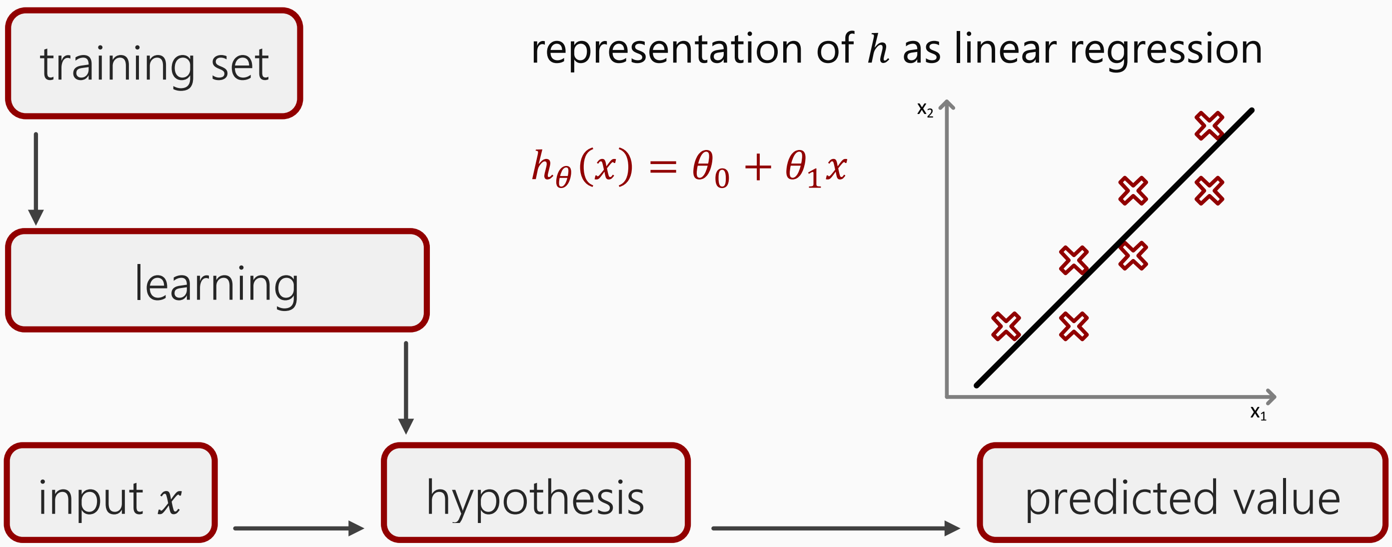
linear regression basics with one variable

model and cost function

model representation

training set definitive notation:

- $m =$ number of training examples
- $x =$ 'input' variables/features
- $y =$ 'output' variables/features
- $(x, y) =$ single training example
- $\left(x^{(i)}, y^{(i)}\right) = i^{th}$ training example



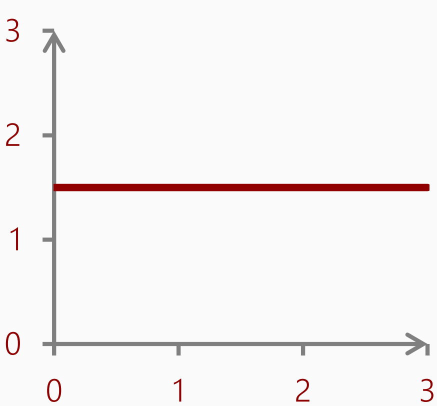
representation of h as linear regression

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

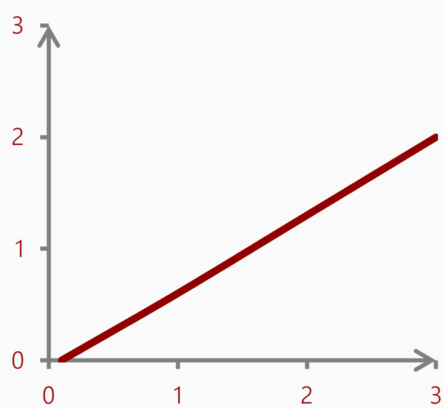
shorthand: $h_\theta(x)$

cost function

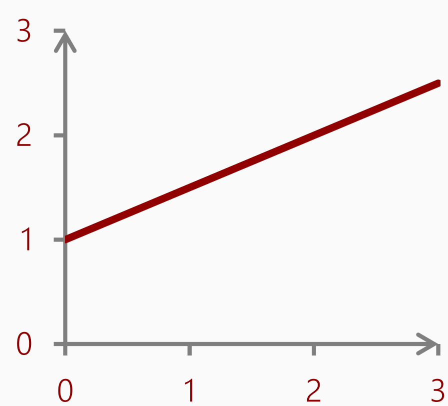
determination of the parameters in $h_\theta(x) = \theta_0 + \theta_1 x$; θ_0 and θ_1:



$\theta_1 = 1.5$
$\theta_0 = 0$



$\theta_1 = 0$
$\theta_0 = 0.5$



$\theta_1 = 1$
$\theta_0 = 0.5$

the motivation is to choose θ_1, θ_0 so that $h_\theta(x)$ is close to y for the training examples $\left(x^{(i)}, y^{(i)}\right)$

the formal expression takes the form of a minimization objective function of the average of the sum of least squared errors (linear regression function):

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

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

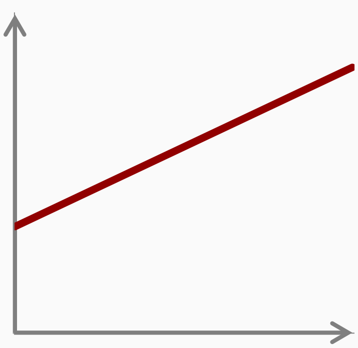
$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)^2$$

$\min_{\theta_0,\theta_1} J(\theta_0, \theta_1) \rightarrow$ **cost function**

cost function – intuition i

hypothesis:

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



parameters:

θ_0, θ_1

cost function:

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

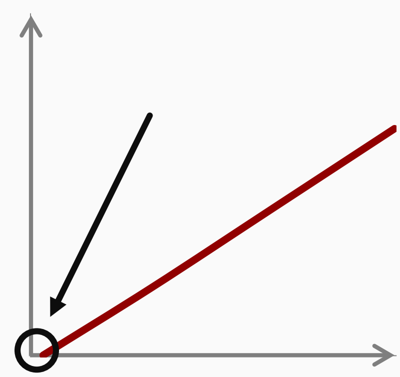
goal:

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

simplified:

hypothesis:

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



parameters:

θ_1

cost function:

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

goal:

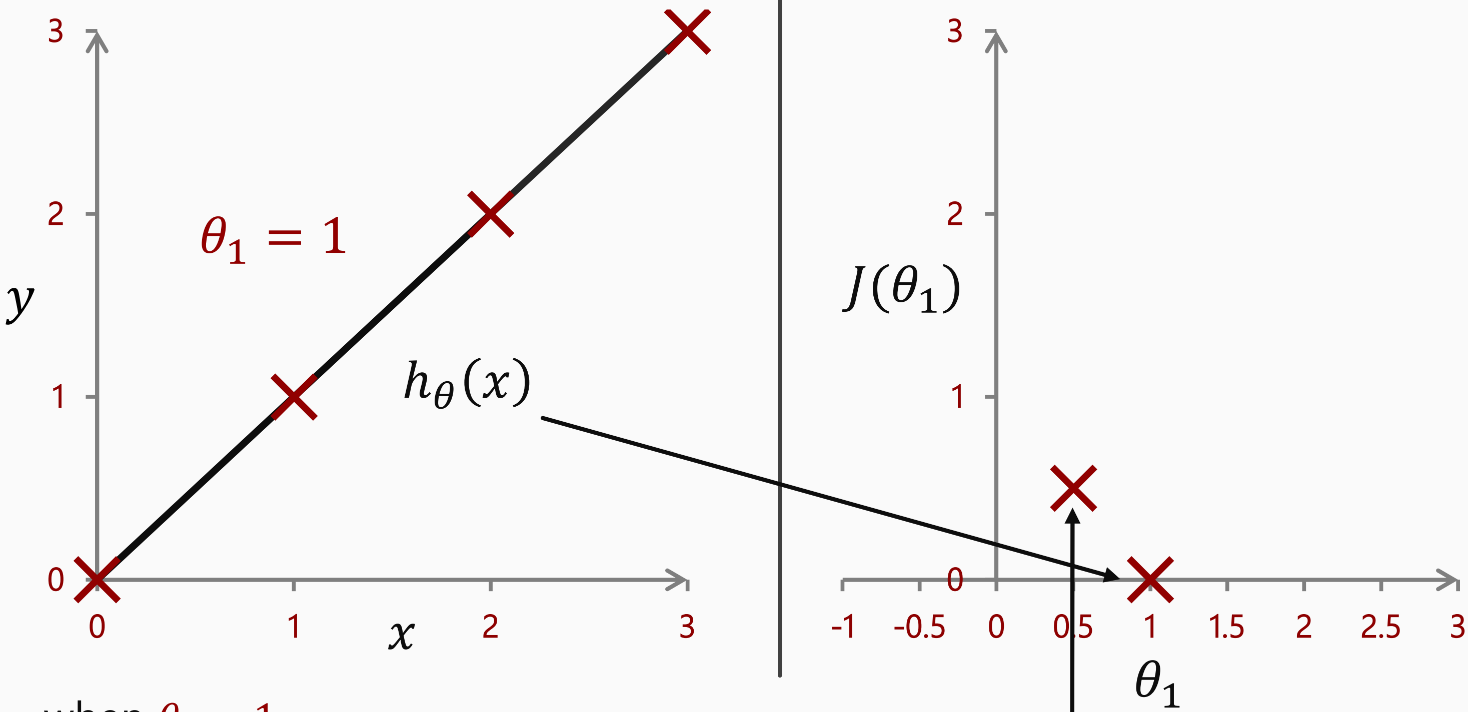
$\underset{\theta_1}{\text{minimize}}\, J(\theta_1)$

$h_\theta(x)$

(for a fixed θ_1, this is a function of x)

$J(\theta_1)$

(function of the parameter θ_1)



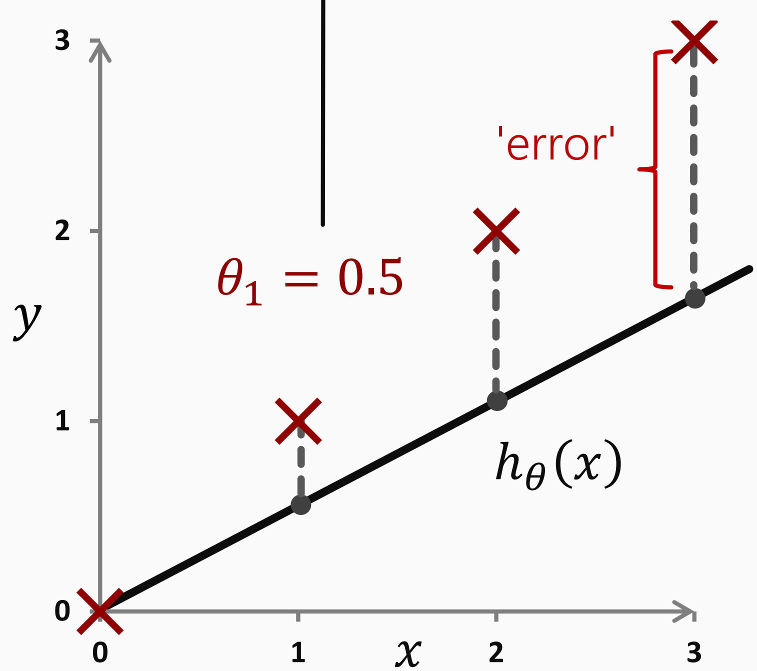
when $\theta_1 = 1$

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

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

$$= \frac{1}{2m}(0^2, 0^2, 0^2) = 0$$

$$J(1) = 0$$

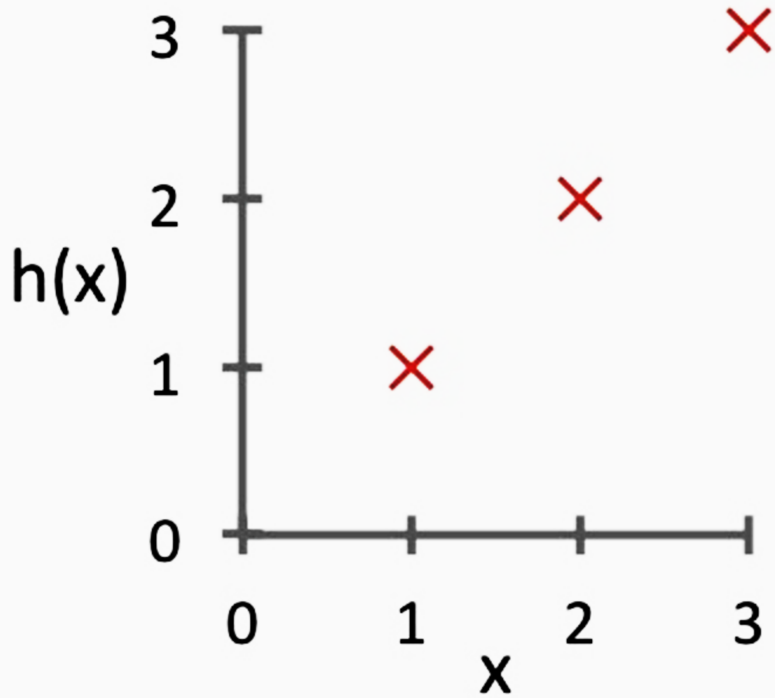


when $\theta_1 = 0.5 \rightarrow J(0.5) \approx 0.58$

$$J(0.5) = \frac{1}{2m}\left[\begin{matrix}(0.5-1)^2 + (1-2)^2 \\ + (1.5-3)^2\end{matrix}\right]$$

$$= \frac{1}{2 \times 3}(3 \times 5) = \frac{3.5}{6} \approx 0.58$$

Suppose we have a training set with m=3 examples, plotted below. Our hypothesis representation is $h_\theta(x) = \theta_1 x$, with parameter θ_1. The cost function $J(\theta_1)$ is $J(\theta_1) = \frac{1}{2m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})^2$. What is $J(0)$?



$$J(0) = \frac{14}{6}$$

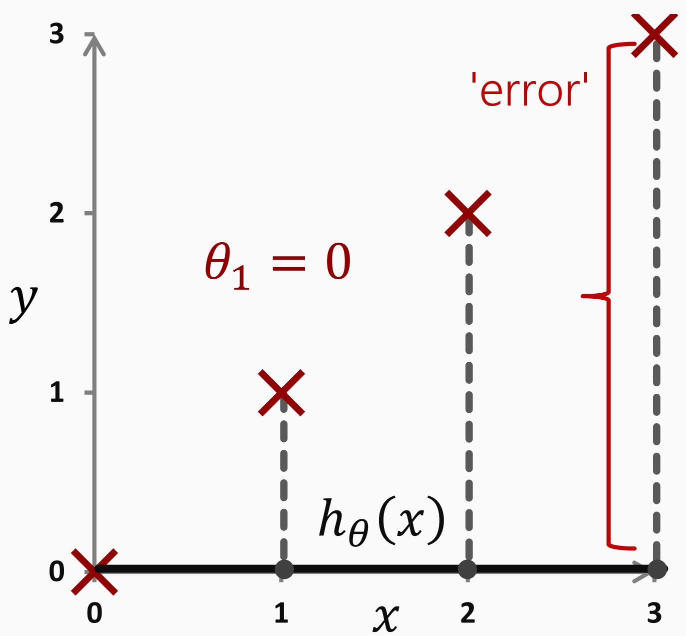
computing all functions to plot the parameters $J(\theta_i)$ ultimately plots an expected polynomial function representative of the cost function

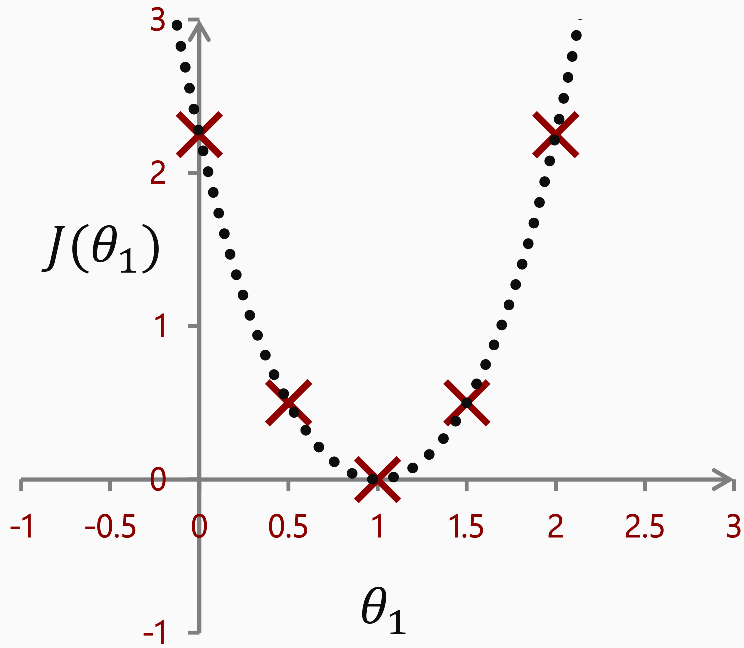
returning to the objective function $\underset{\theta_1}{\text{minimize}}\, J(\theta_1)$: the polynomial function proves the function is minimized when $\theta_1 = 1$. this equally can be seen as fitting the data most accurately in the first illustration

when $\theta_1 = 0 \rightarrow J(0) \approx 2.3$

$$J(0.5) = \frac{1}{2m}[1^2 + 2^2 + 3^2]$$

$$= \frac{1}{6}(14) = \frac{14}{6} \approx 2.3$$

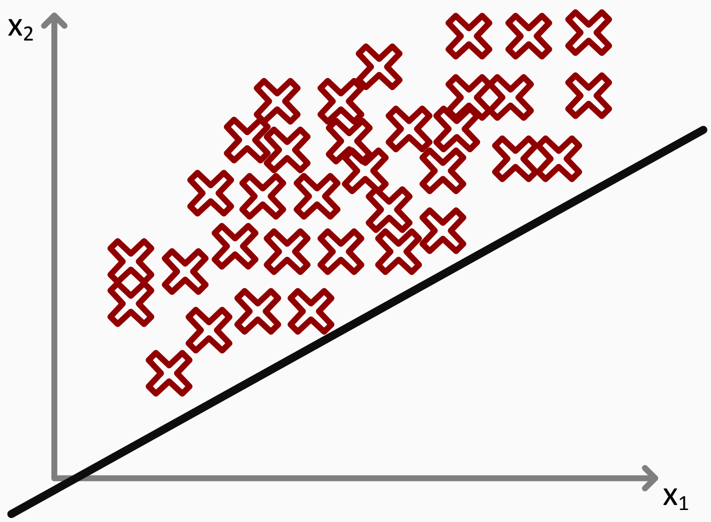




cost function – intuition ii

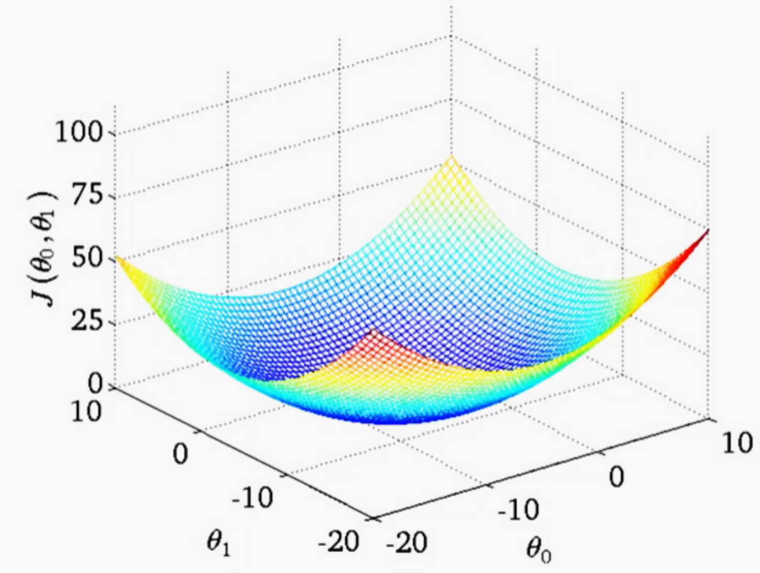
$h_\theta(x)$

(for a fixed θ_0, θ_1, this is a function of x)



$J(\theta_1, \theta_0)$

(function of the parameters θ_0, θ_1)



removing the simplification ($J(\theta_1)$) exercised previously, the function becomes slightly more complicated. the polynomial plotted function illustrated in the last example becomes a 3D contour plot through the inclusion of θ_0 back into the function $J(\theta_1, \theta_0)$

the height of the distance from θ_1, θ_0 are representative of $J(\theta_1, \theta_0)$

the above can be illustrated in 2 dimensions as a contour plot or figure