

Stochastic Gradient descent

↳ picks randomly 1 sample and calculate θ

$$\alpha = 0.1$$

$$\theta_j = \theta_j + \alpha \sum_{i=1}^n (y^{(i)} - h_{\theta}(x^{(i)})) x_j^{(i)}$$

sample(i)	x_1	x_2	y
1	2	1	3
2	2	1	4
3	3	3	5

- ① Compute prediction $- h_{\theta}(x)$
- ② Compute gradients
- ③ Update parameters

$$h_{\theta}(x') = \theta_0 x_0 + \theta_1 x_1 + \theta_2 x_2$$

$$h_{\theta}(x^1) = \theta_0 + 2\theta_1 + 1\theta_2$$

$$h_{\theta}(x^2) = \theta_0 + 2\theta_1 + 1\theta_2$$

$$h_{\theta}(x^3) = \theta_0 + 3\theta_1 + 3\theta_2$$

$$\alpha = 0.1$$

$$1^{st} \text{ iteration, } \theta_0 = \theta_1 = \theta_2 = 0$$

$$\begin{aligned}\theta_0 &= 0 + 0.1(3 - 0)1 + 0.1(4 - 0)1 + 0.1(5 - 0)1 \\ &= 0.3 + 0.4 + 0.5 \\ &= 1.2\end{aligned}$$

$$\theta_1 = 0 + 0.1(3 - 0)2 + 0.1(4 - 0)2 + 0.1(5 - 0)3$$

$$= 0 + 0.6 + 0.8 + 1.5$$

$$= 2.9$$

$$\theta_2 = \theta_2 + \alpha (y_1^3 - h_{\theta} x_1^3) x_1^3 + \alpha (y_2^3 - h_{\theta} x_2^3) x_2^3 + \alpha (y_3^3 - h_{\theta} x_3^3) x_3^3$$

$$= 0 + 0.1 (3-0) \times 1 + 0.1 (4-0) \times 1 + 0.1 (3-0) \times 3$$

$$= 0.3 + 0.4 + 1.5 = 2.2$$

$$= 0 + 1 + 0.5$$

$$\theta_2 = 1.5$$

1st iteration ,

$$\theta_0 = 1.2 \quad \theta_1 = 2.9, \quad \theta_2 = 2.2$$

$$J(\theta) = \frac{1}{2} \sum_{i=1}^n (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

Sample $h_{\theta} X$

$$1 \quad \theta x^1 = (1.2 \times 1) + (2.9 \times 2) + (2.2 \times 1) = 9.2$$

$$2 \quad \theta x^2 = 9.2$$

$$3 \quad \theta x^3 = (1.2 \times 1) + (2.9 \times 3) + (2.2 \times 3) = 16.5$$

$$\begin{array}{r} 6.6 \\ 8.7 \\ 2.2 \\ \hline 16.5 \end{array}$$

$$J(\theta) = \frac{1}{2} [(9.2)^2 + (9.2)^2 + (16.5)^2]$$

$$= 220.765$$

2nd iteration

$$\theta = \theta - \alpha \sum_{i=1}^n (h_{\theta} x - y) x$$

$$\begin{array}{r} 6.2 \\ 8.2 \\ 11.5 \\ \hline 22.9 \end{array}$$

$$\theta_0 = 1.2 - 0.1 [(9.2 - 3) \times 1 + (9.2 - 4) \times 1 + (16.5 - 5) \times 1]$$

$$= 1.2 - 0.1 [22.9]$$

$$= 1.2 - 2.29 = -1.09$$

$$\begin{aligned}\theta_1 &= 2.9 - 0.1 [(6.2)(2) + (5.2)(2) + (11.5)(3)] \\ &= 2.9 - 0.1 [54.3] \\ &= -2.83\end{aligned}$$

$$\begin{aligned}\theta_3 &= 2.2 - 0.1 [(6.2)(1) + 5.2(1) + (11.5)(3)] \\ &= 2.2 - 4.59 \\ &= -2.89\end{aligned}$$

After 2nd iteration

$$\theta_0 = -1.09$$

$$\theta_1 = -2.83$$

$$\theta_2 = -2.89$$

$$J(\theta) = \frac{1}{2} \sum_{i=1}^3 (h_{\theta} X - y)^2$$

Sample	$h_{\theta} X - y$
1	$(-1.09 \times 1) + (-2.83 \times 2) + (-2.89 \times 1) = -9.64$
2	-9.64
3	$(-1.09 \times 1) + (-2.83 \times 3) + (-2.89 \times 3) = -18.25$

$$\begin{aligned}J &= \frac{1}{2} [(-9.64)^2 + (-9.64)^2 + (-18.25)^2] \\ &= \frac{518.92}{2} = 259.46\end{aligned}$$

sample(i)	x_1	x_2	y
1	2	1	3
2	2	1	4
3	3	3	5