

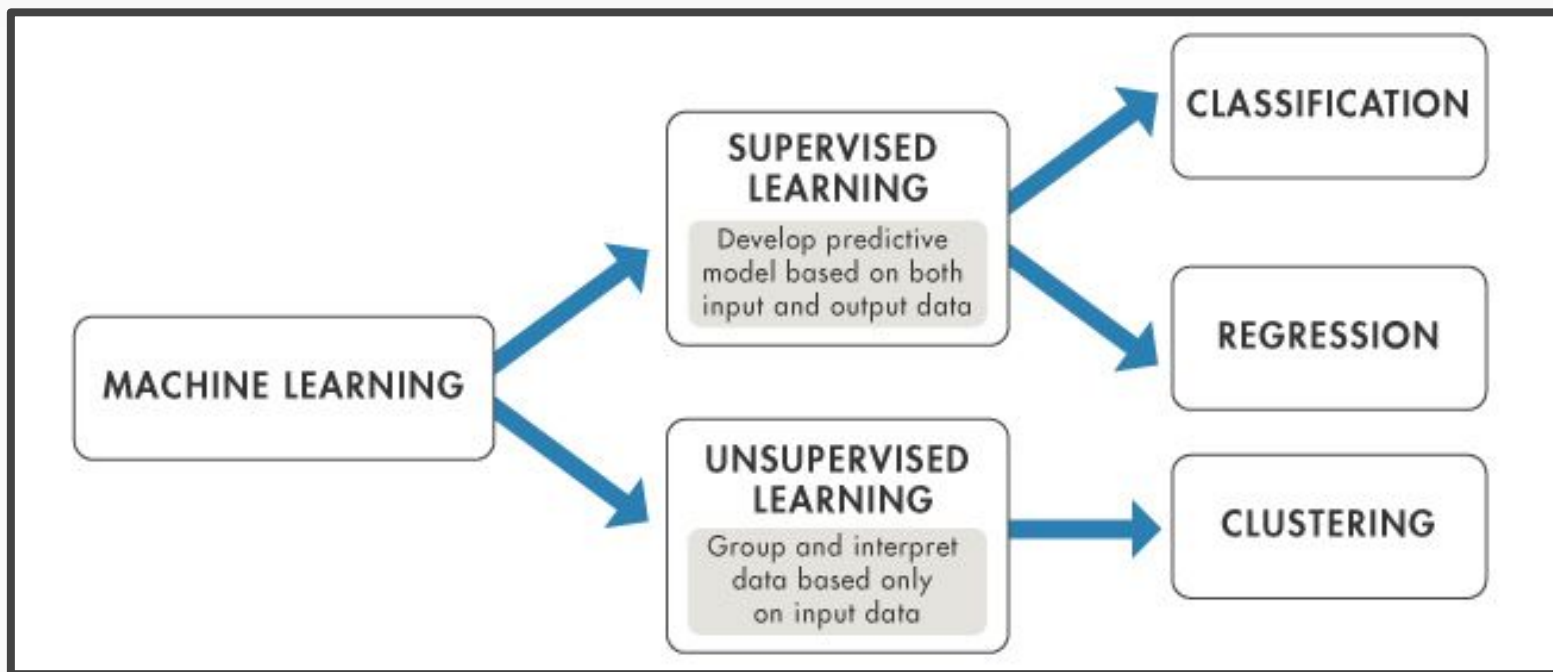


فن تعليم الآلة

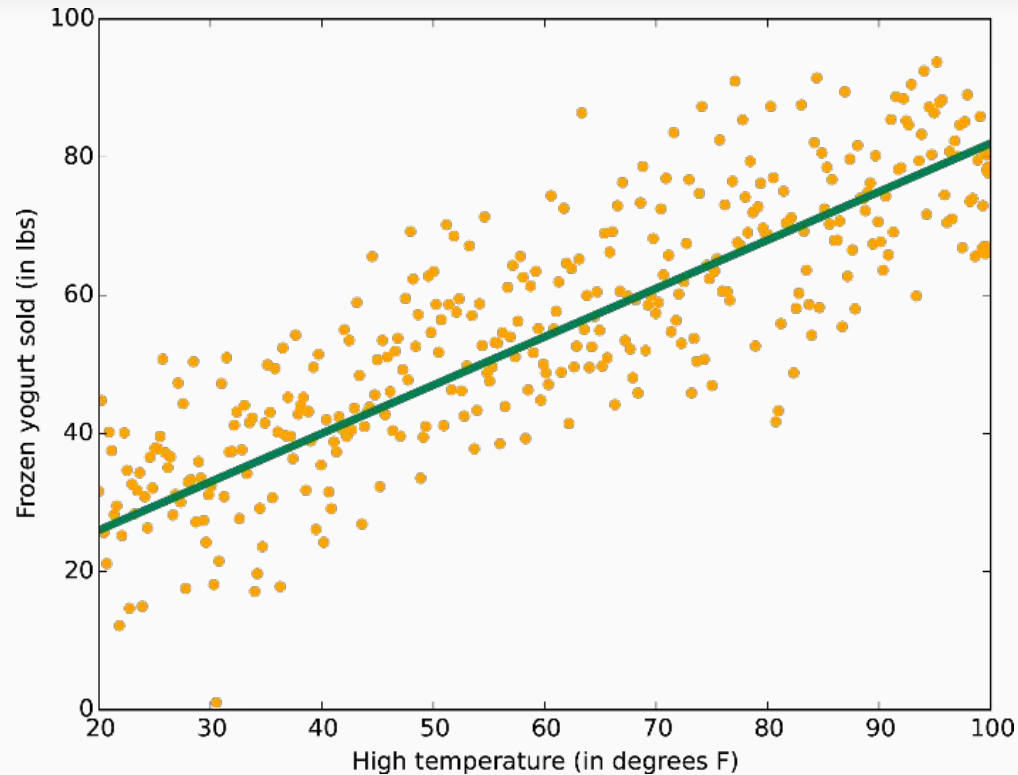
القسم الثاني : التوقع

ختام التوقع

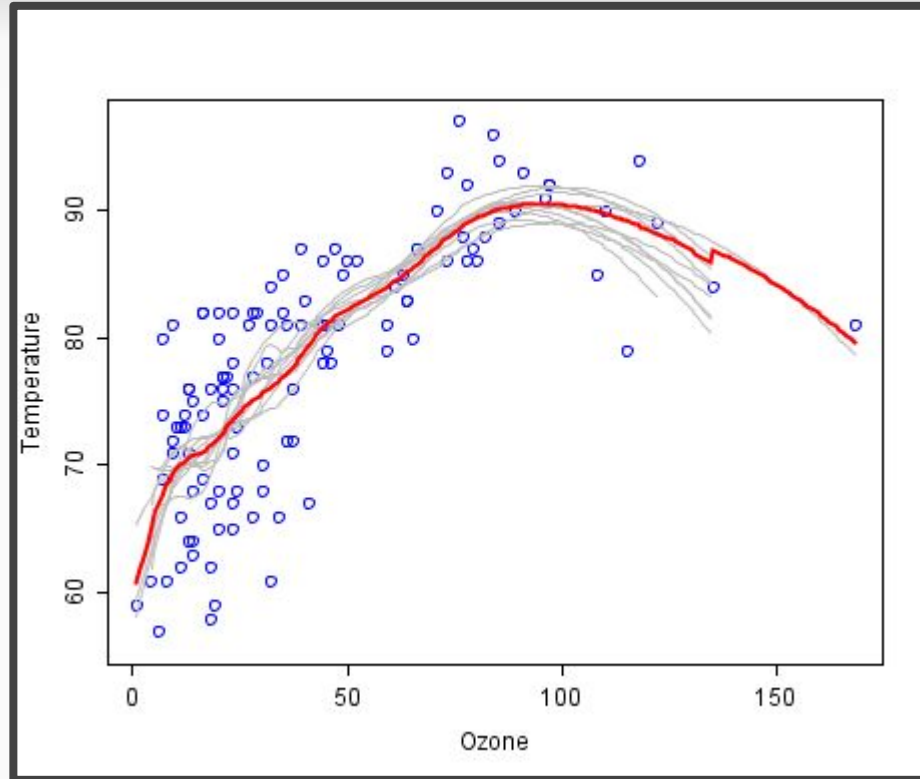
التعليم باشراف و بدون اشراف



Regression التوقع



Regression التوقع



التوقع Regression

تطبيقات التوقع :

- أسعار البيوت
- أسعار الاسهم في البورصة
- حالة الطقس
- المبلغ الذي سيشتري به العميل

Linear Regression التوقع الخطي

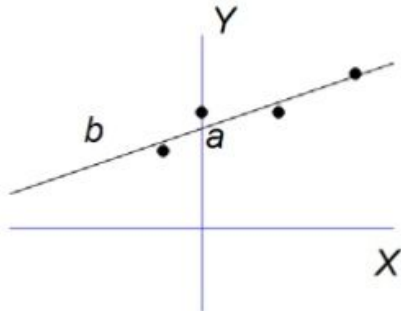
Linear regression equation
(without error)

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

predicted
values of Y

b = slope = rate of
predicted \uparrow/\downarrow for Y
scores for each unit
increase in X

Y-intercept =
level of Y
when X is 0



- ويسمي أيضا (One Variable Regression) أو (Univariate Regression)

التوقع الخطي Linear Regression

| | |
|----------------|-----------------|
| Input X | المدخلات |
| Output Y | المخرجات |
| Rows m | الصفوف |
| Features n | العناصر |
| $h(x)$ | القيمة المتوقعة |
| Cost J | قيمة الخطأ |
| Theta Θ | 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: $J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$

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 θ_0 و θ_1 والتي تجعل من J (نسبة الخطا) اقل ما يمكن
- تسمى 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 |

Linear Regression Equation معادلة التوقع الخطي

Theta0 = 5 , theta 1 = 2

Equation $h(x) = 5 + 2x$

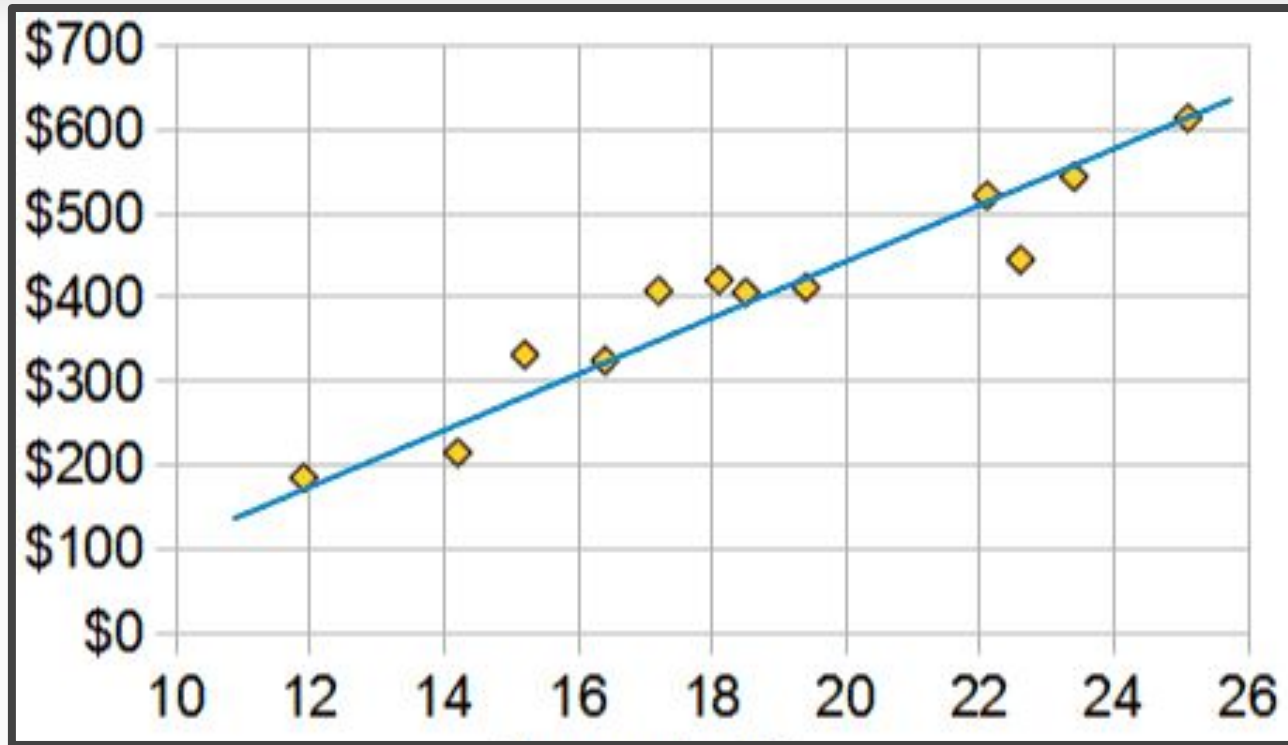
| X | Y | h(x) | h(x) - y | (h(x) - y) ² |
|---|----|------|----------|-------------------------|
| 1 | 7 | 7 | 0 | 0 |
| 2 | 8 | 9 | 1 | 1 |
| 2 | 7 | 9 | 2 | 4 |
| 3 | 9 | 11 | 2 | 4 |
| 4 | 11 | 13 | 2 | 4 |
| 5 | 10 | 15 | 5 | 25 |
| 5 | 12 | 15 | 3 | 9 |

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

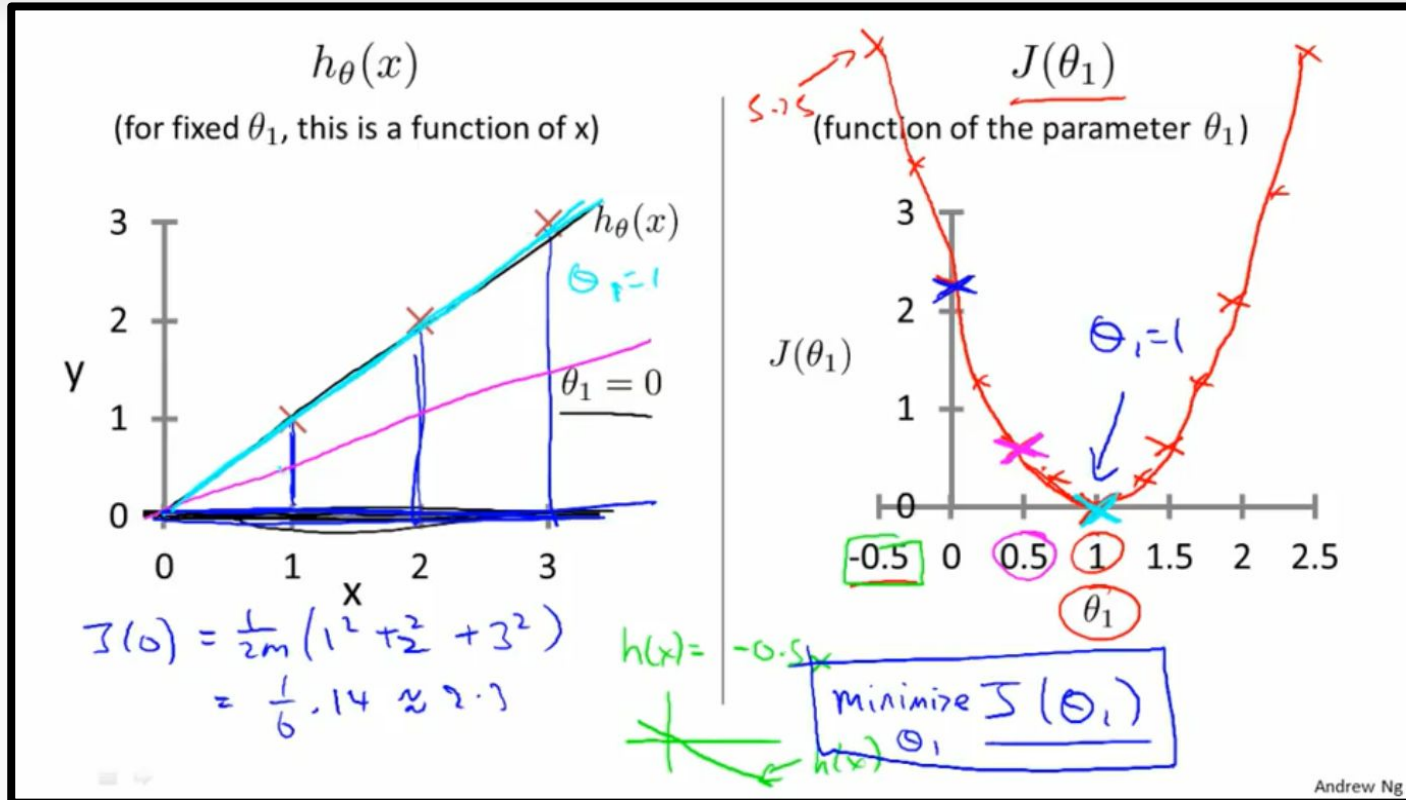
$$J = 1 / 14 (0+1+4+4+4+25+9)$$

$$J = 47/14 = 3.3$$

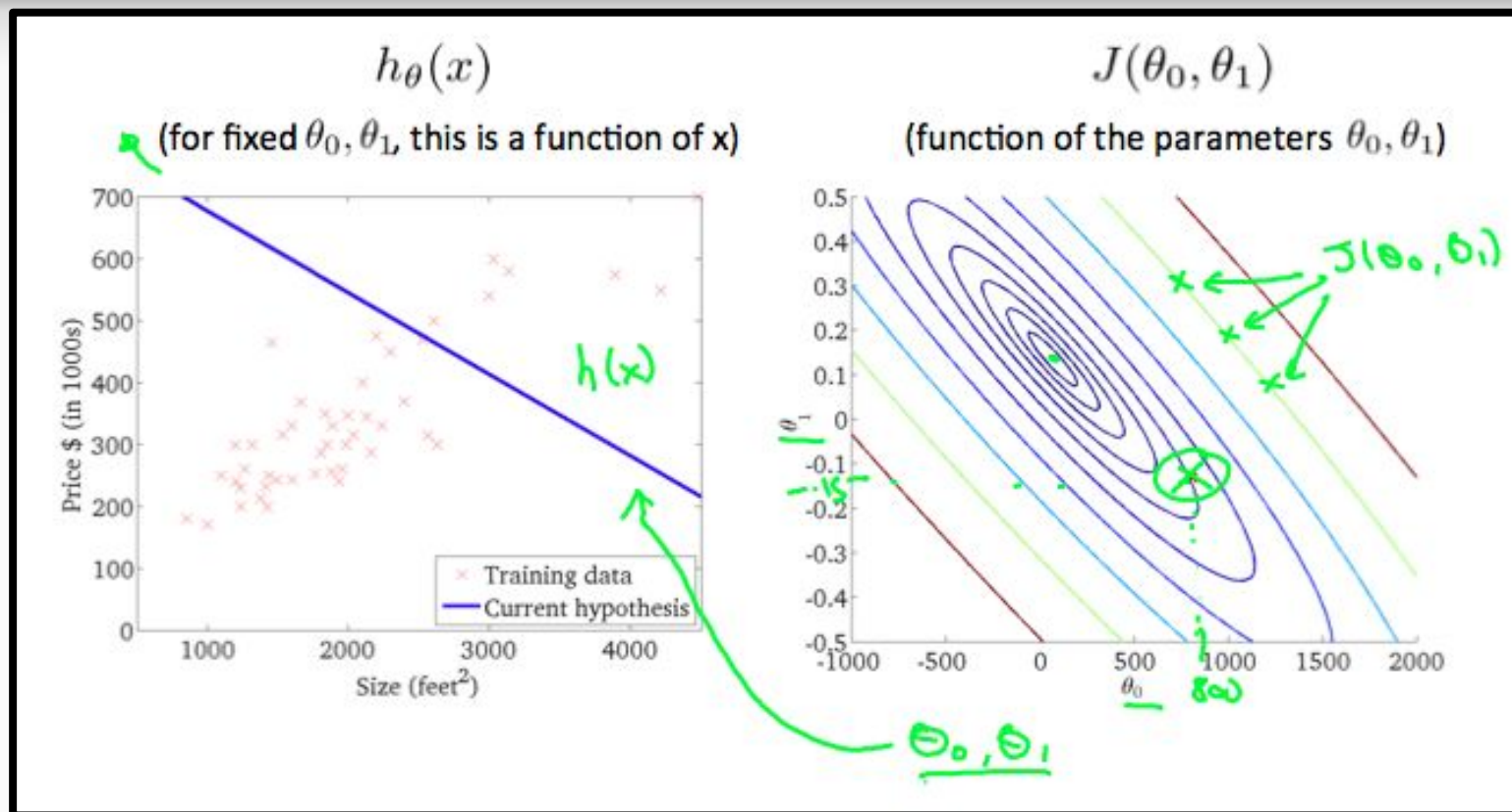
الخط الأكثر ملائمة Best fit line



تحديد قيمة θ_1

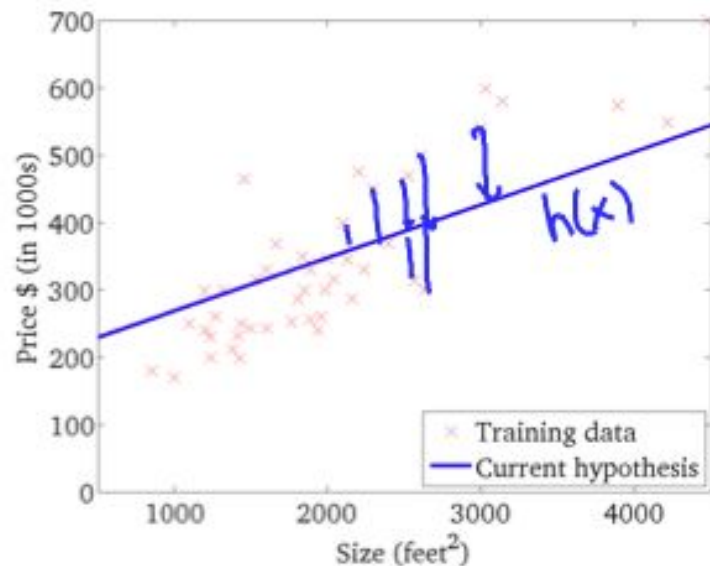


تحديد قيمة ثيتا



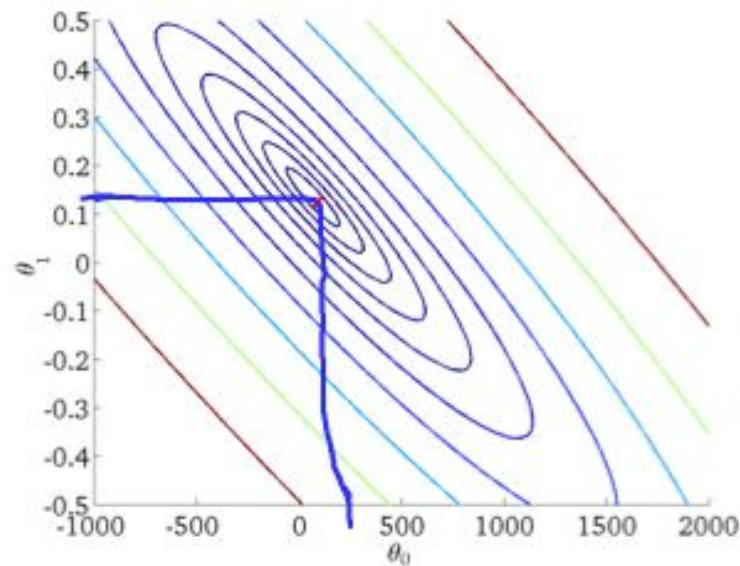
$$h_{\theta}(x)$$

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

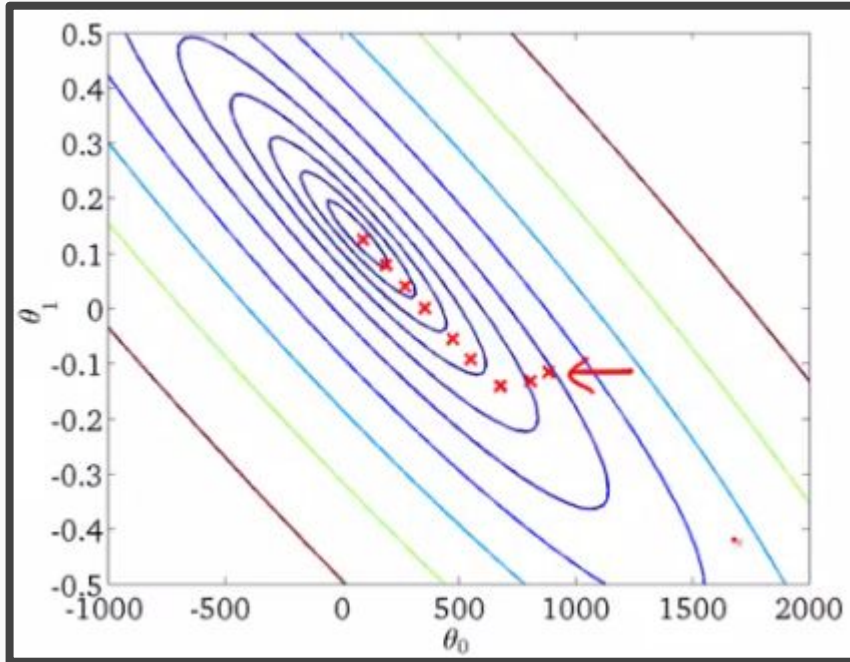


$$J(\theta_0, \theta_1)$$

(function of the parameters θ_0, θ_1)



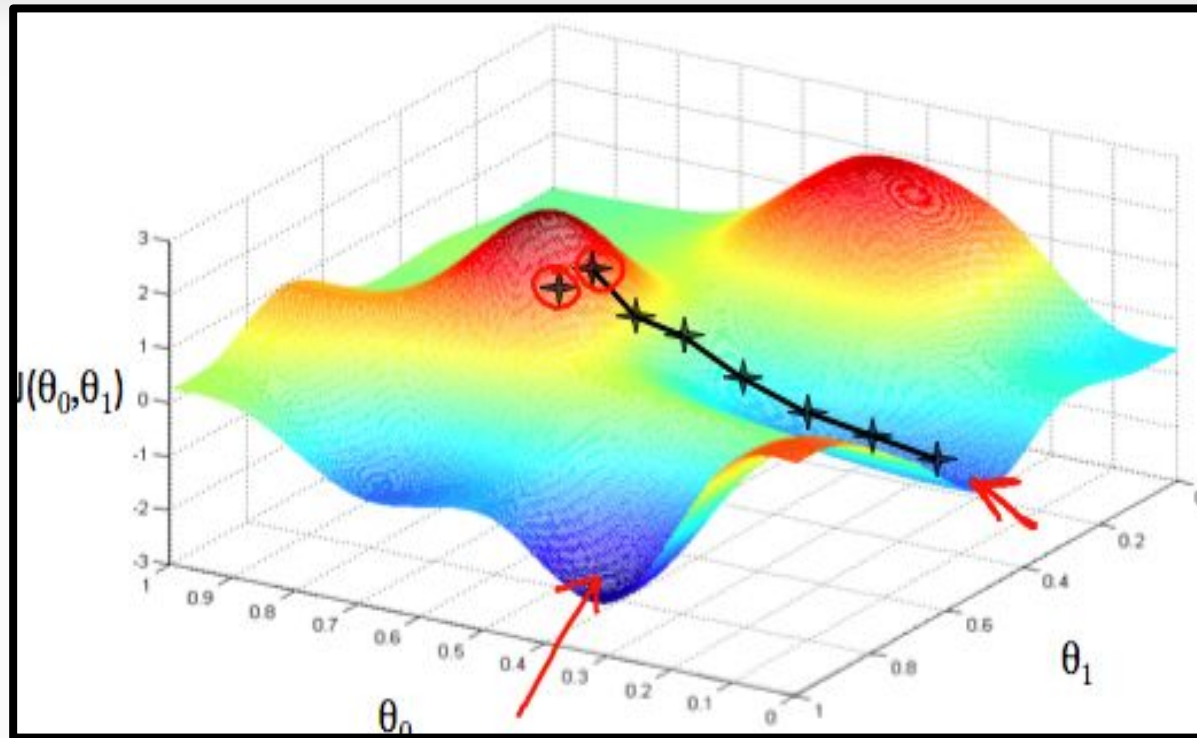
الإنحدار التدريجي Gradient Descent



الانحدار التدريجي :

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

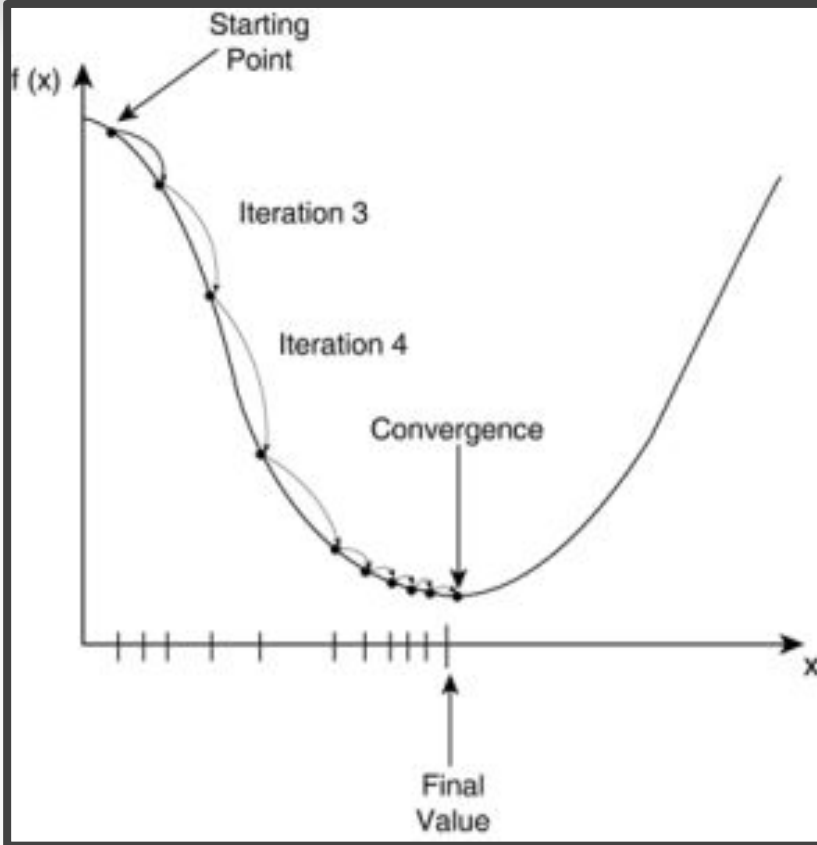
الإنحدار التدريجي Gradient Descent



القيم المحلية :

- فهناك ما يسمى 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 global minimum وهي المطلوبة

معادلة الإنحدار التدريجي



الفكرة :

- لاحظ ان في حالة القيمة الدنيا قيمة التفاضل بصفر (لأن وقتها سيكون الخط شبه مستقيم فالميل سيكون 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 سنقل تدريجيا

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

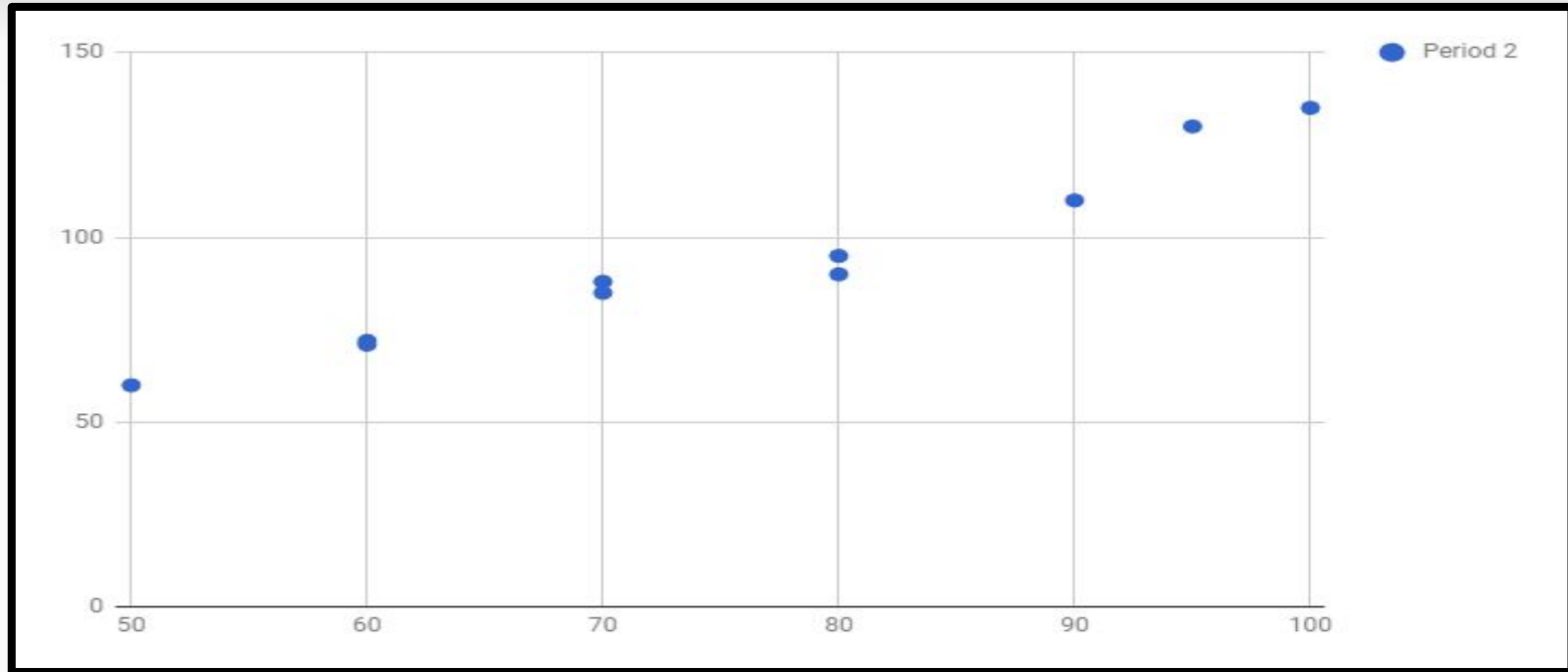
$$\theta_1 := \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^m ((h_{\theta}(x_i) - y_i)x_i)$$

مثال عملي

| السعر (الف \$) Y | مساحة البيت (m^2) X_1 |
|--------------------|-----------------------------|
| 135 | 100 |
| 130 | 95 |
| 110 | 90 |
| 95 | 80 |
| 90 | 80 |
| 85 | 70 |
| 80 | 70 |
| 80 | 60 |

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

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



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 | | |

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

ثبتنا 1 = 0

ثبتنا 3 = 1

المعادلة

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

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 |

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

ثبتنا $1 = 0$

ثبتنا $3 = 1$

المعادلة

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

$$\theta_0 = 1 - ((0.002 / 8) * (28))$$

$$\theta_0 = 1 - 0.007 = 0.993$$

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

$$\theta_0 = 1$$

$$28 = \text{مجموع الفروق}$$

$$0.002 = \alpha$$

$$m = 8 \text{ قيمة}$$

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 |

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

ثبتنا $1 = 0$

ثبتنا $3 = 1$

$$\text{Theta } 1 = 3 - ((0.002 / 8) * (2095))$$

$$\text{Theta } 1 = 3 - 0.52 = 2.48$$

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

$$\text{ثابتنا } 1 = 0$$

$$\text{مجموع الفروق} = 2095$$

$$\text{الفا} = 0.002$$

$$\text{قيمة } m = 8$$

Theta 0 = 1

Theta 1 = 3

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

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

Theta 0 = 1

Theta 1 = 3

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

Theta 0 = 0.993 Theta 1 = 2.48

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

Theta 0 = 1

Theta 1 = 3

Theta 0 = 0.993

Theta 1 = 2.48

Theta 0 = 0.991

Theta 1 = 2.46

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

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

Iteration 1

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

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

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 أكثر من بعد :

- تحدثنا سابقا , عن التعامل مع متغير واحد (قيمة لـ X و نجيب منها قيمة Y) الان نتعامل مع أكثر من متغير
- أكثر من متغير معناها ان البيانات الداخلة لها أكثر 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 تسمى 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 أكثر من بعد :

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

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

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

● 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

● الصيغة المجمعة

repeat until convergence: {

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

}

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

سعر السيارات :

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

| 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 |

Linear Regression with Multivariable 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

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

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

$$x = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n+1}$$

$$\theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{bmatrix} \in \mathbb{R}^{n+1}$$

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

$\downarrow = 1$

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

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

(n+1) x 1 matrix

$\theta^T x$

Multivariate linear regression. \leftarrow

- وقتها الفنكشن ،
هتكون متعددة الحدود زي كدة ،
وهنعمل ماتركس للاكسات ،
وواحدة للثبتات ،
ونضربهم في بعض بعد ما
نعمل ترانزبوس للثبتات

Linear Regression with Multivariables التوقع الخطي لأكثر من متغير

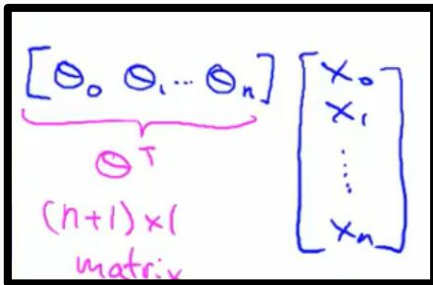
$$h(x) = (\text{Theta})^T X$$

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

$$X_1 = \begin{matrix} 1 \\ 5 \\ 20 \\ 6 \end{matrix}$$

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

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



$[\theta_0 \ \theta_1 \ \dots \ \theta_n]$
 θ^T
(n+1) x 1 matrix

$\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix}$

التوقع الخطي لأكثر من متغير Linear Regression with Multivariables

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

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

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

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

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

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

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 1 = 2.6

Theta 2 = 3.9

Theta 3 = 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

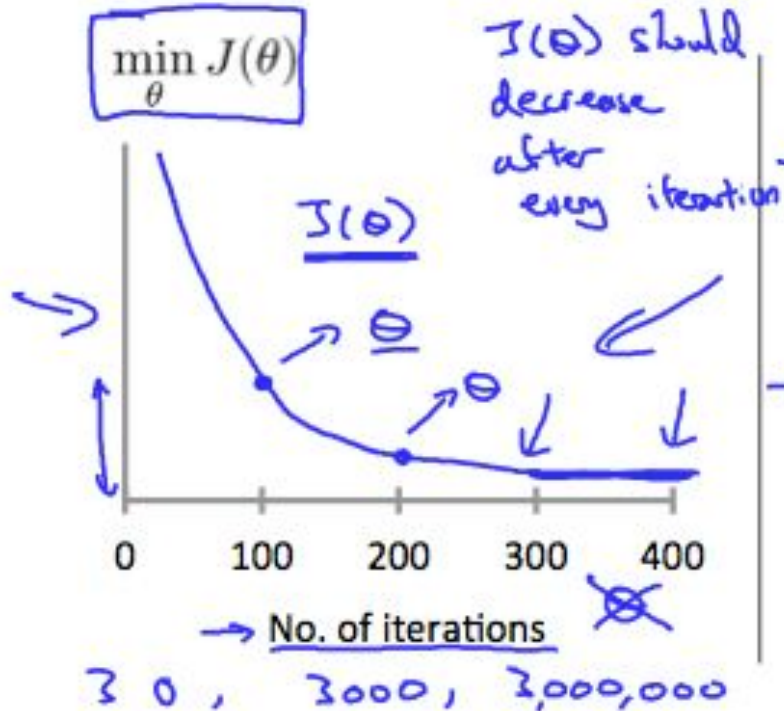
عدد المحاولات Number of Iteration

ما هو العدد المناسب ؟

- من الواضح ان كل ما بنحاول اكثر , قيمة L بتقل و ديه حاجة كويسة
- بس كل ما يزيد عدد المحاولات , كل ما التكلفة و الوقت يزدو . وده عيب كبير
- يبقي نحاول كام مرة ؟ ؟ ؟

عدد المحاولات Number of Iteration

Making sure gradient descent is wo



ما هو العدد المناسب ؟

- الرسمة هنا واضح فيها ان كل ما بنزود عدد المحاولات , كل ما قيمة J هتقل اكثر , بس بعد فترة معينة السلوب بيقترب لصفر , و بيكون فيه عدد ضخم جدا من المحاولات مع فرق بسيط , و هنا لازم نوقف , عشان هيكون ضياع وقت علي الفاضي
- ممكن نوقف بعد 5 او 50 او 5 مليون محاولة , محدش هيقدر يحدد الرقم كام , كل حالة بحالتها

قيمة الفا ؟ ؟

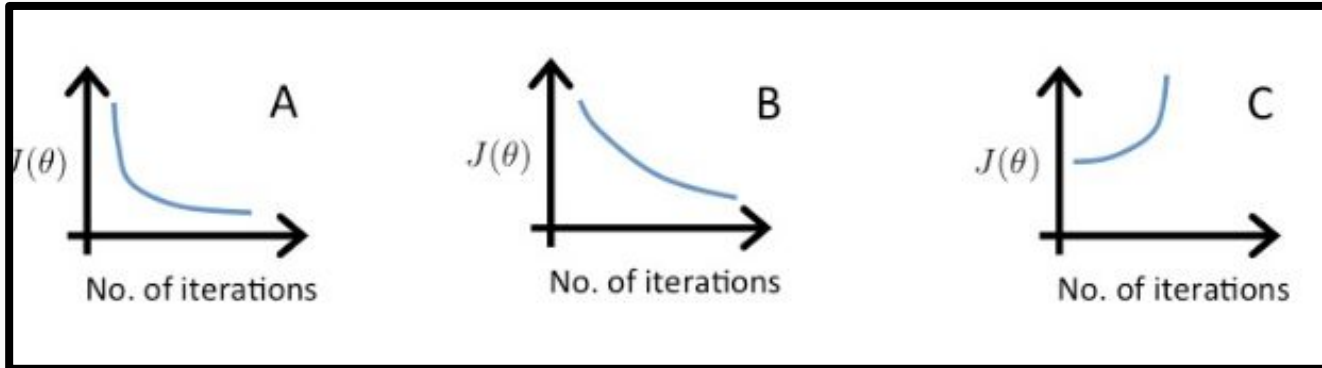
repeat until convergence: {

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

- اختلافها يغير من سرعة التعامل , ودقته

قيمة ألفا ؟ ؟

- لو زادت قيمة ألفا هجري بسرعة , بس ممكن اقع في مشكلة اني ازود قيمة الـ L , ولو مشيت ببطئ , هيكون دقيق بس ببطئ جدا , فلازم اختار قيم مضبوطة

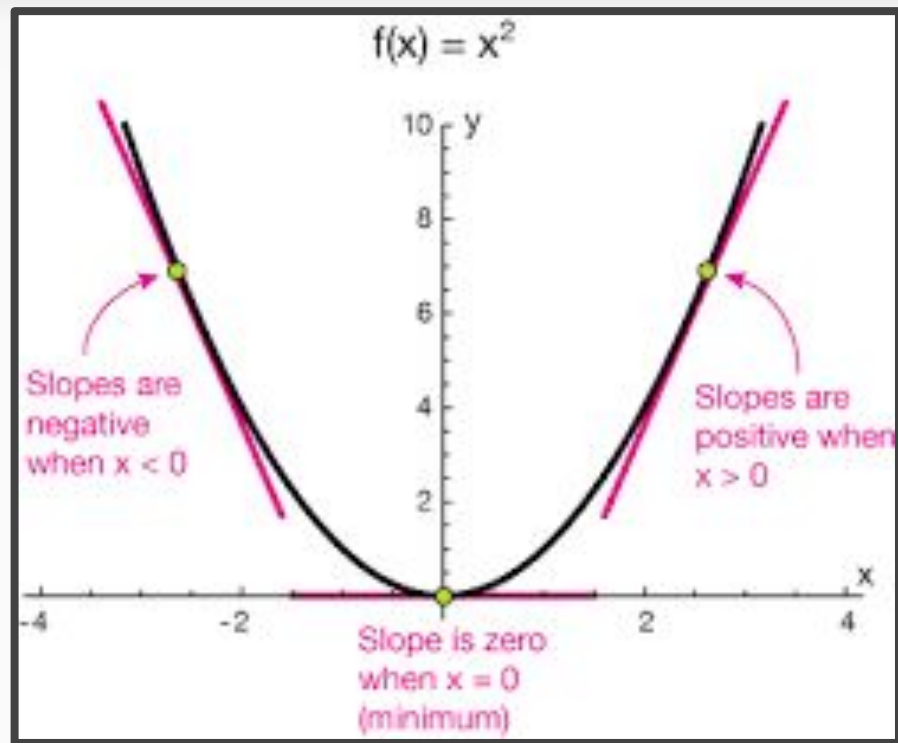


قيمة ألفا ؟ ؟

- اختار قيمة صغيرة , واضرب في 3
- من الممكن اختيار قيمة وسط قيمتين

| | | | | | | | | | |
|-------|-------|------|------|----|----|---|---|----|----|
| 0.001 | 0.003 | 0.01 | 0.03 | .1 | .3 | 1 | 3 | 10 | 30 |
|-------|-------|------|------|----|----|---|---|----|----|

المعادلة العمودية Normal Equation



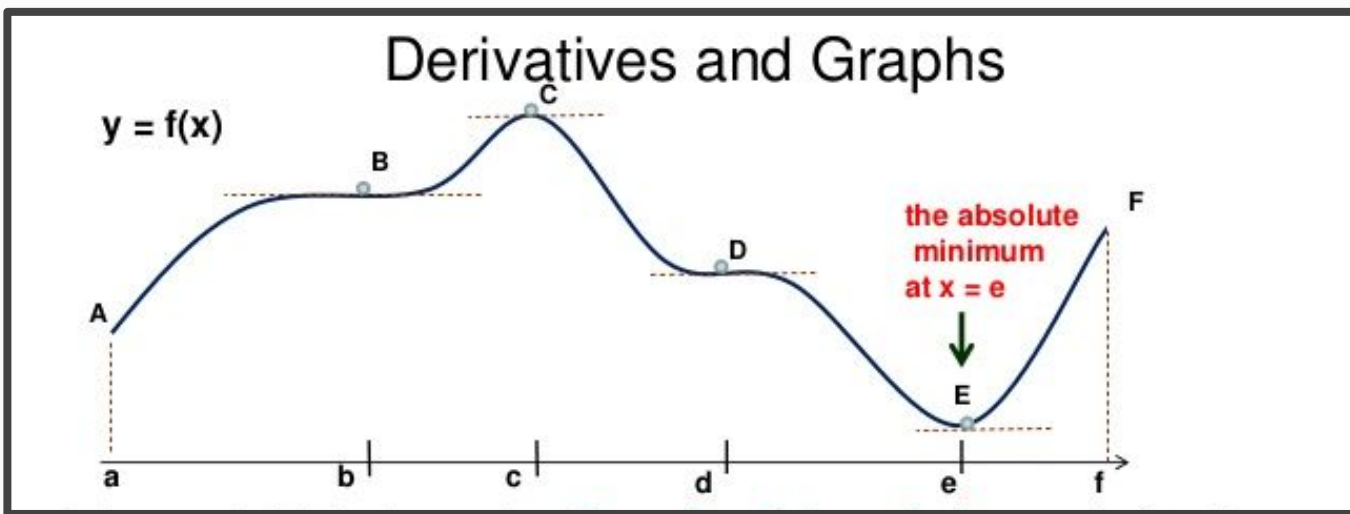
مفهوم القيم العظمي و الصغري

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

المعادلة العمودية Normal Equation

مفهوم القيم العظمي و الصغري

- تتواجد عندما يكون التفاضل يساوي صفر



المعادلة العمودية Normal Equation

طريقة الـ Normal Equation

- وهي عن طريق الاعتماد علي تفاضل الـ J 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

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

المعادلة العمودية Normal Equation

| 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 | Y |
| 6 | 6 | 8 | 114 |
| | | | 120 |
| X_4 | X_5 | | 123 |
| | | | 121 |
| 1 | 1 | | 135 |
| 7 | 7 | | |
| 40 | 46 | | |
| 8 | 10 | | |

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 كثير , 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 n يقل عن 10 الاف خليك في الـ Normal Equation , لو زادت يبقى لازم الـ Gradient Descent
- كمان فيه خوارزميات (زي linear regression) مش هينفع تشتغل الا بالـ Gradient Descent , و خوارزميات تانية ممكن الـ Normal Equation , ف لازم 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 ثاني
- 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 ثاني مضروب في فاكتور , وده هيوذي ان قيمة الماتركس كلها تساوي صفر , فالـ inverse مش هيعمل



محتويات الكورس :

- القسم الأول : مقدمة
- القسم الثاني : التوقع Regression
- القسم الثالث : برنامج بايثون 
- القسم الرابع : التقسيم Classification
- القسم الخامس : الشبكات العصبية NN
- القسم السادس : نظام الدعم الالي SVM
- القسم السابع : التعلم بدون اشراف Unsupervised ML
- القسم الثامن : مواضيع هامة (القيم الشاذة , نظام الترشيحات ...)