Model and cost function:-

Supervised.

Model representation :- Regression | Classification.

Data set → Training set. of housing prices.

Notation:

m = Number of training example

x's = "input" variable | features

y's = output variable | "target" variable

(x, y) - one training example

$(x^{(i)}, y^{(i)})$ - refer i^{th} training example.

Training set

↓

learning algo

↓

[h]

Size of the house as Input → Hypothesis (function) → Estimated price as output

x ... y

linear function

How do we represent h?

$h(x) = \theta_0 + \theta_1 x.$

Linear regression. One variable function

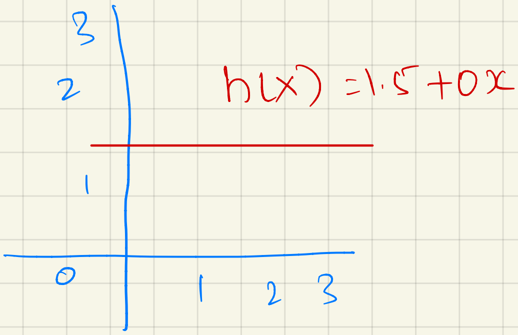
Univariable linear regression

Cost Function :-

Linear Regression

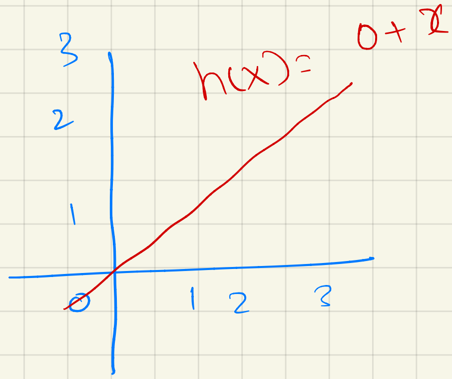
$$h_\theta(x) = \theta_0 + \theta_1 x$$

θ_i is - Parameter of the model



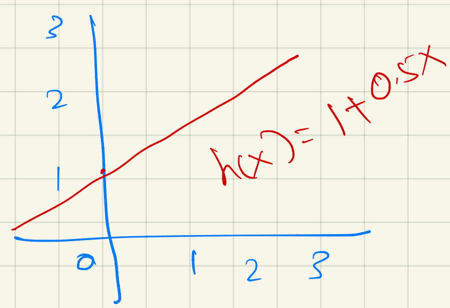
$\theta_0 = 1.5$

$\theta_1 = 1$



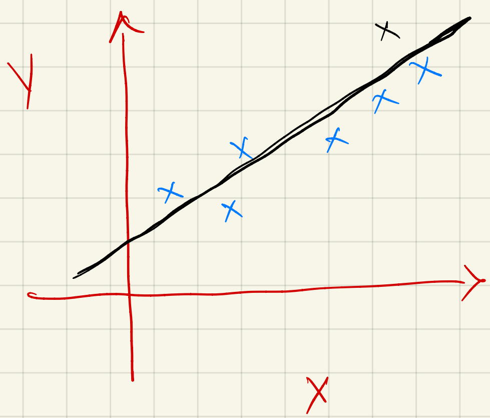
$\theta_0 = 0$

$\theta_1 = 1$



$\theta_0 = 1$

$\theta_1 = 0.5$



(θ_0, θ_1) min close

How to choose θ_0 θ_1 ?

Idea: Choose θ_0, θ_1 so that $h_\theta(x)$ is close to y for our training examples (x, y)

Minimize over θ_0 to θ_1

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

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

Cost Function

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

$$\underset{\theta_0, \theta_1}{\text{minimize}} \; \underbrace{J(\theta_0, \theta_1)}_{\text{Cost Function.}}$$

Cost Function also called Squared Error Cost Function

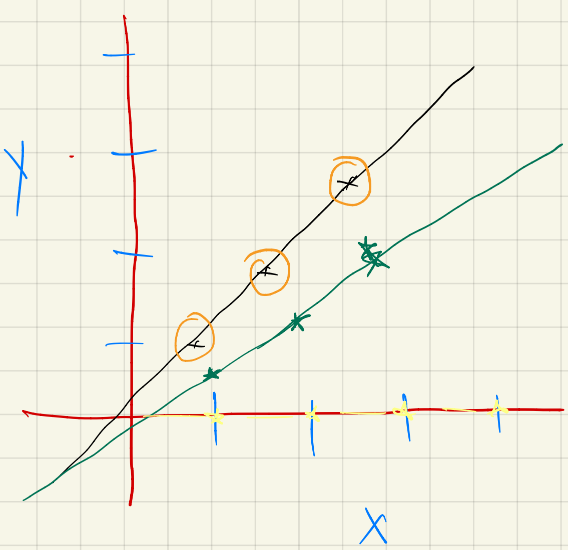
Commonly used for Regression (Linear Problem).

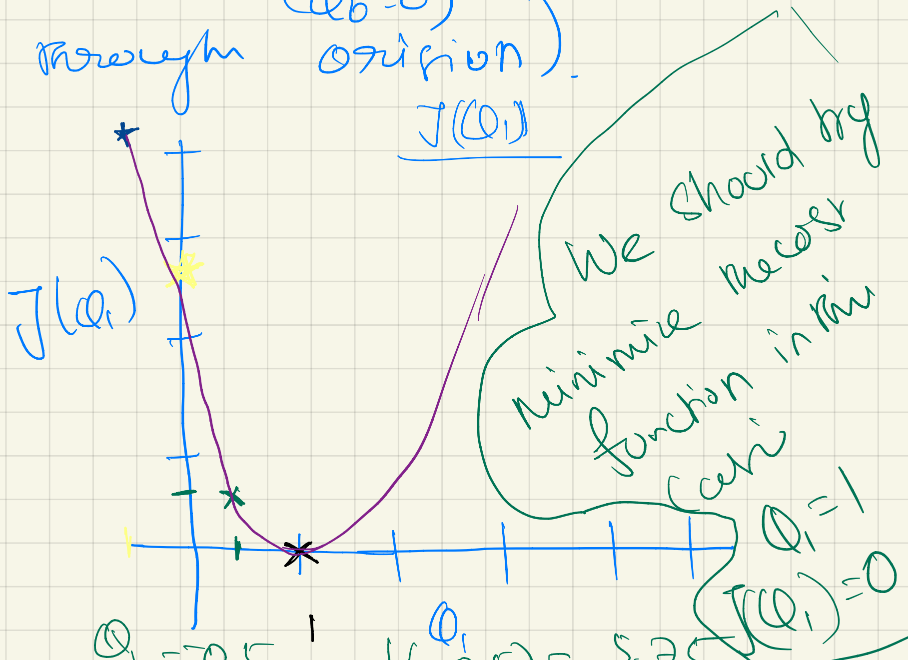
Cost function intuition 1 :-

Simplified hypothesis functio. $h_\theta(x) = \theta_1 x$

$(\theta_0 = 0)$

$h_\theta(x)$ (Pass through origion).





⊚ - Data sets

$\ast$ - $\theta_1 = 0.5$

$\star$ - $\theta_1 = 1$

$\star$ = $\theta_1 = 0$

$\theta_1 = -0.5 = J(-0.5) = 5.75$

$\theta_1 = 0.5 \quad J(0.5) = 0.58$

$\theta_1 = 1 \quad J(1) = 0$

$\theta_1 = 0 \quad J(0) = 2.5$

Cost Function Intuition II

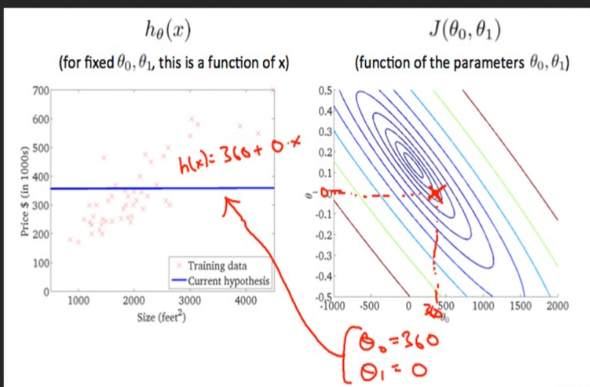
Hypothesis : $h(x) = \theta_0 + \theta_1 x$

Parameter : θ_0, θ_1

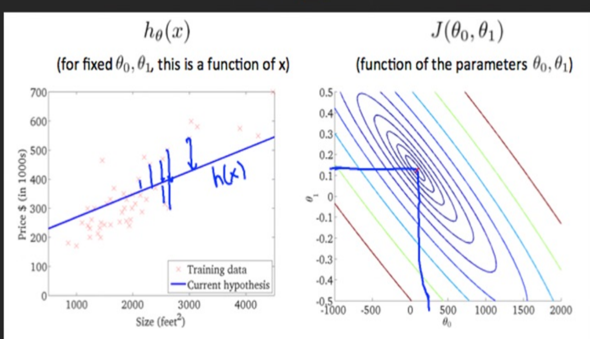
Cost function : $J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$

Goal : $\underset{\theta_0, \theta_1}{\text{minimize}} \; J(\theta_0, \theta_1)$

Taking any color and going along the 'circle', one would expect to get the same value of the cost function. For example, the three green points found on the green line above have the same value for $J(\theta_0,\theta_1)$ and as a result, they are found along the same line. The circled x displays the value of the cost function for the graph on the left when $\theta_0 = 800$ and $\theta_1 = -0.15$. Taking another h(x) and plotting its contour plot, one gets the following graphs:



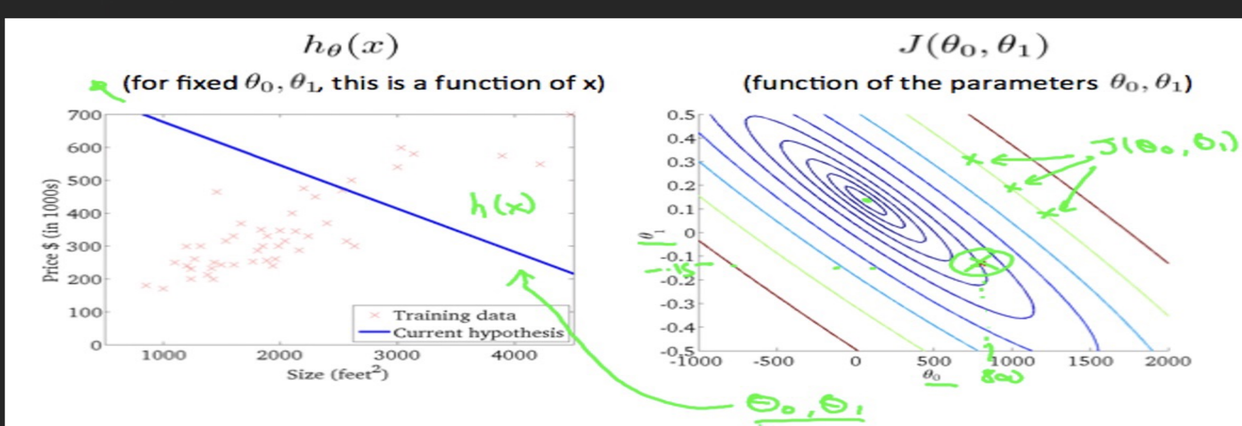
When $\theta_0 = 360$ and $\theta_1 = 0$, the value of $J(\theta_0,\theta_1)$ in the contour plot gets closer to the center thus reducing the cost function error. Now giving our hypothesis function a slightly positive slope results in a better fit of the data.



The graph above minimizes the cost function as much as possible and consequently, the result of θ_1 and θ_0 tend to be aro Screenshot pectively. Plotting those values on our graph to the right seems to put our point in the center of the inner most 'circle'.

Cost Function - Intuition II

A contour plot is a graph that contains many contour lines. A contour line of a two variable function has a constant value at all points of the same line. An example of such a graph is the one to the right below.



Taking any color and going along the 'circle', one would expect to get the same value of the cost function. For example, the three green points found on the green line above have the same value for $J(\theta_0,\theta_1)$ and as a result, they are found along the same line. The circled x displays the value of the cost function for the graph on the left when $\theta_0 = 800$ and