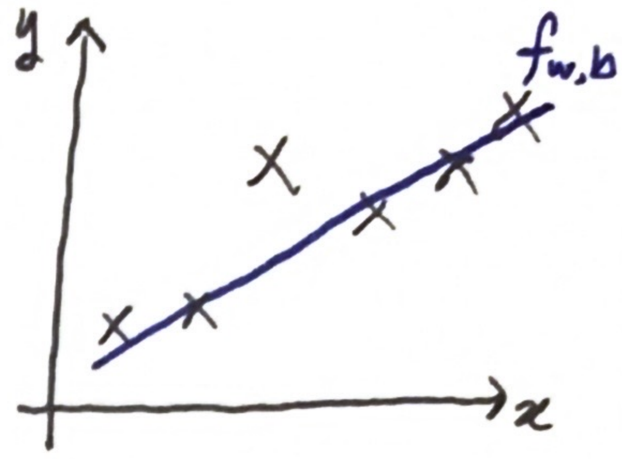
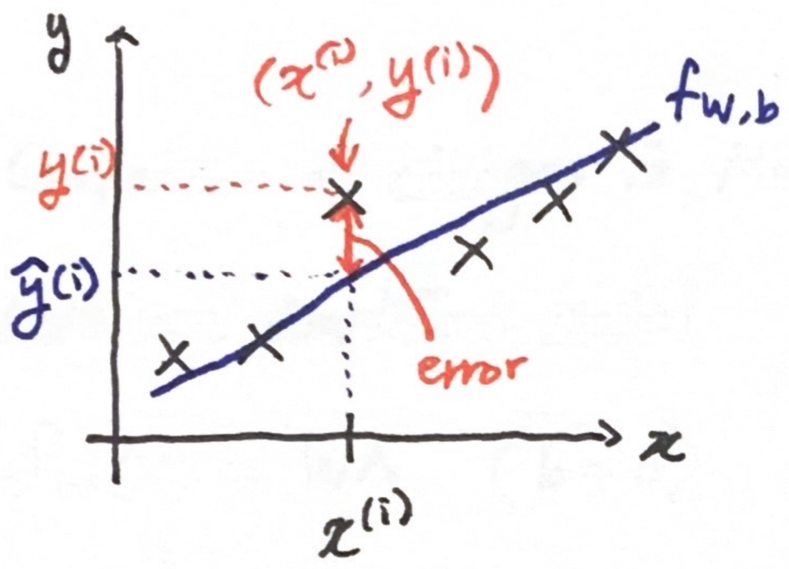
⇒ if training set is here,
what I want to do is
choose values for the parameters
w, b so that the straight line
I get from the function somehow
fits the data well

⇒ like this





For given input $x^{(i)}$ ...

- $(x^{(i)}, y^{(i)})$: y is the target
- $\hat{y}^{(i)} = f_{w,b}(x^{(i)})$: function makes predicted value ($\hat{y}^{(i)}$) based on the input value ($x^{(i)}$)

and $f_{w,b}(x^{(i)}) = wx^{(i)} + b$

* Find values for w, b so that the prediction $\hat{y}^{(i)}$ (prediction) is close to $y^{(i)}$ (true target) for all $(x^{(i)}, y^{(i)})$

⇓

To do this, must construct "Cost Function" (Squared error cost function)

$$J_{(w,b)} = \frac{1}{2m} \sum_{i=1}^{m} \underbrace{\left(\hat{y}^{(i)} - y^{(i)}\right)^2}_{\text{error}}$$

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

$\hat{y}^{(i)}$: prediction equals to the output of the model at $x^{(i)}$

$\hat{y}^{(i)} - y^{(i)}$: errors = how far off the prediction is from the target