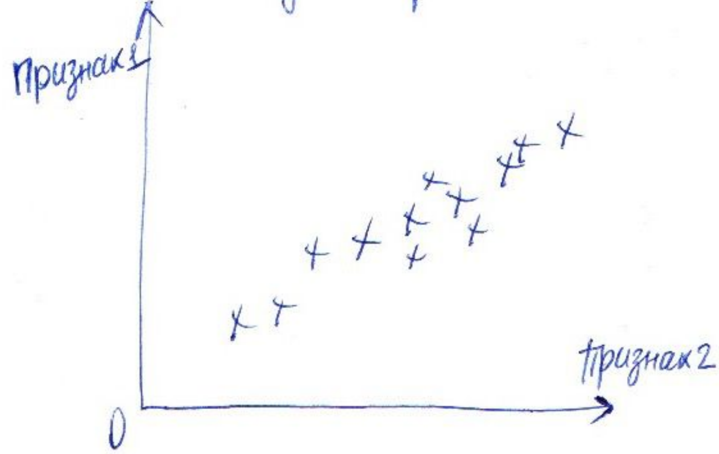
Линейная регрессия на функции одной переменной



Задача: найти такие θ_0 и θ_1, чтобы $J(\theta_0, \theta_1) \to min$.

угол наклона

$$h_\theta(x) = \underline{\underline{\theta_0}} + \theta_1 x$$

смещение

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \Big(\underbrace{h_\theta\left(x^{(i)}\right)}_{= \left(\theta_1 x^{(i)} + \theta_0\right)} - y^{(i)} \Big)^2$$

число точек в пространстве признаков!

Градиентный спуск:

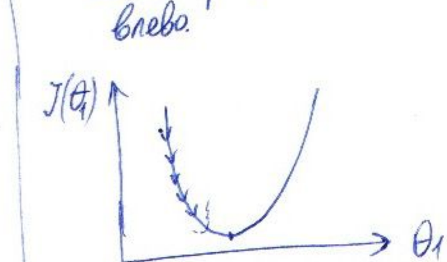
1) $\left.\begin{matrix} \theta_0 = 0 \\ \theta_1 = 0 \end{matrix}\right\}$ ← начальные значения

2) На каждом шаге меняем θ_0 и θ_1 одновременно:

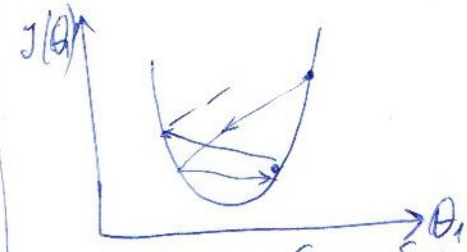
$$\begin{cases} \theta_0 := \theta_0 - \alpha \cdot \boxed{\frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1)} \to 0 \\ \theta_1 := \theta_1 - \alpha \cdot \frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1) \end{cases}$$

коэффициент скорости обучения (learning rate)

Если производная положит., сдвигаемся влево.



α слишком маленький



α слишком большой - перескакиваем минимум

Линейная регрессия на нескольких переменных

Признак 1	Пр 2	Пр 3	Пр 4	y ← Целевая переменная
2104	5	1	45	460
1416	3	2	40	232
1534	3	2	30	315
852	2	1	36	178
...	...	...	...	...

$m = 47$

$$x^{(2)} = \begin{bmatrix} 1416 \\ 3 \\ 2 \\ 40 \end{bmatrix} \leftarrow x_3^{(2)}$$

$$h_\theta(x) = \theta_0 x_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \cancel{\theta_4 x_4} = [\theta_0\ \theta_1\ \theta_2\ \theta_3] \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = \theta^T x$$

x_0 — принято считать $= 1$

Градиентный спуск

Повторять до сходимости {

$$\theta_0 := \theta_0 - \alpha \cdot \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right) \cdot x_0^{(i)}$$

→ номер признака (всего в задаче 4 (от 0-го до 3-го))

$$\theta_1 := \theta_1 - \alpha \cdot \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right) \cdot x_1^{(i)}$$

$$\theta_2 := \theta_2 - \alpha \cdot \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right) \cdot x_2^{(i)}$$

$$\theta_3 := \theta_3 - \alpha \cdot \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right) \cdot x_3^{(i)}$$

← Случай для 4 признаков

}

Нормировка признаков

$$x_i := \frac{x_i - \mu_i}{S_i}$$

μ_i ← среднее всех значений признака i

S_i: СКО = max − min (ат) (ат)