→ Linear regression with one variable

Cost Function

Training Set

Size in feet^2 (x)	Price (\$) in 1000's (y)
2104	460
1416	232
1534	315
852	178
⋮	⋮

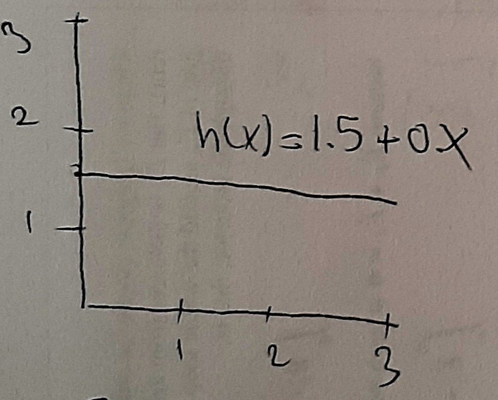
$\}$ $m = 47$

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

linear function

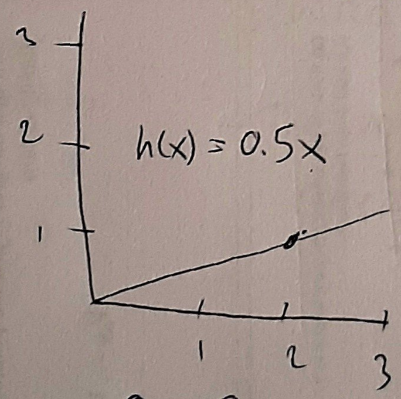
θ_i's : parameters

How to choose θ_i's ?



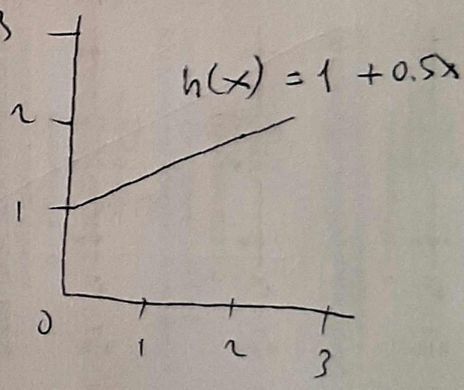
$\theta_0 = 1.5$

$\theta_1 = 0$



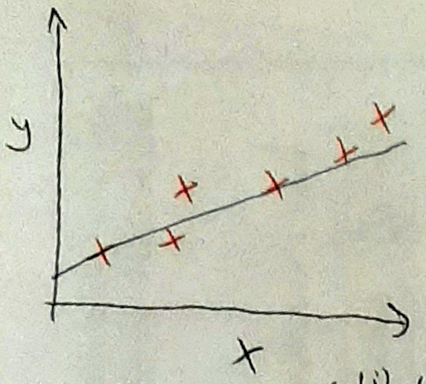
$\theta_0 = 0$

$\theta_1 = 0.5$



$\theta_0 = 1$

$\theta_1 = 0.5$



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

$h_\theta(x)$ meaning the value we predict on implies X that is at least close to the values y for the examples in our training set

$(x^{(i)}, y^{(i)})$

$$\underset{\theta_0\,\theta_1}{\text{minimize}} \; \frac{1}{2m} \sum_{i=1}^{m} \underbrace{\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2}_{h_\theta(x^{(i)}) = \theta_0 + \theta_1 \cdot x^{(i)}}$$

m → # training example

↓ more clear ↓

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

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

Squared error function

→ The most commonly used one for regression problems