

N1а) Выборка  $y_i$  Модель:  $\tilde{y} = \text{const}$ 

$$L = \sum_{i=1}^l (y_i - \tilde{y})^2 \rightarrow \min_{\tilde{y}}$$

$$\frac{\partial L}{\partial \tilde{y}} = -2 \sum_{i=1}^l (y_i - \tilde{y}) = 0 \Rightarrow \sum_{i=1}^l (y_i - \tilde{y}) = 0 \quad ; \quad \sum_{i=1}^l y_i = l \tilde{y} \Rightarrow \tilde{y} = \frac{1}{l} \sum_{i=1}^l y_i = \langle y_i \rangle$$

б)  $\tilde{y} = kx + b$ 

$$L = \sum_{i=1}^l (y_i - kx_i - b)^2$$

$$\begin{cases} \frac{\partial L}{\partial k} = -2 \sum_{i=1}^l (y_i - kx_i - b)x_i = 0 \\ \frac{\partial L}{\partial b} = -2 \sum_{i=1}^l (y_i - kx_i - b) = 0 \end{cases}$$

$$\begin{cases} \sum_{i=1}^l (x_i y_i - kx_i^2 - bx_i) = 0 & | \cdot \frac{1}{l} \\ \sum_{i=1}^l (y_i - kx_i - b) = 0 & | \cdot \frac{1}{l} \end{cases}$$

$$(*) \begin{cases} \langle xy \rangle - k \langle x^2 \rangle - b \langle x \rangle = 0 \\ \langle y \rangle - k \langle x \rangle - b = 0 \quad | \cdot \langle x \rangle \end{cases} \quad \begin{cases} \langle xy \rangle - k \langle x^2 \rangle - b \langle x \rangle = 0 \\ \langle y \rangle \langle x \rangle - k \langle x \rangle^2 - b \langle x \rangle = 0 \end{cases}$$

$$\langle xy \rangle - \langle x \rangle \langle y \rangle - k(\langle x^2 \rangle - \langle x \rangle^2) = 0$$

$$k = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\langle x^2 \rangle - \langle x \rangle^2}$$

$$\langle y \rangle - k \langle x \rangle = b \Rightarrow b = \langle y \rangle - \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\langle x^2 \rangle - \langle x \rangle^2} \langle x \rangle$$

Из (\*) видно, что  $\langle y \rangle = k \langle x \rangle + b \Rightarrow$  аппроксимирующая прямая проходит через точку  $(\langle x \rangle, \langle y \rangle)$

Для многомерного случая:

$$\tilde{y} = Xw \quad \tilde{y}_i = \sum_{\alpha} x_{i\alpha} w_{\alpha} \quad w - \text{столбец весов признаков}$$

$$L = \sum_{i=1}^l (\tilde{y}_i - y_i)^2 \rightarrow \min$$

$$L = \sum_{i=1}^l (\sum_{\alpha} x_{i\alpha} w_{\alpha} - y_i)^2 \quad \frac{\partial L}{\partial w_{\alpha_1}} = 2 \sum_{i=1}^l (\sum_{\alpha} x_{i\alpha} w_{\alpha} - y_i) x_{i\alpha_1} = 0$$

$$\text{Для столбца свободных членов } x_{i\alpha_k} = 1 \quad k: \Rightarrow \frac{\partial L}{\partial w_{\alpha_k}} = 2 \sum_{i=1}^l (\sum_{\alpha} x_{i\alpha} w_{\alpha} - y_i) = 0$$

$$\langle \sum_{\alpha} x_{i\alpha} w_{\alpha} \rangle = \langle y \rangle \Rightarrow \langle y \rangle = w_1 \langle x_1 \rangle + w_2 \langle x_2 \rangle + \dots + w_k \langle x_k \rangle + w_k$$

проходит через "центр масс" выборки.  $\langle x_i \rangle$  - среднее  $i$ -ого столбца мат. X

2

$$y = x_1 w_1 + x_2 w_2 + \dots + x_n w_n + w_0 = Xw + w_0$$

$$\langle y \rangle = \langle x_1 \rangle w_1 + \langle x_2 \rangle w_2 + \dots + \langle x_n \rangle w_n + w_0 = \bar{X} w + w_0$$

$$\underbrace{y - \langle y \rangle}_{\bar{y}} = \underbrace{(x - \bar{x})}_{\tilde{X}} w$$

провели централизацию признаков, или, грубо говоря, сдвинули координаты так, что гиперплоскость проходит через нуль.

Рассмотрим ту же задачу, только введём новый признак

$$f_0(x) = 1 = x_0$$

$$y = x_1 w_1 + x_2 w_2 + \dots + x_n w_n + x_0 w_0$$

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} & 1 \\ x_{21} & x_{22} & \dots & x_{2n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} & 1 \end{pmatrix} \begin{pmatrix} \vec{w} \\ w_0 \end{pmatrix} = \hat{X} \hat{w}$$

Теперь имеем дело со смещённой в нуль задачей

$$\text{Тогда } w = (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T y$$

$$w' = (\hat{X}^T \hat{X})^{-1} \hat{X}^T y$$

$$\tilde{X} \in \text{Mat}_{m \times n}(\mathbb{R}) \Rightarrow (\tilde{X}^T \tilde{X})^{-1} \in \text{Mat}_{n \times n}(\mathbb{R})$$

$$\tilde{X}^T \in \text{Mat}_{n \times m}(\mathbb{R})$$

$$\hat{X} \in \text{Mat}_{m \times (n+1)}(\mathbb{R}) \Rightarrow (\hat{X}^T \hat{X})^{-1} \in \text{Mat}_{(n+1) \times (n+1)}(\mathbb{R})$$

$$\hat{X}^T \in \text{Mat}_{(n+1) \times m}(\mathbb{R})$$

Если вычеркнуть у  $\hat{X}$  последний столбец, а у  $(\hat{X}^T \hat{X})^{-1}$  последние столбец и строку, то получим  $w'$  без последней строки, т.е.  $w$

Таким образом, мы получили два эквивалентных подхода по поиску весов, не стоящих перед свободным членом: либо смещаем начало координат, либо добавляем новый признак (то же самое, что добавить новую координатную ось и по МНК искать гиперплоскость на 1 порядок выше исходной задачи)



№3

$$y = Xw \quad X \in \text{Mat}_{m \times n}(\mathbb{R}) \quad m < n, \quad \text{rg } X = m$$

Будем искать такую псевдообратную матрицу  $Q$ , что

$$L = \|QX - E_n\|^2 \rightarrow \min_{Q \in \text{Mat}_{n \times m}(\mathbb{R})}, \quad \text{где } \|A\|^2 = \sum_{i,j} |a_{ij}|^2$$

$$\frac{\partial L}{\partial Q} = 2(QX - E_n)X^T \Rightarrow QX^T = X^T$$

$$\Rightarrow \boxed{Q = X^T (XX^T)^{-1}} \quad \text{— правая псевдообратная матрица}$$