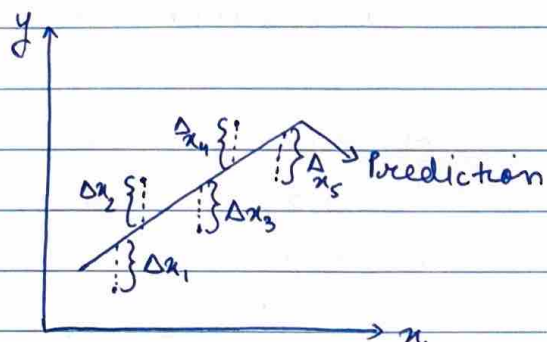


1. Gradient Descent



slope 'intercept

$$y_i = mx_i + c$$

Error / Cost Function

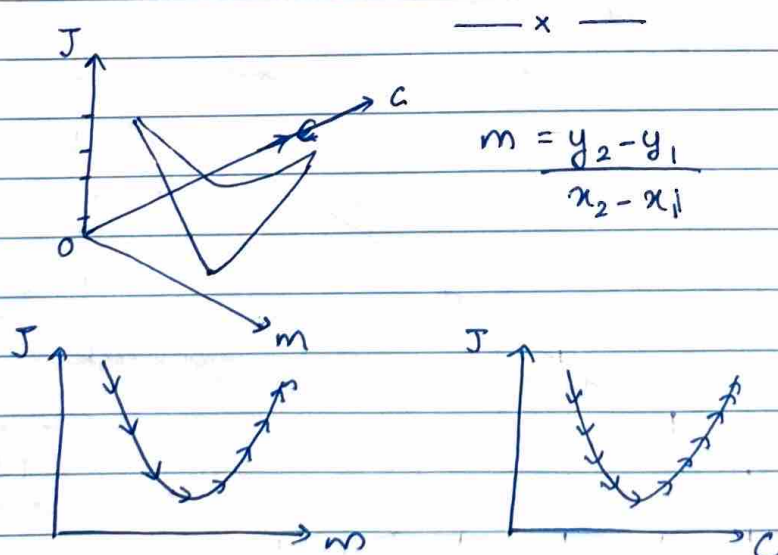
$$MSE = \frac{1}{n} \sum_{i=1}^n [y_i - (mx_i + c)]^2$$

$$J = (y_i - \hat{y})$$

$$J = [y_i - (mx_i + c)]$$

$$J = \sum_{i=1}^n [y_i - (mx_i + c)]^2$$

Sum Squared error (SSE)

learning rate λ, α

$$J = \frac{1}{n} \sum_{i=1}^n [y_i - (mx_i + c)]^2$$

$$\frac{\partial J}{\partial m} = \frac{2}{n} \sum_{i=1}^n (-x_i) [y_i - (mx_i + c)]$$

$$0 = \sum_{i=1}^n -x_i y_i + mx_i^2 + x_i c$$

$$\sum_{i=1}^n mx_i^2 + \sum_{i=1}^n x_i c = \sum_{i=1}^n x_i y_i$$

Diff wrt c, ~~$\frac{\partial J}{\partial c}$~~ $\frac{\partial J}{\partial c} = \frac{2}{n} \sum_{i=1}^n - [y_i - (mx_i + c)]$

$$0 = \sum_{i=1}^n -y_i + mx_i + c$$

$$m \sum_{i=1}^n x_i + \sum_{i=1}^n c = \sum_{i=1}^n y_i$$

$$m = m - \lambda \frac{\partial J}{\partial m}$$

$$c = c - \lambda \frac{\partial J}{\partial c}$$

$$x^n \frac{d}{dx}$$

$$f'(x)$$

$$f''(x) = \frac{d(nx)}{dx} = n$$

$$f'''(x) = 0$$