# Введение в АД

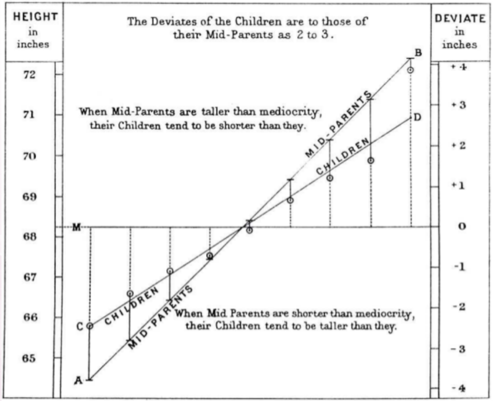
Лекция 3

# Линейная регрессия

# Первое упоминание регрессии

Впервые регрессия упоминается в работе Гальтона "*Регрессия к середине в наследственности роста*", 1885 г.

Пусть $x$ — рост родителей, $y$ — рост детей.
Установлена зависимость $y - \overline{y} \approx \frac{2}{3}(x - \overline{x})$, т.е. регрессия к середине.



# Модель линейной регрессии

Пусть $\mathscr{X} \subset \mathbb{R}^d$ — множество признаков, $\mathscr{Y} = \mathbb{R}$ — таргеты.

Рассматриваем зависимость вида

$$y(x) = \theta_1 x_1 + \ldots + \theta_d x_d,$$

где $x_1, \ldots, x_d$ — признаки,

$\theta = (\theta_1, \ldots, \theta_d)^T$ — вектор параметров.

**Простой пример**

$y = \theta_0 + \theta_1 x,$

$x$ — рост котика,

$y$ — потребление еды,

$\theta_0, \theta_1$ — неизвестные параметры.

Зависимость

- **линейна по параметрам**,
- линейна по аргументу.

**Более сложный пример**

$y = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_2^2,$

$x_1$ — рост котика,

$x_2$ — вес котика,

$y$ — потребление еды,

$\theta_0, \theta_1, \theta_2, \theta_3$ — неизвестные параметры.

Зависимость

- **линейна по параметрам**,
- квадратична по аргументам.

## Нелинейные признаки

Зависимость $y = y(x)$ должна быть **линейна по параметрам**, но не обязана быть линейной по признакам.

Пусть $z_1, ..., z_k$ — набор "независимых" переменных. Можно рассматривать модель $y(x) = \theta_1 x_1(z_1, ..., z_k) + ... + \theta_d x_d(z_1, ..., z_k)$, где $x_j(z_1, ..., z_k)$ — некоторые функции, м.б. нелинейные.

Примеры: $x(z_1, ..., z_k) = 1, \quad x(z_1, ..., z_k) = z_1, \quad x(z_1, ..., z_k) = z_1^2 \ln z_2.$

## Матричная запись

Представим данные в матричном виде

$$Y = \begin{pmatrix} Y_1 \\ ... \\ Y_n \end{pmatrix}, \qquad X = \begin{pmatrix} x_{11} & \dots & x_{1d} \\ ... & & \\ x_{n1} & \dots & x_{nd} \end{pmatrix}.$$

Линейная регрессия предполагает зависимость $Y = X\theta$.

# Пример: Потребление мороженого

Предполагается линейная зависимость потребления мороженого в литрах на человека от среднесуточной температуры воздуха: $ic = \theta_0 + \theta_1 t$.

В этом примере $x_0(t) = 1, \quad x_1(t) = t$,

$$X = \begin{pmatrix} 1 & t_1 \\ \dots & \\ 1 & t_n \end{pmatrix}, \; Y = \begin{pmatrix} IC_1 \\ \dots \\ IC_n \end{pmatrix}, \; \theta = \begin{pmatrix} \theta_0 \\ \theta_1 \end{pmatrix}.$$

Пусть $w = I\{\text{выходной день}\}$, зависимость $ic = \theta_0 + \theta_1 t + \theta_2 t^2 w$.
В этом примере $x_0(t, w) = 1, \quad x_1(t, w) = t, \quad x_2(t, w) = t^2 w$,

$$X = \begin{pmatrix} 1 & t_1 & t_1^2 w_1 \\ \dots & & \\ 1 & t_n & t_n^2 w_n \end{pmatrix}, \; Y = \begin{pmatrix} IC_1 \\ \dots \\ IC_n \end{pmatrix}, \; \theta = \begin{pmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{pmatrix}.$$

# Материал по доске

## Отступление в матричное дифференцирование

Пусть $f : \mathbb{R}^n \to \mathbb{R}$. Тогда

$$\frac{\partial f}{\partial x} = \begin{pmatrix} \frac{\partial f}{\partial x_1} & \dots & \frac{\partial f}{\partial x_n} \end{pmatrix} \text{ — производная (вектор-строка)}$$

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \dots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} \text{ — градиент (вектор-столбец).}$$

**Пример 1**

$f(x) = a^T x$, где $a, x \in \mathbb{R}^n$

$$\frac{\partial f}{\partial x} = a^T \text{ — производная}$$

$$\nabla f = a \text{ — градиент}$$

**Пример 2**

$f(x) = x^T A x$, где $x \in \mathbb{R}^n$,
матрица $A \in \mathbb{R}^{n \times n}$ симметрична

$$\frac{\partial f}{\partial x} = 2x^T A \text{ — производная}$$

$$\nabla f = 2Ax \text{ — градиент}$$

## Пример

Данные:

| | | | |
|---|---|---|---|
| $x$ | 0 | 1 | 2 |
| $y$ | 0 | 4 | 7 |

Модель

$y(x) = \theta_0 + \theta_1 x$

Найдите оценку коэффициентов.

## Решение

Матричный вид: $X = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}$, $Y = \begin{pmatrix} 0 \\ 4 \\ 7 \end{pmatrix}$, $\theta = \begin{pmatrix} \theta_0 \\ \theta_1 \end{pmatrix}$.

Оценка коэффициентов $\widehat{\theta} = \left(X^T X\right)^{-1} X^T Y = \begin{pmatrix} 1/6 \\ 21/6 \end{pmatrix}$

Обученная модель

$$y(x) = \frac{1}{6} + \frac{21}{6} x$$

# Материал по доске


ВСЁ!