

Градиентный спуск.

Частная производная – производная по одной переменной в случае, если функция имеет несколько переменных.

Градиент – вектор, который указывает направление роста функции

Найти градиент для $Z = 7 - 5x^2 - 10y^2$ в точке $M(-5; 7)$

$$\text{grad } z = \frac{\partial z}{\partial x} \vec{i} + \frac{\partial z}{\partial y} \vec{j} = \left(\frac{\partial z}{\partial x}; \frac{\partial z}{\partial y} \right)$$

$$\frac{\partial z}{\partial x} = (7 - 5x^2 - 10y^2)'_x = -10x$$

$$\frac{\partial z}{\partial y} = (7 - 5x^2 - 10y^2)'_y = -20y$$

$$\text{grad } z = -10x\vec{i} - 20y\vec{j} = (-10x; -20y)$$

$$\text{grad}|_M z = -10 * (-5)\vec{i} - 20 * 7 * \vec{j} = (50; -140)$$

$$|\text{grad } z| = \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2} = \sqrt{50^2 + (-140)^2} = 10\sqrt{221} \approx 148,66$$

Вывод: Если движение происходит в направлении градиента функции $(50; -140)$, то получаем скорость максимального изменения функции $148,66$ в точке $M(-5; 7)$.

Зачем всё это?

Линейная регрессия – модель машинного обучения, где предсказываемое значения является суммой взвешенных признаков (факторов).

$$f(x, b) = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_k x_k$$

Функция потерь – характеризует ошибку прогноза на некоем наборе наблюдаемых данных

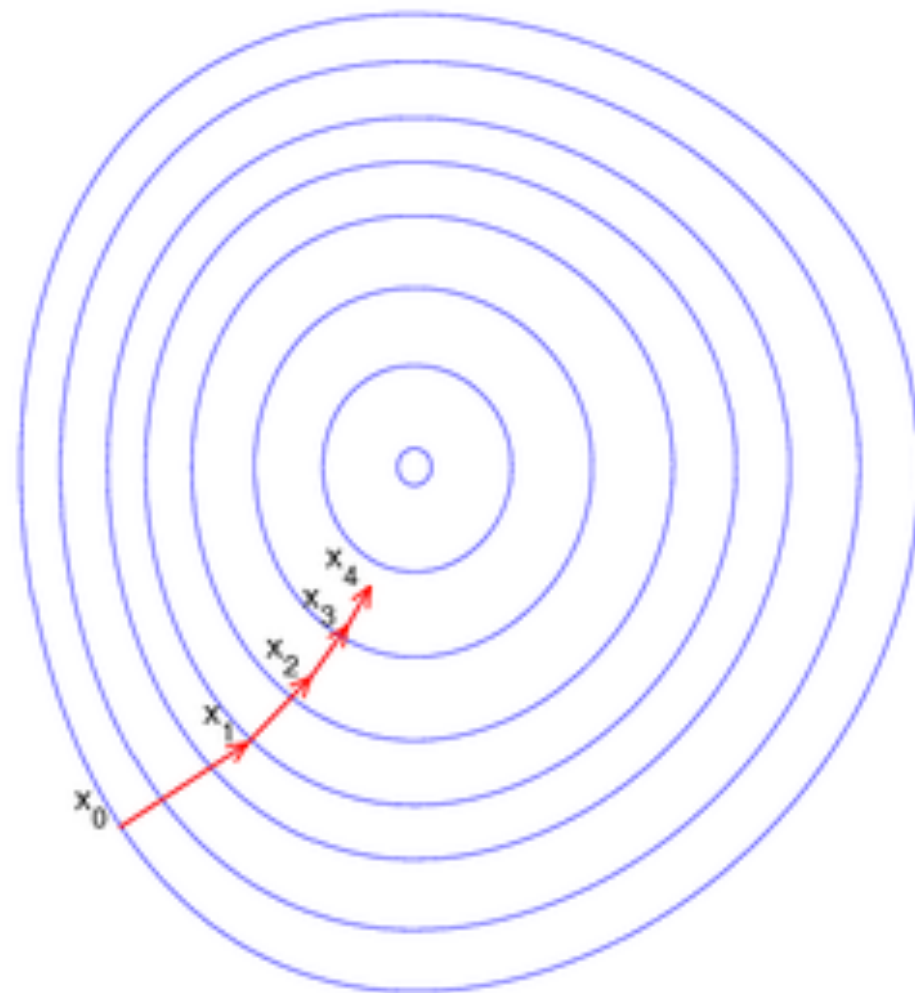
$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

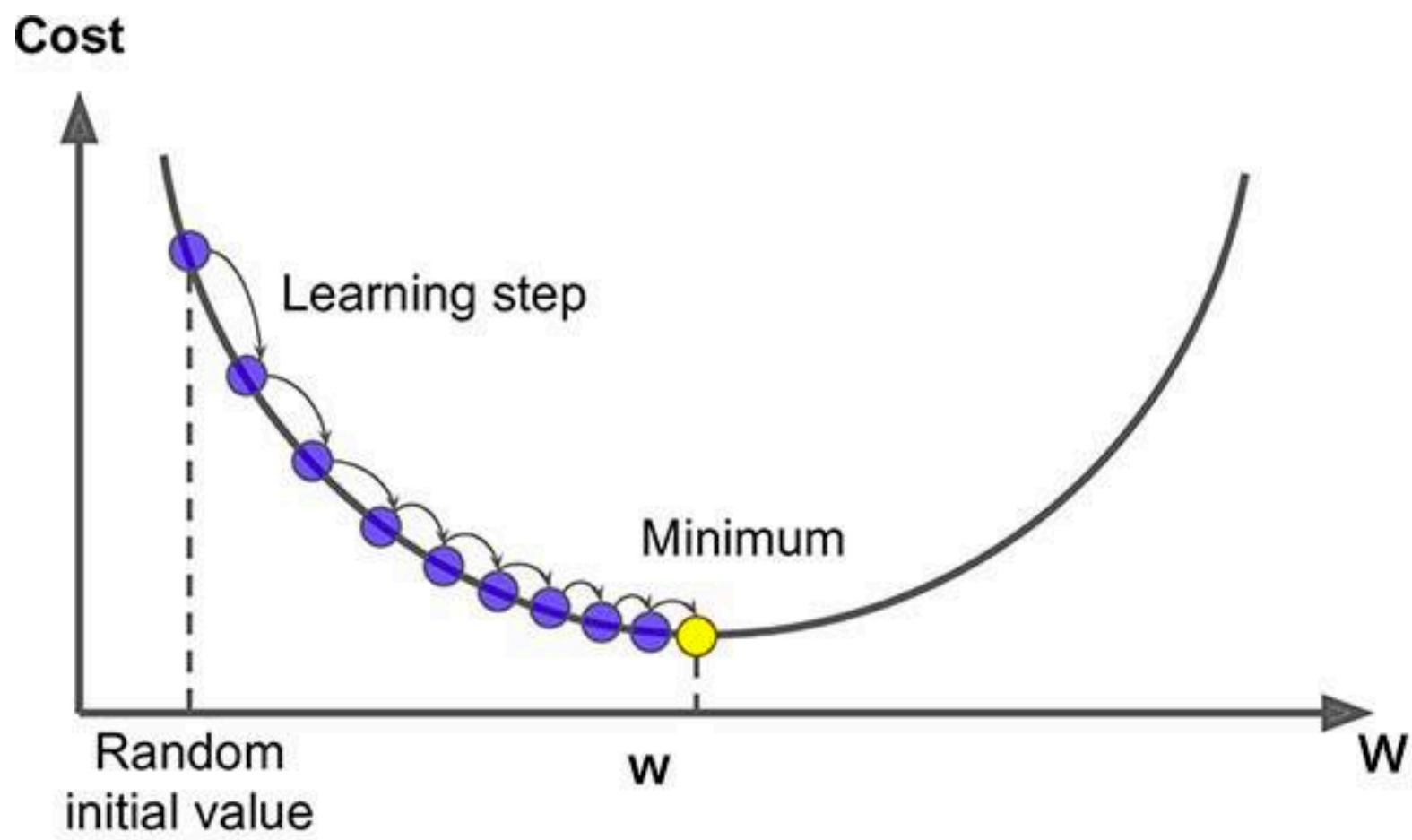
Градиентный спуск – метод нахождения локального минимума или максимума функции.

Repeat until convergence {

$$\theta_j \leftarrow \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$$

}





Rate: 0.01

Step: 0

Func value=59.689

$\theta = -0.400$

