

# Линейная регрессия

Линейная регрессия — метод восстановления зависимости между двумя переменными. Линейная означает, что мы предполагаем, что переменные выражаются через уравнение вида:

$$y = ax + b + \varepsilon. \quad (1)$$

Эпсилон в формуле (1) — это ошибка модели.

## Метод наименьших квадратов

Суть МНК заключается в том, чтобы отыскать такие параметры  $\Theta$ , чтобы предсказанное значение было наиболее близким к реальному. Математически это выглядит так:

$$\sum_{i=0}^n (y_i - \hat{y}_i)^2 \longrightarrow \min_{\Theta}. \quad (2)$$

Требуется найти такой вектор  $\Theta$  (2), при котором выражение:

$$\sum_{i=0}^n (y_i - f(x_i, \Theta))^2, \quad (3)$$

достигает минимума.

Длинные преобразования из статьи. Демонстрация многострочной формулы:

$$\begin{aligned} (Y - A\Theta)^T(Y - A\Theta) &= \\ &= Y^T Y - (A\Theta)^T Y - Y^T A\Theta + (A\Theta)^T A\Theta. \end{aligned} \quad (4)$$

А вот ссылка на статью <https://habr.com/ru/post/307004/>.