



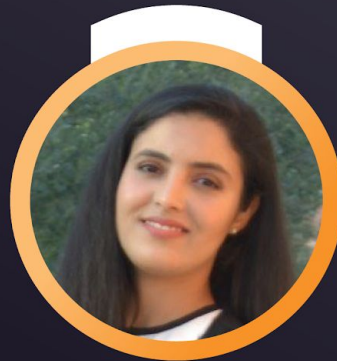
أكاديمية
AL FIHRIYA
الفهرية
ACADEMY

حصص في التعلم الآلي

Linear Regression - Logistic Regression
Regularization - Neural Networks
Clustering - Dimensionality Reduction



رضوان



سكينة



أحمد

فهرس الحصة 2

1. مقدمة

a. شناهو الانحدار الخطي Linear Regression ؟

2. الانحدار الخطي البسيط Simple Linear Regression

a. النموذج الرياضي Model

b. دالة الخطأ Error Function

3. الانحدار الخطي المتعدد Multiple Linear Regression

4. خوارزمية أصل التدرج Gradient Descent

5. أمثلة باستعمال Python و Colab



لغز خفيف ظريف

X	Y
3	6
7	14
9.5	19

- تخيل معايا عطيتك هاد الأزواج
ديال البيانات X و Y
- ايلا قلت ليك را كايئة علاقة
بيناتهم و سولتك شنو قيمة Y
ايلا كان X هو 5 ؟



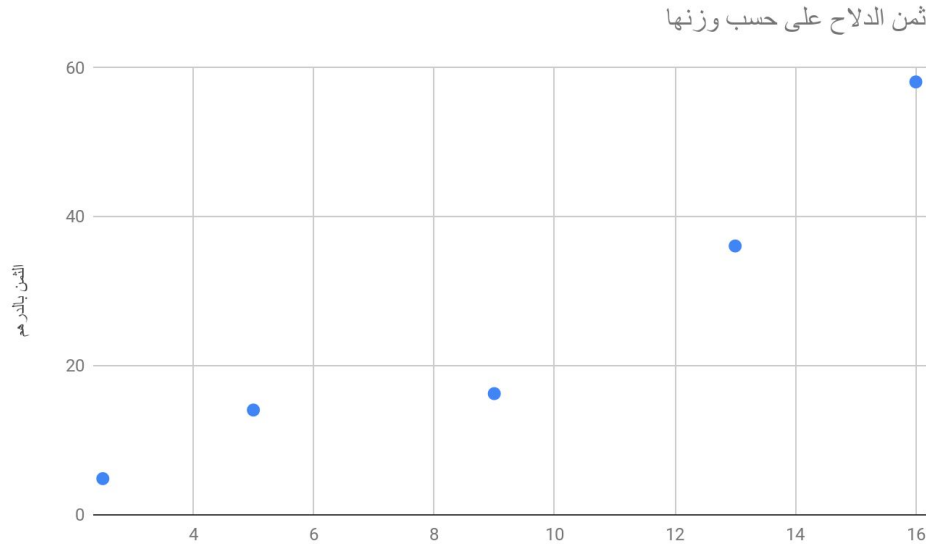
لغز خفيف ظريف

X	Y
3	6
7	14
9.5	19
5	10



- اسهل جواب هو $Y=10$ حيث النموذج الرياضي $Y = 2X$ صالح لذوك البيانات
- التعلم الآلي مبني على هاذ القضية اللي درتي دابا: الخوارزمية كاتشوف البيانات و كاتقلب على النموذج الرياضي الأحسن اللي كايقدر يكون صالح للبيانات

الحياة صعبة



- البيانات فالحياة اليومية ماشي ديما محطوطة على مستقيم !



فهرس الحصة 2

1. مقدمة

a. شناهو الانحدار الخطي Linear Regression ؟

2. الانحدار الخطي البسيط Simple Linear Regression

a. النموذج الرياضي Model

b. دالة الخطأ Error Function

3. الانحدار الخطي المتعدد Multiple Linear Regression

4. خوارزمية أصل التدرج Gradient Descent

5. أمثلة باستعمال Python و Colab



النموذج الرياضي

- أول حاجة كايبدأو بيها الناس اللي كايغيو يدخلو للدومين ديال التعلم الآلي: **الانحدار الخطي Linear Regression**
- النموذج اللي كانفترضو انه مزيان للبيانات كايكون على شكل دالة تألفية

$$Y = \alpha X + \beta \quad (1)$$

- α هو ال **Slope**
- β هو ال **Intercept**



النموذج الرياضي

- X و Y يقدر و يكون متجهات Vectors عامرين بعدة أعداد. و لكن هنا عانبقاو فأسهل مثال اللي هو ملي كايكون البعد Dimension ديا لهم هو 1، يعني فيهم قيمة وحيدة اللي

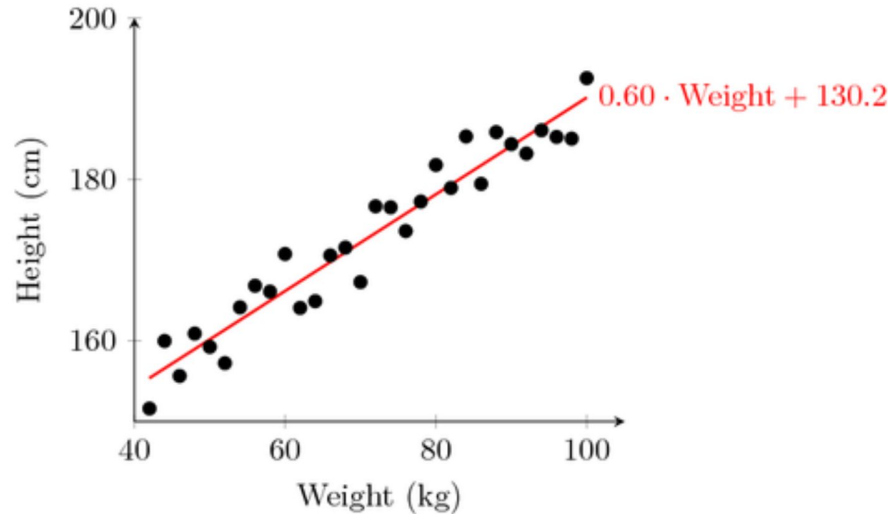
كاتغير فالبيانات: **Simple Linear Regression**

- فهاذ الحالة α و β عددان حقيقيان مستقلان عن X
- ايلا رسمنا النموذج اللي دونا عليه سابقا ، من الواضح انه عا يكون مبيان هذه الدالة مستقيما **معامله الموجه** هو α و β هو **أرتوبه عند الصفر**



Simple Linear Regression الانحدار الخطي البسيط

مثال



هذا مثال ديال بيانات تاع
الوزن و الطول ديال 30
واحد، ايلا رسمناهم فمبيان
و لقينا احسن انحدار خطي،
عانلقاو ان $\alpha = 0.6$ و
 $\beta = 130.2$ و الدالة التآلفية
كاتبان قرية بزاف من
البيانات و كاتأكد أن
النموذج مايبهش



الانحدار الخطي البسيط Simple Linear Regression

كيفاش درنا تا لقينا هاذ الانحدار الخطي البسيط؟

- باش تلقى الانحدار الخطي البسيط اللي يجي مع البيانات ديالك، خاصك تلقى α و β

- ايلا شفتي المثال السابق عا تلاحظ ان الدالة ماكادوزش ديريكث من كُاع البيانات، كايين واحد **التياساع** بين البيانات و الدالة، وهاذ **التياساع** هو **الخطأ Error** اللي دارتو الدالة فذيك القيمة. لنفترض أن فكل نقطة i درنا خطأ ϵ_i يعني:

$$Y_i = \alpha X_i + \beta + \epsilon_i$$

- ايلا بغينا الانحدار الخطي البسيط ديانا يكون مزيان، خاص المجموع ديال الأخطاء **فكُاع البيانات** يكون صغير.



Simple Linear Regression الانحدار الخطي البسيط

كيفاش درنا تا لقينا هاذ الانحدار الخطي البسيط؟

- باستعمال واحد الطريقة سميتها **طريقة المربعات الصغرى** أو **الدنيا** Least Squares عا يكون الهدف هو نلقاو $\hat{\alpha}$ و $\hat{\beta}$ باش نصغرو الجمع ديال المربعات تاع الأخطاء

$$\hat{\alpha}, \hat{\beta} = \min_{\alpha, \beta} \sum_{i=1}^n \hat{\varepsilon}_i^2 = \min_{\alpha, \beta} \sum_{i=1}^n (Y_i - \beta - \alpha X_i)^2$$



Simple Linear Regression الانحدار الخطي البسيط

تمرين: استعمل الاشتقاق باش تلقى الحل

$$\hat{\alpha} = \frac{\sum_{i=1}^n (X_i - \bar{X}) (Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad (2)$$

$$\hat{\beta} = \bar{Y} - \hat{\alpha} \bar{X} \quad (3)$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i \quad (4)$$



فهرس الحصة 2

1. مقدمة

a. شناهو الانحدار الخطي Linear Regression ؟

2. الانحدار الخطي البسيط Simple Linear Regression

a. النموذج الرياضي Model

b. دالة الخطأ Error Function

3. الانحدار الخطي المتعدد Multiple Linear Regression

4. خوارزمية أصل التدرج Gradient Descent

5. أمثلة باستعمال Python و Colab



الانحدار الخطي المتعدد Multiple Linear Regression

مثال

ثمن بيع منزل لا يتعلق بمتغير وحيد...



1112 NW 26TH ST,
CORVALLIS, OR



latitude longitude

44.57808 N 123.2803 W

-123.2803, 44.57808

	lon	lat	price	bedrooms	full_baths
1	-123.2803	44.57808	267500	5	2
2	-123.2330	44.59718	255000	3	2
3	-123.2635	44.56923	295000 ...	3	2 ...
4	-123.2599	44.59453	5000	0	1
5	-123.2632	44.53606	13950	0	2
6	-123.2847	44.59877	233000	3	2

فهاذ الحالة كانوليو فنموذج رياضي ديال **انحدار خطي متعدد**

$$Y_i = \alpha_1 X_{i1} + \alpha_2 X_{i2} + \alpha_3 X_{i3} + \alpha_4 X_{i4} + \dots + \beta + \epsilon_i$$



الانحدار الخطي المتعدد Multiple Linear Regression

كتابة المسألة بطريقة المصفوفات

نقدرو نخطو المعادلات على شكل نظمة

$$\begin{cases} y_1 = \beta + \alpha_1 x_{1,1} + \dots + \alpha_p x_{1,p} + \varepsilon_1 \\ y_2 = \beta + \alpha_1 x_{2,1} + \dots + \alpha_p x_{2,p} + \varepsilon_2 \\ \dots \\ y_n = \beta + \alpha_1 x_{n,1} + \dots + \alpha_p x_{n,p} + \varepsilon_n \end{cases}$$

كتابة بطريقة المصفوفات

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{1,1} & \dots & x_{1,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n,1} & \dots & x_{n,p} \end{pmatrix} \begin{pmatrix} \beta \\ \alpha_1 \\ \vdots \\ \alpha_p \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}$$



الانحدار الخطي المتعدد Multiple Linear Regression

كيفاش تلقى أحسن انحدار خطي متعدد؟

نعاودو ثاني نكتبو طريقة المربعات الصغرى

$$\min \sum_{i=1}^n \hat{\epsilon}_i^2 = \min \hat{\epsilon}^T \hat{\epsilon} = \min_{\hat{\alpha}} [(\mathbf{Y} - \mathbf{X}\hat{\alpha})^T (\mathbf{Y} - \mathbf{X}\hat{\alpha})] \quad (1)$$

$$\min \sum_{i=1}^n \hat{\epsilon}_i^2 = \min_{\hat{\beta}, \hat{\alpha}_1, \dots, \hat{\alpha}_p} \sum_{i=1}^n \left(y_i - \hat{\beta} - \hat{\alpha}_1 x_{i,1} - \dots - \hat{\alpha}_p x_{i,p} \right)^2 \quad (2)$$

$$\hat{\alpha} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad (3)$$



الانحدار الخطي المتعدد Multiple Linear Regression

المعادلات الطبيعية Normal Equations

كأين بعض الملاحظات على المُقدِّر Estimator اللّي لقينا

$$\hat{\alpha} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$$

- المصفوفة مصفوفة $\mathbf{X}^T \mathbf{X}$ مربعة
- المصفوفة $\mathbf{X}^T \mathbf{X}$ ماشي ديمّا عندها Inverse، أحياناً حنا مجبرين نقلبو على Pseudo-Inverse
- حساب المقلوب ديال المصفوفة $\mathbf{X}^T \mathbf{X}$ يقدر يولي معقد جداً حسابياً إيلا عندك عدد كبير جداً ديال البيانات (ملايين أو أكثر...)

كأينة واحد الطريقة أخرى باش تلقّا الحل بلا من المعادلات الطبيعية،
هاذ الطريقة سميتها **خوارزمية أصل التدرج Gradient Descent**



فهرس الحصة 2

1. مقدمة

a. شناهو الانحدار الخطي Linear Regression ؟

2. الانحدار الخطي البسيط Simple Linear Regression

a. النموذج الرياضي Model

b. دالة الخطأ Error Function

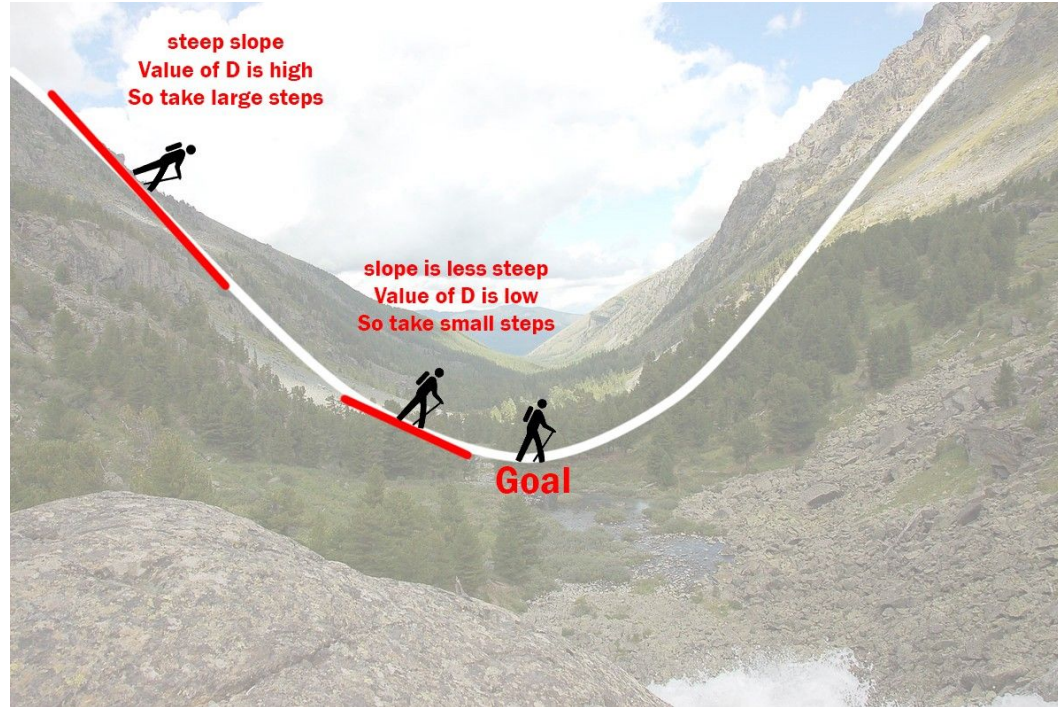
3. الانحدار الخطي المتعدد Multiple Linear Regression

4. خوارزمية أصل التدرج Gradient Descent

5. أمثلة باستعمال Python و Colab



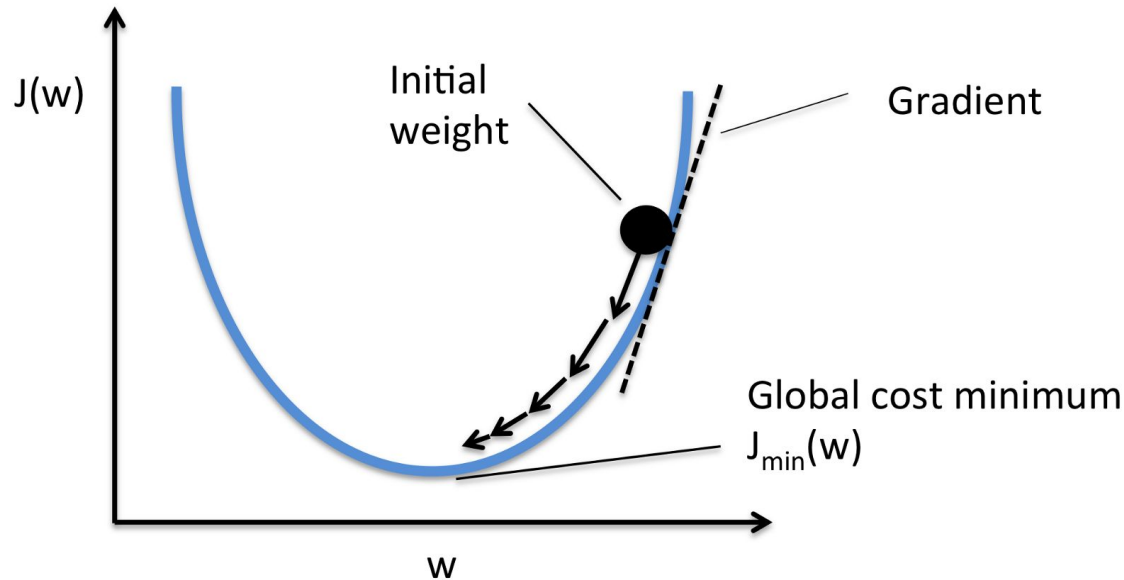
الفكرة العامة



خوارزمية تحسين تكرارية Iterative Optimization Algorithm



الفكرة العامة



$$\mathbf{w}_{k+1} = \mathbf{w}_k - \gamma \nabla J(\mathbf{w}_k)$$

خوارزمية تحسين تكرارية
Optimization Algorithm



أصل التدرج للانحدار الخطي البسيط

$$\hat{\alpha}, \hat{\beta} = \min_{\alpha, \beta} \sum_{i=1}^n \hat{\varepsilon}_i^2 = \min_{\alpha, \beta} \sum_{i=1}^n (Y_i - \beta - \alpha X_i)^2$$

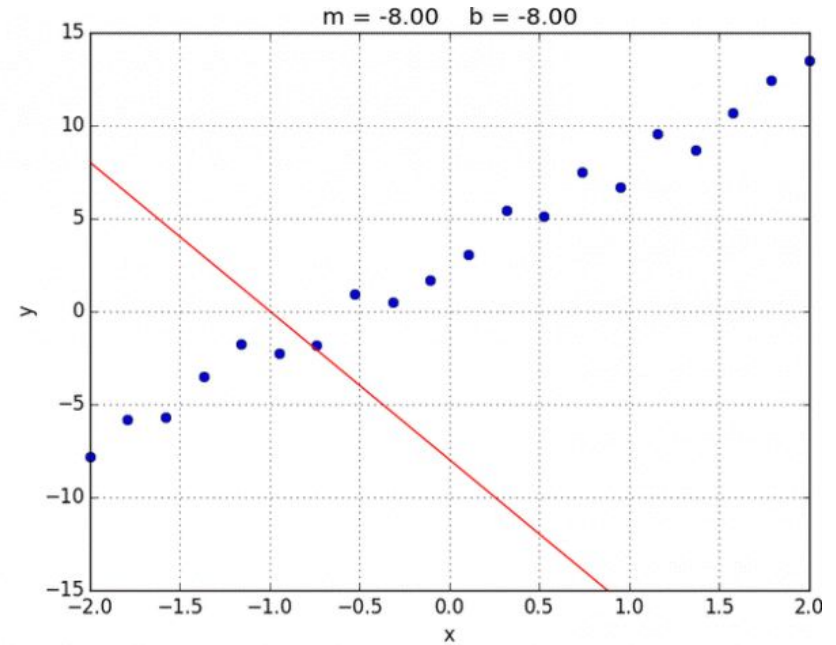
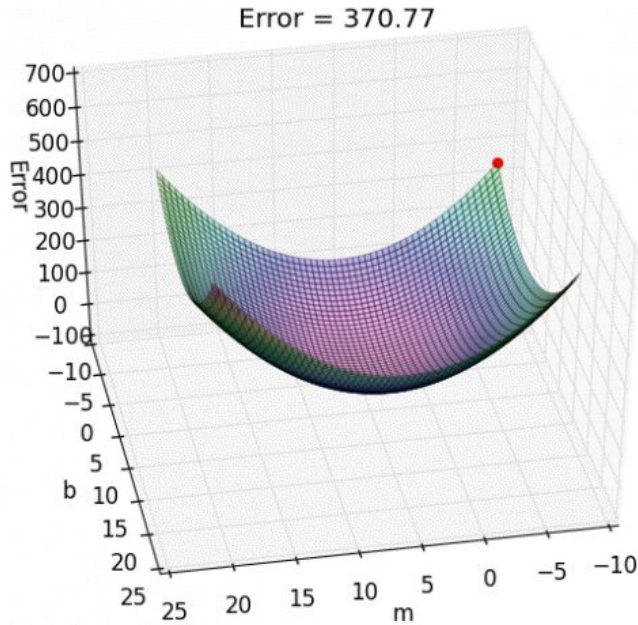
- **الخطوة 1:** بدا من واحد البداية Initialization ديال α و β (مثلا بجوج صفر)
- **الخطوة 2:** عاود واحد K مرة هاد "النزول" باستعمال **معدل التعلم** γ Learning rate

$$\begin{aligned}\beta_k &= \beta_{k-1} - \gamma \frac{1}{n} \sum_{i=1}^n (\beta_{k-1} + \alpha_{k-1} X_i - Y_i) \\ \alpha_k &= \alpha_{k-1} - \gamma \frac{1}{n} \sum_{i=1}^n (\beta_{k-1} + \alpha_{k-1} X_i - Y_i) \cdot X_i\end{aligned}$$



خوارزمية أصل التدرج Gradient Descent

صورة متحركة





ندوزو لشوية دلكود

