Machine Learning Study Notes

Sida Wang

April 24, 2019

This is my personal machine learning study notes for academic purposes only.

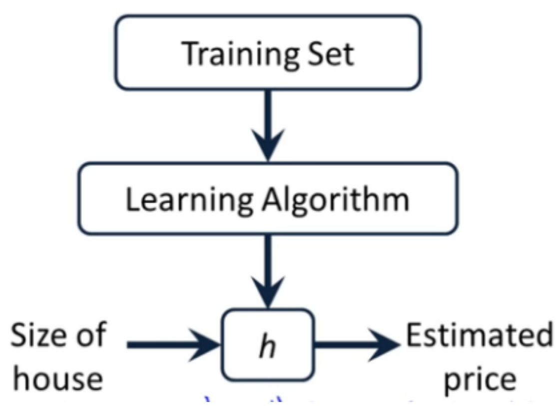
Linear Regression

- **Supervised Learning:** Given the "right answer" for each example in the data.
- **Two types of Supervised Learning:** Regression problem, Classification problem.
- **Regression Problem:** Predict real-valued output.
- **Classification Problem:** Predict dicrete-valued output.

Traning Set

- **Notation:**
- **m**– Number of training examples
- **x**– "input" variable(data)
- **y**– "output" variable(predict)
- **(x,y)**– one training example
- (x^i, y^i)– i^{th} training example

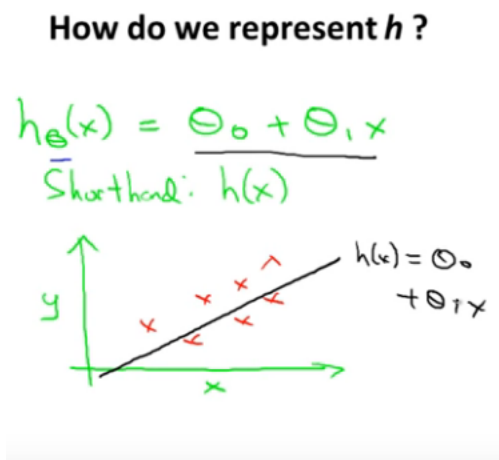
Supervised Learning



The learning algorithm output a function based on the training set.
The function h is a hypothesis that maps from x to y.

- $\boldsymbol{h_\theta(x) = \theta_0 + \theta_1 x}$ (shorthand: h(x))

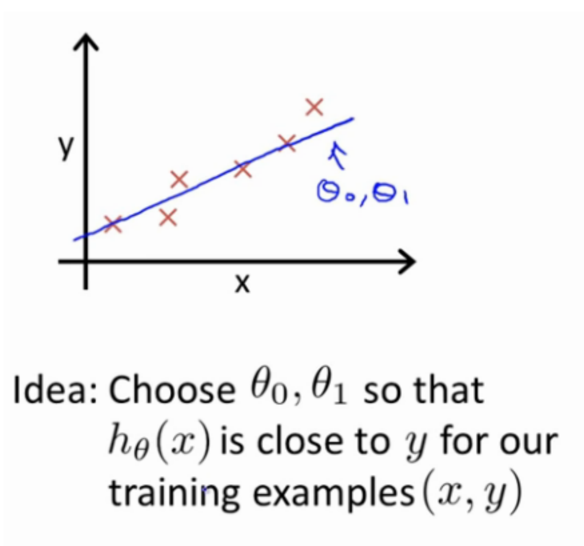
- **h :** predicting y is a linear function of x, $\boldsymbol{\theta_i}$s are parameters.



Univariate Linear Regression

Linear regression with one variable(like above).

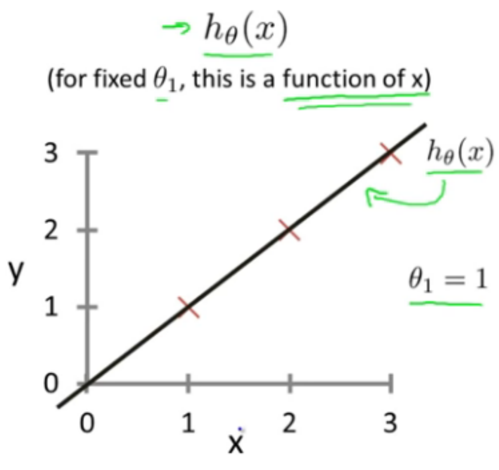
Cost Function



- **Cost Function(Squared error function):**
  $\boldsymbol{J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x^i) - y^i)^2}$

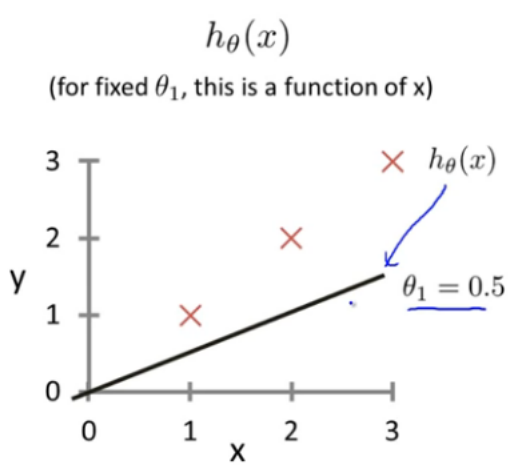
$h_\theta(x) = \theta_0 + \theta_1 x$
Minimize θ_0 and θ_1 will minimize cost function
Goal : Minimize J
most commonly used for linear regression problems

- Example with $\theta_0 = 0$



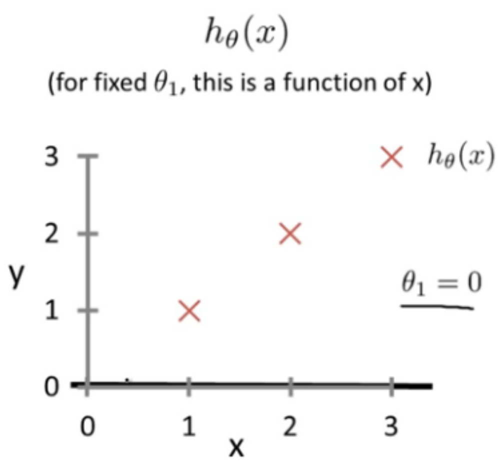
$\theta_1 = 1$

$J(\theta_1) = \dfrac{1}{2m}\sum_{i=1}^{m}(\theta_1(x^i) - y^i)^2$
$= \dfrac{1}{2m}\sum_{i=1}^{m}(0^2 + 0^2 + 0^2)$
$= 0^2$



$\theta_1 = 0.5$

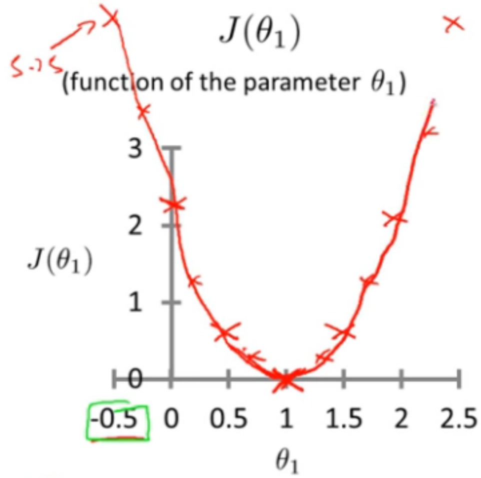
$J(\theta_1) = \frac{1}{2m} \sum_{i=1}^{m} (\theta_1(x^i) - y^i)^2$
$= \frac{1}{2m} \sum_{i=1}^{m} ((0.5 - 1)^2 + (1 - 2)^2 + (1.5 - 3)^2)$
$= \frac{1}{2x3}(3.5)$
$= 0.68$



$\theta_1 = 0$

$J(\theta_1) = \frac{1}{2m} \sum_{i=1}^{m} (\theta_1(x^i) - y^i)^2$
$= \frac{1}{2m} \sum_{i=1}^{m} ((0 - 1)^2 + (0 - 2)^2 + (0 - 3)^2)$
$= \frac{1}{2x3}(14)$
$= 2.3$

Using result plot cost function:



The value of θ_1 that minimize cost function is 1

- **Cost function with two parameters**