№1

1.1. Д-ть, что если $a \in R^n$, $x \in R^n$, то $\frac{\partial(a^T x)}{\partial x} = a$

$$a^T x = \sum_{i=1}^{n} a_i x_i = f$$

$$\frac{\partial f}{\partial x} = \left(\frac{\partial f}{\partial x_1} \dots \frac{\partial f}{\partial x_n}\right) = (a_1 \dots a_n) = a$$

1.2 Д-ть, что если $A \in R^{m \times n}$, $x \in R^n$, то $\frac{\partial(Ax)}{\partial x} = A$

1) $$A \cdot x = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix} = g$$

2) $$\frac{\partial g_i}{x_j} = a_{ij} \rightarrow \frac{\partial g}{\partial x} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

т.е
$$\begin{cases} \frac{\partial(Ax)}{\partial x} \overset{1}{=} \frac{\partial}{\partial x} g \\ \frac{\partial g}{\partial x} \overset{2}{=} A \end{cases} \longrightarrow \frac{\partial(Ax)}{x} = A \quad \boxed{\text{ч.т.д.}}$$

1.3 Если $A \in R^{n \times m}$, $x \in R^n$, то $\frac{\partial(x^T A x)}{\partial x} = (A + A^T)x$

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

$$\blacktriangleleft\ x^T A x = (x_1 \dots x_n)\underbrace{\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn}\end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n\end{pmatrix}}_{\text{из п. 1.2}\ =\ \begin{pmatrix}\sum_{i=1}^n a_{1i}x_i \\ \vdots \\ \sum_{i=1}^n a_{mi}x_i\end{pmatrix}} =$$

$$= (x_1 \dots x_n)\begin{pmatrix}\sum_{i=1}^n a_{1i}x_i \\ \vdots \\ \sum_{i=1}^n a_{mi}x_i\end{pmatrix} = x_1 \sum_i a_{1i}x_i + \dots x_n \sum_i a_{ni}x_i \Rightarrow$$

$$\Rightarrow (x^T A x) = \sum_{j=1}^n \left(x_j \sum_{i=1}^n a_{ji}x_i\right) = g$$

$$\blacktriangleleft\ \frac{\partial g}{\partial x} :$$

$$\begin{cases} \dfrac{\partial g}{\partial x_1} = \sum_i a_{1i}x_i + a_{11}x_1 + a_{21}x_2 + \dots a_{n1}x_n = \sum_{i=1}^n a_{1i}x_i + \sum_{i=1}^n a_{i1}x_i \\ \dfrac{\partial g}{\partial x_n} = \underbrace{\sum_{i=1}^n a_{ni}x_i}_{1.n} + \underbrace{\sum_{i=1}^n a_{in}x_i}_{2.n} \end{cases}$$

$$\to \frac{\partial g}{\partial x} = \begin{pmatrix}\frac{\partial g}{\partial x_1} \\ \vdots \\ n\end{pmatrix} = \underbrace{\begin{pmatrix}\sum_{i=1}^n a_{1i}x_i \\ \vdots \\ \sum_{i=1}^n a_{ni}x_i\end{pmatrix}}_{1} + \underbrace{\begin{pmatrix}\sum_{i=1}^n a_{i1}x_i \\ \vdots \\ \sum_{i=1}^n a_{in}x_i\end{pmatrix}}_{2}$$

$$1 = Ax \qquad 2 = A^T x \longrightarrow \frac{\partial (x^T A x)}{\partial x} = (A + A^T)x$$

2

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

◁ $\|x\|^2 = x_1^2 + x_2^2 \dots + x_n^2$ (*)

$$\frac{\partial \|x\|^2}{\partial x} = \frac{\partial (x_1^2 + \dots + x_n^2)}{\partial x} = 2(x_1 \dots x_n) = 2x \quad \blacksquare$$

⇒

1.5 $x \in R^n$

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

$$◁ \frac{\partial g(x)}{\partial x} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_n}{\partial x_1} & \dots & \frac{\partial g_n}{\partial x_n} \end{pmatrix} = \left\{ \frac{\partial g_i(x_i)}{\partial x_j} = g' \cdot \delta_{ij} \right\} = \begin{pmatrix} \frac{\partial g(x_1)}{\partial x_1} & & 0 \\ & \ddots & \\ 0 & & \frac{\partial g(x_n)}{\partial x_n} \end{pmatrix}$$

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

1.6 Если $h: R^n \to R^m$, $g: R^m \to R^p$, $x \in R^n$, то

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

$$g(h) = \begin{pmatrix} g_1(h_1 \dots h_m) \\ g_2(h_1 \dots h_m) \\ \vdots \\ g_p(h_1 \dots h_m) \end{pmatrix} \qquad h(x) = \begin{pmatrix} h_1(x_1 \dots x_n) \\ \vdots \\ h_m(x_1 \dots x_n) \end{pmatrix}$$

$$◁ \frac{\partial g_i(h(x))}{\partial x_j} = \frac{\partial g_i}{\partial h_1} \frac{\partial h_1}{\partial x_j} + \dots + \frac{\partial g_i}{\partial h_m} \frac{\partial h_m}{\partial x_j} = \sum_{k=1}^{m} \frac{\partial g_i}{\partial h_k} \frac{\partial h_k}{\partial x_j}$$

Т.о $$\frac{\partial g(h(x))}{\partial x} = \begin{pmatrix} \sum\limits_{k=1}^{m} \frac{\partial g_1}{\partial h_k}\frac{\partial h_k}{\partial x_1} & \dots & \sum\limits_{k=1}^{m} \frac{\partial g_1}{\partial h_k}\frac{\partial h_k}{\partial x_n} \\ \vdots & & \vdots \\ \sum \frac{\partial g_p}{\partial h_k}\frac{\partial h_k}{\partial x_1} & \dots & \sum\limits_{k=1}^{m} \frac{\partial g_p}{\partial h_k}\frac{\partial h_k}{\partial x_n} \end{pmatrix} \quad (1)$$

С другой стороны

$$\frac{\partial g}{\partial h} = \begin{pmatrix} \frac{\partial g_1}{\partial h_1} & \frac{\partial g}{\partial h_m} \\ \frac{\partial g_p}{\partial h_1} & \dots \frac{\partial g_p}{\partial h_m} \end{pmatrix} = 1$$

$$\frac{\partial h}{\partial x} = \begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \dots & \frac{\partial h_1}{\partial x_n} \\ \vdots & & \frac{\partial h_m}{\partial x_n} \\ \dots & \dots & \frac{\partial x_n}{} \end{pmatrix} = 2$$

$$\frac{\partial g(h)}{\partial h}\frac{\partial h}{\partial x} = \begin{pmatrix} & & \\ & 1 & \\ & & \end{pmatrix}\begin{pmatrix} & & \\ & 2 & \\ & & \end{pmatrix} = \begin{pmatrix} \sum\limits_{k=1}^{m} \frac{\partial g_1}{\partial h_k}\frac{\partial h_k}{\partial x_1} & \dots & \sum \frac{\partial g_1}{\partial h_k}\frac{\partial h_k}{\partial x_n} \\ \dots & & \sum \frac{\partial g_p}{\partial h_k}\frac{\partial h_k}{\partial x_n} \end{pmatrix} \quad (2)$$

Из (1) и (2) $\Rightarrow$, что $\frac{\partial g(h)}{\partial h}\frac{\partial h}{\partial x} = \frac{\partial (g(h))}{\partial x}$

■

Задание 3

Выборка $\begin{pmatrix} 1 & 4 \\ 1 & 4 \\ 0 & 0 \\ 0 & 2 \\ -1 & 6 \end{pmatrix}$

1) изобразить т.
2) ~~нет~~ Построить модель $f(x) = \beta_0 + \beta_1 x + \beta_2 x^2$, построить график
3) Построить модель того же вида методом ридж регрессии

(2)

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

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

$$X^T y = \begin{pmatrix} 16 \\ 2 \\ 14 \end{pmatrix}$$

Найдем β из системы норм. ур-б $X^T X \beta = X^T y$

$$\begin{cases} 5\beta_0 + \beta_1 + 3\beta_2 = 16 \\ \beta_0 + 3\beta_1 + \beta_2 = 2 \\ 3\beta_0 + \beta_1 + 3\beta_2 = 14 \end{cases} \xrightarrow{I - III} \begin{cases} 2\beta_0 = 2 \\ \Downarrow \\ \beta_0 = 1 \end{cases} \rightarrow \begin{cases} 3\beta_1 + \beta_2 = 1 \\ \beta_1 + 3\beta_2 = 11 \end{cases}$$

$$\rightarrow \beta_2 = 1 - 3\beta_1 \Rightarrow -8\beta_1 = 8 \rightarrow \beta_1 = -1 \Rightarrow \beta_2 = 4$$

$$\beta = \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix} \Rightarrow \boxed{f(x) = 1 - x + 4x^2}$$

5

③ Построить модель методом ридж-регрессии для $\lambda = 1$

$$X^T X + \lambda I = X^T X + I = \begin{pmatrix} 6 & 1 & 3 \\ 1 & 4 & 1 \\ 3 & 1 & 4 \end{pmatrix}$$

$$(X^T X + \lambda I)\beta = X^T y$$

$$\begin{cases} 6\beta_0 + \beta_1 + 3\beta_2 = 16 \\ \beta_0 + 4\beta_1 + \beta_2 = 2 \\ 3\beta_0 + \beta_1 + 4\beta_2 = 14 \end{cases} \xrightarrow{I - III} \left\{ \begin{array}{l} 3\beta_0 - \beta_2 = 2 \\ \Downarrow \\ \beta_2 = 3\beta_0 - 2 \end{array} \right\}^{(*)} \longrightarrow$$

[redacted] $\xrightarrow[+(*)]{III - II} \begin{cases} \beta_2 = 3\beta_0 - 2 \\ \beta_0 + 4\beta_1 + \beta_2 = 2 \\ 2\beta_0 - 3\beta_1 + 3\beta_2 = 12 \end{cases} \longrightarrow$

β_2 из $(I') \to II$ $\longrightarrow \begin{cases} 4\beta_0 + 4\beta_1 = 4 \\ 2\beta_0 - 3\beta_1 + 3\beta_2 = 12 \end{cases} \xrightarrow{\beta_2 \text{ в } (II')} \begin{cases} \beta_0 + \beta_1 = 1 \\ 11\beta_0 - 3\beta_1 = 18 \end{cases} \to$

$\to 14\beta_0 = 21 \Rightarrow \beta_0 = \frac{3}{2} \Rightarrow \beta_1 = -\frac{1}{2} \Rightarrow \beta_2 = \frac{5}{2}$

$$\Rightarrow \quad \beta = \begin{pmatrix} 3/2 \\ -1/2 \\ 5/2 \end{pmatrix}$$

$$f_1(x) = 3/2 - 1/2\, x + 5/2\, x^2$$

Задание 9

Обучающая выборка
$$\begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 1 & 2 \\ 1 & 0 & 2 \\ 1 & 1 & 4 \\ 1 & 2 & 3 \end{pmatrix}^T$$

1) Оценим вероятности классов

$$\hat{P_r}\{Y=0\} = \frac{5}{8}$$
$$\hat{P_r}\{Y=1\} = \frac{3}{8}$$

Оцениваем средние для классов

$$\hat{\mu}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$
$$\hat{\mu}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \Longrightarrow \begin{cases} X_0 = \begin{pmatrix} 0 & 1 & 0 & 2 & 2 \\ -1 & 0 & 0 & 0 & 1 \end{pmatrix} \\ \tilde{X}_0 = \begin{pmatrix} -1 & 0 & -1 & 1 & 1 \\ -1 & 0 & 0 & 0 & 1 \end{pmatrix} \\ X_1 = \begin{pmatrix} 2 & 4 & 3 \\ 0 & 1 & 2 \end{pmatrix} \Rightarrow \tilde{X}_1 = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \end{cases}$$

Выборочные матрицы ковариации для каждого класса

$$\hat{\Sigma}_0 = \frac{1}{N_0 - 1} \sum_{y^{(i)}=0} (x^{(i)} - \hat{\mu}_0)(x^{(i)} - \hat{\mu}_0)^T = \{N=5\}$$

7
$$= \frac{1}{4} \begin{pmatrix} -1 & 0 & -1 & 1 & 1 \\ -1 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & -1 \\ 0 & 0 \\ -1 & 0 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}$$

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

Оценка матрицы ковариации:

$$\hat{\Sigma} \underset{\substack{N=8 \\ k=2}}{} \frac{1}{N-k} \sum_k \sum_{y^{(i)}=k} (x^{(i)} - \hat{\mu}_k)(x^{(i)} - \hat{\mu}_k)^T = \frac{1}{6}\begin{pmatrix} 6 & 3 \\ 3 & 4 \end{pmatrix}$$

Обратные матрицы

$$\hat{\Sigma}_0^{-1} = \frac{\hat{\Sigma}_0^T}{\Delta\hat{\Sigma}_0} = \frac{\begin{pmatrix} \frac{1}{2} & -1/2 \\ -1/2 & 1 \end{pmatrix}}{\left(\frac{1}{2} - \frac{1}{4}\right)} = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}$$

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

$$\hat{\Sigma}^{-1} = \begin{pmatrix} 8/5 & -6/5 \\ -6/5 & 12/5 \end{pmatrix}$$

- Линейные дискриминантные ф-ии

$$\delta_0(x) = \underbrace{x^T}_{(x_1, x_2)} \hat{\Sigma}^{-1} \hat{\mu}_0 - \frac{1}{2}\hat{\mu}_0^T \hat{\Sigma}^{-1} \hat{\mu}_0 + \underbrace{\ln \hat{P}_1\{Y=0\}}_{\ln(5/8)}$$

$$\Downarrow$$

$$\delta_0(x) = \frac{8}{5}x_1 - \frac{6}{5}x_2 - \frac{4}{5} + \ln 5 - 3\ln 2 \quad (1)$$

$$\delta_1(x) = x^T \hat{\Sigma}^{-1} \hat{\mu}_1 - \frac{1}{2}\hat{\mu}_1^T \hat{\Sigma}^{-1} \hat{\mu}_1 + \ln \hat{P}_2\{Y=1\}$$

$$\hookrightarrow \delta_1 = \frac{18}{5}x_1 - \frac{6}{5}x_1 - \frac{6}{5}x_2 - \frac{24}{5} + \ln 3 - 3\ln 2 \quad (2)$$

Разделяющая поверхность ~~парабола~~ прямая с уравнением, которое получается при прир. $\delta_0(x) = \delta_1(x)$

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

• Квадратичные дискриминантные ф-ии:

1. $$\delta_0(x) = -\frac{1}{2}\ln \underbrace{\det \hat{\Sigma}_0}_{I} - \underbrace{\frac{1}{2}(x-\hat{\mu}_0)^T \hat{\Sigma}_0^{-1}(x-\hat{\mu}_0)}_{II} + \underbrace{\ln \hat{P}_2\{Y=0\}}_{III}$$

(I) $\det \hat{\Sigma}_0 = \frac{1}{4}$

(II) $$-\frac{1}{2}(x-\hat{\mu}_0)^T \hat{\Sigma}_0^{-1}(x-\hat{\mu}_0) = -\frac{1}{2}(x_1-1, x_2)\begin{pmatrix}2 & -2\\ -2 & 4\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}$$

$$= -(x_1-1, x_2)\begin{pmatrix}x_1-1-x_2\\ -x_1+2x_2+1\end{pmatrix} =$$

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

(III) $\ln 5/8$

$$\delta_0 = \ln 5 - 2\ln 2 - (x_1^2 - 2x_1x_2 + 2x_2^2 + 2x_2 - 2x_1 + 1)$$

2. $$\delta_1(x) = \underbrace{-\frac{1}{2}\ln\det\hat{\Sigma}_1}_{I} - \underbrace{\frac{1}{2}(x-\hat{\mu}_1)^T\hat{\Sigma}_1^{-1}(x-\hat{\mu}_1)}_{II} + \underbrace{\ln\hat{P}_2\{Y=1\}}_{I}$$

(I) $\left.\begin{matrix}\det\hat{\Sigma}_1 = 3/4\\ \ln P_2\{Y=1\} = 3/8\end{matrix}\right\} \rightarrow$

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

(II) $$-\frac{1}{2}(x-\hat{\mu}_1)^T\hat{\Sigma}_1^{-1}(x-\hat{\mu}_1) = -\frac{1}{3}(x_1-3\quad x_2-1)\begin{pmatrix}2 & -1\\ -1 & 2\end{pmatrix}\begin{pmatrix}x_1-3\\ x_2-1\end{pmatrix} =$$

$$= -\frac{1}{3}\left(2x_1^2 - 2x_1x_2 + 2x_2^2 - 10x_1 + 2x_2 + 14\right)$$

$$\delta_1(x) = -\frac{1}{3}\left(2x_1^2 - 2x_1x_2 + 2x_2^2 - 10x_1 + 2x_2 + 14\right) + \frac{1}{2}\ln 3 - 2\ln 2$$

Разделяющая пов-ть — парабола, прир. δ_0 и δ_1:

$$(x_1 - 2x_2)^2 + 4x_1 + 4x_2 - 11 - 3\ln\frac{5}{\sqrt{3}} = 0$$

Задание 15

Дана обучающая выборка

$$\begin{array}{c} x_1 \\ x_2 \\ y \end{array} \begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

С помощью баесова классификатора оценить вероятность $P_2(Y=0 \mid X_1=1, X_2=1)$

$P_2(Y=1 \mid X_1, X_2=1)$

Оценим априорные вероят.

$\hat{P}_2\{Y=0\} = \frac{1}{2}$ $\quad \hat{P}_2\{Y=1\} = \frac{1}{2}$

Оценим условные вероятности

$\hat{P}_2\{X_1=1 \mid Y=0\} = \frac{2}{5}$ $\quad \hat{P}_2\{X_2=1 \mid Y=0\} = \frac{3}{5}$

$\hat{P}_2\{X_1=1 \mid Y=1\} = \frac{3}{5}$ $\quad \hat{P}_2\{X_2=1 \mid Y=1\} = 1$

Основное предположение баесова классификатора

$$\text{I} \quad \hat{P}_2\{Y=0 \mid X_1=1, X_2=1\} = \frac{P_2\{X_1=1 \mid Y=0\} P_2\{X_2=1 \mid Y=0\} P_2\{Y=0\}}{P_2\{X_1=1, X_2=1\}}$$

$$\text{II} \quad \hat{P}_2\{Y=1 \mid X_1=1, X_2=1\} = \frac{P_2\{X_1=1 \mid Y=1\} P_2\{X_2=1 \mid Y=1\} P_2\{Y=1\}}{P_2\{X_1=1, X_2=1\}}$$

$$\text{I} = \frac{2/5 \cdot 3/5 \cdot 1/2}{P_2\{X_1=1, X_2=1\}} = \frac{3}{25} \frac{1}{P_2\{\ldots\}}$$

$$\text{II} = \frac{3}{10} \cdot \frac{1}{P_2\{\ldots\}}$$

$$P_2\{\ldots\} = P_2\{X_1=1, X_2=1\} = \frac{3}{25} + \frac{3}{10} = \frac{21}{50}$$

$$\boxed{\text{I} = \frac{2}{7} \qquad \text{II} = \frac{5}{7}}$$

Упражнение 4.1

Д-во $v_0 = \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x^{(i)}$

$W_k = v_0 + L(v_1 \ldots v_k)$ — k-мерное лин. многообр.

$$v_0 = \arg\min_{a_0}\left(\sum_{i=1}^{N} \mathrm{dist}^2(x_i, a_0)\right) = \arg\min_{a_0 \in R^n}\left(\sum_{i=1}^{N} \|x^{(i)} - a_0\|^2\right) =$$

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

$\blacksquare$