Lecture 2: Linear Regression With One Variable

Model Representation

Notation:

m = *Number of training examples*

x = *"input" variable / features*

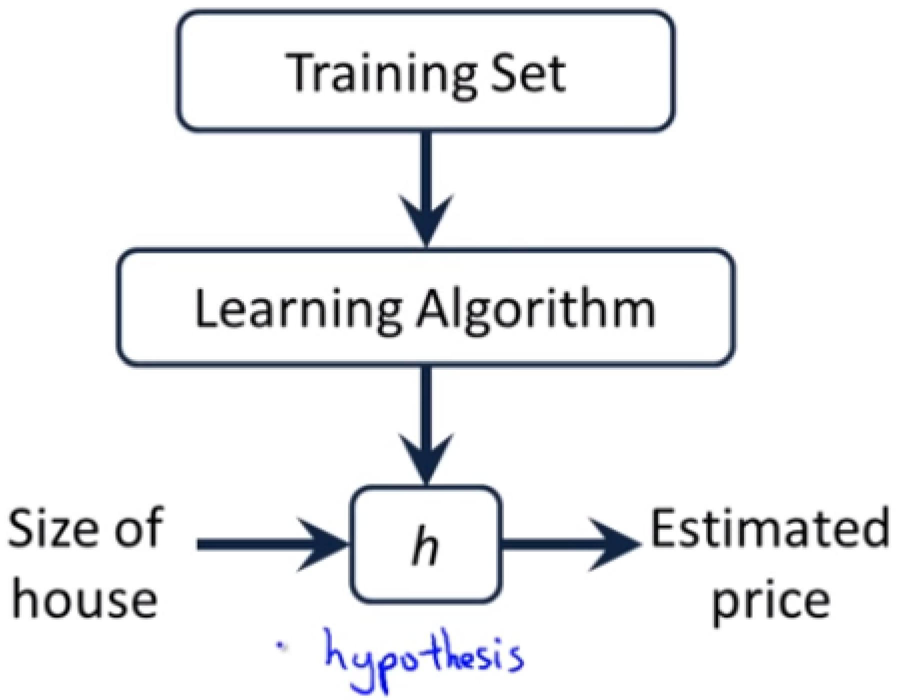
y = *"output" variable / "target" varible*

(x,y) = *one training example*

$(x^{(i)}, y^{(i)}) = i^{th}$ *training example*

h = *hypothesis / function of x --> y*

θ = *parameters of the model*

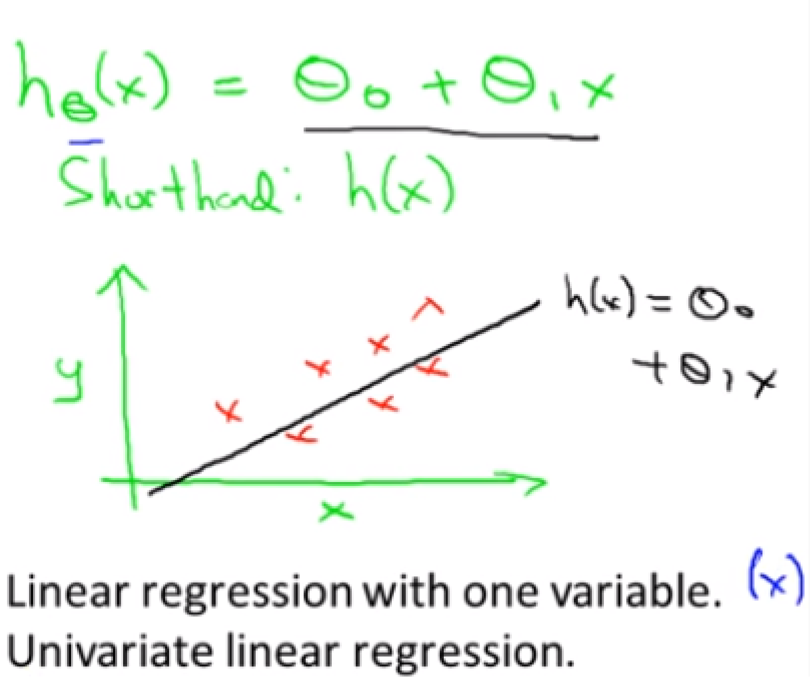


h *is a function that maps from x's to y's*

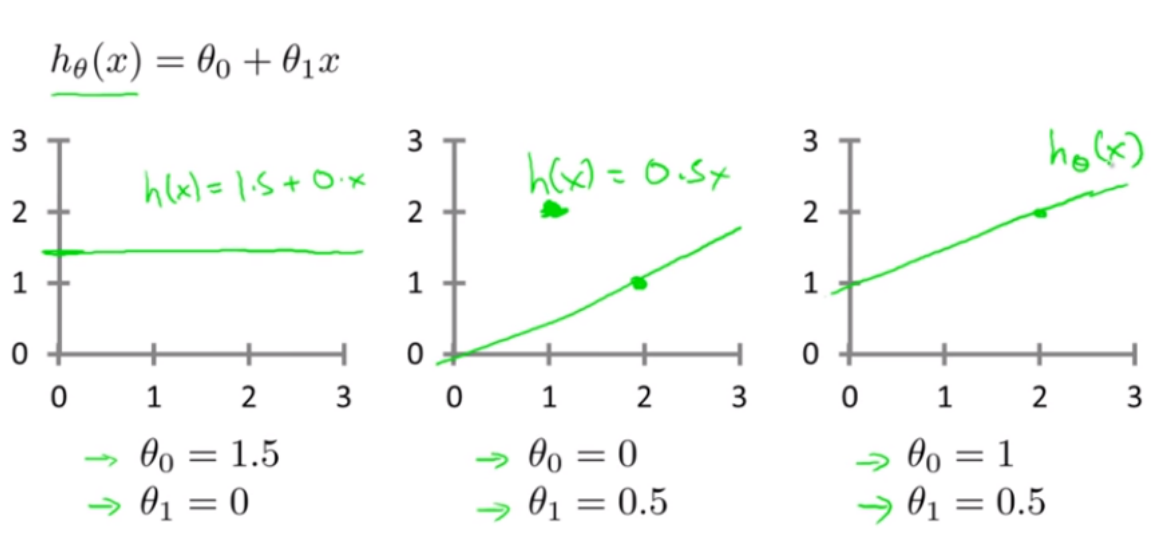
How to represent h?

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

Shorthand for *Shorthand for* $h_\theta(x) \rightarrow \boldsymbol{h(x)}$



In the above picture, an example of house prices. This is a Linear regression with one variable.



In the above picture, an example of how $h_\theta(x)$ *changes with* θ_0 *and* θ_1

Cost Function: To get the accurate prediction

Idea: Choose ***training parameters*** *so that the* ***predicted value*** *is closer to the* ***actual answer*** *for our training examples (x,y).*

Or in other words: Choose θ_0, θ_1 *so that* $\boldsymbol{h_Q(x)}$ *is close to* ***y*** *for our training examples (x,y).*

Minimization problem:

$$\text{minimise } \theta_0, \theta_1 : (h_\theta(x) - y)^2$$

$$\text{minimise } \theta_0, \theta_1 : \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2$$

$$\text{minimise } \theta_0, \theta_1 : \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2$$

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

Note: *The $\frac{1}{m}$ comes from taking average of the sum and the $\frac{1}{2}$ comes from "making the math bit easier".*

Rewriting this function cleanly: the conventional way

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

Cost Function is also called Squared error function.

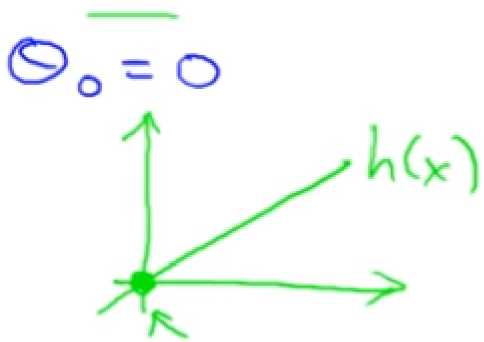
Why do we Square our error function?

Because it is found to work really well with most problems and is widely used. Though there are more error functions/ cost function

Understanding what the math expression means:

- *Taking θ_1 as 0:*
  The hypothesis becomes:

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



*Here you see that $\theta_0 = 0$ simply means that the $h(x)$ **function passes through origin**.*

- *so,*

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

- *replacing $h_\theta(x^{(i)})$*

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