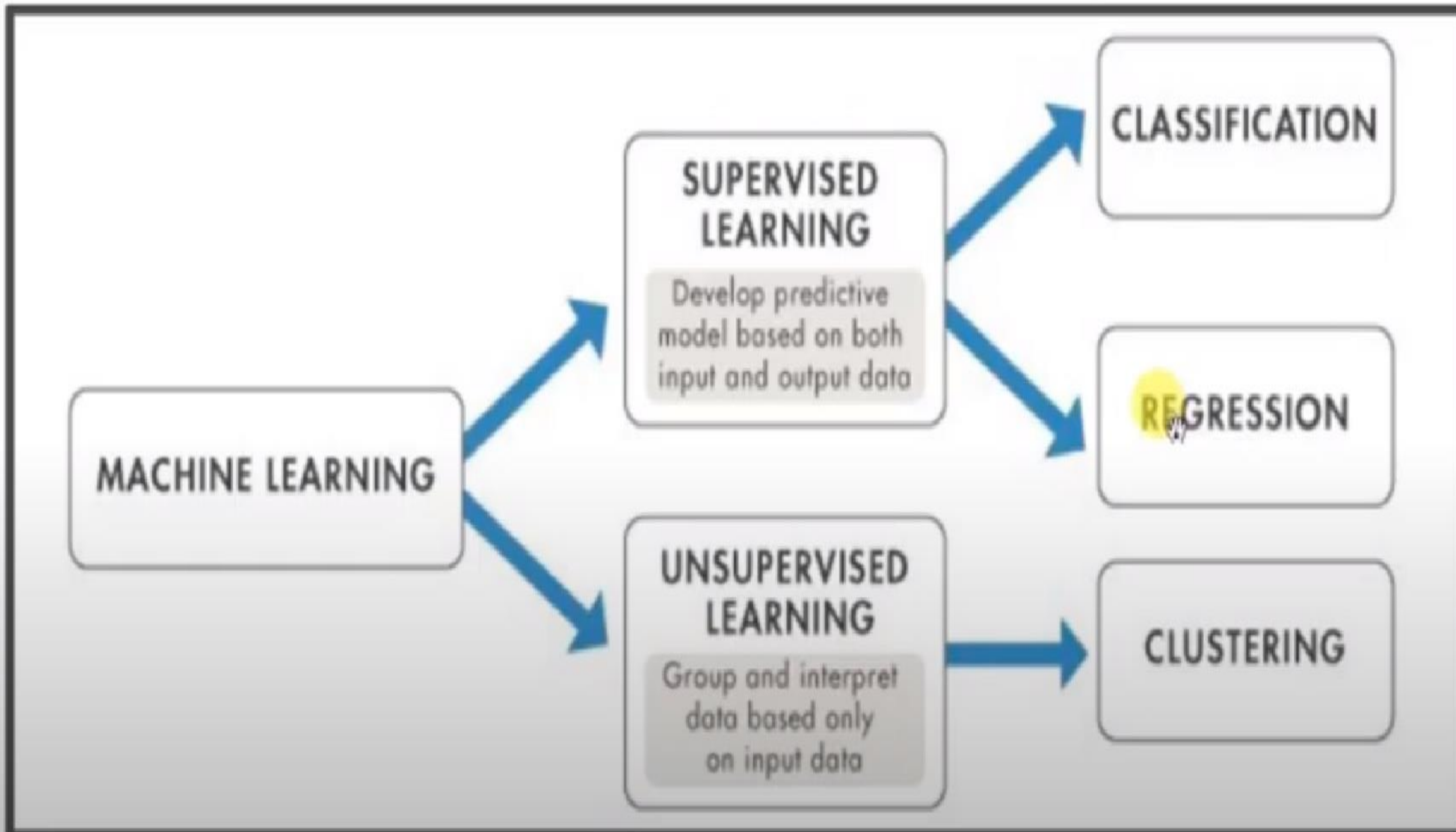
Lec. 9

Linear Regression

Assist. Prof. Dr. Saad Albawi

B-02 -التوقع الخطي Linear Regression

التعليم باشراف و بدون اشراف



0:52 / 17:17

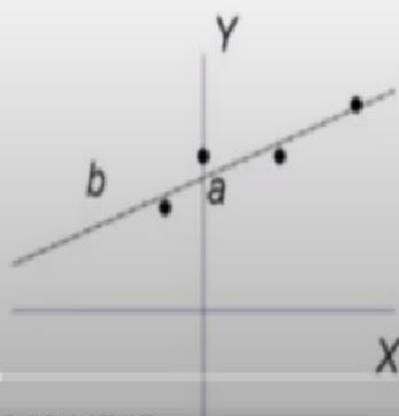
Linear Regression التوقع الخطي

Linear regression equation

(without error)

$$\hat{Y} = bX + a$$

predicted values of Y	b = slope = rate of predicted ↑/↓ for Y scores for each unit increase in X	Y-intercept = level of Y when X is 0



- و يسمي أيضا (One Varilable Regression) او (Unvariate Regression)

Linear Regression التوقع الخطي

# of Favourites (X)	# of Posts (Y)
36	14
21	12
47	22
11	11
72	33
95	46
58	25
81	34
9	3
18	12
2	0
15	4
29	10
66	19
31	20

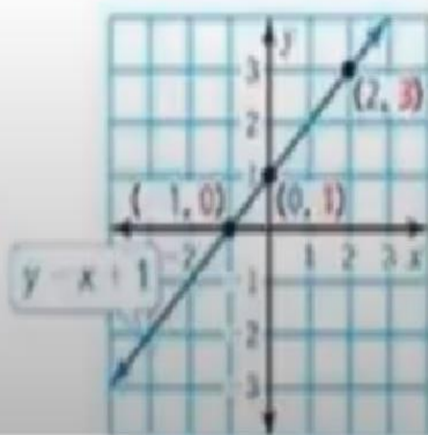
B-02 -التوقع الخطي Linear Regression

المعادلة الخطية

Linear Equations

- A **linear equation** is an equation whose graph is a line.
- The points on the line are **solutions** of the equation.

x	y	(x, y)
−1	0	(−1, 0)
0	1	(0, 1)
2	3	(2, 3)

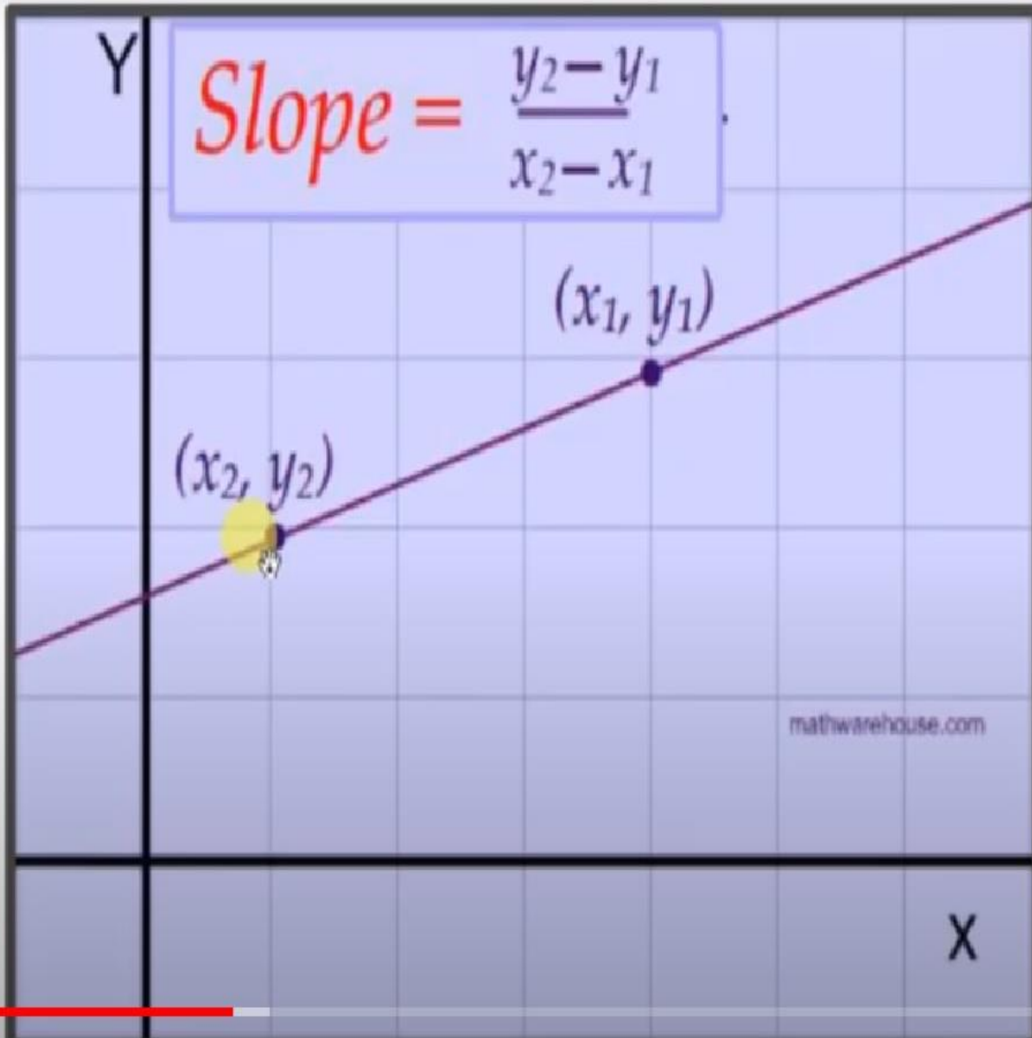


معناها :

- علاقة بين متغيرين , حين يكون كلا من Y , X لهما أس 1

4:24 / 17:17

Slope الميل



B-02 -التوقع الخطي Linear Regression

اضغط على Esc للخروج من وضع ملء الشاشة

الميل Slope

$y_2 = 1 \qquad y_1 = -7$

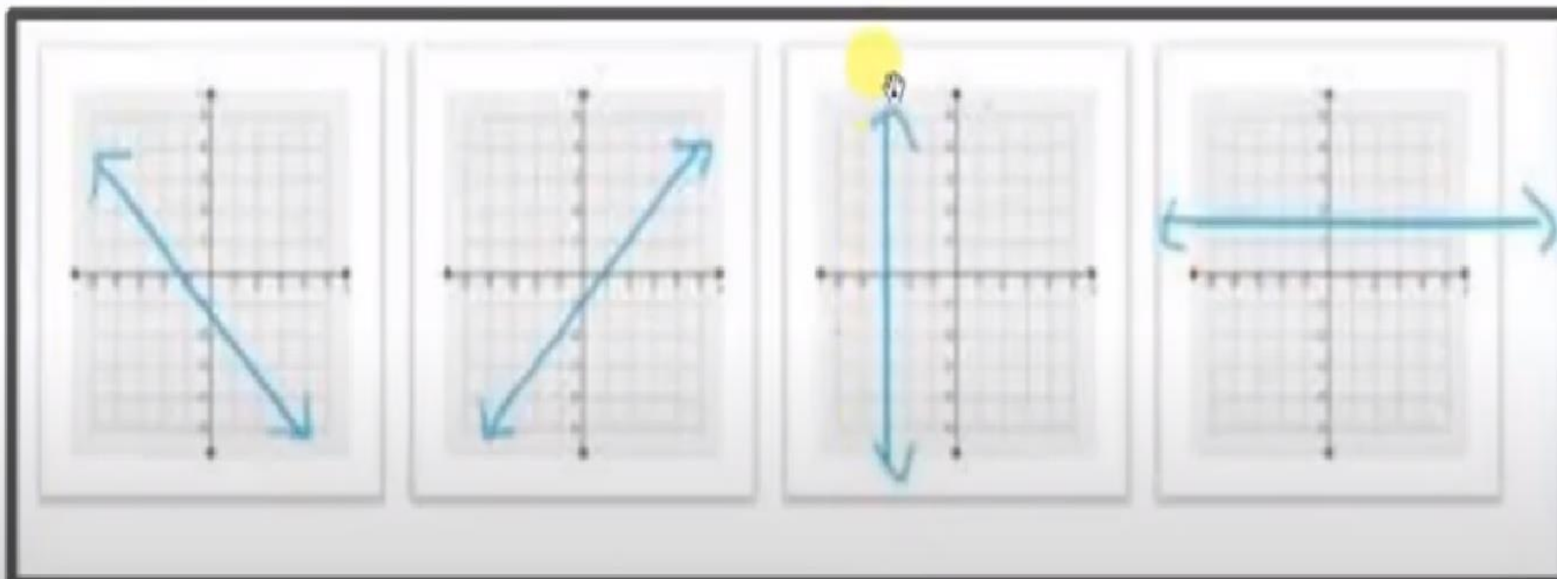
$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{1-(-7)}{12-(-4)}$$

$x_2 = 12 \qquad x_1 = -4$

7:21 / 17:17

B-02 التوقع الخطي- Linear Regression

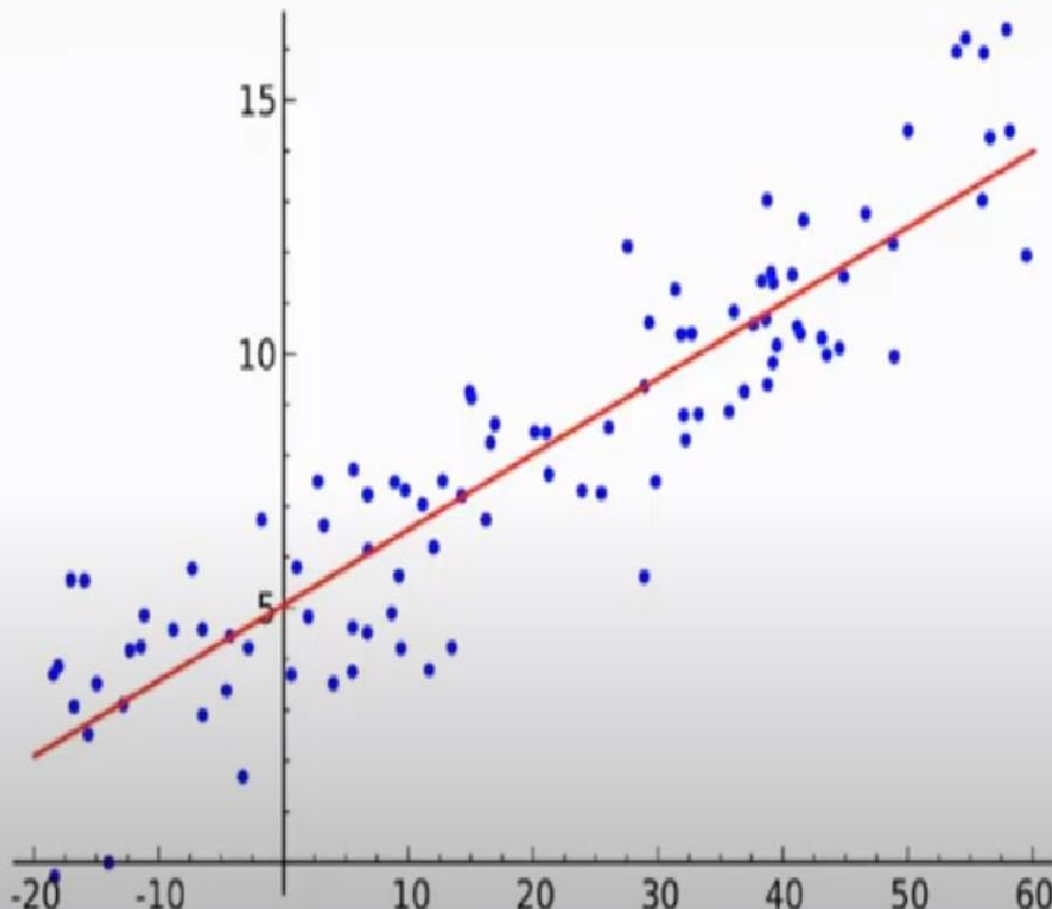
Slope الميل



7:48 / 17:17

B-02 -التوقع الخطي Linear Regression

التوقع الخطي Linear Regression



9:24 / 17:17

Linear Regression التوقع الخطي

Input X	المدخلات
Output Y	المخرجات
Rows m	الصفوف
Features n	العناصر
h(x)	القيمة المتوقعة
Cost J	قيمة الخطأ
Theta ⊖	معاملات الـ X

Linear Regression Equation معادلة التوقع الخطي

Hypothesis: $h_\theta(x) = \theta_0 + \theta_1 x$

Parameters: θ_0, θ_1

Cost Function: $J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)} \right)^2$

Goal: $\underset{\theta_0, \theta_1}{\text{minimize}}\ J(\theta_0, \theta_1)$

- الهدف تقليل الفارق بين قيمة h(x) و هي القيمة المتوقعة من المعادلة الخطية و قيمة y و هي القيمة الحقيقية
- يتم القسمة علي 2m لربط قيمة الخطا بعدد القيم بالعينة
- الهدف ايجاد قيم ثيتا 1 و ثيتا 2 , والتي تجعل من J (نسبة الخطا) اقل ما يمكن
- تسمي احيانا Cost error function

Linear Regression Equation معادلة التوقع الخطي

Theta0 = 5 , theta 1 = 2 Equation $h(x) = 5 + 2x$

X	Y	h(x)	h(x) - y	$(h(x) - y)^2$
1	7			
2	8			
2	7			
3	9			
4	11			
5	10			
5	12			

Linear Regression Equation معادلة التوقع الخطي

Theta0 = 5 , theta 1 = 2 Equation $h(x) = 5 + 2x$

X	Y	h(x)	h(x) - y	$(h(x) - y)^2$
1	7	7		
2	8	9		
2	7	9		
3	9	11		
4	11	13		
5	10	15		
5	12	15		

Linear Regression Equation معادلة التوقع الخطي

Theta0 = 5 , theta 1 = 2   Equation $h(x) = 5 + 2x$

X	Y	h(x)	h(x) - y	$(h(x) - y)^2$
1	7	7	0	
2	8	9	1	
2	7	9	2	
3	9	11	2	
4	11	13	2	
5	10	15	5	
5	12	15	3	

Linear Regression Equation معادلة التوقع الخطي

Theta0 = 5 , theta 1 = 2 Equation $h(x) = 5 + 2x$

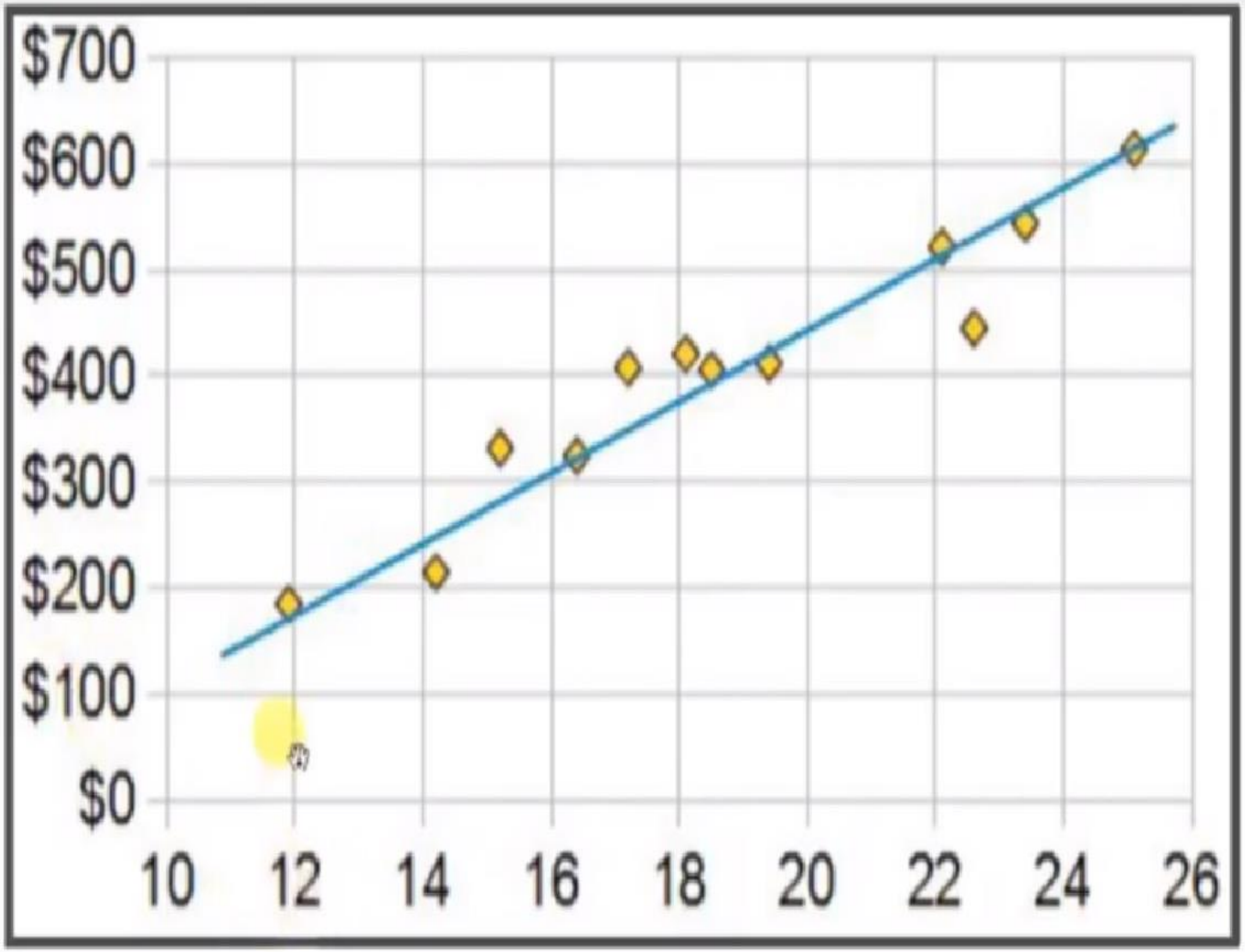
X	Y	h(x)	h(x) - y	$(h(x) - y)^2$
1	7	7	0	0
2	8	9	1	1
2	7	9	2	4
3	9	11	2	4
4	11	13	2	4
5	10	15	5	25
5	12	15	3	9

$$\boxed{J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2}$$

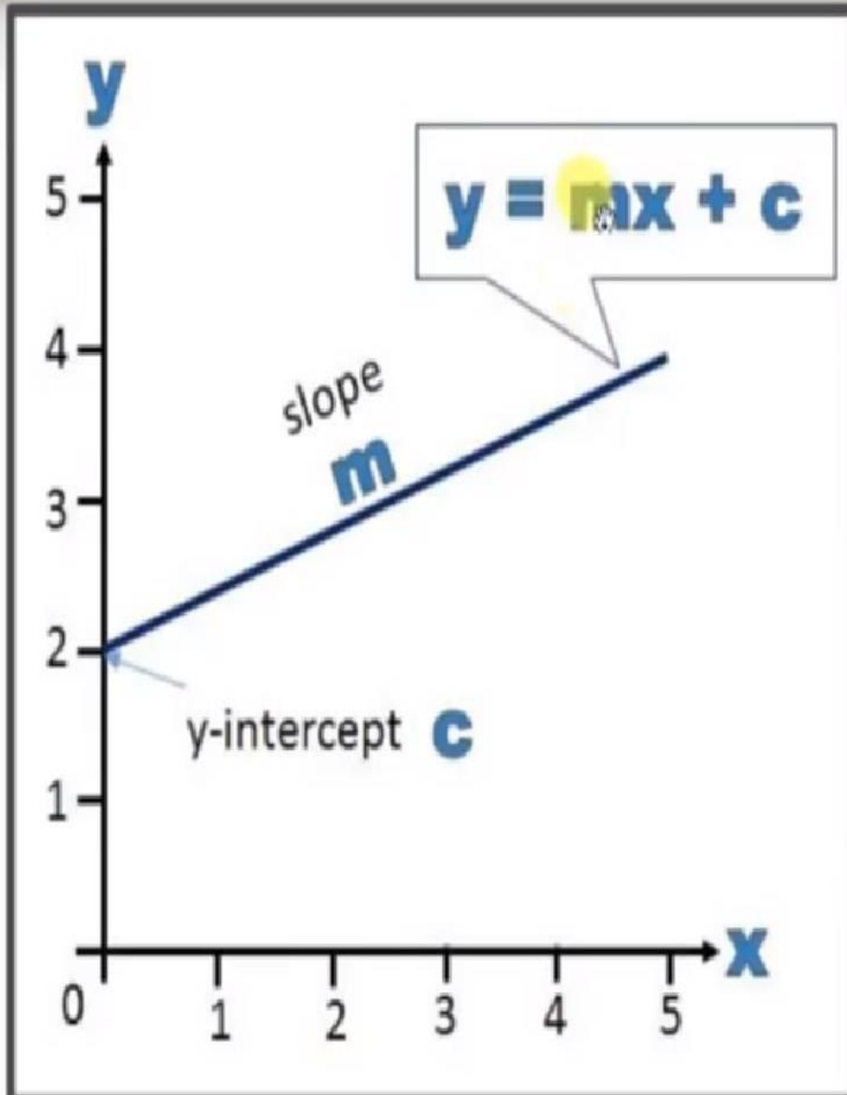
J = 1 / 14 (0+1+4+4+4+25+9)

J = 47/14 = 3.3

Best fit line الخط الأكثر ملائمة



معادلة الخط المستقيم



Linear Regression Equation معادلة التوقع الخطي

Hypothesis: $h_\theta(x) = \theta_0 + \theta_1 x$

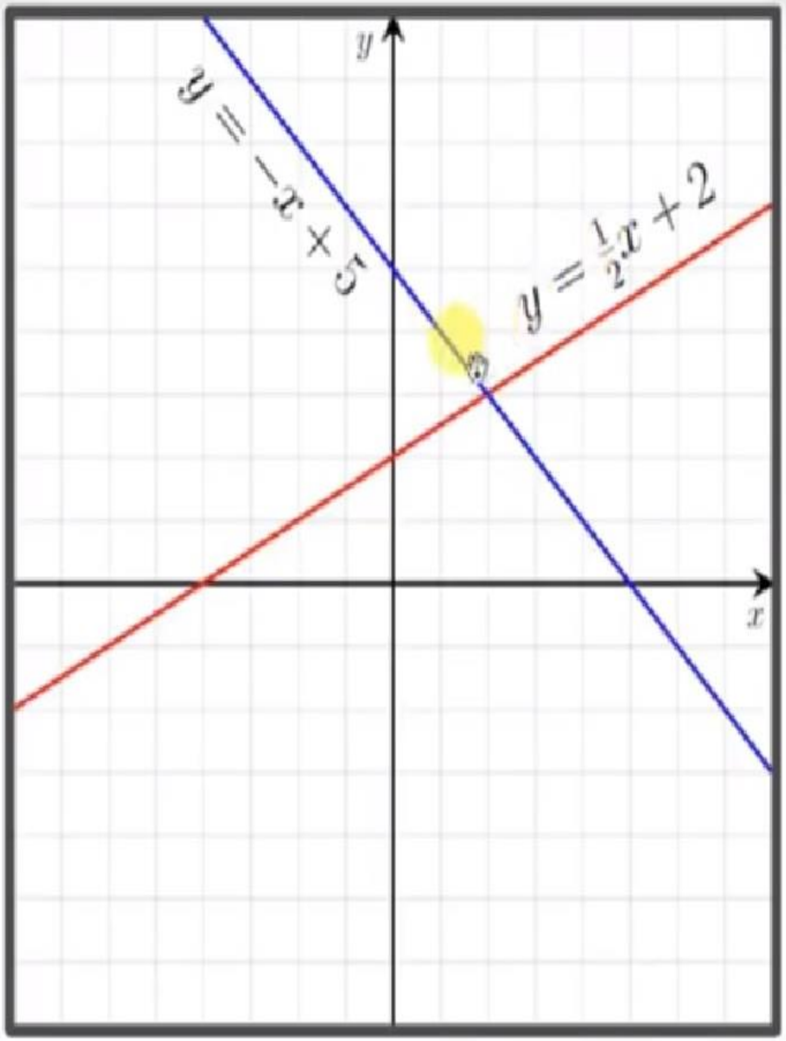
Parameters: θ_0, θ_1

Cost Function: $J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)} \right)^2$

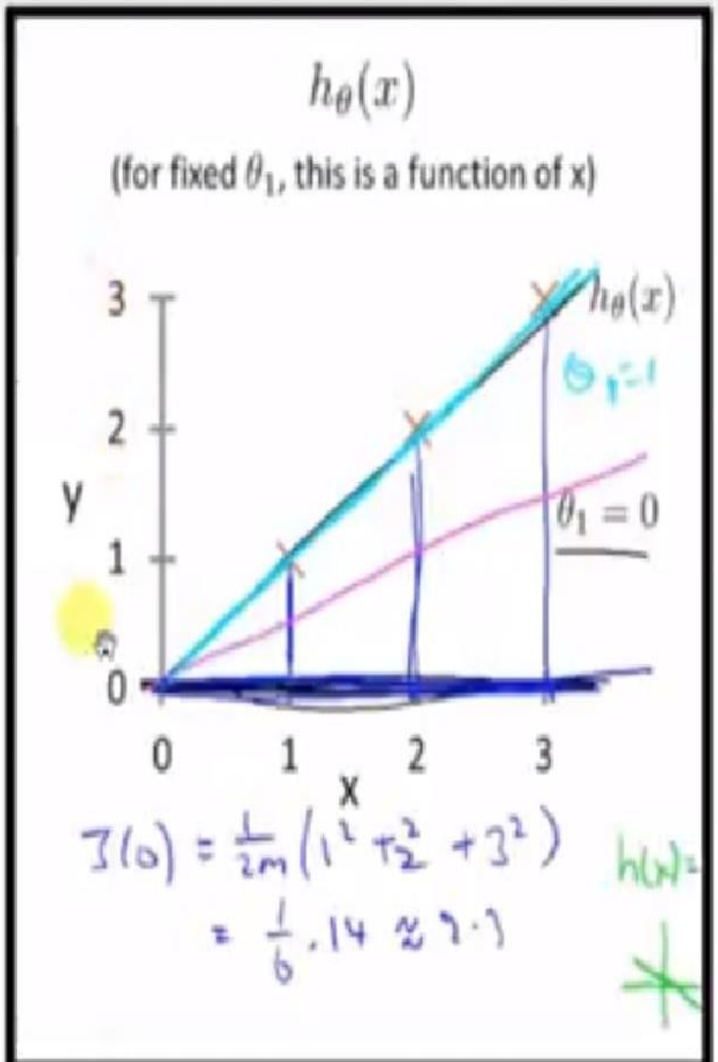
Goal: $\underset{\theta_0, \theta_1}{\text{minimize}}\ J(\theta_0, \theta_1)$

- الهدف تقليل الفارق بين قيمة h(x) و هي القيمة المتوقعة من المعادلة الخطية و قيمة y و هي القيمة الحقيقية
- يتم القسمة علي 2m لربط قيمة الخطا بعدد القيم بالعينة
- الهدف ايجاد قيم ثيتا 1 و ثيتا 2 , والتي تجعل من J (نسبة الخطا) اقل ما يمكن
- تسمي احيانا Cost error function

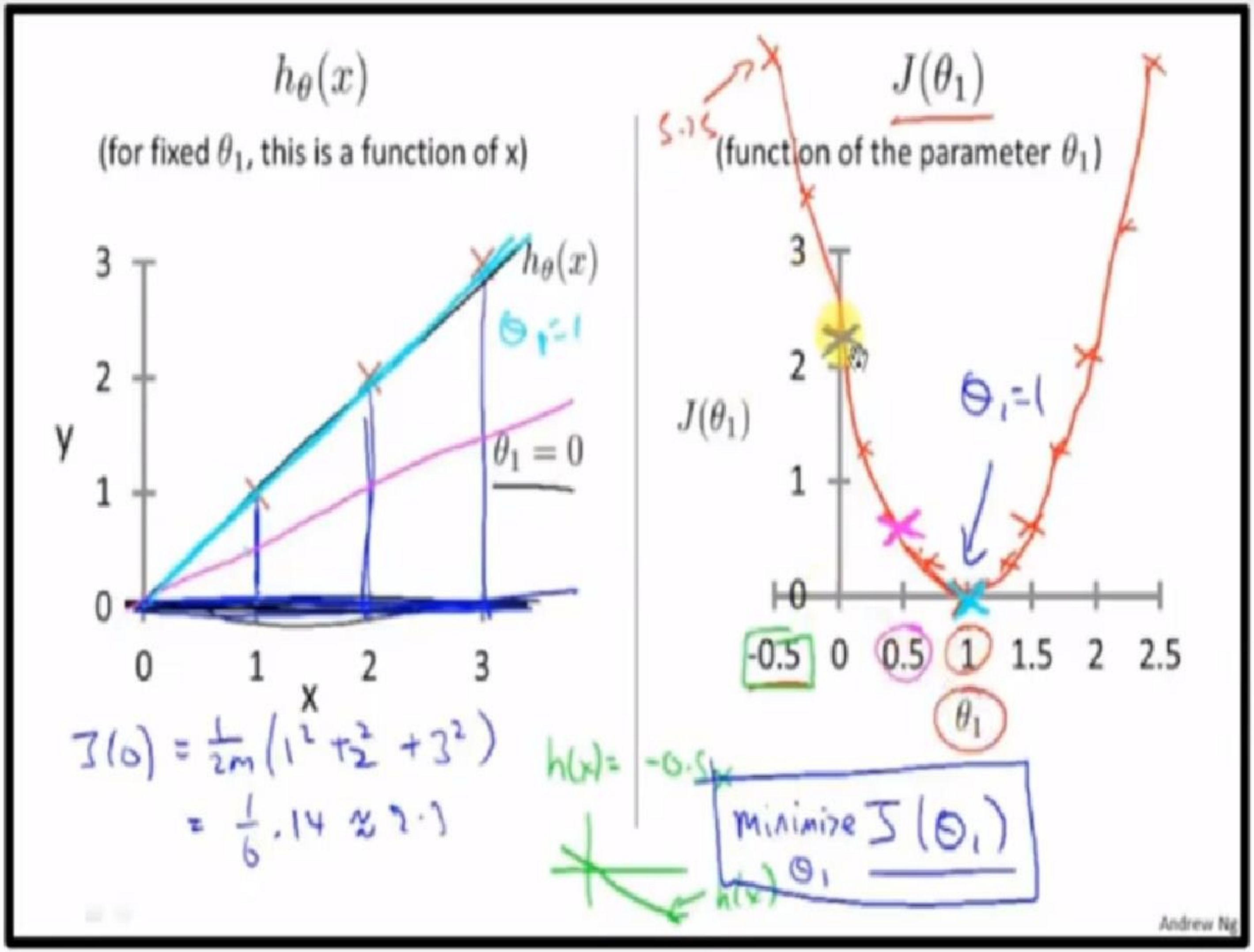
معادلة الخط المستقيم



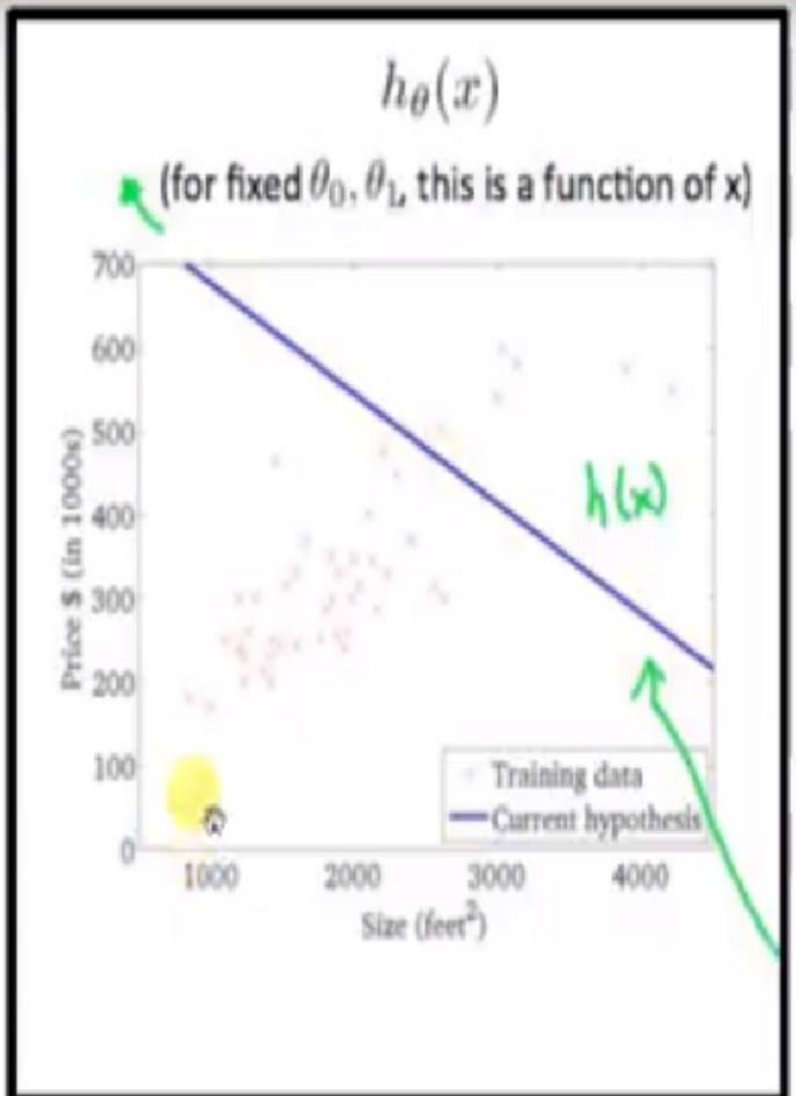
تحديد قيمة ثيتا



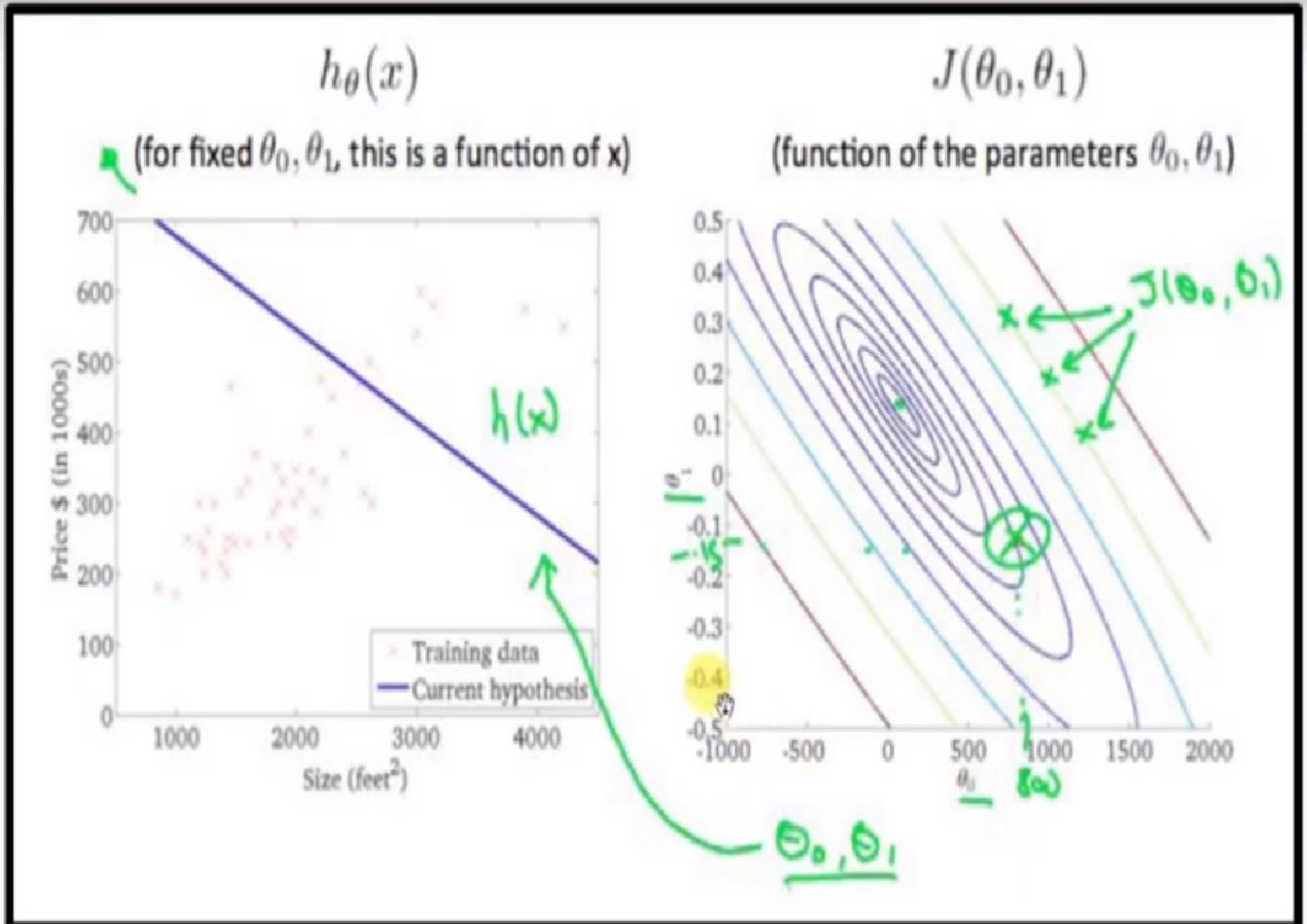
تحديد قيمة ثيتا



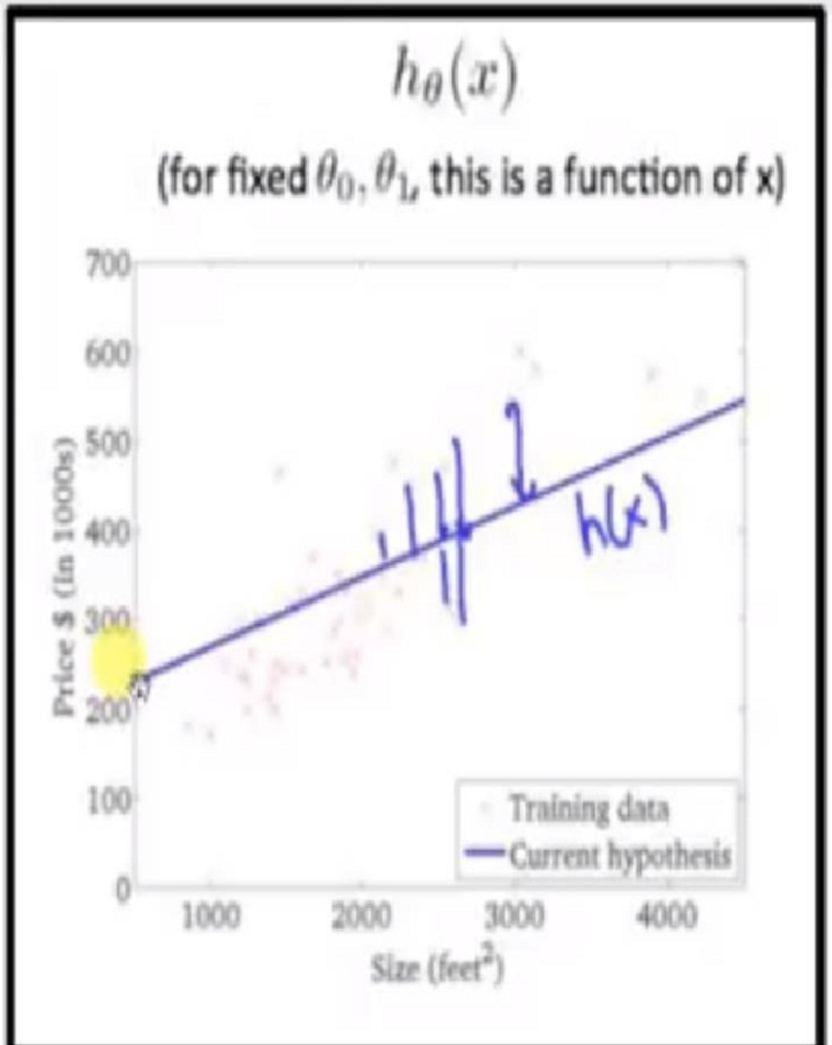
تحديد قيمة ثيتا



تحديد قيمة ثيتا



تحديد قيمة ثيتا



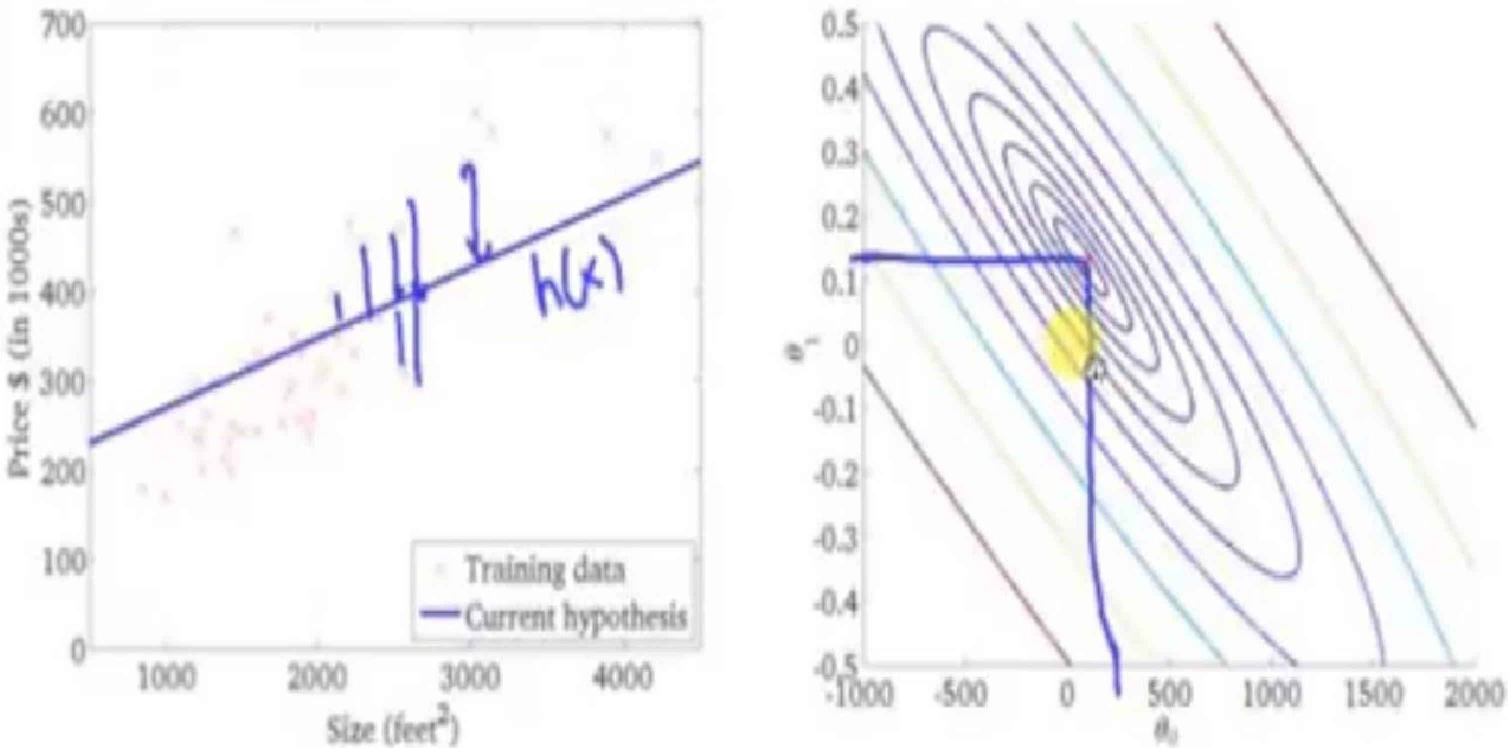
تحديد قيمة ثيتا

$h_\theta(x)$

(for fixed θ_0, θ_1, this is a function of x)

$J(\theta_0, \theta_1)$

(function of the parameters θ_0, θ_1)



Linear Regression Equation معادلة التوقع الخطي

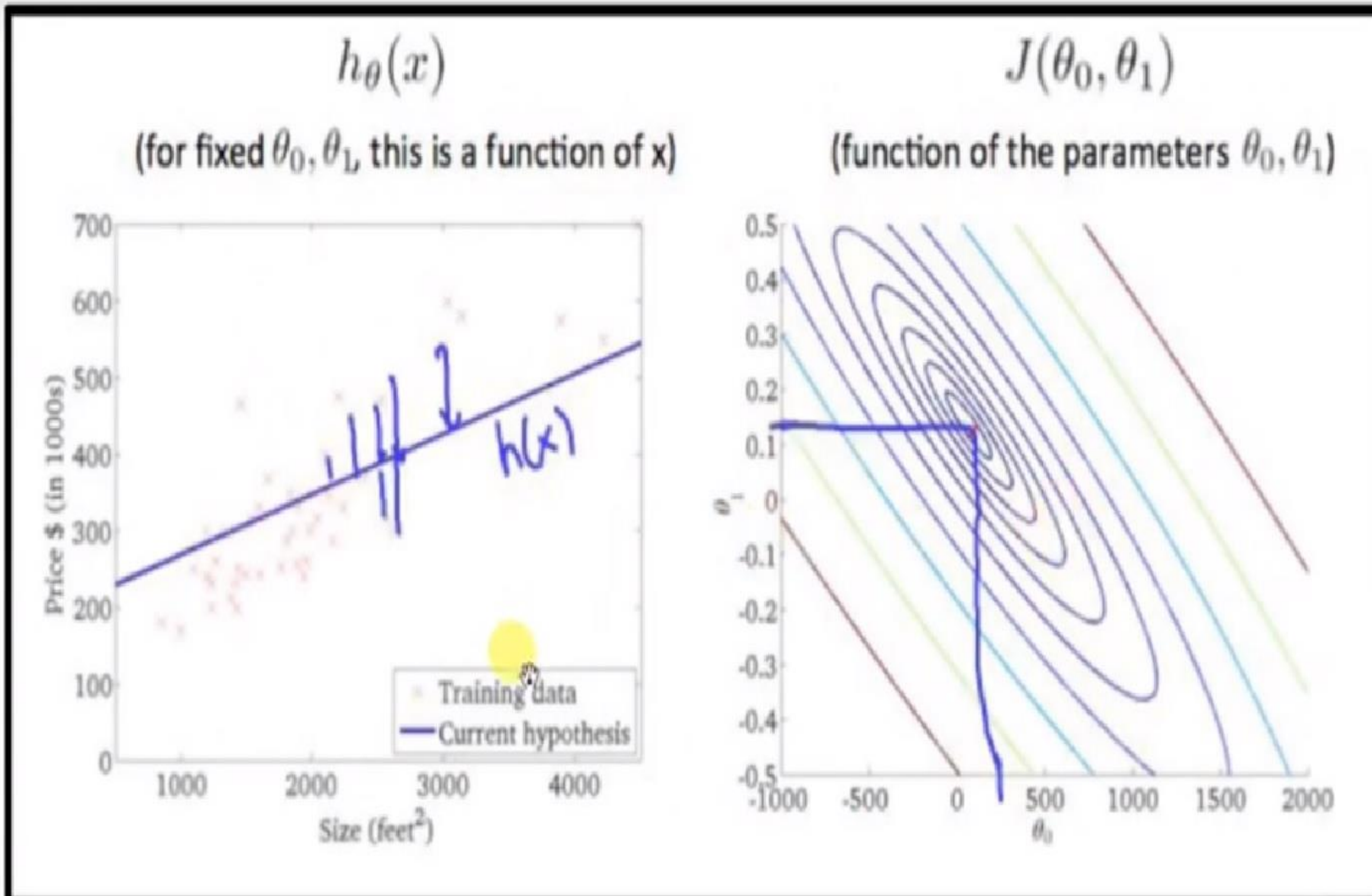
Hypothesis: $h_\theta(x) = \theta_0 + \theta_1 x$

Parameters: θ_0, θ_1

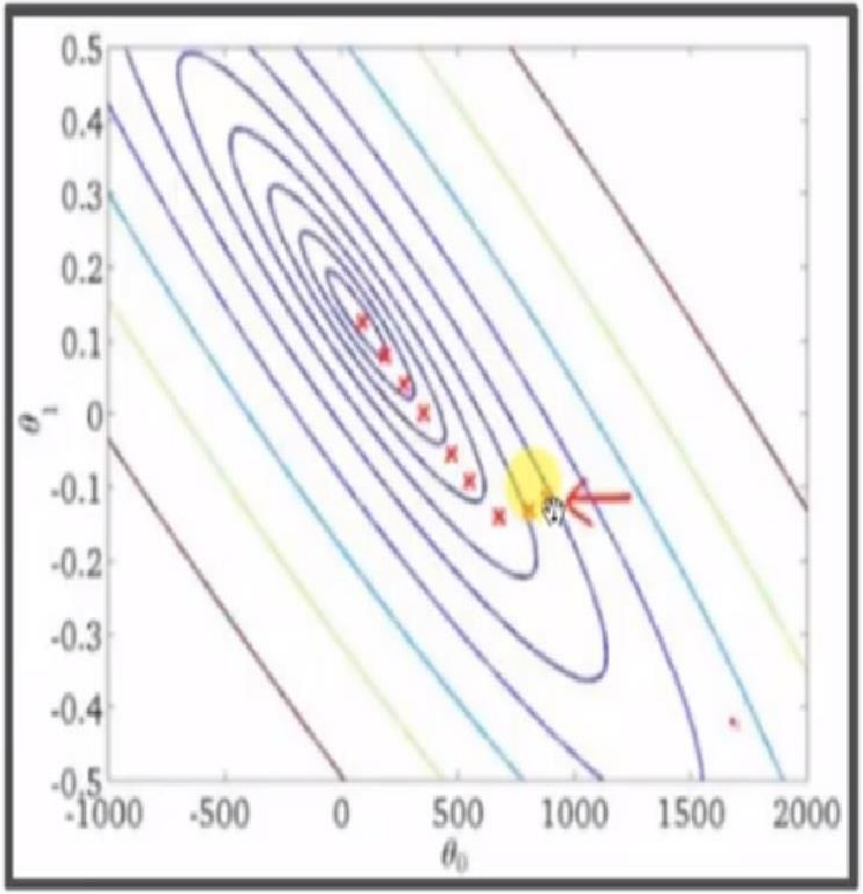
Cost Function: $J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$

Goal: $\underset{\theta_0, \theta_1}{\text{minimize}}\; J(\theta_0, \theta_1)$

تحديد قيمة ثيتا



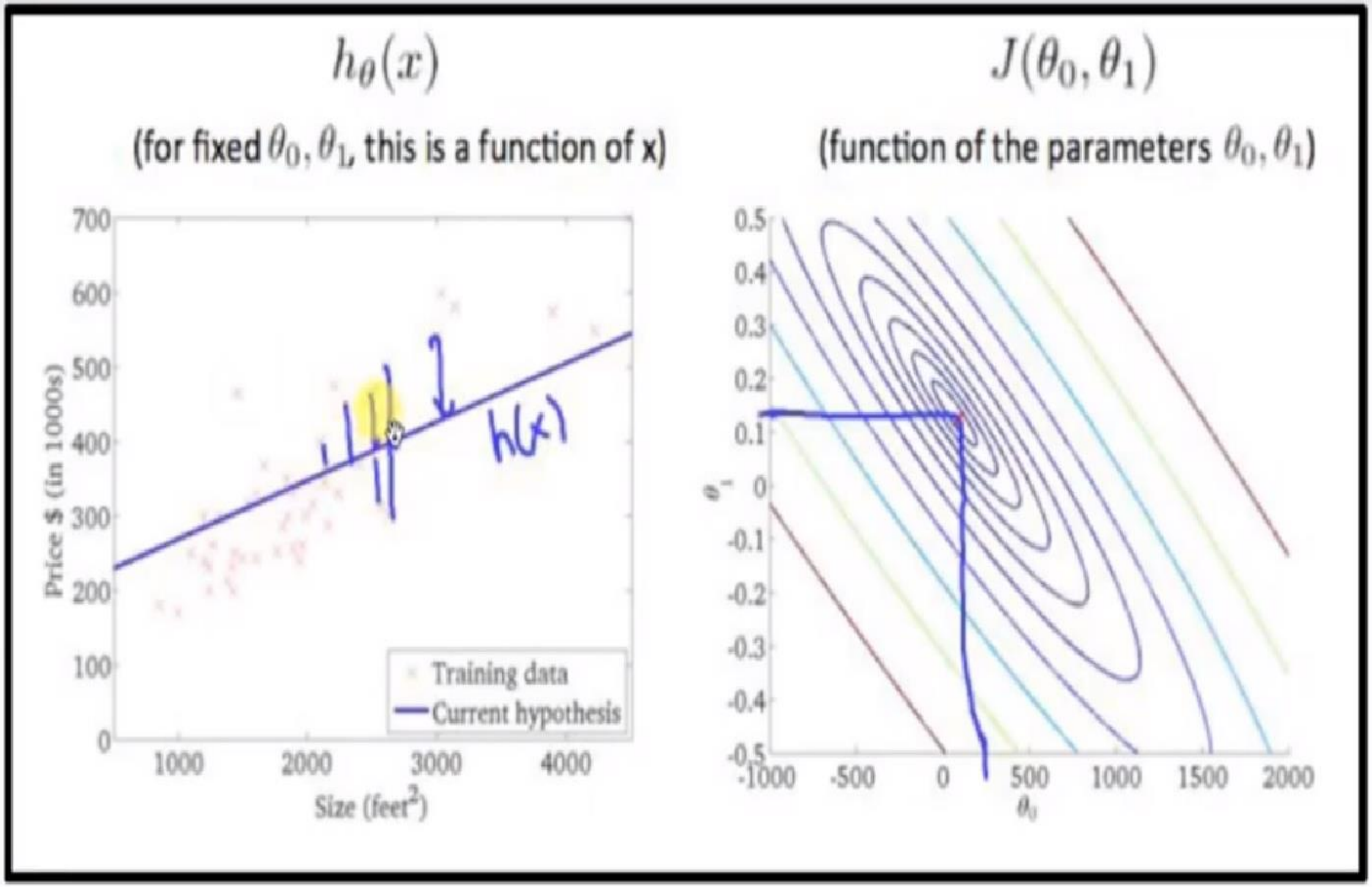
الإنحدار التدريجي Gradient Descent



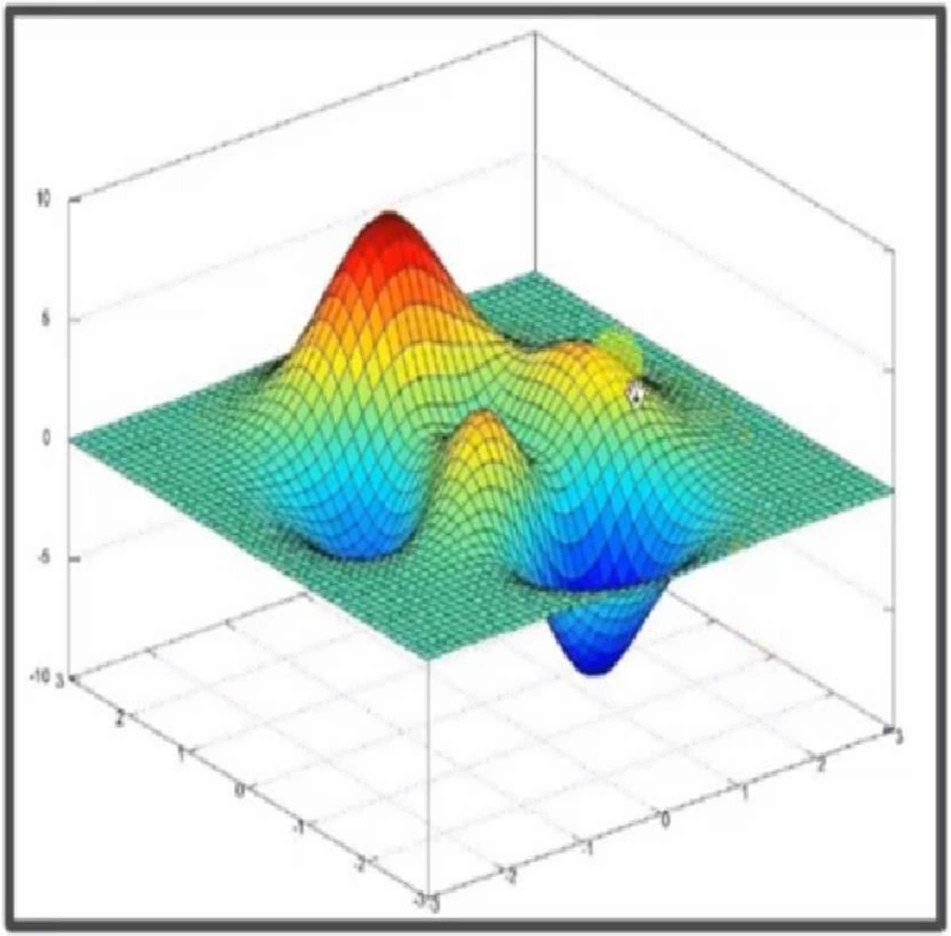
الانحدار التدريجي :

- طالما نحن نبحث عن قيم ثيتا 0 و 1 التي ستقلل قيمة J باقصي قدر , فسنفرض قيم لثيتا 1 و 2 , ثم نقوم بتقليلها تدريجيا حتي نصل لاقل قيمة للـ J

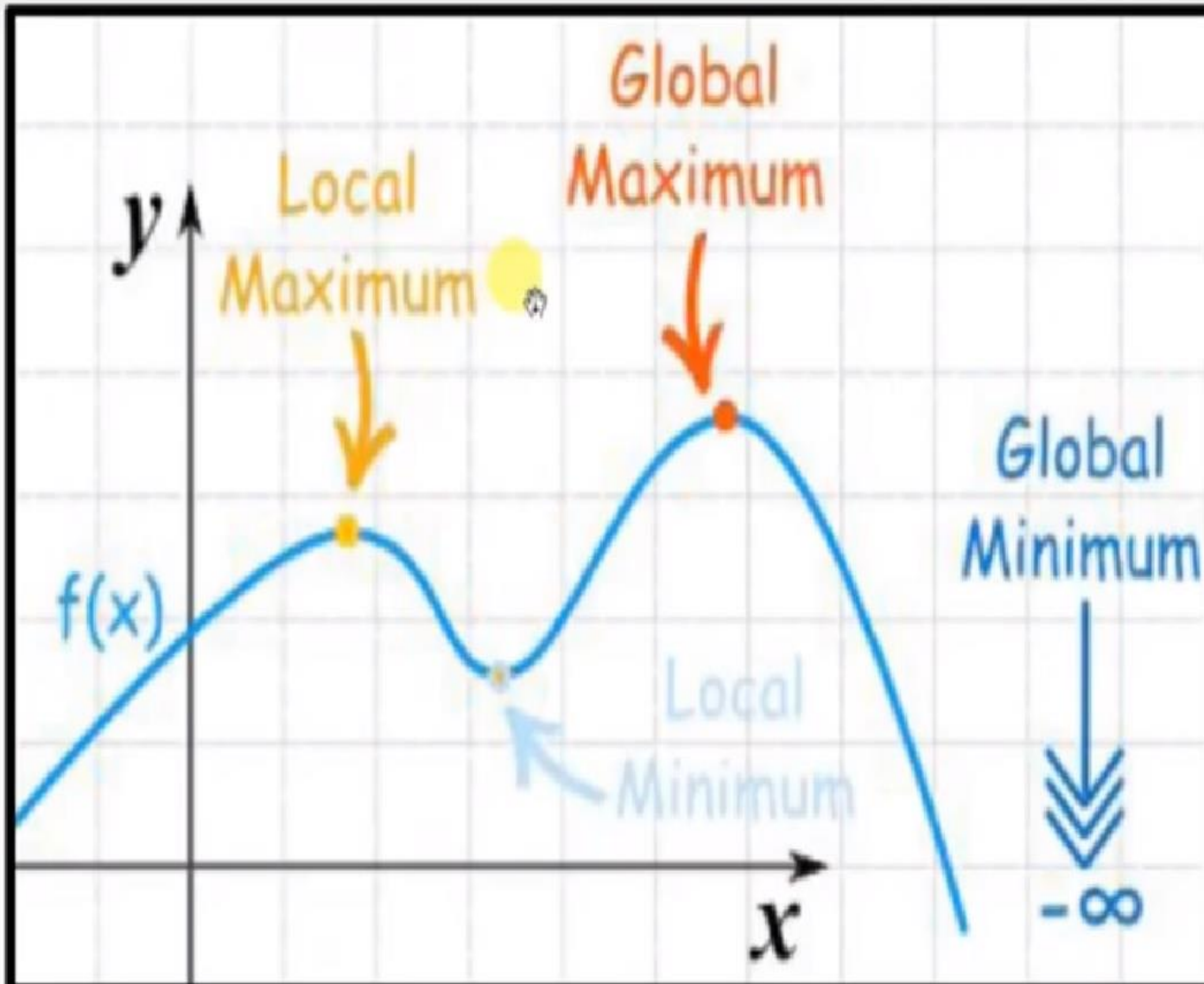
تحديد قيمة ثيتا



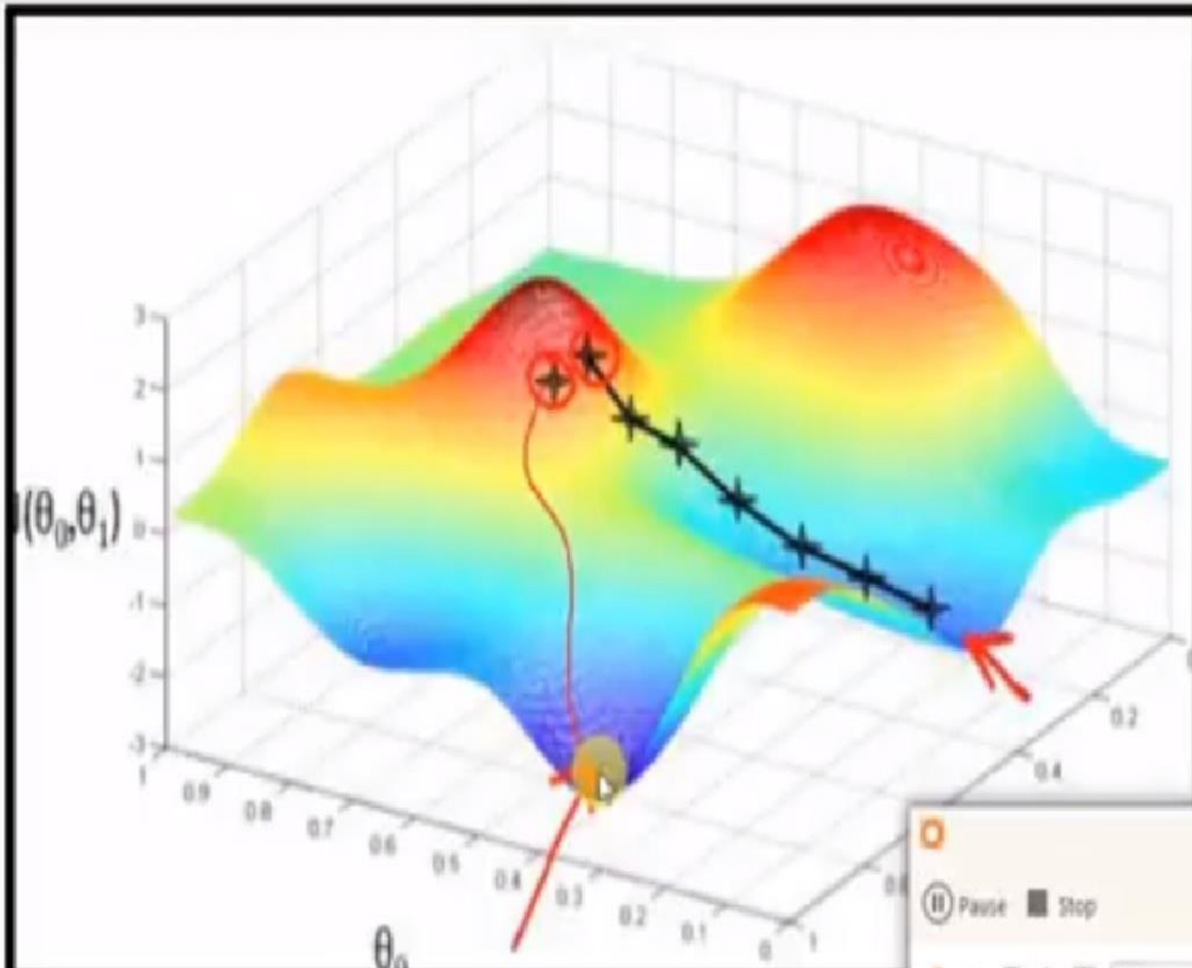
الإنحدار التدريجي Gradient Descent



Local Max & Min القيم المحلية

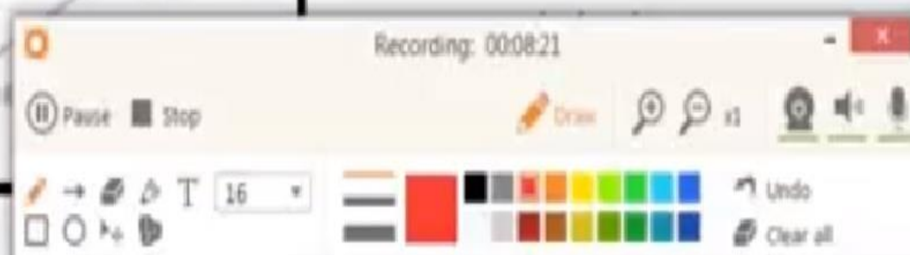


الإنحدار التدريجي Gradient Descent



القيم المحلية :

- فهناك ما يسمي local minimum يعني قيمة دنيا , لكن محلية (علي اليمين) و لا نري جوارها اي قيم دنيا اخري , لكن في الحقيقة هناك قيم اقل منها لكن ابعد
- والقيمة الأقل جميعا اسمها



معادلة الإنحدار التدريجي

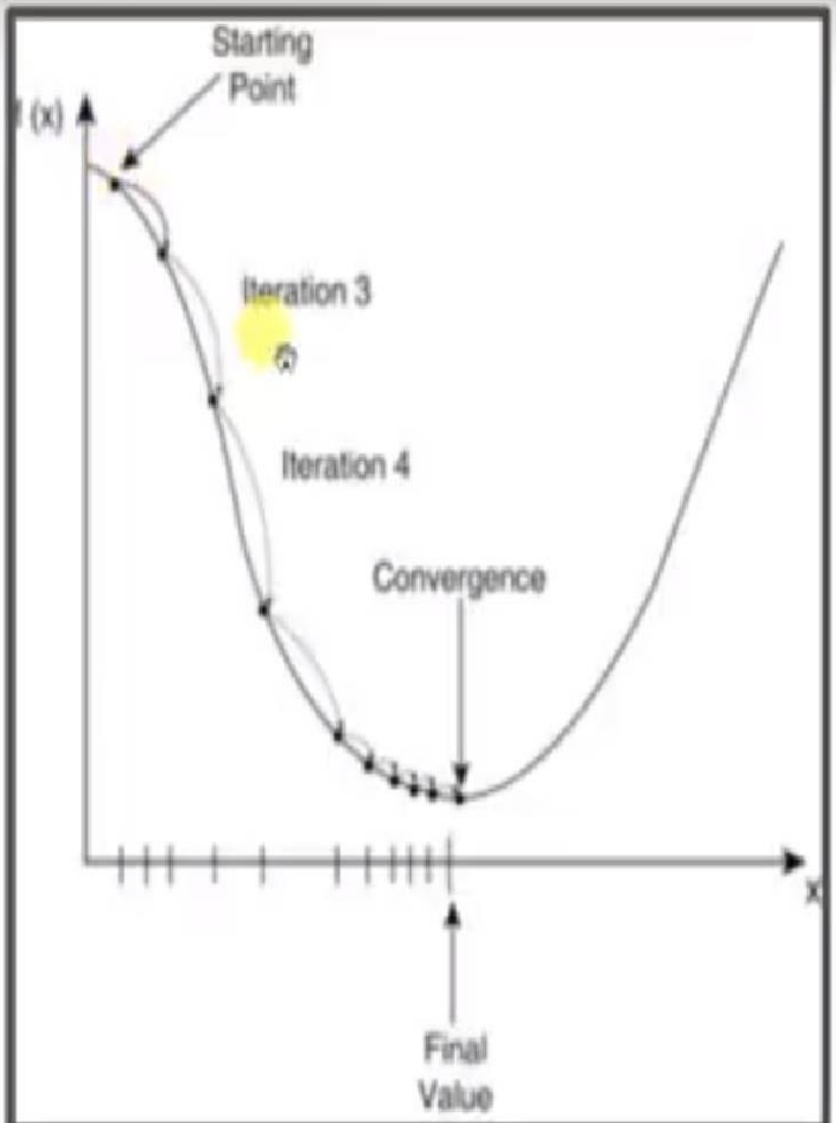
الفكرة :

- اولا فكرة ان =: معناها ان يا برنامج , اعمل overwrite للقيمة اليسري بقيمة اليمني
- الفا هو معامل , ان زاد , ستكون الخطوات اسرع (و غالبا اقل دقة) و ان قل ستكون الخطوات ابطئ و ادق
- ما هو علي يمين الفا , هو اشتقاق جزئي للدالة , بالنسبة لثيتا
- لاحظ ان هذه المعادلة تكون مرة لثيتا صفر , ومرة لثيتا 1
- كلا المعادلتين يمشيان بالتوازي معا , ويتم تكرارهما معا
- ولاحظ ان الابديت لازم يتم زي اللي علي الشمال , مش اليمين
-
-

repeat until convergence:

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$$

معادلة الإنحدار التدريجي



الفكرة :

- لاحظ ان في حالة القيمة الدنيا قيمة التفاضل بصفر (لأن وقتها هيكون الخط شبه مستقيم فالميل هيكون تقريبا 0)
- لاحظ ان قيمة التفاضل تقل كلما قل الميل (التفاضل هو ميل الخط المستقيم ,فتدريجيا هيقل قيمة التفاضل لتغير الميل) , وكلما اقترب من القيمة الدنيا , فلا داعي لتقليل الالفا , فالقيمة نفسها ستقل تدريجيا

معادلة الإنحدار التدريجي

Correct: Simultaneous update	Incorrect:
$\text{temp0} := \theta_0 - \alpha \frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1)$	$\text{temp0} := \theta_0 - \alpha \frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1)$
$\text{temp1} := \theta_1 - \alpha \frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1)$	$\theta_0 := \text{temp0}$
$\theta_0 := \text{temp0}$	$\text{temp1} := \theta_1 - \boxed{\alpha \frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1)}$
$\theta_1 := \text{temp1}$	$\theta_1 := \text{temp1}$

الفكرة :

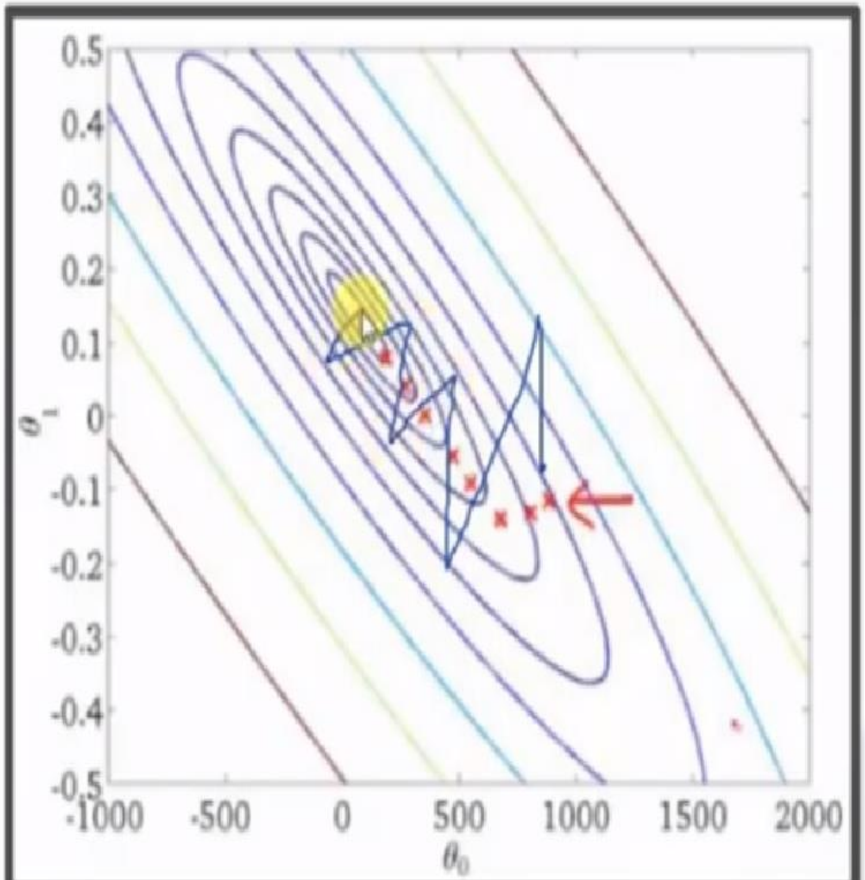
- كلا المعادلتين يمشيان بالتوازي معا , ويتم تكرارهما معا
- ولاحظ ان الابديت لازم يتم زي اللي على الشمال , مش اليمين

معادلة الإنحدار التدريجي

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i)$$

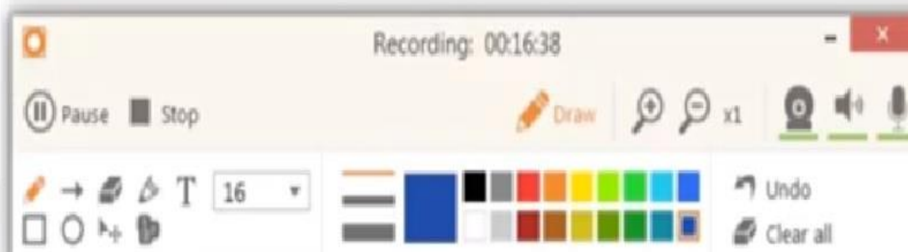
$$\theta_1 := \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^{m} ((h_\theta(x_i) - y_i) x_i)$$

الإنحدار التدريجي Gradient Descent

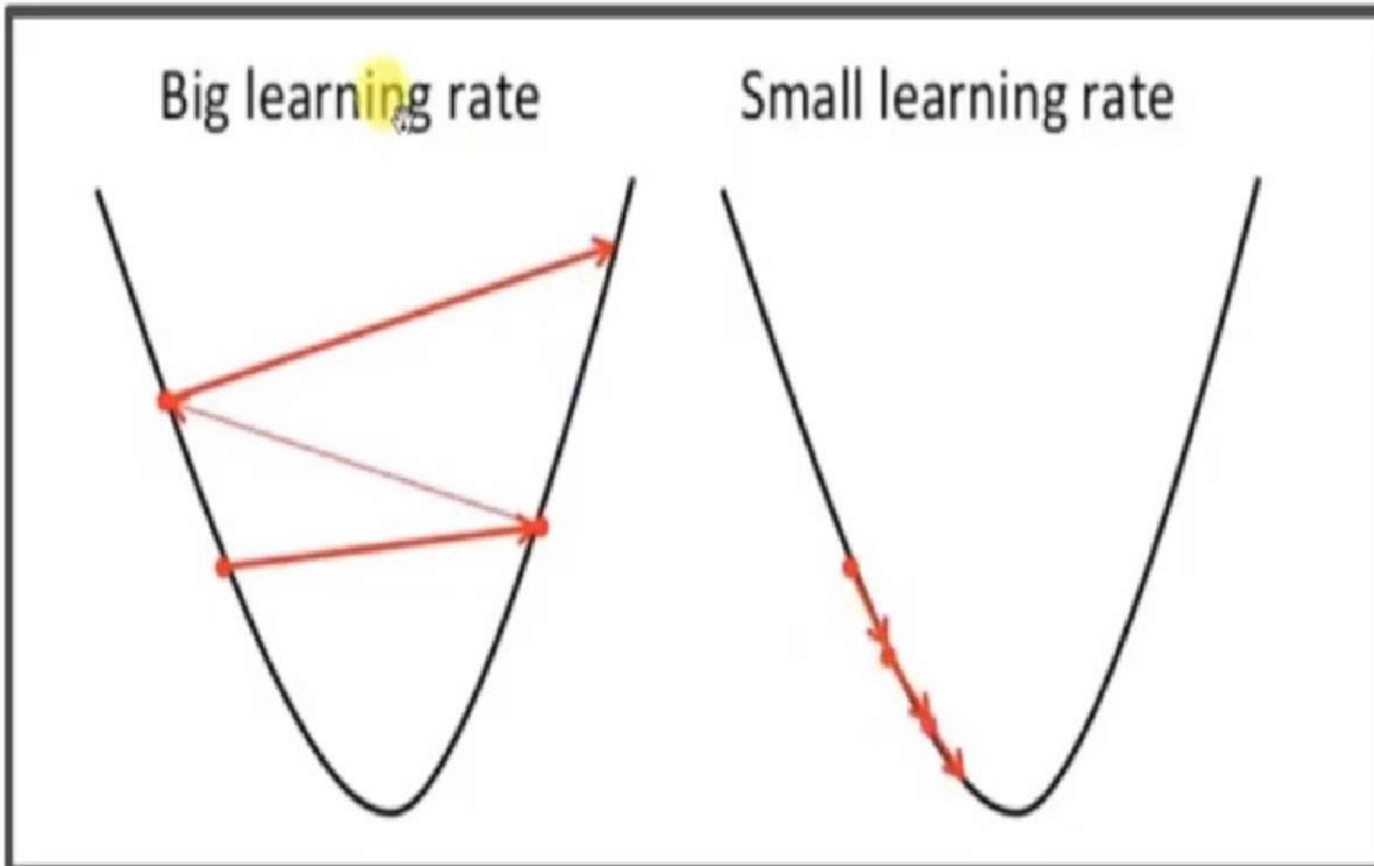


الانحدار التدريجي :

- طالما نحن نبحث عن قيم ثيتا 0 و 1 التي ستقلل قيمة J باقصي قدر , فسنفرض قيم لثيتا 1 و 2 , ثم نقوم بتقليلها تدريجيا حتي نصل لاقل قيمة للـ J



قيمة ألفا



مثال عملي

مساحة البيت (م²) X_1	السعر (الف $) Y
100	135
95	130
90	110
80	95
80	90
70	85
70	80
60	80

- لاحظ ان المساحة اكس , بينما السعر هو واي
- عشان اعمل best fit line هنفرض الثيتات قيم معينة , وليكن ثيتا 0 = 1 و ثيتا 1= 3
- المعادلة هتكون :

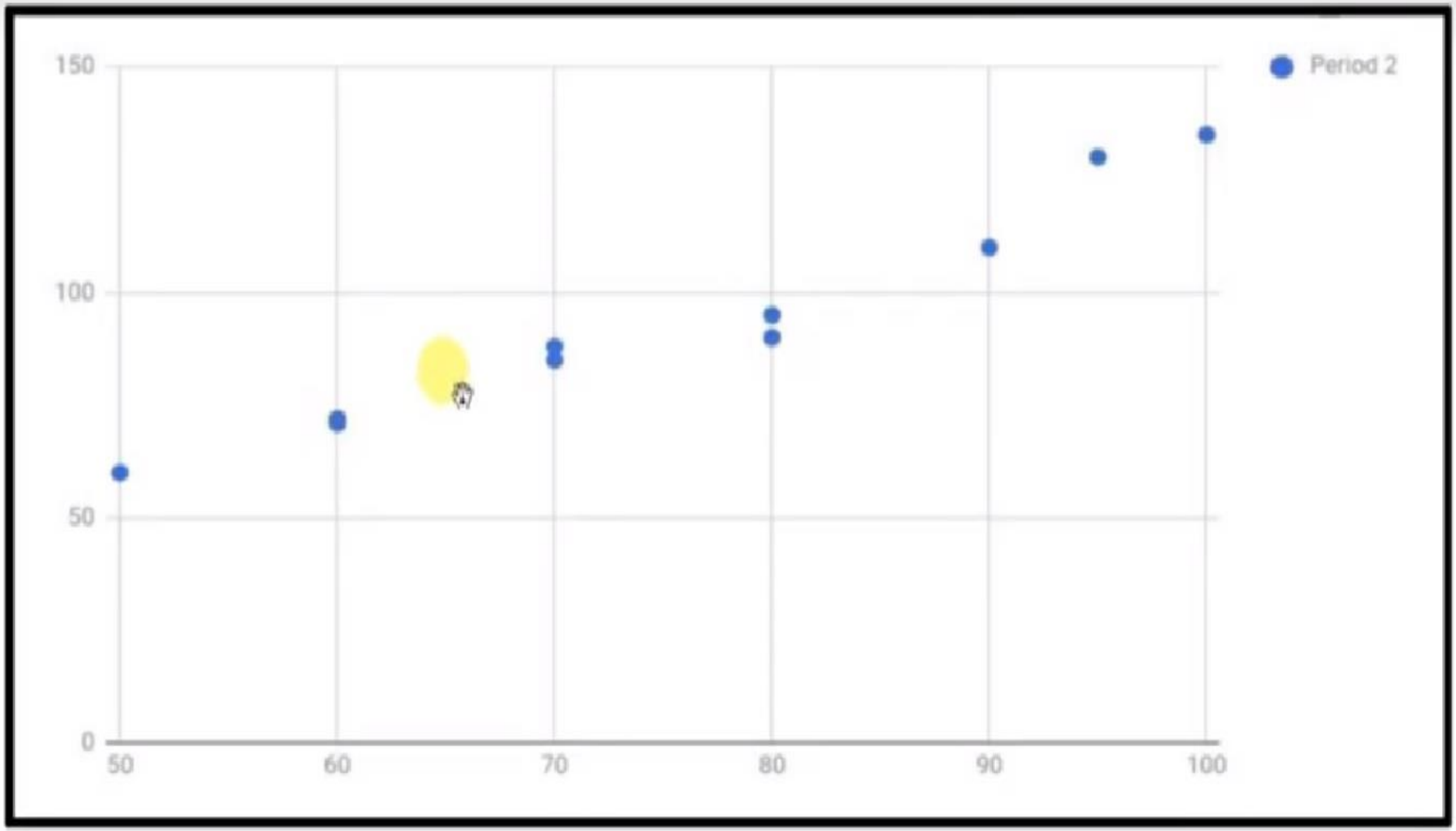
$$h(x) = 1 + 3X$$

معادلة الإنحدار التدريجي

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i)$$

$$\theta_1 := \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^{m} ((h_\theta(x_i) - y_i) x_i)$$

مثال عملي



مثال عملي

X_1	Y	h(x)	h(x) -y
100	300		
95	285		
90	270		
80	240		
80	235		
70	200		
70	205		
60	180		

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i)$$

ثيتا 0 = 1

ثيتا 1 = 3

المعادلة

h(x) = 1 + 3 X

مثال عملي

X_1	Y	h(x)	h(x) -y
100	300	301	
95	285	286	
90	270	271	
80	240	241	
80	235	241	
70	200	211	
70	205	211	
60	180	181	

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i)$$

ثيتا 0 = 1

ثيتا 1 = 3

المعادلة

h(x) = 1 + 3 X

مثال عملي

X_1	Y	h(x)	h(x) -y
100	300	301	1
95	285	286	1
90	270	271	1
80	240	241	1
80	235	241	6
70	200	211	11
70	205	211	6
60	180	181	1

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i)$$

ثيتا 0 = 1

ثيتا 1 = 3

المعادلة

h(x) = 1 + 3 X

مثال عملي

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i)$$

ثيتا 0 = 1

مجموع الفروق = 28

الفا = 0.002

قيمة m = 8

Theta 0 = 1 - ((0.002 / 8) * (28))

Theta 0 = 1 - 0.007 = 0.993

معادلة الإنحدار التدريجي

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i)$$

$$\theta_1 := \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^{m} ((h_\theta(x_i) - y_i) x_i)$$

مثال عملي

X_1	Y	h(x)	h(x) -y	(h(x) -y)x
100	300	301	1	100
95	285	286	1	95
90	270	271	1	90
80	240	241	1	80
80	235	241	6	480
70	200	211	11	770
70	205	211	6	420
60	180	181	1	60

$$\theta_1 := \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^{m} ((h_\theta(x_i) - y_i) x_i$$

ثيتا 0 = 1

ثيتا 1 = 3

مثال عملي

$$\theta_1 := \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^{m} ((h_\theta(x_i) - y_i)x_i$$

ثيتا 0 = 1

مجموع الفروق = 2095

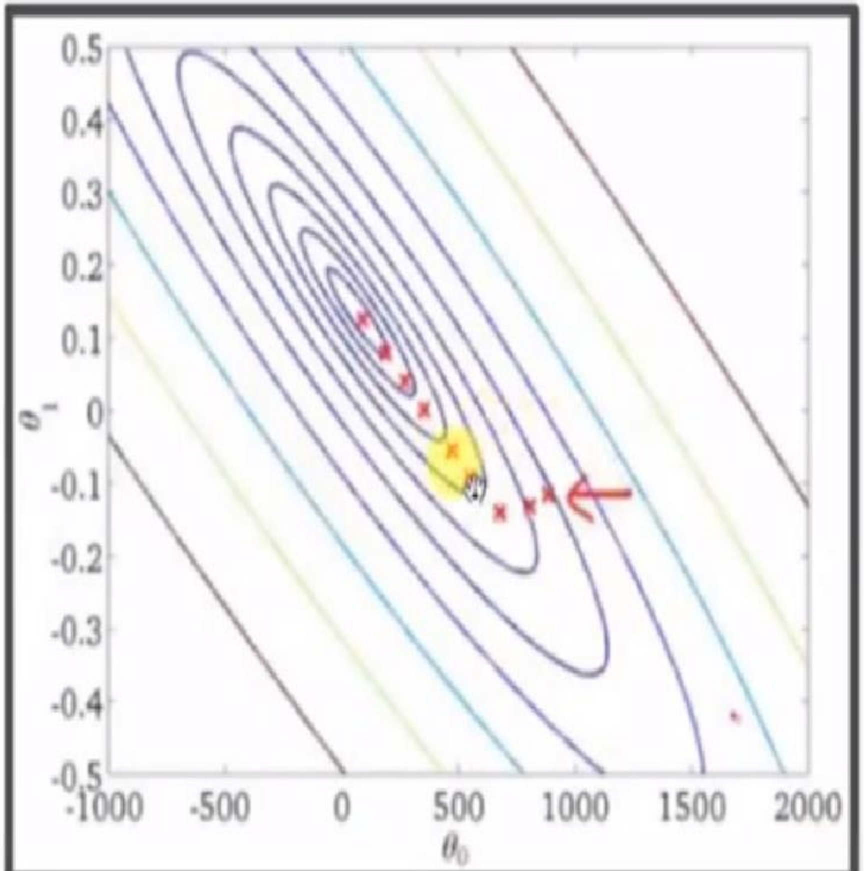
الفا = 0.002

قيمة m = 8

Theta 1 = 3 - ((0.002 / 8) * (2095)

Theta 1 = 3 - 0.52 = 2.48

الإنحدار التدريجي Gradient Descent



الانحدار التدريجي :

- طالما نحن نبحث عن قيم ثيتا 0 و 1 التي ستقلل قيمة L بأقصي قدر , فسنفرض قيم لثيتا 1 و 2 , ثم نقوم بتقليلها تدريجيا حتي نصل لاقل قيمة للـ L

مثال عملي

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i)$$

$$\theta_1 := \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^{m} ((h_\theta(x_i) - y_i) x_i)$$

Theta 0 = 1 Theta 1 = 3

Theta 0 = 0.993 Theta 1 = 2.48

Theta 0 = 0.991 Theta 1 = 2.46

..

..

Theta 0 = 0.825 Theta 1 = 1.772

التوقع الخطي لأكثر من متغير Linear Regression with Multivariables

التعامل مع اكثر من بعد :

- تحدثنا سابقا , عن التعامل مع متغير واحد (قيمة لـ X و نجيب منها قيمة Y) الان نتعامل مع اكثر من متغير
- أكثر من متغير معناها ان البيانات الداخلة لها اكثر معلومة لكل صف , فبدلا من ادخال مساحة البيت لمعرفة سعره (X واحدة) , نقوم بادخال مساحة البيت و عدد غرفه , وعمره, و موقعه , وحالته , ولونه , لتحديد سعره , وهذه الأشياء تسمي features

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

Multiple features (variables).

Size ($feet^2$)	Number of bedrooms	Number of floors	Age of home (years)	Price ($1000)
x_1	x_2	x_3	x_4	y
2104	5	1	45	460
1416	3	2	40	232
1534	3	2	30	315
852	2	1	36	178
…	…	…	…	…

$m = 47$

Notation:

- n = number of features $n = 4$
- $x^{(i)}$ = input (features) of i^{th} training example.
- $x_j^{(i)}$ = value of feature j in i^{th} training example.

التعامل مع اكثر من بعد :

- فنري ان سعر البيت (Y) يتأثر بعدد من العوامل (Features) (Xs
- عدد الاكسات نسميه n, بينما عدد الصفوف لازال m
-
-

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

$$x_j^{(i)} = \text{value of feature } j \text{ in the } i^{th} \text{ training example}$$

- الرقم اللي فوق يكون رقم الصف (انهي ريكورد فيهم m) و الرقم اللي تحت هيكون رقم العمود (انهي معلومة فيهم n)

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n$$

For convenience of notation, define $\boxed{x_0 = 1.}$ $\quad (x_0^{(i)} = 1)$

$$x = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n+1} \qquad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{bmatrix} \in \mathbb{R}^{n+1}$$

$$\underbrace{\begin{bmatrix} \theta_0 & \theta_1 & \cdots & \theta_n \end{bmatrix}}_{\substack{\theta^T \\ (n+1)\times 1 \\ \text{matrix}}} \underbrace{\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix}}_{x}$$

$$\theta^T x$$

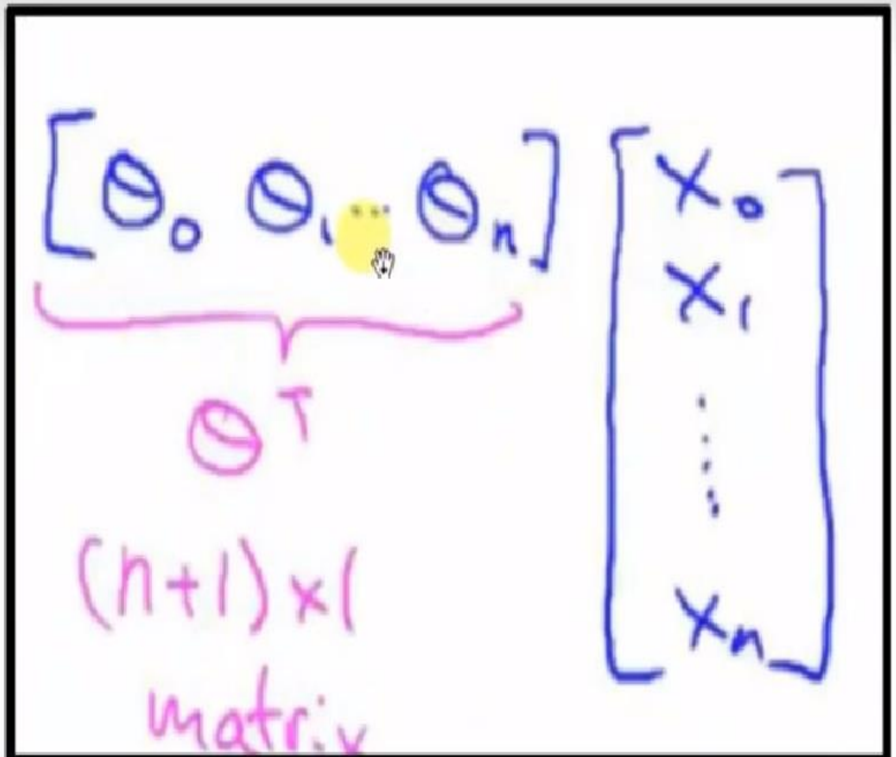
$$h_\theta(x) = \theta_0 \overset{=1}{x_0} + \theta_1 x_1 + \cdots + \theta_n x_n$$

$$= \boxed{\theta^T x.}$$

Multivariate linear regression.

- وقتها الفنكشن , هتكون متعددة الحدود زي كدة , و هنعمل ماتركس للاكسات , وواحدة للثيتات , ونضربهم في بعض بعد ما نعمل ترانزبوس للثيتا

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير



- ليه بنعمل ترانزبوس ؟
  لان الثيتا و الاكس اصلا هما فيكتور (عمود واحد في كذا صف) , فلازم اعمل ترانزبوس لواحد فيهم و اضربه في التاني , عشان تكون المصفوفة الاولي صف واحد في 5 عواميد مثلا , والتانية زي ما هي 5 صفوف في عمود واحد , يتضربو يبقو رقم واحد بس

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n$$

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

$$\text{Cost Function:}\quad J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

- وخد بالك الصيغة القديمة اللي كانت للـ J هيكون فيها شوية تعديل , عشان مبقاش عامل واحد , نفس المعادلة , لكن دلوقتي H بقت فيها ثيتات كتيرة
- لما اعمل تفاضل , هتظل اتش زي هي , وهتختفي كل الثيتات التانية عدا الثيتا اللي باعمل تفاضل علي اساسها اللي هتتبقي في الاخر

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

- القانون الجديد

repeat until convergence: {

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)}) \cdot x_0^{(i)}$$

$$\theta_1 := \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)}) \cdot x_1^{(i)}$$

$$\theta_2 := \theta_2 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)}) \cdot x_2^{(i)}$$

$\cdots$

}

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

- الصيغة المجمعة

repeat until convergence: {

$$\theta_j := \theta_j - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)}) \cdot x_j^{(i)} \qquad \text{for j := 0...n}$$

}

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

سعر السيارات :

X_1	X_2	X_3	Y
العمر	القدرة	الاسطوانات	السعر
5	20	6	12
5	35	6	14
6	38	8	16
7	40	8	15
7	46	10	20

- عدد السيارات 5 (m)
- المعلومات عن كل سيارة (features n) 3

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

X_1	X_2	X_3	Y
العمر	القدرة	الاسطوانات	السعر
5	20	6	114
5	35	6	120
6	38	8	123
7	40	8	121
7	46	10	135

X_1	X_2	X_3
1	1	1
5	5	6
20	35	38
6	6	8

X_4	X_5
1	1
7	7
40	46
8	10

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

X_1	X_2	X_3	Y
العمر	القدرة	الاسطوانات	السعر
5	20	6	114
5	35	6	120
6	38	8	123
7	40	8	121
7	46	10	135

X_1	X_2	X_3
1	1	1
5	5	6
20	35	38
6	6	8

X_4	X_5
1	1
7	7
40	46
8	10

Theta

Theta0
Theta1
Theta2
Theta3

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

X_1	X_2	X_3	Y
العمر	القدرة	الاسطوانات	السعر
5	20	6	114
5	35	6	120
6	38	8	123
7	40	8	121
7	46	10	135

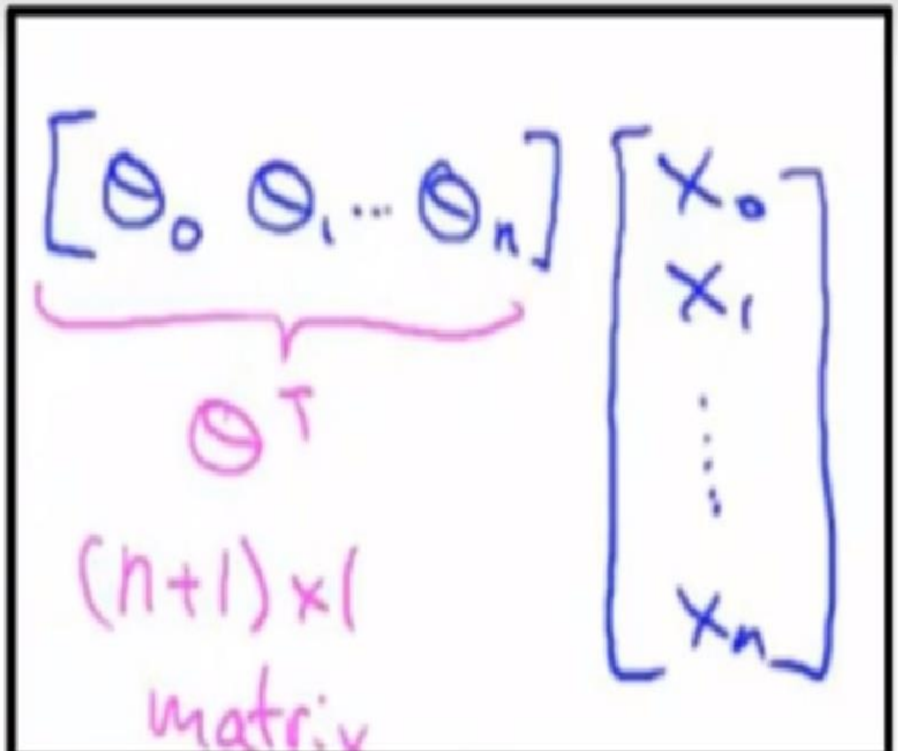
X_1	X_2	X_3
1	1	1
5	5	6
20	35	38
6	6	8

X_4	X_5
1	1
7	7
40	46
8	10

Theta

Theta0	5
Theta1	2
Theta2	3
Theta3	6

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير



- ليه بنعمل ترانزبوس ؟

لان الثيتا و الاكس اصلا هما فيكتور (عمود واحد في كذا صف) , فلازم اعمل ترانزبوس لواحد فيهم و اضربه في التاني , عشان تكون المصفوفة الاولي صف واحد في 5 عواميد مثلا , والتانية زي ما هي 5 صفوف في عمود واحد , يتضربو يبقو رقم واحد بس

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n$$

B-11 - مثال علي أكثر من متغير Multivariable Example

اضغط على Esc للخروج من وضع ملء الشاشة

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

$h(x) = (\text{Theta})^T X$

$$(\text{Theta})^T = \begin{matrix} 5 \\ 2 \\ 3 \\ 6 \end{matrix}^{\,T} = (5\ 2\ 3\ 6) \qquad X_1 = \begin{matrix} 1 \\ 5 \\ 20 \\ 6 \end{matrix}$$

9:23 / 20:41

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

$h(x) = (\text{Theta})^T X$

$$(\text{Theta})^T = \begin{matrix} 5 \\ 2 \\ 3 \\ 6 \end{matrix}^T = (5\ 2\ 3\ 6) \qquad X_1 = \begin{matrix} 1 \\ 5 \\ 20 \\ 6 \end{matrix}$$

$$h(x)_1 = (5\ 2\ 3\ 6) \begin{matrix} 1 \\ 5 \\ 20 \\ 6 \end{matrix} = 5*1 + 2*5 + 3*20 + 6*6 = 111$$

$$\underbrace{[\theta_0\ \theta_1 \cdots \theta_n]}_{\theta^T} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$(n+1) \times 1$ matrix

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

$h(x) = (Theta)^T X$

$$(Theta)^T = \begin{matrix} 5 \\ 2 \\ 3 \\ 6 \end{matrix}^{T} = (5\ 2\ 3\ 6) \qquad X_1 = \begin{matrix} 1 \\ 5 \\ 20 \\ 6 \end{matrix}$$

$$\underbrace{[\theta_0\ \theta_1 \cdots \theta_n]}_{\theta^T} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$(n+1) \times 1$ matrix

$$h(x)_1 = (5\ 2\ 3\ 6) \begin{matrix} 1 \\ 5 \\ 20 \\ 6 \end{matrix} = 5*1 + 2*5 + 3*20 + 6*6 = 111$$

$h(x)_1$ = 111 $h(x)_2$ = 119 $h(x)_3$ = 127 $h(x)_4$ = 122 $h(x)_5$ = 140

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

- القانون الجديد

repeat until convergence: {

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)}) \cdot x_0^{(i)}$$

$$\theta_1 := \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)}) \cdot x_1^{(i)}$$

$$\theta_2 := \theta_2 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)}) \cdot x_2^{(i)}$$

...

}

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)}) \cdot x_0^{(i)}$$

Suppose Alpha = 0.01 **m = 5**

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)}) \cdot x_0^{(i)}$$

Suppose Alpha = 0.01   m = 5

Theta 0 = 5 - (0.01 /5) [(111-114) + (119-120) + (127-123)+ (122-121)+ (140-135) (1)] = 4.9

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)}) \cdot x_0^{(i)}$$

Suppose Alpha = 0.01   m = 5

Theta 0 = 5 - (0.01 /5) [(111-114) + (119-120) + (127-123)+ (122-121)+ (140-135) (1)] = 4.9

Theta 1 = 2 - (0.01 /5) [(111-114) + (119-120) + (127-123)+ (122-121)+ (140-135) (5)] = 2.6

Theta 2 = 3 - (0.01 /5) [(111-114) + (119-120) + (127-123)+ (122-121)+ (140-135) (20)] = 3.9

Theta 3 = 6 - (0.01 /5) [(111-114) + (119-120) + (127-123)+ (122-121)+ (140-135) (6)] = 6.4

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

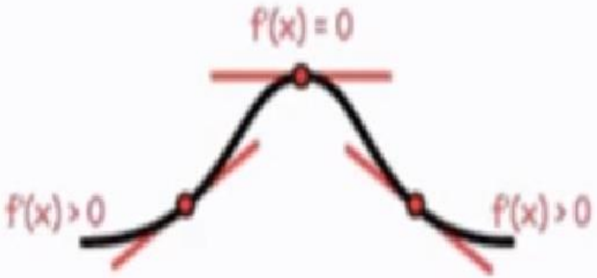
$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)}) \cdot x_0^{(i)}$$

Suppose Alpha = 0.01 m = 5

Theta 0 = 4.9	Theta 0 = 4.7	Theta 0 = 4.68	Theta 0 = 4.6236
Theta 1 = 2.6	Theta 1 = 2.55	Theta 1 = 2.542	Theta 1 = 2.5398
Theta 2 = 3.9	Theta 2 = 3.87	Theta 2 = 3.863	Theta 2 = 3.8605
Theta 3 = 6.4	Theta 3 = 6.36	Theta 3 = 6.357	Theta 3 = 6.35721

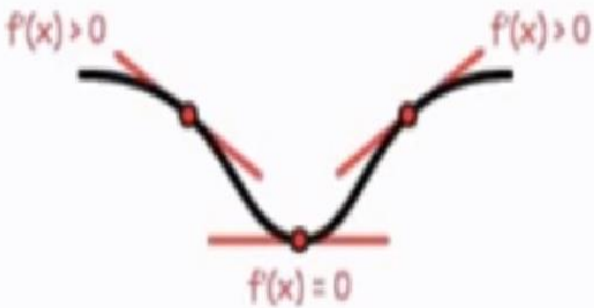
Normal Equation المعادلة العمودية

مفهوم القيم العظمي و الصغري



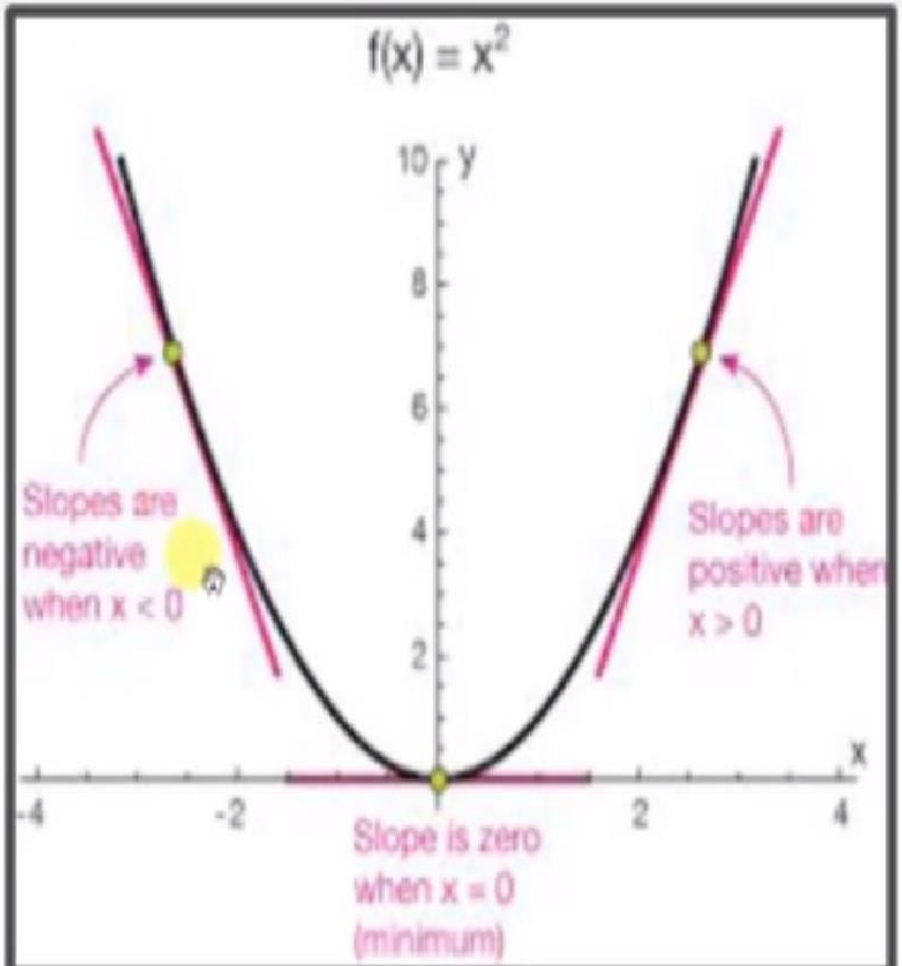
- اي منحني تصاعدي , تكون قيمة التفاضل موجبة

- اي منحني هابط , تكون قيمة التفاضل سالبة



- اي قيمة صغري او كبري تكون قيمة التفاضل صفر

Normal Equation المعادلة العمودية



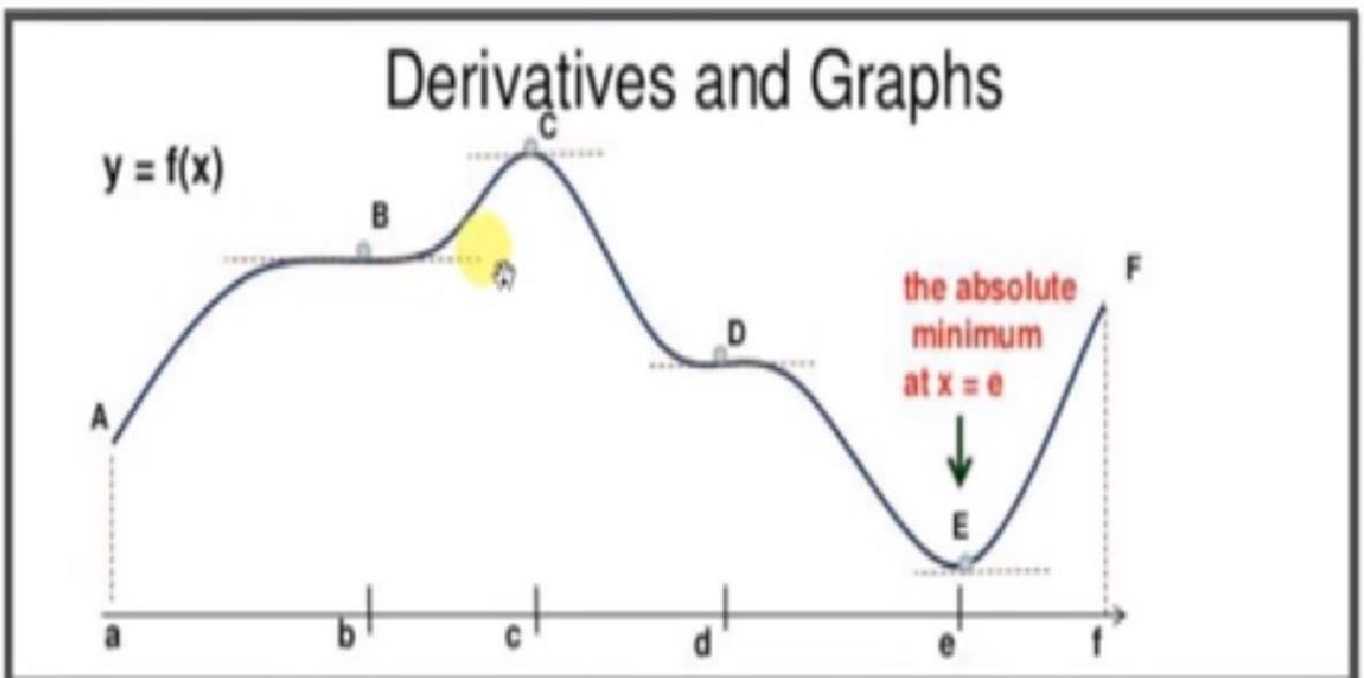
مفهوم القيم العظمي و الصغري

- اي منحني تصاعدي , تكون قيمة التفاضل موجبة
- اي منحني هابط , تكون قيمة التفاضل سالبة
- اي قيمة صغري او كبري تكون قيمة التفاضل صفر

Normal Equation المعادلة العمودية

مفهوم القيم العظمي و الصغري

- تتواجد عندما يكون التفاضل يساوي صفر



Linear Regression Equation معادلة التوقع الخطي

Theta0 = 5 , theta 1 = 2 Equation $h(x) = 5 + 2x$

X	Y	h(x)	h(x) - y	$(h(x) - y)^2$
1	7	7	0	0
2	8	9	1	1
2	7	9	2	4
3	9	11	2	4
4	11	13	2	4
5	10	15	5	25
5	12	15	3	9

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

J = 1 / 14 (0+1+4+4+4+25+9)

J = 47/14 = 3.3

المعادلة العمودية Normal Equation

طريقة الـ Normal Equation

- وهي عن طريق الاعتماد علي تفاضل الـ ل و مساوتها بالصفر لايجاد قيمة الثيتا المطلوبة
- و اذا كان لدينا جدول مثل هذا لاكثر من متغير , فبعد مفاضلتها و مساوتها بالصفر ستكون الثيتا كالتالي

$$\boxed{\theta = (X^T X)^{-1} X^T y}$$

Normal Equation المعادلة العمودية

X_1	X_2	X_3	Y
العمر	القدرة	الاسطوانات	السعر
5	20	6	114
5	35	6	120
6	38	8	123
7	40	8	121
7	46	10	135

X_1	X_2	X_3
1	1	1
5	5	6
20	35	38
6	6	8

X_4	X_5
1	1
7	7
40	46
8	10

Y
114
120
123
121
135

Normal Equation المعادلة العمودية

$$\theta = (X^T X)^{-1} X^T y$$

$$X = \begin{matrix} 1 & 5 & 20 & 6 \\ 1 & 5 & 35 & 6 \\ 1 & 6 & 38 & 8 \\ 1 & 7 & 40 & 8 \\ 1 & 7 & 46 & 10 \end{matrix} \qquad X^T = \begin{matrix} 1 & 1 & 1 & 1 & 1 \\ 5 & 5 & 6 & 7 & 7 \\ 20 & 35 & 38 & 40 & 46 \\ 6 & 6 & 8 & 8 & 10 \end{matrix}$$

Normal Equation المعادلة العمودية

$$X^T X = \begin{matrix} 5 & 30 & 179 & 38 \\ 30 & 184 & 1105 & 234 \\ 179 & 1105 & 6785 & 1414 \\ 38 & 234 & 1414 & 300 \end{matrix}$$

$$(X^T X)^{-1} = \begin{matrix} 10.4 & -2.6 & 0.05 & 0.4 \\ -2.6 & 1.3 & -0.02 & -0.5 \\ 0.05 & -0.02 & 0.008 & -0.02 \\ 0.46 & -0.5 & -0.02 & 0.54 \end{matrix}$$

Norma

$$(X^T X)^{-1} \; X^T$$

$$
\begin{matrix}
10.4 & -2.6 & 0.05 & 0.4 \\
-2.6 & 1.3 & -0.02 & -0.5 \\
0.05 & -0.02 & 0.008 & -0.02 \\
0.46 & -0.5 & -0.02 & 0.54
\end{matrix}
\qquad
\begin{matrix}
1 & 1 & 1 & 1 & 1 \\
5 & 5 & 6 & 7 & 7 \\
20 & 35 & 38 & 40 & 46 \\
6 & 6 & 8 & 8 & 10
\end{matrix}
$$

Normal Equation المعادلة العمودية

X_1	X_2	X_3	Y
العمر	القدرة	الاسطوانات	السعر
5	20	6	114
5	35	6	120
6	38	8	123
7	40	8	121
7	46	10	135

X_1	X_2	X_3
1	1	1
5	5	6
20	35	38
6	6	8

X_4	X_5
1	1
7	7
40	46
8	10

Y
114
120
123
121
135

Norma

$$(X^T X)^{-1} X^T Y$$

$$\begin{matrix} 0.8 & 1.55 & -0.1 & -2.6 & -1.5 \\ 0.5 & 0.2 & 0.4 & 1.7 & 0.58 \\ -0.01 & 0.11 & 0.074 & 0.07 & 0.078 \\ 0.8 & 0.5 & 1.02 & 0.48 & 1.44 \end{matrix}$$

$$\begin{matrix} 114 \\ 120 \\ 123 \\ 121 \\ 135 \end{matrix}$$

4×5 5×1

4×1

Normal Equation المعادلة العمودية

$(X^T X)^{-1} \; X^T \; Y$

0.8	1.55	-0.1	-2.6	-1.5	114	-252.2
0.5	0.2	0.4	1.7	0.58	120	414.2
-0.01	0.11	0.074	0.07	0.078	123	40.162
0.8	0.5	1.02	0.48	1.44	121	529.14
					135	

المعادلة العمودية Normal Equation

أيهما أفضل : الـ **Gradient Descent** ولا الـ **Normal Equation** :

- الـ **Normal Equation** ميزتها ان مش محتاج تحسب قيمة الفا , و مش هتعمل خطوات كتير , هي خطوة واحدة
- بس عيبها انها بتكون صعبة و بطيئة جدا مع عدد كبير للـ features اللي هي n لان الماتركس هتكون مخيفة خاصة لعمل الـ inverse , فلو عدد الـ ى يقل عن 10 الاف خليك في الـ **Normal Equation** ,لو زادت يبقي لازم الـ **Gradient Descent**
- كمان فيه خوارزميات (زي linear regression) مش هينفع تشتغل الا بالـ **Gradient Descent** , و خوارزميات تانية ممكن الـ **Normal Equation**, فلازم تكون عارف الاتنين و تشوف مين مناسب لايه

المعادلة العمودية Normal Equation

أحيانا بتحصل مشكلة في نوع النورمال , ان مصفوفة $X^T X$ في اكس تكون singular و معناها ان مش هينفع يتعمل لها inverse , وده هيعمل مشكلة

غالبا بيكون سبب انها singular حاجة من اتنين

- ان عدد الـ m (عدد الصفوف) اقل من عدد الـ n (العواميد او المعلومات عن كل بيت) خاصة لو الفرق كبير , فا اما تمسح شوية عواميد مش مهمة , او تزود بيانات و صفوف , او تشوف نوع تاني
- اما يكون فيه عمودين معتمدين علي بعض , يعني فيه مثلا مساحة البيت بالمتر المربع , ومساحة البيت بالقدم المربع ,وده معناه ان فيه عمود كامل يساوي عمود تاني مضروب في فاكتور , وده هيؤدي ان قيمة الماتركس كلها تساوي صفر , فالـ inverse مش هيظهر