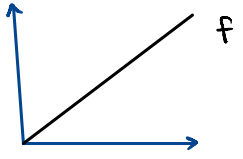
Goal: To plot a graph of all cost functions and choose the w that suits our data (minimizes J)

model:

$f_{w,b}(x) : wx + b$ ← we're getting rid of the parameter b for simplicity for the sole purpose of easily understanding how 'w' is selected.

parameters:

w

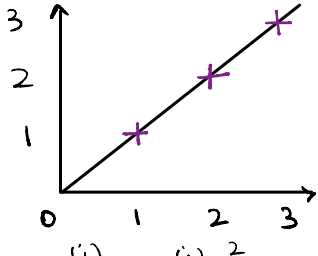


cost function:

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

We have to find out the value of $J(w)$ $\forall w (w \in Z)$.

For eg. for $w = 1$,

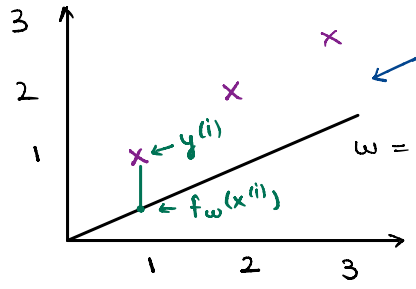


← $\forall x^{(i)} \forall y^{(i)} : x^{(i)} = y^{(i)}$

$$\Rightarrow J(1) = \frac{1}{2m}\left(wx^{(i)} - y^{(i)}\right)^2 = \frac{1}{2m}\left(0^2 + 0^2 + 0^2 + \dots\right)$$

$$= 0$$

If $w = 0.5$



$$J(0.5) = \frac{1}{2m}\left[(\underset{w}{0.5}\,\underset{x}{(1)} - \underbrace{1}_{y^{(i)}})^2 + (1-2)^2 + (1.5-3)^2 \ldots \right]$$

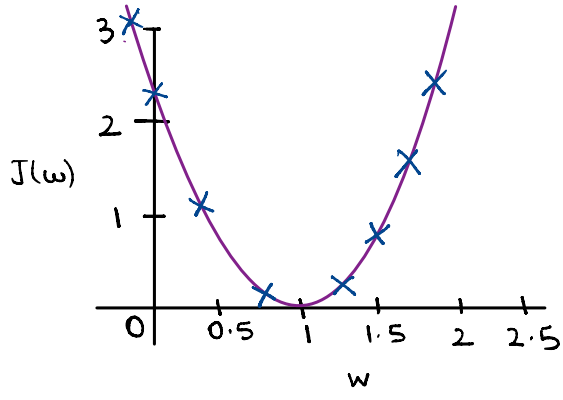
For simplicity, let's take the datasize to be 3 ($m = 3$).

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

$$= \frac{1}{6}\left[3.5 \right] \approx 0.58$$

Similarly, $J(0) = 2.3$, $J(1.5) \approx 0.58$, etc.

By plotting all of these points on a graph we see a parabola is formed.



cost

Our job is to find the w for which J(w) is the least (i.e. the error is the least) with the help of a graph.

In this case, the answer is $w = 1$.