = \frac{18}{5} x_1 - \frac{6}{5} x_2 - \frac{29}{5} + \ln \frac{3}{8}$$

Разделяющая поверхность

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

$$\frac{8}{5} x_1 - \cancel{\frac{6}{5} x_2} - \frac{4}{5} + \ln \frac{5}{8} = \frac{18}{5} x_1 - \cancel{\frac{6}{5} x_2} - \frac{29}{5} + \ln \frac{3}{8}$$

$$x_1 = 2 + \frac{1}{2} \ln \left(\frac{5}{3}\right)$$

2) Методом квадратичного дискрим. анализа построить дискриминантные ф-и. Изобразить точки и разделяющие поверхности.

$\delta_0(x) = \frac{1}{2} \ln \det M_0 - \frac{1}{2}(x-\mu_0)^T M^{-1}$

$(x-\mu_0) + \ln P\{Y=0\} = -x_1^2 + 2x_1 + 2x_1 x_2 -$

$-2x_2 - 2x_1^2 + \ln \frac{5}{4} - 1$

$\delta_1(x) = -\frac{1}{2} \ln \frac{3}{4} - \frac{2}{3} x_1^2 - \frac{2}{3} x_2^2 + \frac{2}{3} x_1 x_2 +$

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

Разделяющая поверхность:

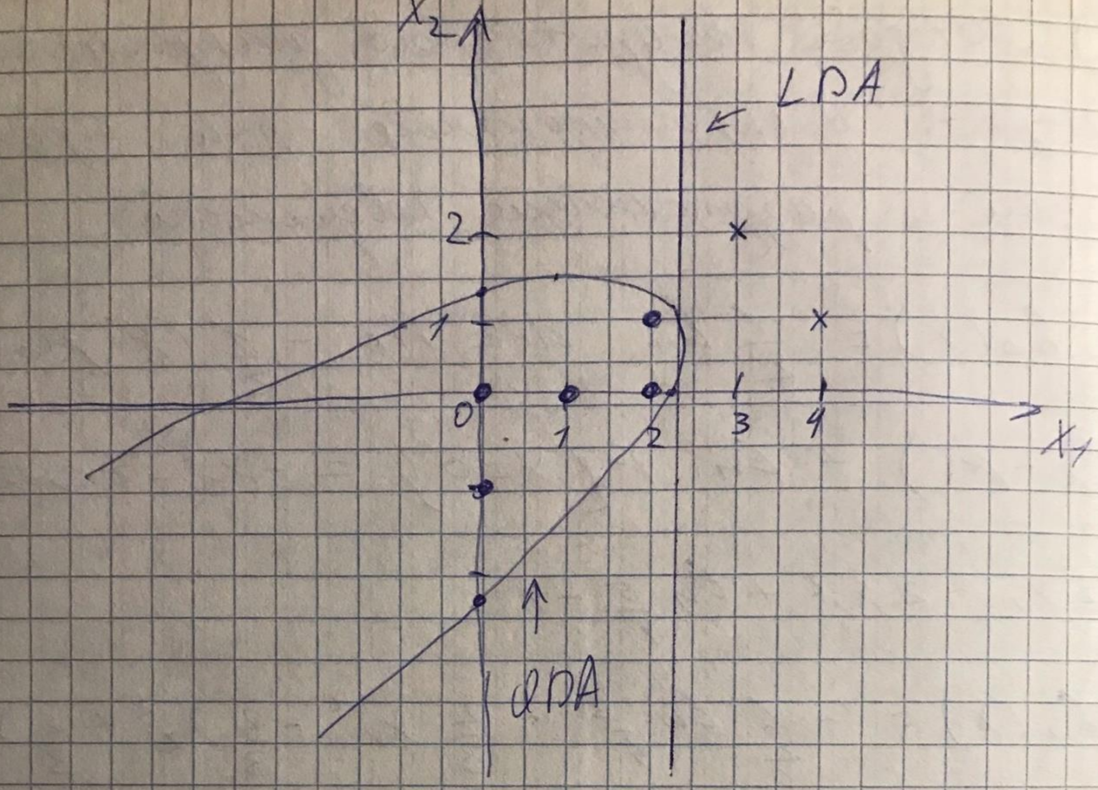
$$\delta_0(x) = \delta_1(x)$$

$-x_1^2 + 2x_1 + 2x_1 x_2 - 2x_2 - 2x_1^2 + \ln\frac{5}{4} - 1 = -\frac{1}{2} \ln \frac{3}{4} -$

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

$-\frac{1}{3} x_1^2 - \frac{4}{3} x_1 + \frac{4}{3} x_1 x_2 - \frac{4}{3} x_2^2 - \frac{4}{3} x_2 + \frac{11}{3} +$

$+ \frac{1}{2} \ln \frac{25}{3} = 0$


X_2
LDA
2
1
0
1
2
3
4
X_1
QDA