

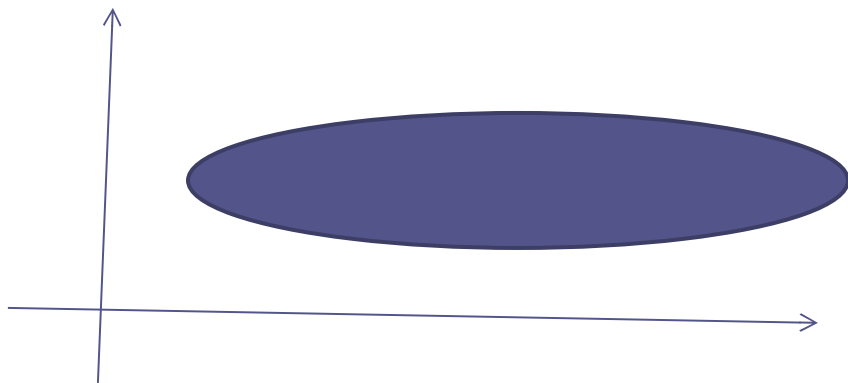
Урок 2. Масштабирование признаков. Регуляризация. Стохастический градиентный спуск.

2021

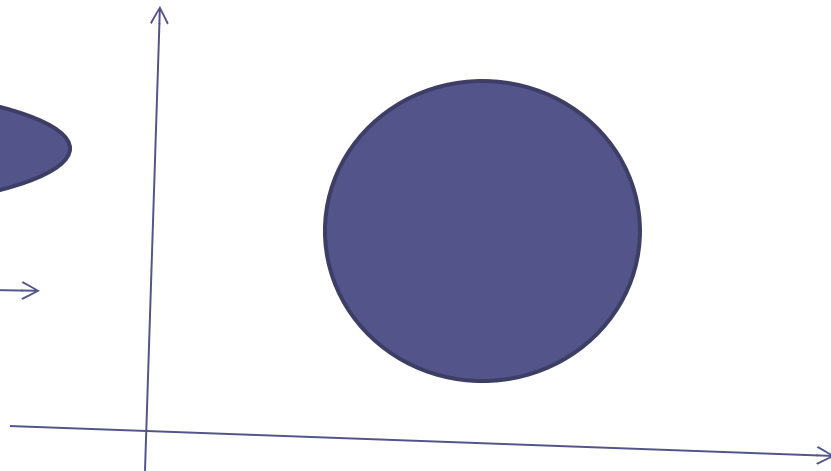
Мария Корлякова

Масштабирование признаков

- Без масштаба

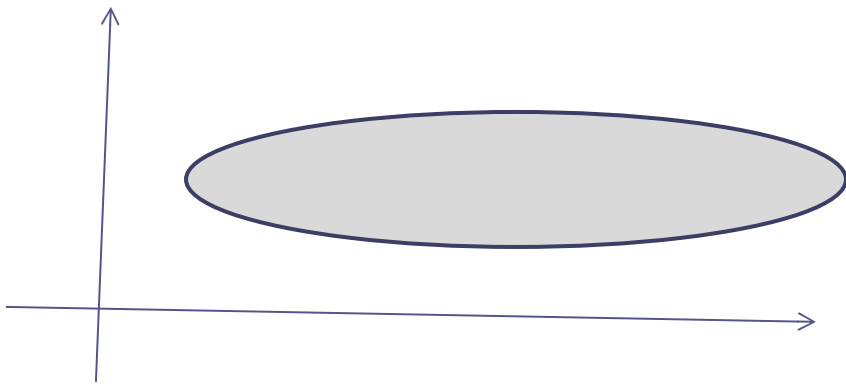


С масштабом

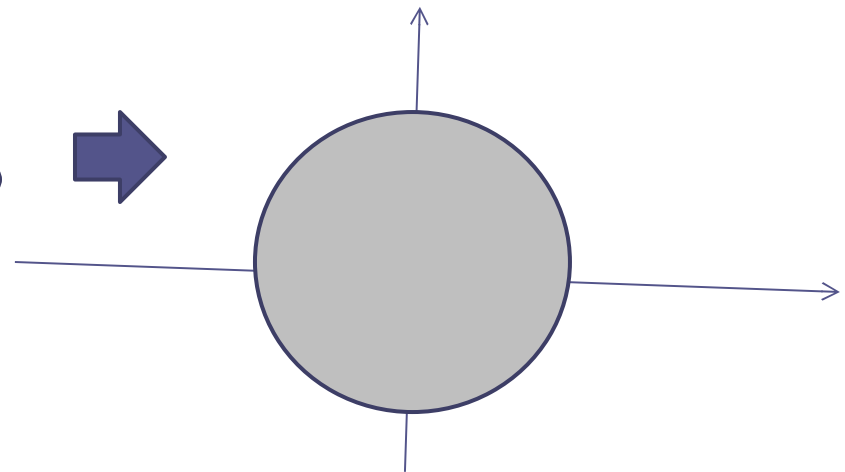


Масштабирование признаков

- Без масштаба



- С масштабом



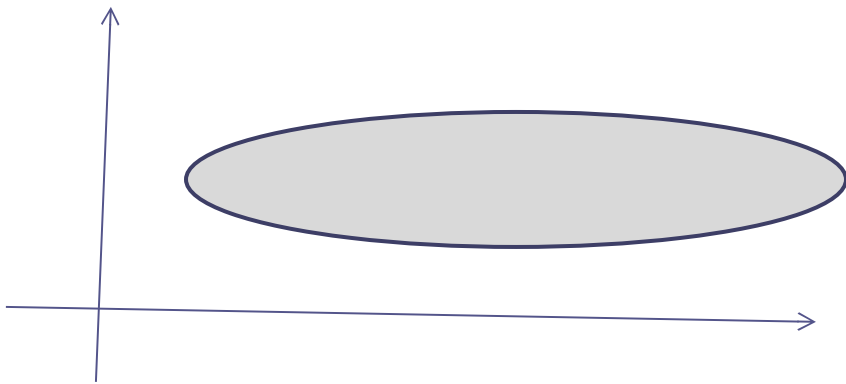
- Стандартизация

$$\underline{x_i^j = \frac{x_i^j - \mu_j}{\sigma_j} .}$$

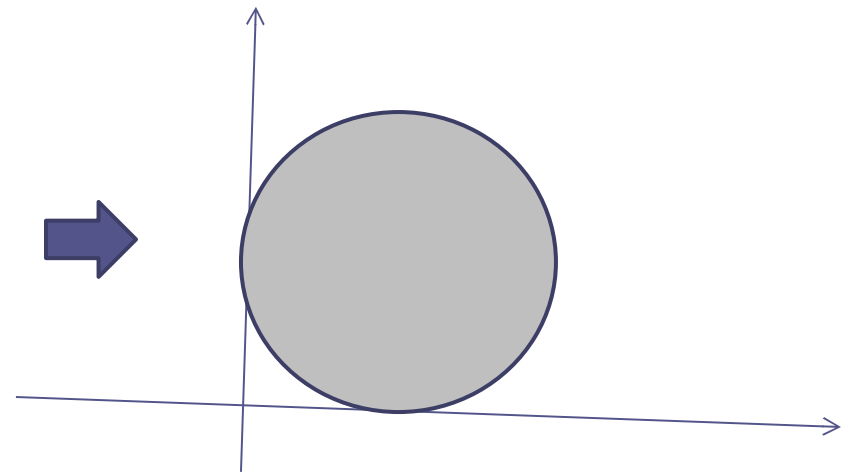
$$\mu_j = \frac{1}{l} \sum_{i=1}^l x_i^j$$
$$\sigma_j = \sqrt{\frac{1}{l} \sum_{i=1}^l (x_i^j - \mu_j)^2}$$

Масштабирование признаков

- Без масштаба



- С масштабом



- Нормализация

$$x_i^j = \frac{x_i^j - \min_j(x_i^j)}{\max_j(x_i^j) - \min_j(x_i^j)}.$$

Стохастический градиентный спуск

$$w^k = w^{k-1} - \eta_k \nabla Q(w^{k-1}, X).$$

$$\nabla_w Q(w, X) = \frac{2}{l} X^T (Xw - y).$$

$$\frac{\partial Q}{\partial w_j} = \frac{2}{l} \sum_{i=1}^l x_i^j (\langle w, x_i \rangle - y_i).$$

- Экономим ресурсы
- Stochastic gradient descent, SGD

Стохастический градиентный спуск

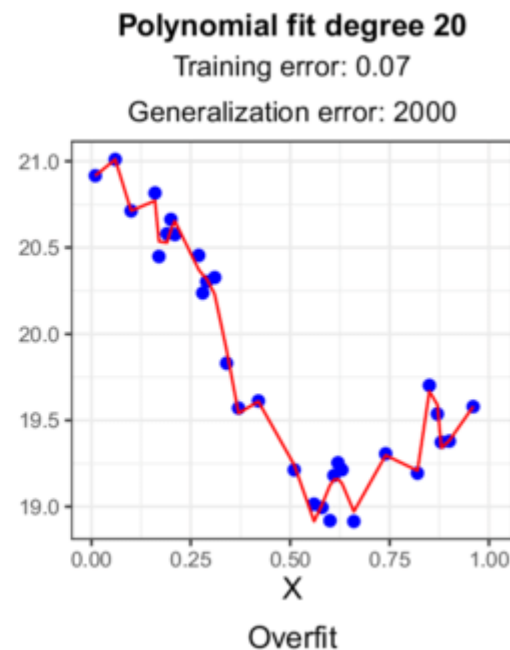
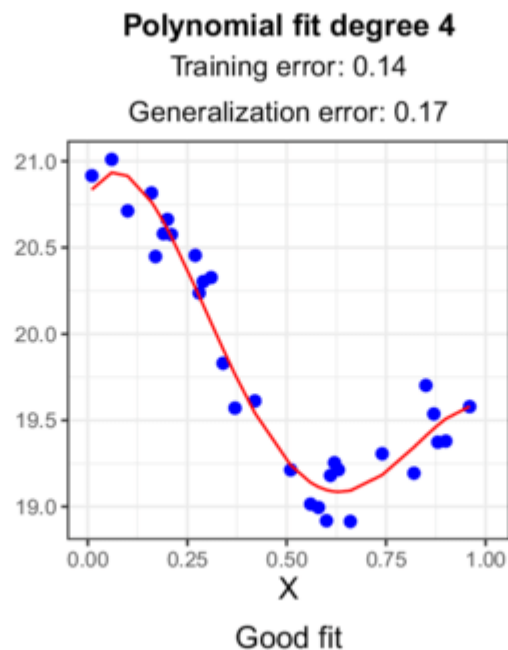
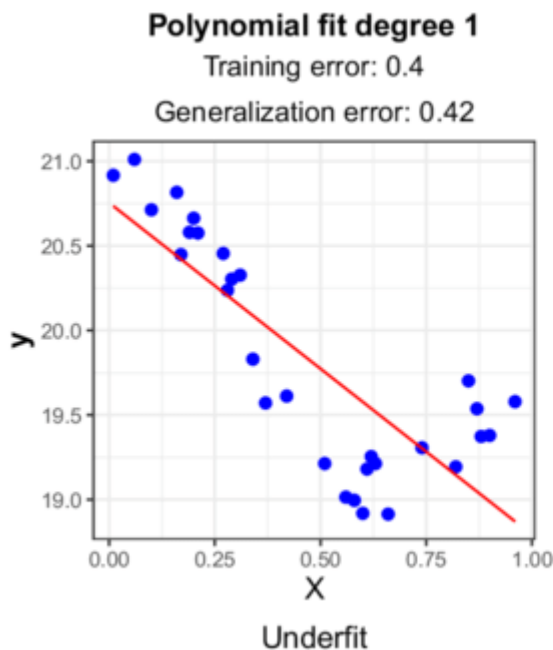
- Stochastic gradient descent, SGD

$$w^k = w^{k-1} - \eta_k \nabla Q(w^{k-1}, \{x_i\}),$$

$$Q := (1 - \lambda)Q + \lambda Q_i$$

Переобучение

- Overfitting
- Underfitting



Переобучение

- Cross-validation
- Hold Out
- Регуляризация

Регуляризация

- Штраф сложности

- L2

$$Q(w, X) + \lambda ||w||^2 \rightarrow \min_w.$$

- L1

$$||w||_1 = \sum_{j=1}^d |w_j|.$$

R² - детерминация

$$R^2 = 1 - \frac{\sum_{i=1}^l (a(x_i) - y_i)^2}{\sum_{i=1}^l (y_i - \bar{y})^2}$$

