Julian Tsang

Question 6: Apply gradient descent by hand.

x	y
3	2
1	2
0	1
4	3

Apply Gradient Descent

Let $\alpha = 1/2$ $\quad \theta_0 = 1 \quad \theta_1 = .5$ $\quad h_\theta(x) = \theta_0 + \theta_1 x = 1 + .5x$

$$\theta_j := \theta_j - \alpha \frac{d}{d\theta_j} J(\theta)$$

$$= \theta_j - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right) x_j^{(i)}$$

Find $J(1, 1/2)$:

$$\frac{1}{8}\left[\left((1+\tfrac{3}{2}) - 2\right)^2 + \left[(1+\tfrac{1}{2}) - 2\right]^2 + \left((1+0) - 1\right)^2 + \left[(1+2) - 3\right]^2 \right]$$

$$\frac{1}{8}\left[\left(\tfrac{1}{2}\right)^2 + \left(-\tfrac{1}{2}\right)^2\right] = \frac{1}{16}$$

$$J(1, 1/2) = .0625$$

Iteration 1 $\quad \theta_0 = 1 \quad \theta_1 = .5$

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

$$= 1 - .5\left(\tfrac{1}{4}\right)\left[\left[(1+\tfrac{3}{2}) - 2\right] + \left[(1+\tfrac{1}{2}) - 2\right] + \left[(1+0) - 1\right] + \left[(1+\tfrac{4}{2}) - 3\right]\right](1)$$

$$= 1 - \frac{1}{8}\left[\frac{1}{2} - \frac{1}{2} + 0 + 0\right](1) = 1$$

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

$$.5 - .5\left(\frac{1}{4}\right)\left[\frac{1}{2}(3) - \frac{1}{2}(1) + 0(0) + 0(4)\right]$$

$$\frac{1}{2} - \frac{1}{8}(1) = 3/8$$

$$J(1, 3/8) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2 \qquad h_\theta(x) = 1 + \frac{3}{8}x$$

$$\frac{1}{8}\left(\left(1 + \frac{3}{8}(3)\right) - 2\right)^2 + \left(\left(1 + \frac{3}{8}(1)\right) - 2\right)^2 + \left(\left(1 + \frac{3}{8}(0)\right) - 1\right)^2 + \left(\left(1 + \frac{3}{8}(4)\right) - 3\right)^2$$

$$\frac{1}{8}\left(\left(1 + \frac{9}{8}\right) - 2\right)^2 + \left(\left(1 + \frac{3}{8}\right) - 2\right)^2 + (1 - 1)^2 + \left(\left(1 + \frac{12}{8}\right) - 3\right)^2$$

$$\frac{1}{8}\left[\left(\frac{1}{8}\right)^2 + \left(-\frac{5}{8}\right)^2 + 0^2 + \left(-\frac{1}{2}\right)^2\right]$$

$$\frac{1}{8}\left[\frac{1}{64} + \frac{25}{64} + \frac{1}{4}\right]$$

$$\frac{1}{8}\left[\frac{21}{32}\right] \approx .082 > J(1, .5) \times$$

α too large

Iteration 2:

let $\alpha = .1 \quad \theta_0 = 1 \quad \theta_1 = .5$

$$\theta_0 = 1 - .1\left(\tfrac{1}{4}\right)\left[\tfrac{1}{2} - \tfrac{1}{2} + 0 + 0\right] = 1$$

$$\theta_1 = .5 - .1\left(\tfrac{1}{4}\right)\left[\tfrac{1}{2}(3) - \tfrac{1}{2}(1) + 0(0) + 0(4)\right]$$

$$.5 - .1\left(\tfrac{1}{4}\right)(1) = .475$$

$J(1, .475)$ $\qquad\qquad$ $h_\theta(x) = 1 + .475x$

$$\frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

$$\frac{1}{8}\Big[\left[(1 + .475(3)) - 2\right]^2 + \left[(1 + .475(1)) - 2\right]^2 + \left[(1 + .475(0) - 1\right]^2 + \left[(1 + .475(4) - 3\right]^2$$

$$\frac{1}{8}\left[.180625 + .275625 + 0 + .01\right]$$

$$\frac{1}{8}\left[.46625\right]$$

$$= .05828125 < J(1, 1/2) \checkmark$$

Continue gradient descent with $\alpha = .1$