Мешалкин И.А
группа: 3822М1ФИин1

№1.

$\exists\ g: \mathbb{R}^n \to \mathbb{R}^m,\ x \in \mathbb{R}^n$. Матрицей Якоби назыв. матрица

$$\frac{\partial g}{\partial x} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_n} \end{pmatrix}$$

Док-ть:

1) если $a \in \mathbb{R}^n,\ x \in \mathbb{R}^n$, то $\frac{\partial (a^T x)}{\partial x} = a$

Док-во:

$$a^T x = [a_1 \ldots a_n] \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = (a_1 x_1 + \ldots + a_n x_n) = \sum_{i=1}^{n} a_i x_i$$

$$\forall j: \frac{\partial}{\partial x_j}\left(\sum_{i=1}^{n} a_i x_i\right) = a_j \Rightarrow \frac{d(\sum a_i x_i)}{d x} = \begin{bmatrix} \frac{\partial(\sum a_i x_i)}{\partial x_1} \\ \vdots \\ \frac{\partial(\sum a_i x_i)}{\partial x_n} \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = a$$

\#

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

Док-во:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix};\quad Ax = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \times \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} x_1 + \ldots + a_{1n} x_n \\ \vdots \\ a_{m1} x_1 + \ldots + a_{mn} x_n \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{i=1}^{n} a_{1i} x_i \\ \vdots \\ \sum_{i=1}^{n} a_{mi} x_i \end{pmatrix}$$

$$\forall j: \frac{\partial (Ax)}{\partial x_j} = \begin{bmatrix} \frac{\partial}{\partial x_j}(\sum a_{1i} x_i) \\ \vdots \\ \frac{\partial}{\partial x_j}(\sum a_{mi} x_i) \end{bmatrix} = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

↓

$$\frac{\partial(Ax)}{\partial x} = \begin{bmatrix} \frac{\partial(\sum_{i=1}^{n} a_{1i}x_i)}{\partial x_1} & \cdots & \frac{\partial(\sum a_{1i}x_i)}{\partial x_n} \\ \vdots & & \\ \frac{\partial(\sum_{i=1}^{n} a_{mi}x_i)}{\partial x_1} & \cdots & \frac{\partial(\sum^{n} a_{mi}x_i)}{\partial x_n} \end{bmatrix} =$$

$$= \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = A \quad \#$$

3) 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:

из п.2: $x^T A = \begin{pmatrix} a_{11}x_1 & \cdots & a_{1n}x_1 \\ \vdots & & \\ a_{n1}x_n & \cdots & a_{nn}x_n \end{pmatrix} = T$

$$Tx = \begin{pmatrix} a_{11}x_1^2 & \cdots & a_{1n}x_n \\ a_{n1}x_nx_1 & \cdots & a_{nn}x_n^2 \end{pmatrix} \to \frac{\partial(Tx)}{\partial x} = \begin{pmatrix} 2a_{11}x_1 & \cdots & (a_{1n}+a_{n1})x_n \\ (a_{n1}+a_{1n})x_1 & \cdots & (2a_{nn})x_n \end{pmatrix}$$

$$\| \\ (A^T + A)x$$

Если $A^T = A$, то $A + A^T = 2A$ #

4) Если $x \in \mathbb{R}^n$, то $\frac{\partial \|x\|^2}{\partial x} = 2x$

Док-во:

$\|x\|^2 = (x,x) = \sum x_i^2$

$$\frac{\partial(\sum x_i^2)}{\partial x} = \begin{bmatrix} \frac{\partial \sum x_i^2}{\partial x_1} \\ \vdots \\ \frac{\partial \sum x_i^2}{\partial x_n} \end{bmatrix} = \begin{bmatrix} 2x_1 \\ \vdots \\ 2x_n \end{bmatrix} = 2x \quad \#$$

$\Downarrow$

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

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

Док-во:

$$g(x) = (g(x_1) \dots g(x_n))^T \Rightarrow \frac{\partial g(x)}{\partial x} = \begin{pmatrix} g'_{x_1}(x_1) & \dots & \overset{0}{g'_{x_1}(x_n)} \\ \vdots & & \\ \underset{=0}{g'_{x_n}(x_1)} & \dots & g'_{x_n}(x_n) \end{pmatrix} =$$

$$= \begin{pmatrix} g'_{x_1}(x_1) & & 0 \\ & \ddots & \\ 0 & & g'_{x_n}(x_n) \end{pmatrix} = \operatorname{diag} g'(x) \quad \#$$

6) Если $h: \mathbb{R}^n \to \mathbb{R}^m$, $g: \mathbb{R}^m \to \mathbb{R}^p$, $x \in \mathbb{R}^n$, то $\frac{\partial g(h(x))}{\partial x} = \frac{\partial g(h(x))}{\partial h} \frac{\partial h(x)}{\partial x}$

Док-во:

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

$$\forall j: \frac{\partial g_i(h(x))}{\partial x_j} = \frac{\partial g_i(h(x))}{\partial h_1} \cdot \frac{\partial h_1(x)}{\partial x_j} + \frac{\partial g_i(h(x))}{\partial h_2} \cdot \frac{\partial h_2(x)}{\partial x_j} + \dots + \frac{\partial g_i(x)}{\partial h_m} \cdot \frac{\partial h_m(x)}{\partial x_j} \Downarrow$$

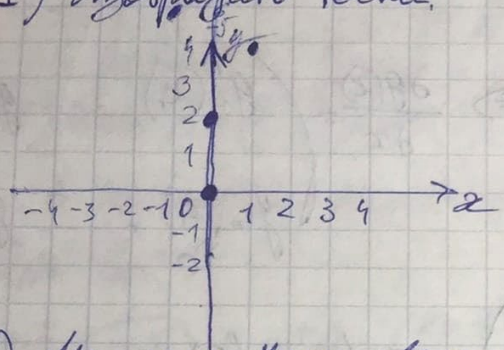
$$\begin{pmatrix} \frac{\partial g_1(h(x))}{\partial h_1} & \dots & \frac{\partial g_1(h(x))}{\partial h_m} \\ \vdots & & \\ \frac{\partial g_p(h(x))}{\partial h_1} & \dots & \frac{\partial g_p(h(x))}{\partial h_m} \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial h_1(x)}{\partial x_1} & \dots & \frac{\partial h_m(x)}{\partial x_n} \\ \vdots & & \\ \frac{\partial h_m(x)}{\partial x_1} & \dots & \frac{\partial h_m(x)}{\partial x_n} \end{pmatrix} = \frac{\partial g(h(x))}{\partial h(x)} \cdot \frac{\partial h(x)}{\partial x} \quad \#$$

№3.

Дана обучающая выборка

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

1) Изобразить точки.



2) Методом наим. квадратов построить модель вида $f(x)=\beta_0+\beta_1 x+\beta_2 x^2$. Построить графики

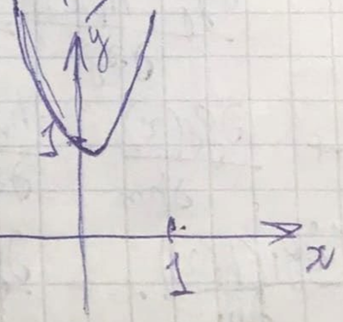
$$X^T=\begin{pmatrix}1&1&1&1&1\\1&1&0&0&-1\\1&1&0&0&1\end{pmatrix}\qquad y=\begin{pmatrix}4\\4\\0\\2\\6\end{pmatrix}$$

$$X^TX=\begin{pmatrix}5&1&3\\1&3&1\\3&1&3\end{pmatrix}\qquad X^Ty=\begin{pmatrix}16\\2\\14\end{pmatrix}$$

Решим сист. ур-ий $X^TX\beta=X^Ty$

$$\begin{pmatrix}5&1&3\\1&3&1\\3&1&3\end{pmatrix}\cdot\begin{pmatrix}\beta_0\\\beta_1\\\beta_2\end{pmatrix}=\begin{pmatrix}16\\2\\14\end{pmatrix}\Rightarrow\begin{pmatrix}\beta_0\\\beta_1\\\beta_2\end{pmatrix}=\begin{pmatrix}1\\-1\\4\end{pmatrix}$$

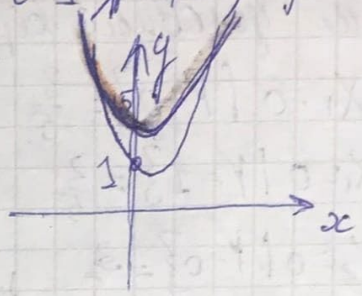
$$\Rightarrow f(x)=1-x+4x^2$$



3) Построить модель ридж-регрессии с пар-ом $\lambda=1$

$$x^T x + \lambda I = \begin{pmatrix} 6 & 1 & 3 \\ 7 & 4 & 1 \\ 3 & 1 & 4 \end{pmatrix} \to \left(\begin{array}{ccc|c} 0 & 1 & 0 & -1/2 \\ 2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 5/2 \end{array}\right) \to \beta = (3/2 \;\; -1/2 \;\; 5/2)$$

$$\Rightarrow g(x) = 3/2 - 1/2\,x + 5/2\,x^2$$



№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 дискриминац. анализа для каждого класса построить дискриминантную ф-ю и запис. ур-е разделяющ. пов-ти

- оценка вер-ти классов:

$\hat{P}_r\{Y=0\} = \frac{5}{8}$, $\hat{P}_r\{Y=1\} = \frac{3}{8}$

Средние для классов:

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

Объекты распределены по нормал. закону:

$$p(x|y) = \frac{1}{\sqrt{(2\pi)^d \det \Sigma_y}}\, e^{-\frac{1}{2}(x-\mu_y)^T \Sigma_y^{-1} (x-\mu_y)}$$

мат. ожид. в d-мерном пр-ве

$\mathrm{cov}(X_i, X_j) = E\,\underbrace{(x_i - Ex_j)(x_j - Ex_i)}_{\text{центрир}}$

$\mathrm{cov}(x_i, x_i)$ – дисперсия

Мат. ковариации для каждого класса:

$$\hat{\Sigma}_0 = \frac{1}{N_0 - 1} \sum_{y^{(i)}=0} (x^{(i)} - \hat{\mu}_0)(x^{(i)} - \hat{\mu}_0)^T = \frac{1}{5-1}\left(\left(\begin{pmatrix}0\\-1\end{pmatrix} - \begin{pmatrix}1\\0\end{pmatrix}\right) \cdot \left(\begin{pmatrix}0\\-1\end{pmatrix} - \begin{pmatrix}1\\0\end{pmatrix}\right)^T + \right.$$

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

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

$$\hat{\Sigma}_1 = \frac{1}{N_1 - 1} \sum_{y^{(i)}=1} (x^{(i)} - \hat{\mu}_1)(x^{(i)} - \hat{\mu}_1) = \frac{1}{2}\left(\begin{pmatrix}1&1\\1&1\end{pmatrix} + \begin{pmatrix}1&0\\0&0\end{pmatrix} + \begin{pmatrix}0&0\\0&1\end{pmatrix}\right) =$$

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

Оценим матрицу ковариации:

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

$$= \begin{pmatrix}1 & 2/3\\ 2/3 & 5/6\end{pmatrix}$$

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

Линейные дискриминантные ф-ии:

$$\delta_0(x) = x^T \hat{\Sigma}^{-1}\hat{\mu}_0 - \frac{1}{2}\hat{\mu}_0^T \hat{\Sigma}^{-1}\hat{\mu}_0 + \ln \hat{P}r\{Y_1 = 0\} = (x_1\ x_2)\begin{pmatrix}8/5 & -6/5\\ -6/5 & 12/5\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}$$

$$- \frac{1}{2}(1\ 0)\begin{pmatrix}8/5 & -6/5\\ -6/5 & 12/5\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} + \ln\frac{5}{8} = \frac{8}{5}x_1 - \frac{6}{5}x_2 - \frac{4}{5} + \ln\frac{5}{8}$$

$$\delta_1(x) = x^T \hat{\Sigma}^{-1}\hat{\mu}_1 - \frac{1}{2}\hat{\mu}_1^T \hat{\Sigma}^{-1}\hat{\mu}_1 + \ln \hat{P}r\{Y = 1\} = (x_1\ x_2)\begin{pmatrix}8/5 & -6/5\\ -6/5 & 12/5\end{pmatrix}\begin{pmatrix}3\\1\end{pmatrix}$$

$$= \frac{18}{5}x_1 - \frac{6}{5}x_2 - \frac{24}{5} + \ln\frac{3}{8}$$

Разделяющая пов-ть – это прямая с ур-ем:

$$\delta_0(x) = \delta_1(x) \Rightarrow 2x_1 - 4 + \ln\frac{3}{8} - \ln\frac{5}{8} = 0$$

2) Методом квадратичного дискриминантного анализа построить дискриминантные ф-ии. Изобразить точки и раздел. пов-ть.

• Квадратич. дискриминантные ф-ии:

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

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

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

$$- \left((x_1-1)(x_1-x_2-1) + x_2(-x_1+2x_2+1)\right) = -\frac{1}{2}\ln 4 + \ln\frac{5}{8} -$$

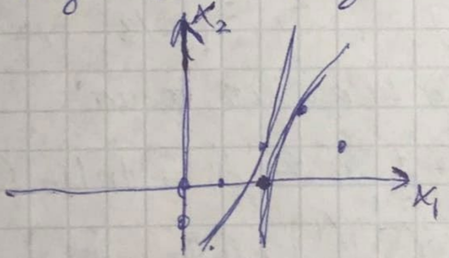
$$- \left(x_1^2 - 2x_1x_2 + 2x_2 - 2x_1 + 1 + 2x_2^2\right) = -(x_1-1)^2 - 2(x_2^2 - x_1x_2 + x_2)$$

$$\delta_1(x) = -\frac{1}{6}\left(4x_1^2 + 4x_2^2 - 4x_1x_2 - 20x_1 + 4x_2 + 28\right) - \frac{1}{2}\ln\frac{12}{9} + \ln\frac{3}{8}$$

Разделяющ. пов-ть:

$$\delta_0(x) = \delta_1(x) = \frac{1}{3}x_1^2 - \frac{4}{3}x_2^2 - \frac{4}{3}x_1x_2 + \frac{4}{3}x_1 + \frac{4}{3}x_2 - \frac{11}{3} -$$

$$- \frac{1}{2}\ln\frac{4}{3} + \ln\frac{3}{8} + \frac{1}{2}\ln 4 - \ln\frac{5}{8}$$



№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 пом-ю байесовского классификатора оценить вер-ти $P_r(Y=0; X_1=1, X_2=1)$ $P_r(Y=1, X_1=1, X_2=1)$

$\widehat{Pr}\{Y=0\} = \frac{1}{2}$, $\widehat{Pr}\{Y=1\} = \frac{1}{2}$

Оценим условные вер-ти:

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

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

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

$\widehat{Pr}\{X_2=1 \mid Y=1\} = 0$, $\widehat{Pr}\{X_2=1 \mid Y=1\} = 1$

$$Pr\{Y=0 \mid X_1=1, X_2=1\} = \frac{Pr\{X_1 \mid Y=0\} Pr\{X_2=1 \mid Y=0\} Pr\{Y=0\}}{Pr\{X_1=1, X_2=1\}} =$$

$$= \frac{\frac{2}{5} \cdot \frac{3}{5} \cdot \frac{1}{2}}{\frac{3}{25} + \frac{1}{5}} = \frac{3}{8}$$

$$Pr\{Y=1 \mid X_1=1, X_2=1\} = \frac{Pr\{X_1=1 \mid Y=1\} Pr\{X_2=1 \mid Y=1\} Pr\{Y=1\}}{Pr\{X_1=1, X_2=1\}} =$$

$$= \frac{5}{8}$$

Ответ: $\frac{3}{8}$ и $\frac{5}{8}$

Упражнение 4.1.

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x^{(i)}$$

$W_k = V_0 + L(v_1 \dots v_k)$ – k-мерное лин. многообразие

• Будем искать V_0 как реш-е задачи minimize L_0:

$$V_0 = \underset{a_0 \in \mathbb{R}^n}{\operatorname{argmin}} \left(\sum_{i=1}^{N} dist^2(x_i, L_0) \right) = \underset{a_0 \in \mathbb{R}^n}{\operatorname{argmin}} \sqrt{\sum_{i=1}^{N} \| x^{(i)} - a_0 \|^2} =$$

$$= \frac{1}{N} \sum_{i=1}^{N} x^{(i)} = \bar{x}$$