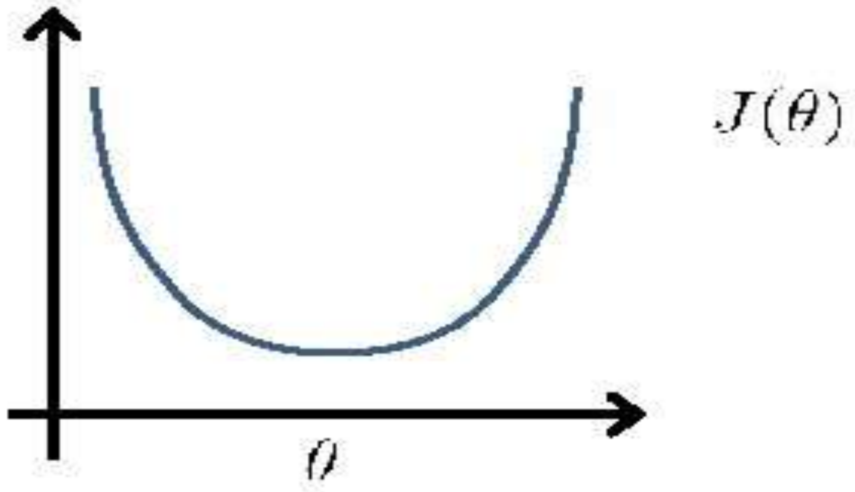
FORTH LECTURE
NORMAL EQUATION

formula

$$\hat{\theta} = (X^T \cdot X)^{-1} \cdot X^T \cdot y$$

We use the normal equation to calculate the optimal weights which leeds us to the global minimum in one step

Gradient Descent



Normal equation: Method to solve for θ analytically.

Andrew Ng

example

$$X = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 1 & 0 & 6 \end{bmatrix} \qquad y = \begin{bmatrix} 5 \\ 20 \\ 15 \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 1 & 0 & 6 \end{bmatrix} \quad X.t = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 4 & 0 \\ 3 & 5 & 6 \end{bmatrix}$$

$$X.t * X = \begin{bmatrix} 2 & 2 & 9 \\ 2 & 20 & 26 \\ 9 & 26 & 70 \end{bmatrix}$$

$$(X*X.t)^{-1} = \frac{1}{||A||} * adj(A)$$

$$||A|| = 1(4*6) - 2(0-5)+3(0-4) = 22$$

calculate the adj by using detirmnant for each number in matrix A

$$adj = \begin{bmatrix} 24 & -5 & -4 \\ 12 & 3 & -2 \\ -2 & 5 & 4 \end{bmatrix} \quad \text{put the sign} = \begin{bmatrix} 24 & 5 & -4 \\ -12 & 3 & 2 \\ -2 & -5 & 4 \end{bmatrix} \text{transpose} = \begin{bmatrix} 24 & -12 & -2 \\ 5 & 3 & -5 \\ -4 & 2 & 4 \end{bmatrix}$$

$$(X*X.t)^{-1} = \frac{1}{||A||} * adj(A)$$

$$\frac{1}{22} * \begin{bmatrix} 24 & -12 & -2 \\ 5 & 3 & -5 \\ -4 & 2 & 4 \end{bmatrix} = \begin{bmatrix} \frac{24}{22} & \frac{-12}{22} & \frac{-2}{22} \\ \frac{5}{22} & \frac{3}{22} & \frac{-5}{22} \\ \frac{-4}{22} & \frac{2}{22} & \frac{4}{22} \end{bmatrix}$$

$$X.T * y = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 4 & 0 \\ 3 & 5 & 6 \end{bmatrix} * \begin{bmatrix} 5 \\ 20 \\ 15 \end{bmatrix} = \begin{bmatrix} 20 \\ 90 \\ 205 \end{bmatrix}$$

$$(X*X.t)^{-1}\, X.T * y = \begin{bmatrix} \frac{24}{22} & \frac{-12}{22} & \frac{-2}{22} \\ \frac{-5}{22} & \frac{3}{22} & \frac{-5}{22} \\ \frac{-4}{22} & \frac{2}{22} & \frac{4}{22} \end{bmatrix} * \begin{bmatrix} 20 \\ 90 \\ 205 \end{bmatrix}$$

$$\begin{bmatrix} -45.909 \\ -29.772 \\ 41.8181 \end{bmatrix}$$

now we can use these parameter to make our prediction

$$\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3$$