Задача 35

Задача восстановления регрессии с квадратичн. функцией потерь $L(y',y)=(y'-y)^2$

① Д-ть если $f^*(x)=\underset{c}{\operatorname{argmin}}\, E((y-c)^2 \mid x=x)$, то
$f^*(x)=E(y \mid x=x)$

② Средний риск $R(f^*)$?

$R(f)=E\,L(f(x),y)=\int_{XY} L(f(x),y)\,p(x,y)\,dx\,dy=$ {для задачи восстан. регрессии

$=\underbrace{\int_{XY} L(f(x),y)\,p(y|x)\,dy\,p(x)\,dx}_{\to \min}=\underbrace{\int_X E((f(x)-y)^2 \mid x)\,p(x)\,dx}_{*}$

$\Rightarrow f^*(x)=\underset{f}{\operatorname{argmin}}\, E((y-c)^2 \mid x=x)$

① $\frac{\partial}{\partial c}\left(\int_Y (c-y)^2 p(y|x)\,dy\right)=2\int (c-y)\,p(x|y)\,dy=2\left(\int_Y c\,p(y|x)\,dy\right.$
$\left.-\int_Y y\,p(y|x)\,dy\right)=2\left(c\underbrace{\int_Y p(y|x)\,dy}_{=1}-\int_Y y\,p(y|x)\,dy\right)=2c-2\int_Y y\,p(y|x)\,dy=0$

$\longrightarrow \quad c=\int_Y y\,p(y|x)\,dy$

$f^*(x)=\int y\,p(y|x)\,dy=E(y \mid x=x)$ ■ чтд

② $R(f^*)=E((E(y|x)-y)^2 \mid x)\,p(x)=E(D(y|x))$
из ⊛ и ①

$R(f^*)=E(D(y|x))$ ①

Задача 36

Задача восстановления регрессии с функцией потерь $L(y', y) = |y' - y|$

Д-ть, что min среднему риску доставляет при этом условная медиана $f(x) = median(Y | X = x)$

$$R(f) = \int_{X \cdot Y} \left(|f(x) - y| \, p(y|x) \, dy \right) p(x) \, dx$$

$$f(x) = \arg\min_{y'} \int_Y |y' - y| \, p(y|x) \, dy$$

$$\frac{\partial}{\partial y'} \int_Y |y' - y| \, p(y|x) \, dy = \int_Y sign(y' - y) \, p(y|x) \, dy = ① - ② = 0$$

① $\int_{y' > y} p(y|x) \, dy$

② $\int_{y' < y} p(y|x) \, dy$

При этом ① + ② = 1 $\Rightarrow$ $\begin{cases} ① - ② = 0 \\ ① + ② = 1 \end{cases}$ $\Rightarrow$ ① = ② = $\frac{1}{2}$

$\Rightarrow y' = f^*(x) = median(Y | X = x)$

Задание 07

Дана выборка

Выборка: $\begin{pmatrix} 4 & 0 & -1 & 3 & 4 \\ 2 & -3 & -2 & 1 & 2 \\ 3 & 2 & 2 & 1 & -3 \end{pmatrix}$

Находим $\overline{x} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$

1. Центрируем данные: $X_c = \begin{pmatrix} 2 & 2 & 2 \\ -2 & -3 & 1 \\ -3 & -2 & 1 \\ 1 & 1 & 0 \\ 2 & 2 & -4 \end{pmatrix}$

2. Находим $C = X_c^T X_c$

$$C = \begin{pmatrix} 2 & -2 & -3 & 1 & 2 \\ 2 & -3 & -2 & 1 & 2 \\ 2 & 1 & 1 & 0 & -4 \end{pmatrix} \begin{pmatrix} 2 & 2 & 2 \\ -2 & -3 & 1 \\ -3 & -2 & 1 \\ 1 & 1 & 0 \\ 2 & 2 & -4 \end{pmatrix} = \begin{pmatrix} 22 & 21 & -9 \\ 21 & 22 & -9 \\ -9 & -9 & 22 \end{pmatrix}$$

3. Собственные числа:

$$\begin{pmatrix} 22-\lambda & 21 & -9 \\ 21 & 22-\lambda & -9 \\ -9 & -9 & 22-\lambda \end{pmatrix} = -(\lambda-1)(\lambda-16)(\lambda-49)=0$$

Собственные вектора

$$\left.\begin{array}{l} \lambda_1 = 1 \to v_1 = \frac{\sqrt{2}}{2}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \\ \lambda_2 = 16 \to v_2 = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \cdot \frac{\sqrt{11}}{11} \\ \lambda_3 = 49 \to v_3 = \frac{\sqrt{22}}{11}\begin{pmatrix} -3/2 \\ -3/2 \\ \end{pmatrix} \end{array}\right\} \text{главные компоненты}$$

Продолжение 17 задачи

$\frac{1}{N-1}\lambda = \frac{1}{N-1}\sigma^2$ ⟵ дисперсия по главным компонентам

Ответ: а) $\vartheta_1 = \frac{\sqrt{2}}{2}\begin{pmatrix}-1\\1\\0\end{pmatrix}$

$\vartheta_2 = \frac{\sqrt{11}}{11}\begin{pmatrix}1\\1\\3\end{pmatrix}$

$\vartheta_3 = \frac{\sqrt{22}}{22}\begin{pmatrix}-3\\-3\\2\end{pmatrix}$

} Главные компоненты

б) $\frac{1}{N-1}\sigma_1^2 = 0{,}25$

$\frac{1}{N-1}\sigma_2^2 = 4$

$\frac{1}{N-1}\sigma_3^2 = 12{,}25$

} дисперсия по главным компонентам

Задача 37

Как должна выглядеть ф-я потерь, чтобы минимуму среднему риску давала условная ...

$$L(y', y) = \mathbb{I}(y' \neq y)$$

$$R(f) = \int\limits_{Y}\int\limits_{X} L(f(x), y)\, p(y|x)\, p(x)\, dy\, dx = \sum_{y=1}^{K} \int \mathbb{I}(f(x), y) \times P_r(y|x)$$

$$R(f) = \int\limits_{X} \left(1 - P_r(y = f(x)|x)\right) p(x)\, dx$$

$$f^*(x) = \underset{y \in Y}{\arg\min} \left(1 - P_r(y|x)\right) = \underset{y \in Y}{\arg\max} P_r(y|x)$$