№ 9 [illegible] Г. К

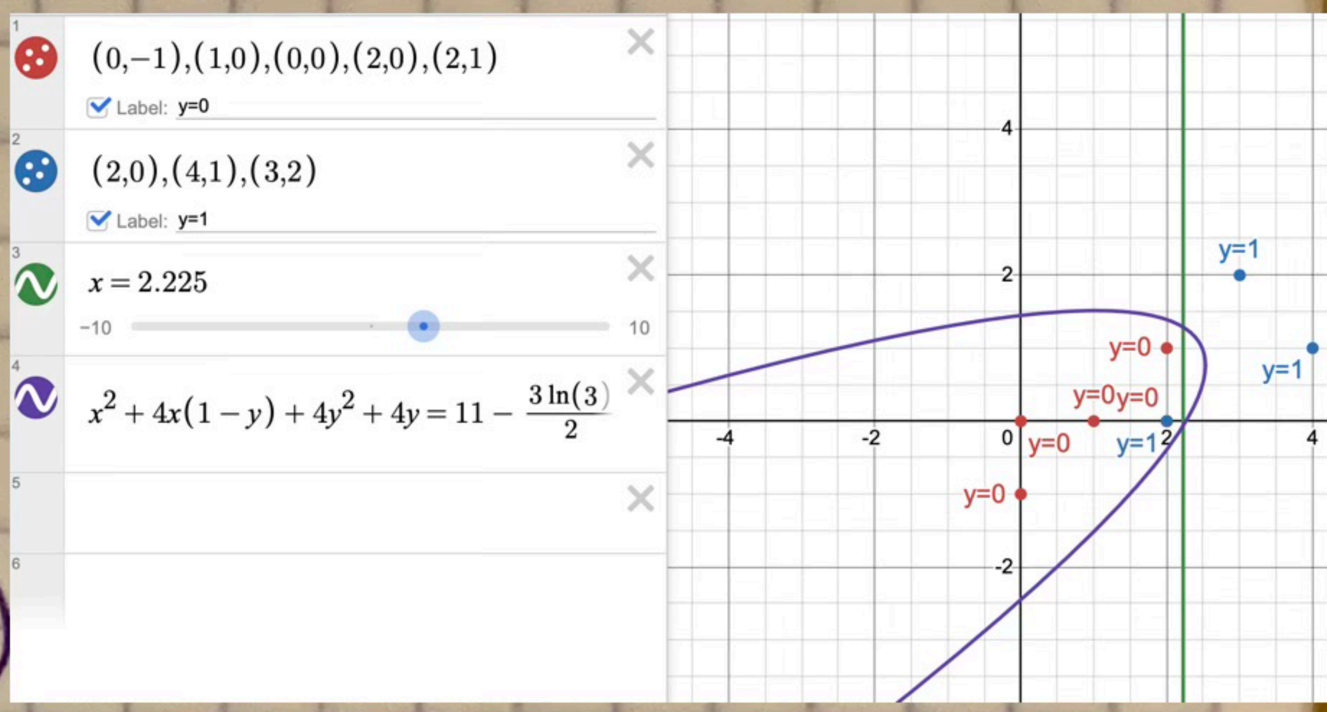
3824 М. ПМкн

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 для каждого класса построить дискриминантную функцию и записать уравнение разделяющей поверхности

2) Методом квадратичного дискриминантного анализа построить дискриминантные функции. Изобразить точки и разделяющие поверхности (кривые)



1) $\hat{P}r\{Y=0\} = 5/8$, $\hat{P}r\{Y=1\} = 3/8$

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

$$\hat{\Sigma}_0 = \frac{1}{N_0 - 1} \sum_{y^{(i)}=0} (x^{(i)} - \hat{\mu}_0)(x^{(i)} - \hat{\mu}_0)^T = \frac{1}{4}\begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}$$

$$\hat{\Sigma}_1 = \frac{1}{N_1 - 1} \sum_{y^{(i)}=1} (x^{(i)} - \hat{\mu}_1)(x^{(i)} - \hat{\mu}_1)^T = \frac{1}{2}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

$$\hat{\Sigma} = \frac{1}{N-K} \sum_{K} \sum_{y^{(i)}=K} (x^{(i)} - \hat{\mu}_K)(x^{(i)} - \hat{\mu}_K)^T = \frac{1}{6}\left(\begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\right) = \frac{1}{6}\begin{pmatrix} 6 & 3 \\ 3 & 4 \end{pmatrix}$$

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

$$\delta_0(x) = x^T \hat{\Sigma}^{-1} \hat{\mu}_0 - \frac{1}{2} \hat{\mu}_0^T \hat{\Sigma}^{-1} \hat{\mu}_0 + \ln \hat{P}r\{Y=0\} = 1.6x_1 - 1.2x_2 - 0.8 + \ln 5/8$$

$$\delta_1(x) = x^T \hat{\Sigma}^{-1} \hat{\mu}_1 - \frac{1}{2} \hat{\mu}_1^T \hat{\Sigma}^{-1} \hat{\mu}_1 + \ln \hat{P}r\{Y=1\} = 3.6x_1 - 1.2x_2 - 4.8 + \ln 3/8$$

$\delta_0(x) = \delta_1(x)$: $x_1 = 2.255$

2) $$\delta_0(x) = -\frac{1}{2}\ln\det\hat{\Sigma}_0 - \frac{1}{2}(x - \hat{\mu}_0)^T \hat{\Sigma}_0^{-1}(x - \hat{\mu}_0) + \ln\hat{P}r\{Y=0\} = -1 - x_1^2 - 2x_2 - 2x_2^2 + 2x_1(1+x_2) + \ln\frac{5}{8}$$

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

$$= \frac{1}{6}\left(-28 - 4x_1^2 - 4x_2 - 4x_2^2 + 4x_1(5 + x_2)\right) - \ln\frac{4096}{27}$$

$$\delta_0(x) = \delta_1(x): \; x_1^2 + 4x_1(1 - x_2) + 4x_2^2 + 4x_2 = 11 - \frac{3\ln 3}{2} + 3\ln 5$$

№15 Курабелиани Г.К

3824М1ПМки

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 байесовского классификатора оценить вероятность

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

$\hat{Pr}\{Y=0\} = 1/2$, $\hat{Pr}\{Y=1\} = 1/2$

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

$\hat{Pr}\{X_2=1 \mid Y=0\} = 3/5$, $\hat{Pr}\{X_2=1 \mid Y=1\} = 1$

$$Pr\{Y=0 \mid X_1=1, X_2=1\} = \frac{Pr\{X_1=1 \mid Y=0\} \cdot Pr\{X_2=1 \mid Y=0\} \cdot Pr\{Y=0\}}{Pr\{X_1=1, X_2=1\}} = 2/7$$

$$Pr\{Y=1 \mid X_1=1, X_2=1\} = \frac{Pr\{X_1=1 \mid Y=1\} \cdot Pr\{X_2=1 \mid Y=1\} \cdot Pr\{Y=1\}}{Pr\{X_1=1, X_2=1\}} = 5/7$$

№16

Дана выборка:

x_1	4	0	-2	2
x_2	3	1	-3	-1

. Найти PC и Var. Изобразить точки и PC.

$$X = \begin{pmatrix} 4 & 3 \\ 0 & 1 \\ -2 & -3 \\ 2 & -1 \end{pmatrix}, \quad \bar{x} = (1, 0) \Rightarrow X_c = \begin{pmatrix} 3 & 3 \\ -1 & 1 \\ -3 & -3 \\ 1 & -1 \end{pmatrix}$$

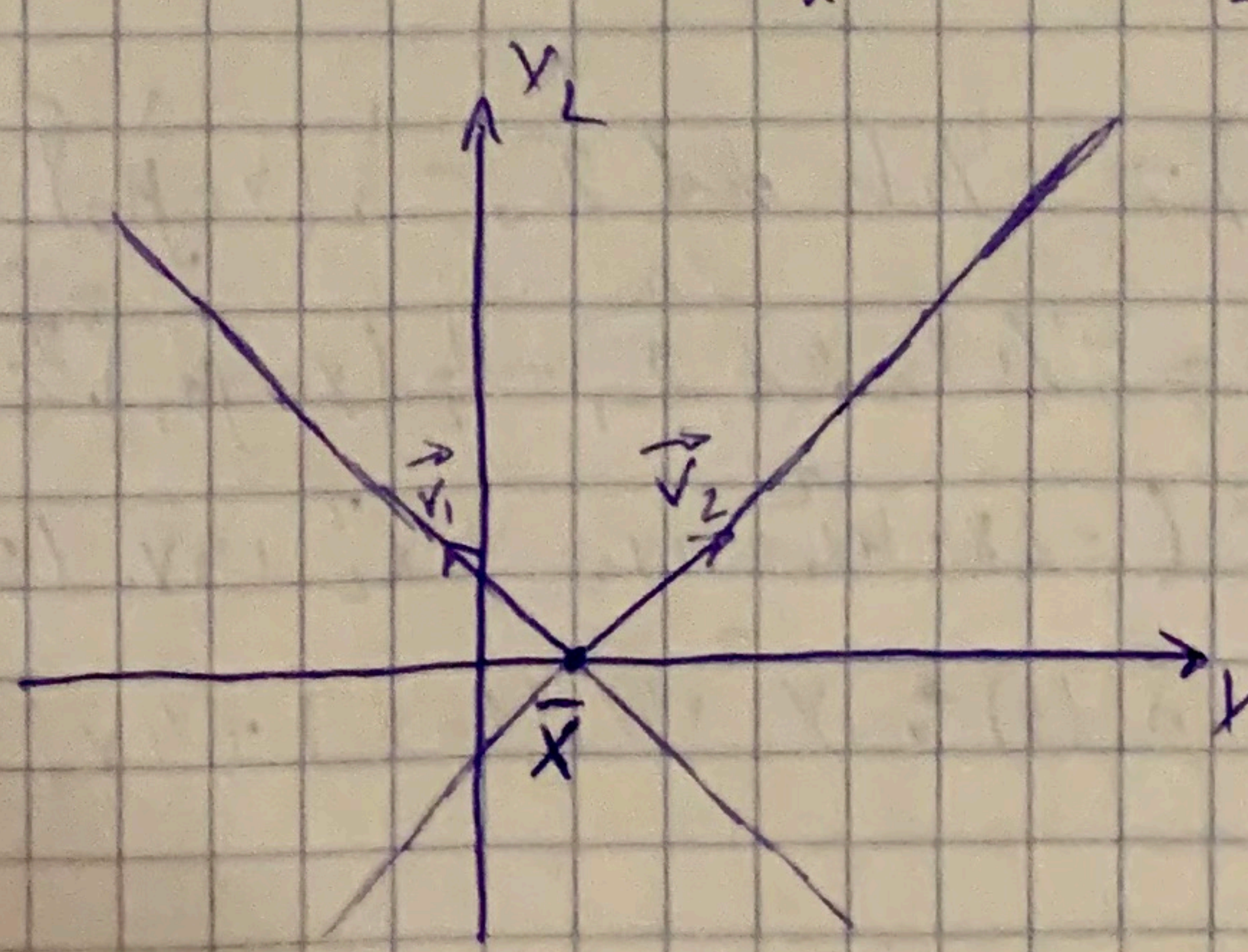
$$C = X_c^T X_c = \begin{pmatrix} 20 & 16 \\ 16 & 20 \end{pmatrix}$$

$$\det(C - \lambda I) = \begin{vmatrix} 20-\lambda & 16 \\ 16 & 20-\lambda \end{vmatrix} = \lambda^2 - 40\lambda + 144 \Rightarrow \lambda_1 = 4, \lambda_2 = 36$$

$$V_1 = \frac{\sqrt{2}}{2}\begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad V_2 = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$\frac{1}{N-1}\lambda_1 = 4/3$ $\qquad$ $\frac{1}{N-1}\lambda_2 = \frac{36}{3}$

$\frac{\lambda_1}{\lambda_1+\lambda_2} = \frac{1}{10}$ $\qquad$ $\frac{\lambda_2}{\lambda_1+\lambda_2} = \frac{9}{10}$



№13 Курасбедиани Г.К

3824М1ПМки

В задаче классификации на K классов $\{\overline{1,K}\}$ последний слой нейронной сети вычисляет softmax-функцию:

$g_k(s_1, s_2, \ldots, s_K) = e^{s_k} / \sum_{l=1}^{K} e^{s_l}$. В качестве штрафа используется кросс-энтропия: $R^{(i)} = -\sum_{k=1}^{K} I(y^{(i)} = k) \ln g_k(s_1, s_2, \ldots, s_K)$, где $g_k(s_1, s_2, \ldots, s_K)$ – softmax-функция. Доказать:

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

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

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

1) $$\frac{\partial g_k}{\partial s_l} = -\frac{e^{s_k} e^{s_l}}{\left(\sum_{l=1}^{K} e^{s_l}\right)^2} = -g_k g_l \qquad [l \neq k]$$

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

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

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

3) $$\frac{\partial R^{(i)}}{\partial s_l} = -\sum_{k=1}^{K} \frac{I(y^{(i)} = k)}{g_k} \cdot g_k(I(k=l) - g_l) = g_l - I(l = y^{(i)})$$