# Linear Regression

**Hypothesis:**

$$h_\theta(x) = \theta_0 + \theta_1 X$$

**Parameters:**

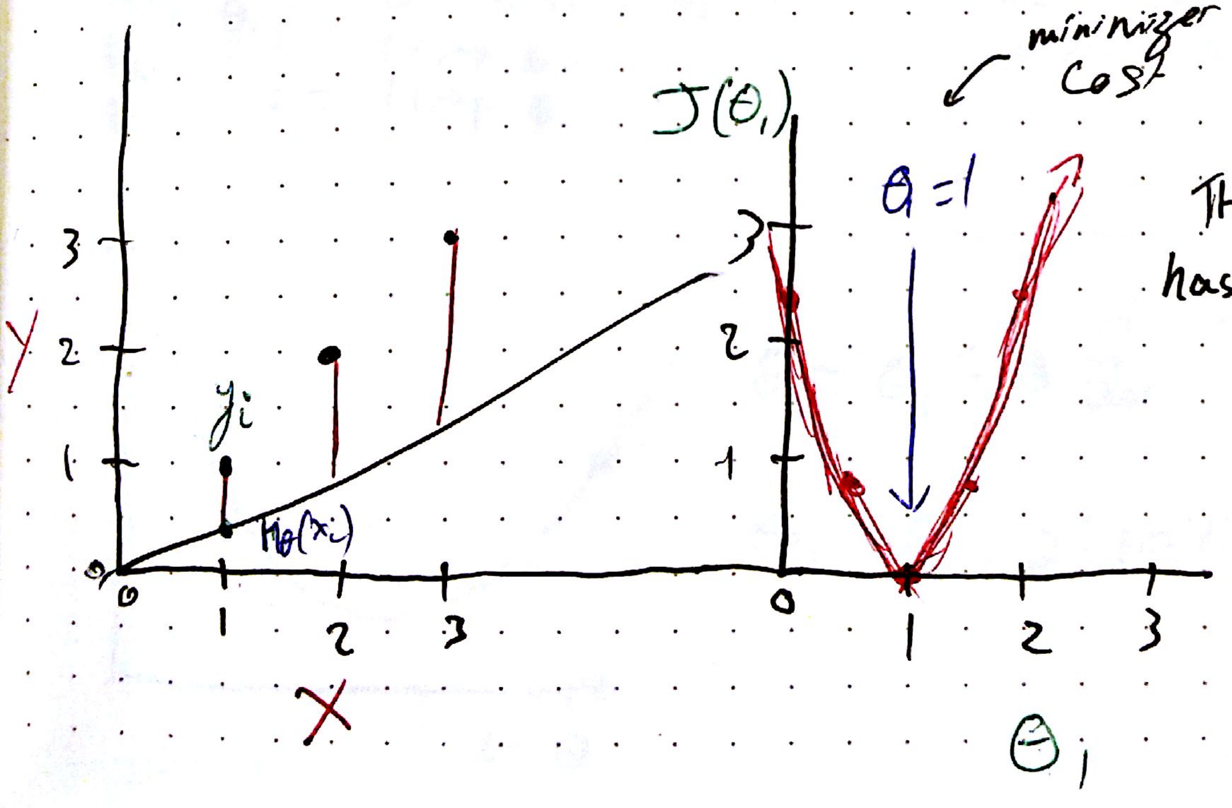
$\theta_0, \theta_1$

**Cost Function:**

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left( h_\theta(x_i) - y_i \right)$$

**Goal:**

$$\underset{\theta_0, \theta_1}{\text{Minimize}} \; J(\theta_0, \theta_1)$$




This parabola has a 3d shape "bowl"

# Linear Gradient Descent

Assignment

## Gradient Descent

$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$

(for $j=0, j=1$)

Have some func: $J(\theta_0, \theta_1)$

Want: $\min_{\theta_0, \theta_1} J(\theta_0, \theta_1)$

Outline:

- Start w/ some $\theta_0, \theta_1$
- Change $\theta_0, \theta_1$ to reduce $J(\theta_0, \theta_1)$ until a minimum is reached

$\alpha$ - learning rate (step size)

## $\alpha$

If $\alpha$ is too small, gradient descent is slow. If too big it can fail to converge or even diverge.

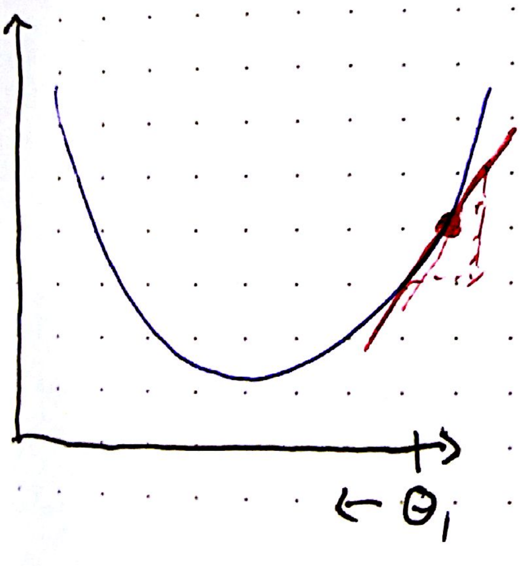
Simultaneous Update:

$temp\text{-}\theta_0 := \theta_0 - \alpha \frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1)$

$temp\text{-}\theta_1 := \theta_1 - \alpha \frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1)$

$\theta_0 := temp\text{-}\theta_0$

$\theta_1 := temp\text{-}\theta_1$




$\theta_1 := \theta_1 - \alpha \frac{d}{d\theta_1} J(\theta_1)$

$\theta_1 := \theta_1 - \alpha$ (positive value)

# Deriving Gradient Descent for Regression

## Gradient Descent

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$$

## Linear Regression Model

$$h_\theta(x) = \theta_0 + \theta_1 x$$

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left( h_\theta(x_i) - y_i \right)^2$$

---

$$\frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1) = \frac{\partial}{\partial \theta_j} \cdot \frac{1}{2m} \sum_{i=1}^{m} \left( h_\theta(x_i) - y_i \right)^2$$

$$= \frac{\partial}{\partial \theta_j} \cdot \frac{1}{2m} \sum_{i=1}^{m} \left( \theta_0 + \theta_1 x_i - y_i \right)^2$$

$\theta_0, j=0$: $\frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1) = \frac{1}{m} \sum_{i=1}^{m} \left( h_\theta(x_i) - y_i \right)$

$\theta_1, j=1$: $\frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1) = \frac{1}{m} \sum_{i=1}^{m} \left( h_\theta(x_i) - y_i \right) \cdot x_i$