# Linear Regression

Note: change $\hat{y}$ for different curves

$$\hat{y} \begin{cases} f_{\vec{w}b}(\vec{x}) = \vec{w} \cdot \vec{x} + b \\ = w_1x_1 + w_2x_2 + \cdots + w_nx_n + b, \end{cases}$$

gradient Descent

$$w_j = w_j - \alpha \underbrace{\frac{\partial}{\partial w} J(\vec{w}, b)}_{\frac{1}{m}\sum_{i=1}^{m}\left(f_{\vec{w}b}(\vec{x}^{(i)}) - y^{(i)}\right)x_j^{(i)}}$$

$x_j^{(i)}$ changes with each iteration $(j < n)$

$$\sum_{j=1 \to n}$$

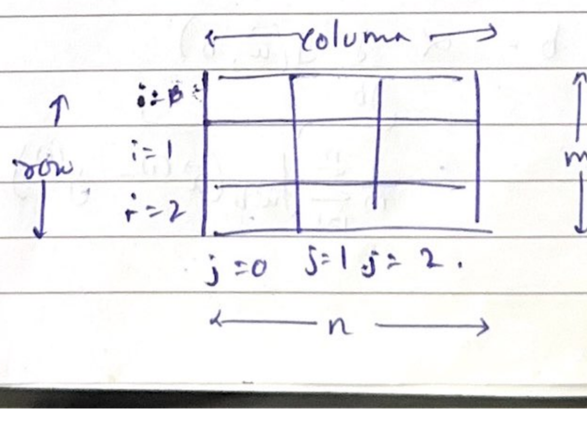
for $w_1, x_1$
$w_2, x_2$
$w_3, x_3$
$\vdots$
$w_n, x_1$.

$\alpha \to$ Learning rate

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

cost function $\to$ error

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




m = no. of examples = rows

n = no. of features = columns.

# Logistic Regression

use only for logistic for this scenario

same as $\hat{y} = f_{\vec{w}b}$ $\leftarrow \hat{z}$

$$z = \vec{w}\vec{x} + b$$

$$g(z) = \frac{1}{1+e^{-z}}$$

Note: change z for different decision boundary

closer the value is to 1, more possible it is true
" " " " " 0, " " " " false

cost function

$$\text{Cost function:} \frac{1}{m}\sum_{i=1}^{m} \underbrace{\boxed{\frac{1}{2}\left(f_{\vec{w}b}(\vec{x}^{(i)}) - y^{(i)}\right)^2}}_{\text{Loss function}}$$

$$L\left(f_{\vec{w}b}(\vec{x}^{(i)}), y^{(i)}\right) = -y^{(i)}\log\left(f_{\vec{w}b}(\vec{x}^{(i)})\right) - \left(1-y^{(i)}\right)\log\left(1-f_{\vec{w}b}(\vec{x}^{(i)})\right)$$

or

$$\therefore J(\vec{w}, b) = \frac{1}{m}\sum_{i=1}^{m}\left[L\left(f_{\vec{w}b}(\vec{x}^{(i)}), y^{(i)}\right)\right]$$

$$= -\frac{1}{m}\sum_{i=1}^{m}(-L)$$

gradient descent

$$w_j = w_j - \alpha \underbrace{\frac{\partial}{\partial w_j} J(\vec{w}, b)}_{\frac{1}{m}\sum_{i=1}^{m}\left(f_{\vec{w}b}(\vec{x}^{(i)}) - y^{(i)}\right) x_j^{(i)}}$$

$j = 1 \ldots n$

$$b = b - \alpha \underbrace{\frac{\partial}{\partial b} J(\vec{w}, b)}_{\frac{1}{m}\sum_{i=1}^{m} f_{\vec{w}b}(\vec{x}^{(i)} - y^{(i)})}$$

# Regularization - logistic regression.

## Cost function with Regularization

$$\min_{\vec{w},b} J(\vec{w},b) = \min_{\vec{w},b} \frac{1}{2m}\sum_{i=1}^{m} \overbrace{\left(f_{\vec{w}b}(\vec{x}^{(i)}) - y^{(i)}\right)^2}^{\text{mean squared error}} +$$

~~$j = 1, 2 \ldots n$~~

$$\underbrace{\frac{\lambda}{m}\sum_{j=1}^{n} w_j^2}_{\text{regulization term}}$$

## Regulized linear regression

$$w_j = w_j - \alpha \underbrace{\frac{\partial}{\partial w} J(\vec{w},b)}$$

$$\hookrightarrow \; j=1 \ldots n \left[\frac{1}{m}\sum_{i=1}^{m}\left[\left(f_{\vec{w}b}(\vec{x}^{(i)}) - y^{(i)}\right)x_j^{(i)}\right] + \overbrace{\frac{\lambda}{m} w_j}^{\text{added to each value of } \frac{\partial}{\partial w} \text{ after 'm'}}\right]$$

Note: j goes from 1 to n for each iteration of m/i and is stored as a seperate sum in an array

$j = 1 \ldots n$

$$b = b - \alpha \underbrace{\frac{\partial}{\partial b} J(\vec{w},b)}$$

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

## Regularization — logistic regression

Cost function using Regularization

$$J(\vec{w}, b) = \frac{-1}{m} \sum_{i=1}^{m} \left[ y^{(i)} \log\left(f_{\vec{w},b}(\vec{x}^{(i)})\right) + (1-y^{(i)}) \log\left(1 - f_{\vec{w},b}(\vec{x}^{(i)})\right) \right] + \frac{\lambda}{2m} \sum_{j=1}^{n} w_j^2$$

Gradient descent is same as linear regression

Note:- use proper $\left[ f_{\vec{w},b}(\vec{x}^{(i)}) \right]$ formula.