Linear Regression

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

Cost Function

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

it is used to find appropriate w, b for linear regression model.

appropriate w, b is found by repeated usage of gradient descent

repeat until convergence {

$$w = w - \alpha \boxed{\frac{\partial}{\partial w} J(w,b)} \rightarrow \frac{1}{m}\sum_{i=1}^{m}\left(f_{w,b}(x^{(i)}) - y^{(i)}\right)x^{(i)}$$

$$b = b - \alpha \boxed{\frac{\partial}{\partial b} J(w,b)} \rightarrow \frac{1}{m}\sum_{i=1}^{m}\left(f_{w,b}(x^{(i)}) - y^{(i)}\right)$$

}

$$f_{w,b}(x^{(i)})$$

$$\frac{\partial}{\partial w}\frac{1}{2m}\sum_{i=1}^{m}\left(\underline{wx^{(i)} + b} - y^{(i)}\right)^2$$

$$\Rightarrow \frac{1}{\cancel{2}m} \times \cancel{2} \sum_{i=1}^{m}\left(wx^{(i)} + b - y^{(i)}\right)x^{(i)}$$

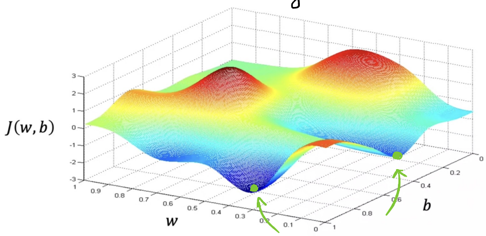
$$\Rightarrow \frac{1}{m}\sum_{i=1}^{m}\left(f_{w,b}(x^{(i)}) - y^{(i)}\right)x^{(i)}$$

$$\frac{\partial}{\partial b}\frac{1}{2m}\sum_{i=1}^{m}\left(wx^{(i)} + b - y^{(i)}\right)^2$$

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

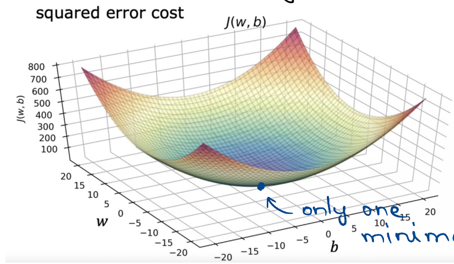
In a linear regression model, there is only one minima unlike other models because linear regression is a convex function..

not linear regression



local minima

linear regression



only one minima

convex function

→ soup bowl

if the learning rate is chosen correctly then the single minima in linear regression (global minima) can be reached