Machine Learning

Dr. Mehran Safayani

safayani@iut.ac.ir

safayani.iut.ac.ir

https://www.aparat.com/mehran.safayani

https://github.com/safayani/machine_learning_course

Department of Electrical and computer engineering, Isfahan university of technology, Isfahan, Iran

Supervised Learning

- Regression
- Classification

example

Notation:

M: number of training samples

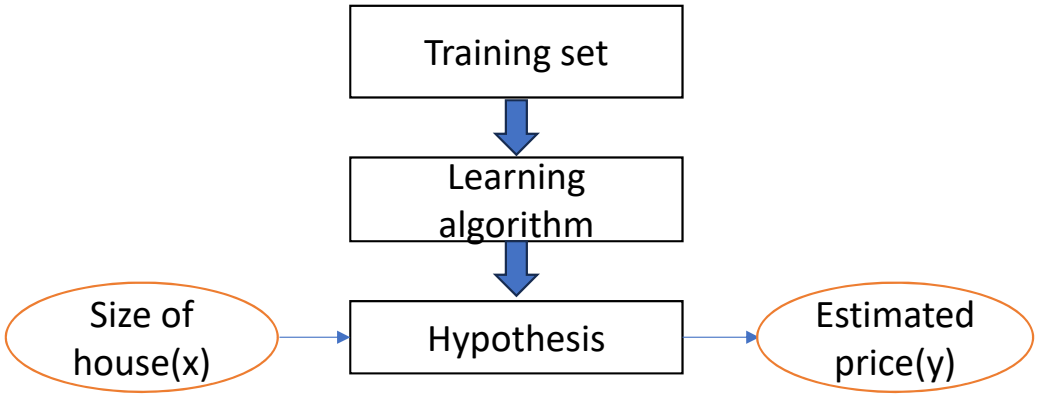
x: input variable

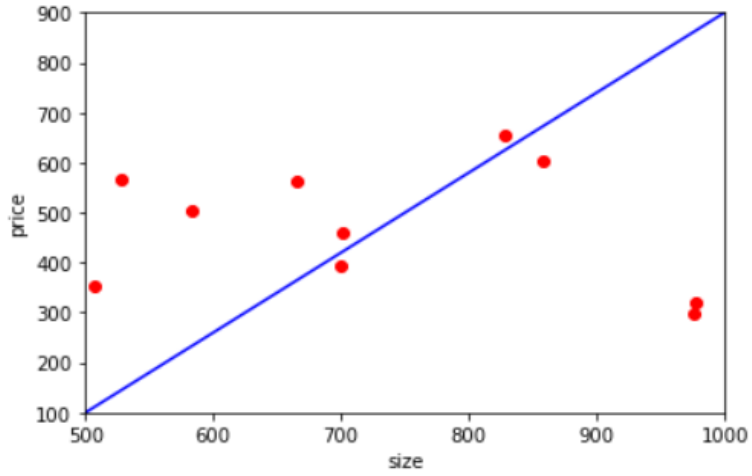
y: output variable
Or
target variable

(x_i, y_i): i th training sample

number	Size (x variable)	Price (y variable)	
1	100	500	(x_1, y_1)
2	750	2000	(x_2, y_2)
3	852	178	(x_3, y_3)
	…	…	
M	3210	870	(x_m, y_m)

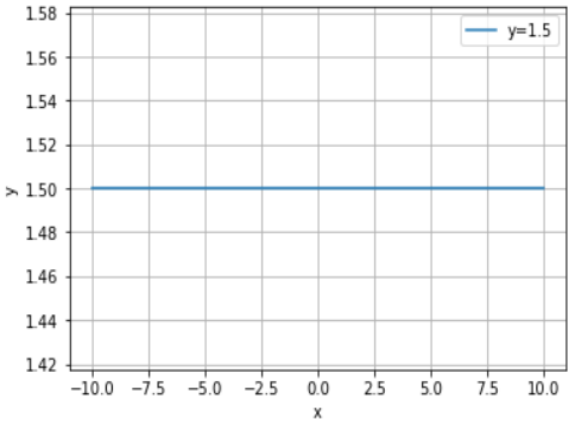
example





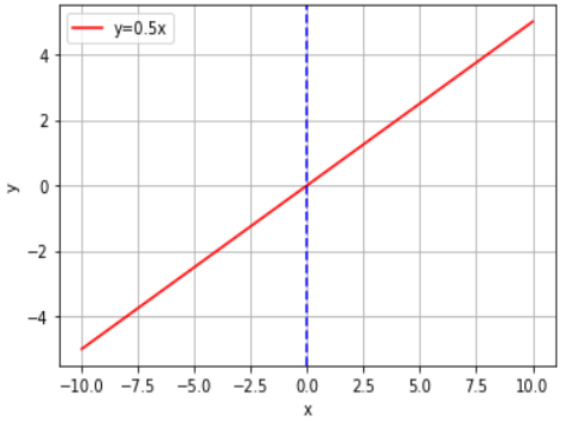
$h(x) = \theta_0 + \theta_1 x$

parameters= $\{\ \theta_0 , \theta_1\ \}$



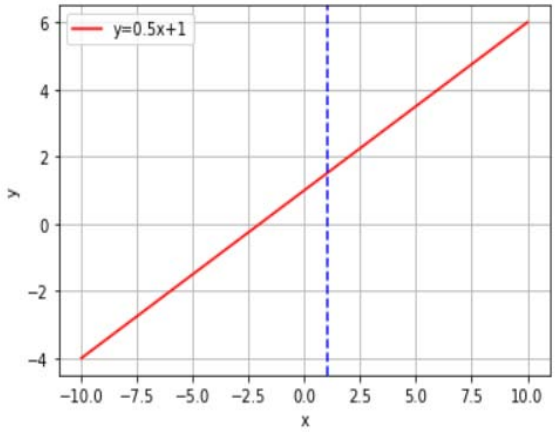
h(x) = 1.5

θ_0 = 1.5
θ_1 = 0



h(x) = 0.5x

θ_0 = 0
θ_1 = 0.5



h(x)=0.5x+1

θ_0 = 1
θ_1 = 0.5

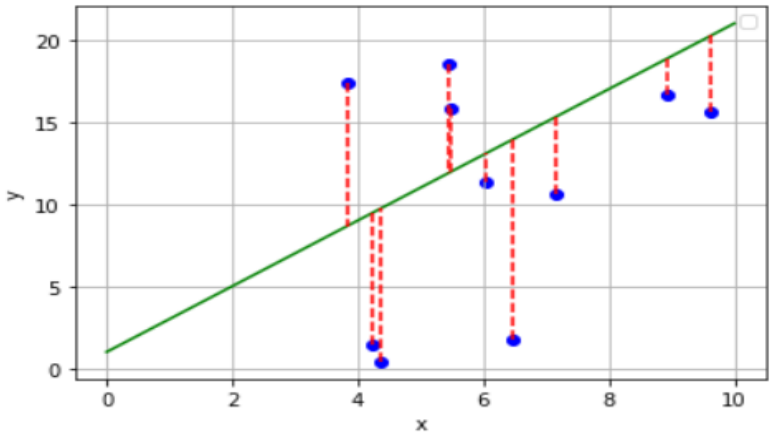
Cost Function

$$J(\theta_0 , \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} (h(x_i) - y_i)^2$$

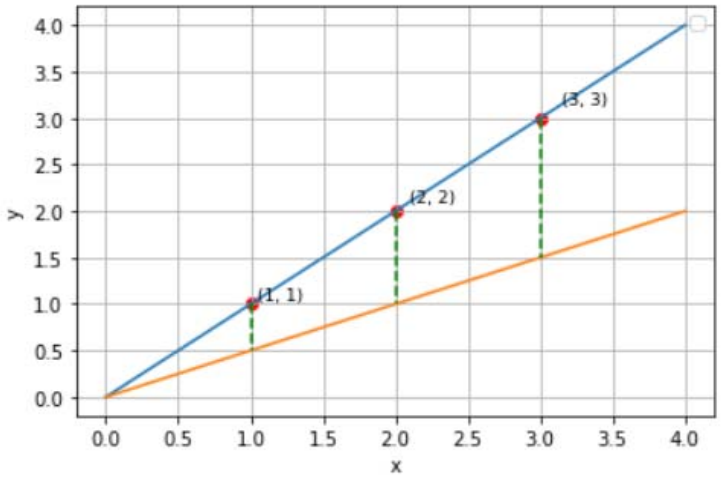
Mean square error(MSE)

Minimize $J(\theta_0 , \theta_1)$

θ_0 , θ_1



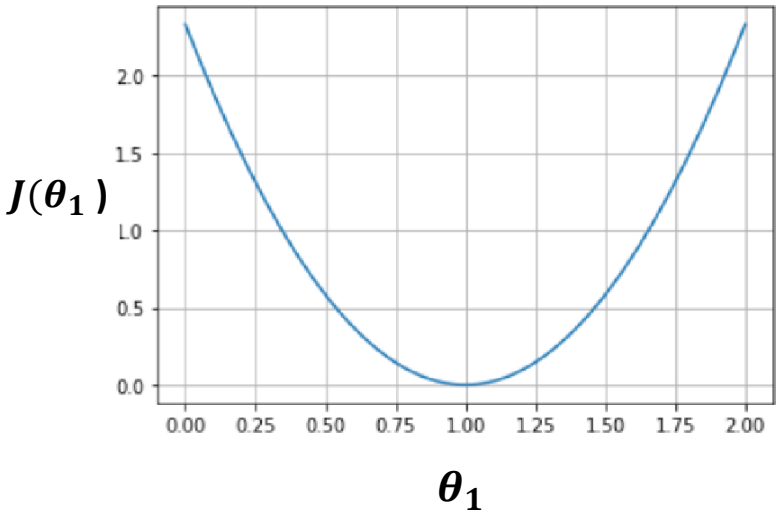
example



$$J(\theta_0=0\ ,\ \theta_1=0.5\ ) = \frac{1}{2m}\sum_{i=1}^{m}(0.5x_i - y_i)^2$$
$$= \frac{1}{2*3}\ [(0.5-1)^2 + (1-2)^2 + (1.5-3)^2]$$
$$= \frac{1}{6}\,(3.5) = 0.58$$

$$J(\theta_0=0\ ,\ \theta_1=1\ ) = \frac{1}{2m}\sum_{i=1}^{m}(x_i - y_i)^2$$
$$= \frac{1}{2*3}\ [(1-1)^2 + (2-2)^2 + (3-3)^2]$$
$$= \frac{1}{6}\,(0) = 0$$

example



θ_1	$J(\theta_1)$
0	14/6
0.5	0.58
1	0
1.5	0.58
2	14/6

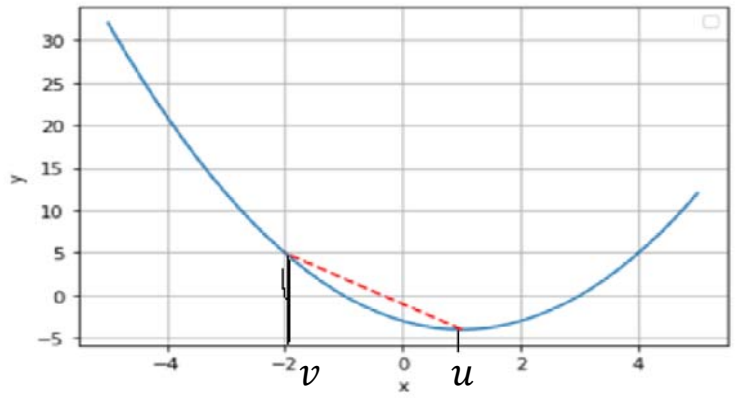
- Plotting the cost for each value of θ_1
- The minimum point: θ_1=1
- Using Grid Search to find best values of parameters

Cost Function

- $J(\theta_0 , \theta_1) = \frac{1}{m} \sum_{i=1}^{m} |h(x_i) - y_i|$ Mean absolute error(MAE)

- Better for outliers compared with MSE

Convexity



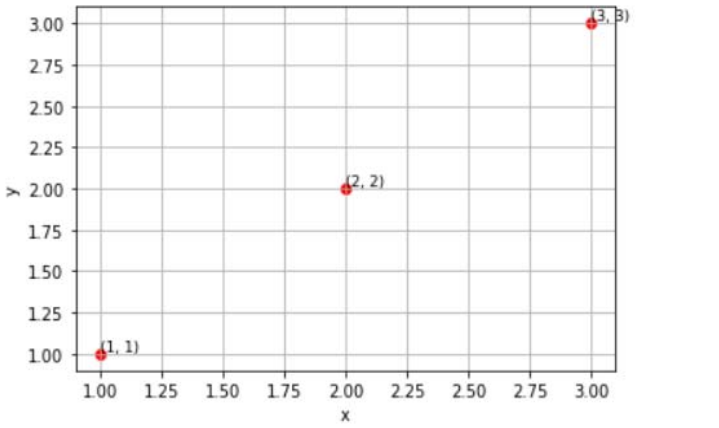
Function h(u) with u∈ X is convex if for any u, v ∈ X and for any $0 \leq \lambda \leq 1$ we have:

$$\mathbf{h(\lambda u + (1-\lambda)v) \leq \lambda\, h(u) + (1-\lambda)\, h(v)}$$

برای توابع محدب هر بهینه محلی یک بهینه سراسری است.

مجموع چند تابع محدب محدب است

example



$if\ \theta_1 = -1$:

MAE= $\frac{1}{3}$ [|1 - (-1)| + |2 - (-2)| +|3 - (-3)|] = 4

$if\ \theta_1 = 0$:

MAE= $\frac{1}{3}$ [|1 - 0| + |2 - 0| +|3 - 0|] = 2

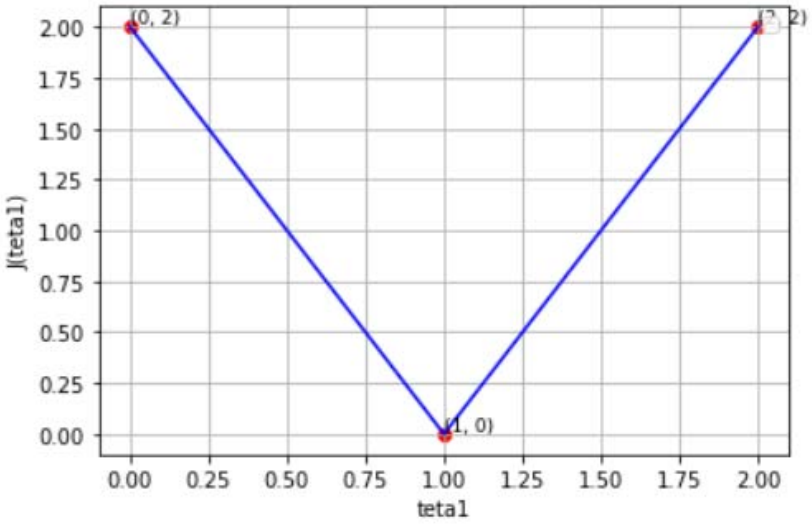
$if\ \theta_1 = 1$:

MAE= $\frac{1}{3}$ [|1 - 1| + |2 - 2| +|3 - 3|] = 0

$if\ \theta_1 = 2$:

MAE= $\frac{1}{3}$ [|1 - 2| + |2 - 4| +|3 - 6|] = 2

example



MAE is convex

$\boldsymbol{\theta_1}$	$\boldsymbol{J(\theta_1)}$
-1	4
-0.5	3
0	2
0.5	1
1	0
1.5	1
2	2
2.5	3
3	4