Université de M'Sila
Faculté des mathématiques et de l'informatique
Département d’informatique
2025

Linear regression

Linear regression using the Gradient Descent method

Introduction

What is linear regression?

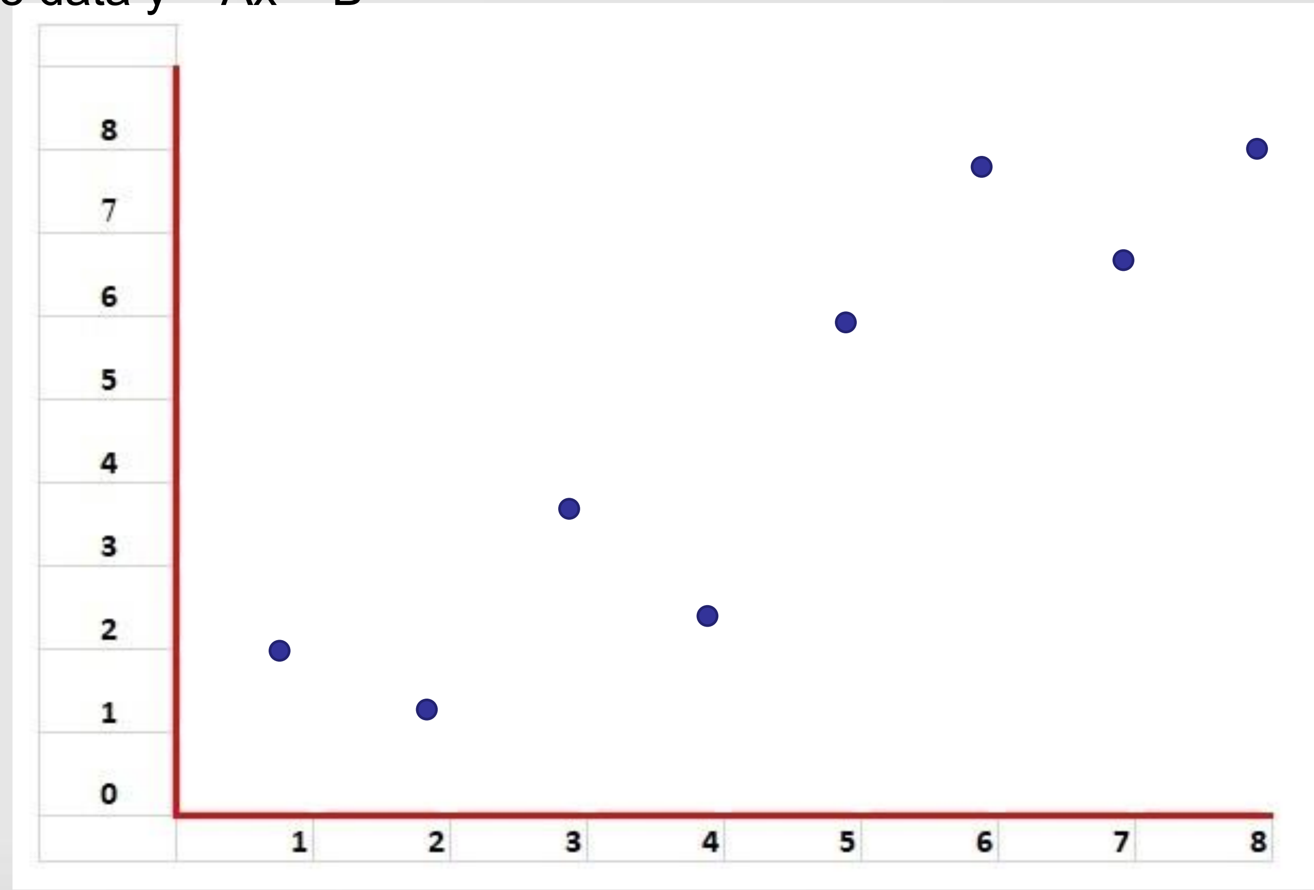
In statistics, a linear regression model is a model that seeks to establish a linear relationship between an independent variable and another dependent variable.

Example

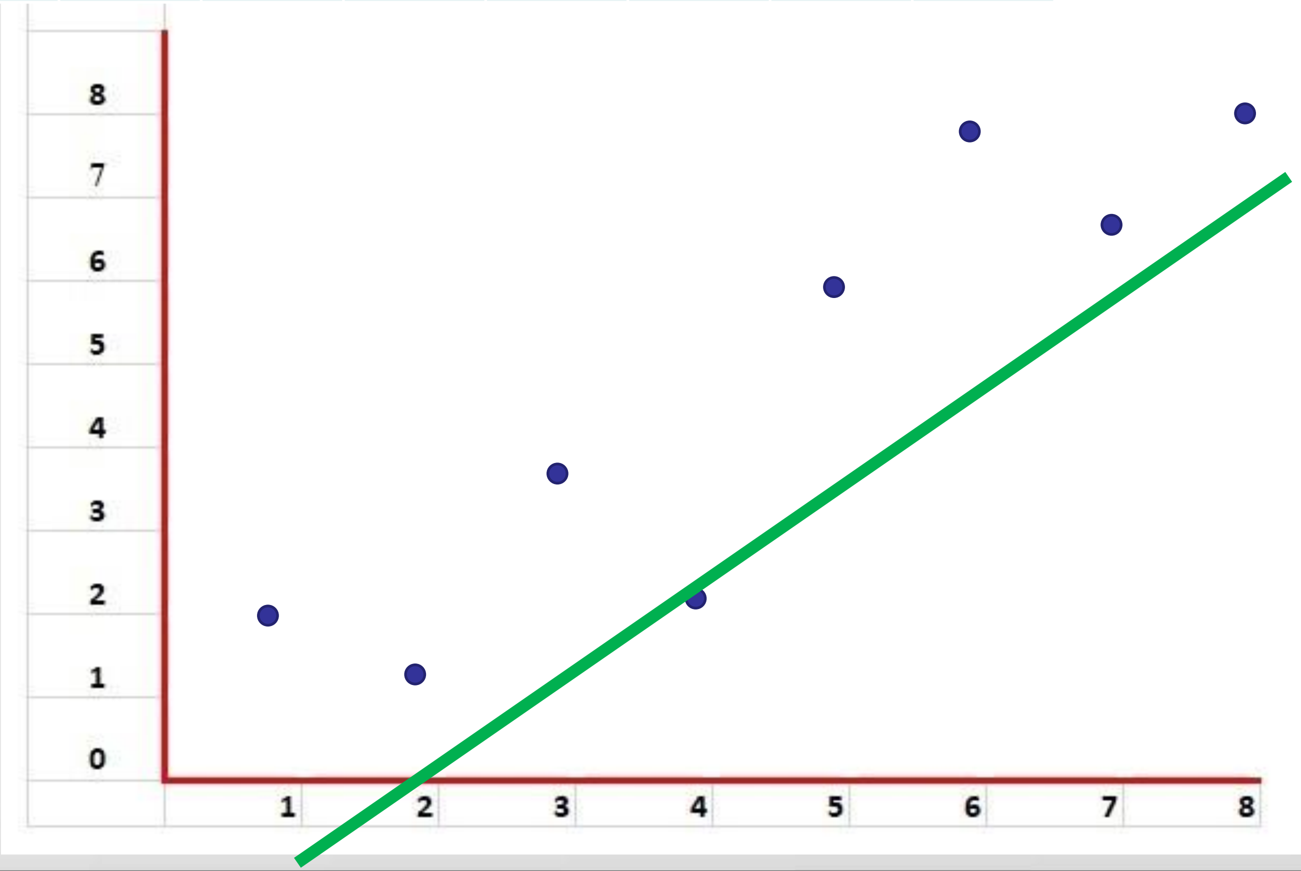
X	1	2	3	4	5	6	7	8
Y	1.9	1.1	3.5	2	5.7	7.5	6.8	7.9

L'objectif?

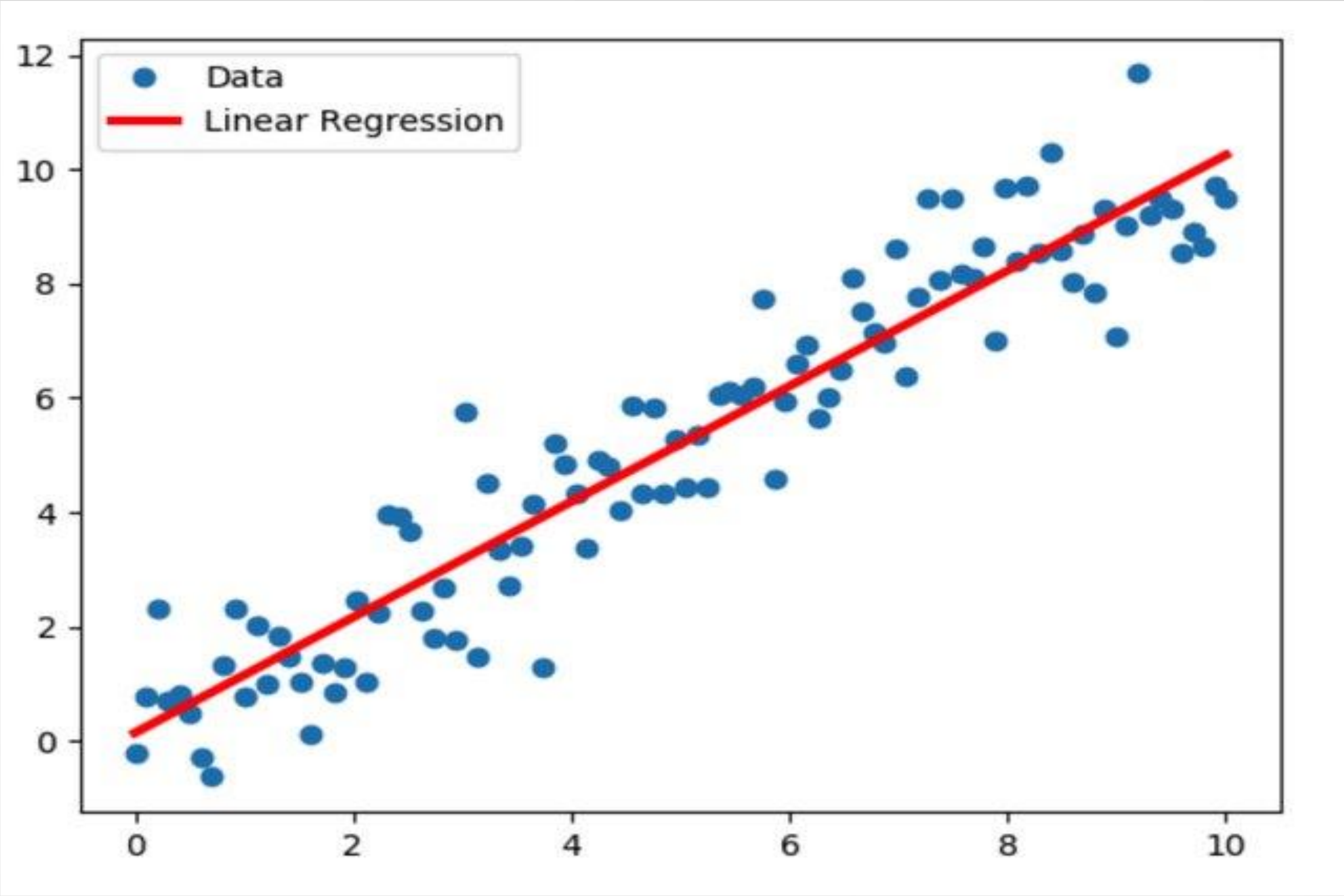
The goal in linear regression is to find the equation of the line that represents the best fit for the data y = Ax + B



X	1	2	3	4	5	6	7	8
Y	1.1	1.7	3.5	2	5.7	7.5	6.8	7.9



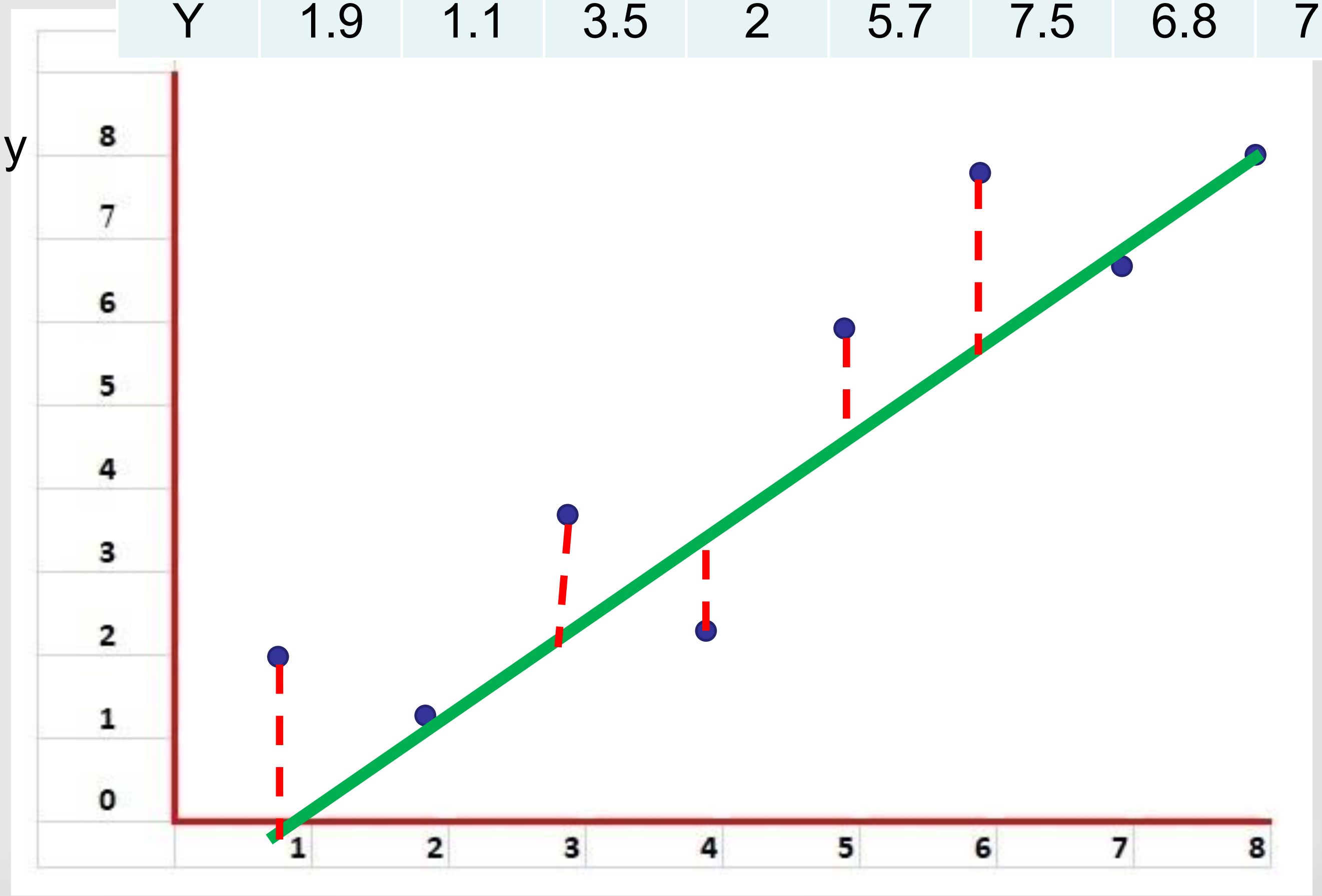
Exemple



X	1	2	3	4	5	6	7	8
Y	1.9	1.1	3.5	2	5.7	7.5	6.8	7.9

$h(x) = y$

h′x) = Ax+B = y



The cost function

The cost function

$$\text{Loss} = \sum_{i=1}^{n}(y_i - h(x_i))$$

So our objective is :

$$\text{minimiser} \sum_{i=1}^{n}(y_i - h(x_i))$$

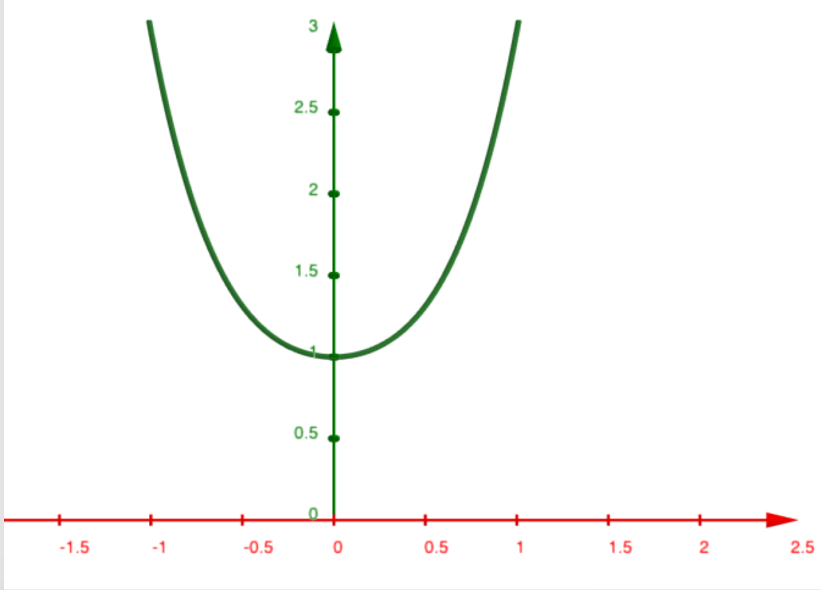
Sum of Squared Errors (SSE)

But because of a problem in this function, we use

$$\text{minimiser} \sum_{i=1}^{n}(y_i - h(x_i))^2 \qquad 1/2$$

So our objective is :

$$\text{minimiser} \sum_{i=1}^{n}(x)^2$$



Descente de la gradient

$$\text{loss} = \sum_{i=1}^{n}(y_i - h(x_i))^2$$

$$\text{loss} = \sum_{i=1}^{n}(y_i - (Ax_i + B))^2$$

The derivative with respect to A:

$$\frac{d\ loss}{d\ A} = \frac{d}{d\ A}(\sum_{i=1}^{n}(y_i - (Ax_i + B))^2$$

$$= \sum_{i=1}^{n} -2x_i(y_i - (Ax_i + B))$$

Descente de la gradient

$$\text{loss} = \sum_{i=1}^{n}(y_i - h(x_i))^2$$

$$\text{loss} = \sum_{i=1}^{n}(y_i - (Ax_i + B))^2$$

The derivative with respect to B:

$$\frac{d\ loss}{d\ B} = \frac{d}{d\ B}\left(\sum_{i=1}^{n}(y_i - (Ax_i + B))^2\right)$$

$$= \sum_{i=1}^{n} -2\,(y_i - (Ax_i + B)))$$

We will start from initial and random values of A and B, denoted respectively A_0 and B_0 and we calculate: :

$$slope_A = \frac{d\ loss\ (A_0, B_0)}{d\ A}$$

$$slope_B = \frac{d\ loss\ (A_0, B_0)}{d\ B}$$

$Step_A$ and $Step_B$ are two variables used to update the two values A and B.And they also define the learning rate.

$$Step_A = Slope_A * Learning\ rate$$
$$Step_B = Slope_B * Learning\ rate$$

The new values of A and B are calculated as follows:

$$new_A = \ old_A - Step_A$$
$$new_B = \ old_B - Step_B$$

We repeat this process until we obtain very small values for steps, or we reach the limit of the number of iterations.

Summary

1- Calculate the derivative of the **Loss** function for each parameter (A and B).

2- Give random initial values for A and B

3- Calculate **$\mathbf{slope_A}$** and **$\mathbf{slope_B}$**

4- Calculate **$\mathbf{step_A}$** and **$\mathbf{step_B}$** such as ***step*** = ***slope * learning rate***

5- Calculate **$\mathbf{new_A}$** and **$\mathbf{new_B}$** such as **$\mathbf{new_A}$** = **$\mathbf{old_A}$ - *$step_A$***

6- Repeat from 3 down to a very small number of steps or a maximum number of iterations.

Example

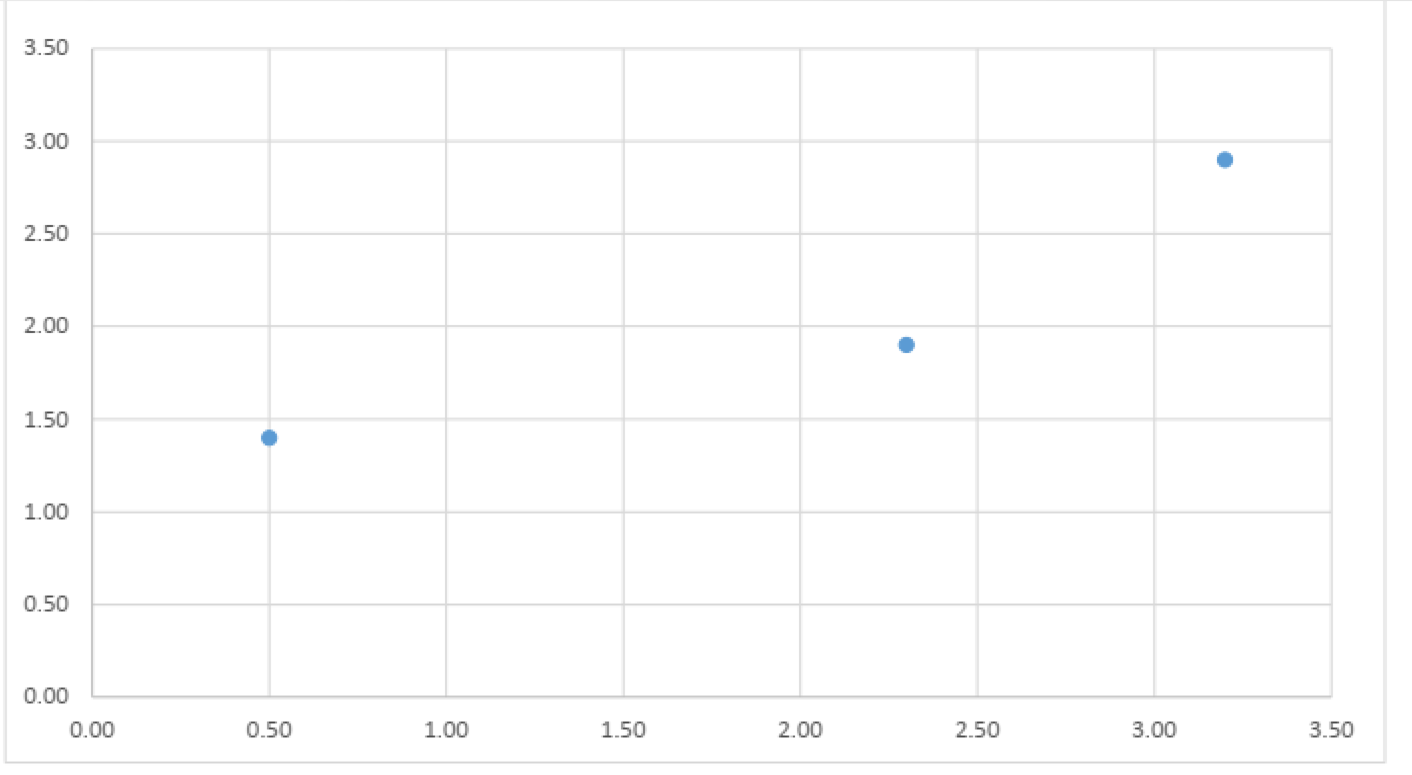
x	0.5	2.3	3.2
y	1.4	1.9	2.9

A =0
B =0

$$slope_A = \frac{d\ loss\ (A_0, B_0)}{d\ A} = \sum_{i=1}^{n} -2x_i(y_i - (Ax_i + B))$$

$$slope_B = \frac{d\ loss\ (A_0, B_0)}{d\ B} = \sum_{i=1}^{n} -2\ (y_i - (Ax_i + B)))$$

Example



Example

x	0.5	2.3	2.9
y	1.4	1.9	3.2

A =0
B =0

$$slope_A = -2 * 0.5(1.4 - (0 * 0.5 + 0)) - 2 * 2.3\left(1.9 - (0 * 2.3 + 0)\right) - 2 * 3.2(2.9 - (0 * 3.2 + 0)) = -4.78$$

$$slope_B = -2(1.4 - (0 * 0.5 + 0)) - 2\left(1.9 - (0 * 2.3 + 0)\right) - 2(2.9 - (0 * 3.2 + 0)) = -2.066$$

Example

$$Step_A = \ Slope_A \ * Learning\ rate$$
$$Step_B = \ Slope_B \ * Learning\ rate$$

$$Step_A = -4.78 * 0.01 = -0.0478$$
$$Step_B = -2.06 * 0.01 = -0.0206$$

$$\boldsymbol{new_A} = \mathbf{0} - (-0.0478) = 0.0478$$
$$\boldsymbol{new_B} = \mathbf{0} - (-0.0206) = 0.0206$$

A =0
B =0

A = 0.0478
B =- 0.0206

Example

x	0.5	2.3	3.2
y	1.4	1.9	2.9

A = 0.0478
B =- 0.0206

$$slope_A = -2 * 0.5(1.4 - (0.0478 * 0.5 + 0.0206) -2 * 2.3\big(1.9 - (0.0478 * 2.3 + 0.0206)\big) -2 * 3.2(2.9 - (0.0478 * 3.2 + 0.0206) = -4.490$$

$$slope_B = -2(1.4 - (0.0478 * 0.5 + 0.0206) -2\big(1.9 - (0.0478 * 2.3 + 0.0206)\big) -2(2.9 - (0.0478 * 3.2 + 0.0206) = -1.950$$

Example

$$Step_A = Slope_A * Learning\ rate$$
$$Step_B = Slope_B * Learning\ rate$$

A =0
B =0

A = 0.0478
B =- 0.0206

$$Step_A = \mathbf{-4.490} * 0.01 = -0.0490$$
$$Step_B = \mathbf{-1.950} * 0.01 = -0.0195$$

$$\boldsymbol{new_A} = 0.0478 - (-0.0490) = \mathbf{0.092}$$
$$\boldsymbol{new_B} = 0.0206 - (-0.0195\mathbf{8}) = \mathbf{0.0.092}$$

A	**0**	0.0478	0.092	0.134	0.174	….	0.51587
B	**0**	0.0206	0.092	0.058	0.075	…..	1.03492

$Step_A$	-0.287	-0.1815	-0.11439	….	7.379491e-05
$Step_B$	-0.1239	-0.08212	-0.05540	……	-0.000186

Example

x	y	A	B		Ax+B
0.5	1.4	0.51587	1.03492		**1.292855**
2.3	1.9				**2.221421**
3.2	2.9				**2.685704**

If x = 4.1 then the prediction is y = **3.149987**

Méthode de calcule numérique

$$\frac{d\ loss}{d\ B} = \sum_{i=1}^{n} -2\left(y_i - (Ax_i + B)\right) = 0$$

$$= \sum_{i=1}^{n}\left(y_i - (Ax_i + B)\right) = 0$$ Because -2 does not equal 0

$$\sum_{i=1}^{n}\left(y_i - (Ax_i + B)\right) = 0$$

$$\sum y_i - \sum Ax_i - \sum B = 0$$

$$\sum y_i - A\sum x_i - nB = 0$$

$$\sum y_i - A\sum x_i = \text{nB}$$

$$\boldsymbol{B = \frac{\sum y_i}{n} - \frac{A\sum x_i}{n}}$$

$$\mathbf{B} = \boldsymbol{\bar{y} - A\,\bar{x}}$$

$$\mathbf{B} = \ \overline{\boldsymbol{y}} \ - \boldsymbol{A}\,\overline{\boldsymbol{x}}$$

$$\frac{d\ loss}{d\ A} = \textstyle\sum_{i=1}^{n} -2x_i\big(y_i - (Ax_i + B)\big) = 0$$

$$-2\sum_{i=1}^{n} x_i\big(y_i - (Ax_i + B)\big) = 0$$

$$\sum_{i=1}^{n} x_i\big(y_i - (Ax_i + \overline{\boldsymbol{y}} \ - \boldsymbol{A}\,\overline{\boldsymbol{x}})\big) = 0$$

$$\sum_{i=1}^{n} x_i\big(y_i - \bar{y} \ - (Ax_i \ - \boldsymbol{A}\,\overline{\boldsymbol{x}})\big) = 0$$

$$\sum_{i=1}^{n} x_i\left(y_i - \bar{y} - (Ax_i - \boldsymbol{A}\,\overline{\boldsymbol{x}})\right) = 0$$

$$\sum_{i=1}^{n} x_i\left((y_i - \overline{y}) - A(x_i - \overline{\boldsymbol{x}})\right) = 0$$

$$\sum x_i(y_i - \overline{y}) - A\sum x_i(x_i - \overline{x}) \qquad \sum x_i(y_i - \overline{y}) = A\sum x_i(x_i - \overline{x})$$

$$\boldsymbol{A} = \frac{\sum \boldsymbol{x_i}(\boldsymbol{y_i} - \overline{\boldsymbol{y}})}{\sum \boldsymbol{x_i}(\boldsymbol{x_i} - \overline{\boldsymbol{x}})} \qquad \boldsymbol{A} = \frac{\sum(\boldsymbol{x_i} - \overline{\boldsymbol{x}})(\boldsymbol{y_i} - \overline{\boldsymbol{y}})}{\sum(\boldsymbol{x_i} - \overline{\boldsymbol{x}})^2}$$

$$\mathbf{B} = \bar{y} - A\,\bar{x}$$

$$A = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sum (x_i - \overline{x})^2}$$

Example

x	y
0.5	1.4
2.3	1.9
3.2	2.9

$$\mathbf{B} = \bar{y} - A\,\bar{x}$$

$$A = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$$

Example

x	y
0.5	1.4
2.3	1.9
3.2	2.9
$\overline{x}$ = 2	$\overline{y}$ = 2.066

$$\mathbf{B} = \overline{y} - A\,\overline{x}$$

$$A = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sum (x_i - \overline{x})^2}$$

Example

x	y	$x_i - \overline{x}$
0.5	1.4	-1.5
2.3	1.9	0.3
3.2	2.9	1.2
$\overline{x}$ = 2	$\overline{y}$ = 2.066	

$$A = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sum(x_i - \overline{x})^2}$$

$$\mathrm{B} = \overline{y} - A\,\overline{x}$$

Example

x	y	$x_i - \bar{x}$	$y_i - \bar{y}$
0.5	1.4	-1.5	-0.666
2.3	1.9	0.3	-0.166
3.2	2.9	1.2	0.8333
$\bar{x}$ = 2	$\bar{y}$ = 2.066		

$$A = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$\mathrm{B} = \bar{y} - A\,\bar{x}$$

Example

x	y	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})*(y_i - \bar{y})$
0.5	1.4	-1.5	-0.666	0.999
2.3	1.9	0.3	-0.166	-0.0498
3.2	2.9	1.2	0.8333	1
$\bar{x}$ = 2	$\bar{y}$ = 2.066			$\sum = 1.9492$

$$A = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$$

$$\mathrm{B} = \bar{y} - A\,\bar{x}$$

Example

x	y	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})*(y_i - \bar{y})$	$(x_i - \bar{x})^2$
0.5	1.4	-1.5	-0.666	0.999	2.25
2.3	1.9	0.3	-0.166	-0.0498	0.09
3.2	2.9	1.2	0.8333	1	1.44
$\bar{x}$ = 2	$\bar{y}$ = 2.066			$\sum = 1.9492$	$\sum = 3.78$

$$A = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \frac{1.9492}{3.78} = 0.5157$$

$$B = \bar{y} - A\,\bar{x} = 2.066 - 0.5157 * 2 = 1.064$$

Example

$$A = 0.5157, B = 1.064$$

x	y	$Ax + B$
0.5	1.4	1.32
2.3	1.9	2.25
3.2	2.9	2.71

Regression Loss Functions

Mean Squared Error (MSE)

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Regression Loss Functions

Sum of Squared Errors (SSE)

$$\mathrm{SSE} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

Regression Loss Functions

Root Mean Squared Error (RMSE)

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

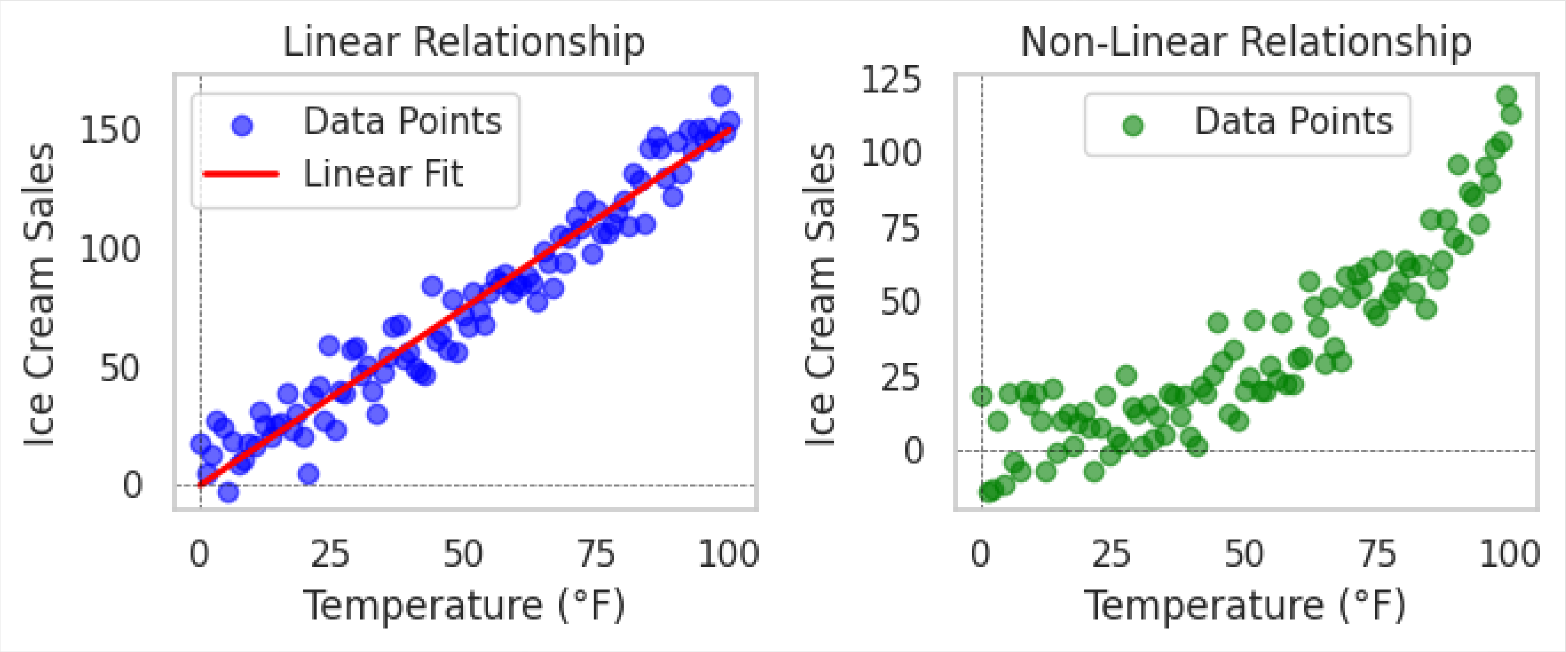
Regression Loss Functions

Mean Absolute Error (MAE)

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

The main problems of linear regression

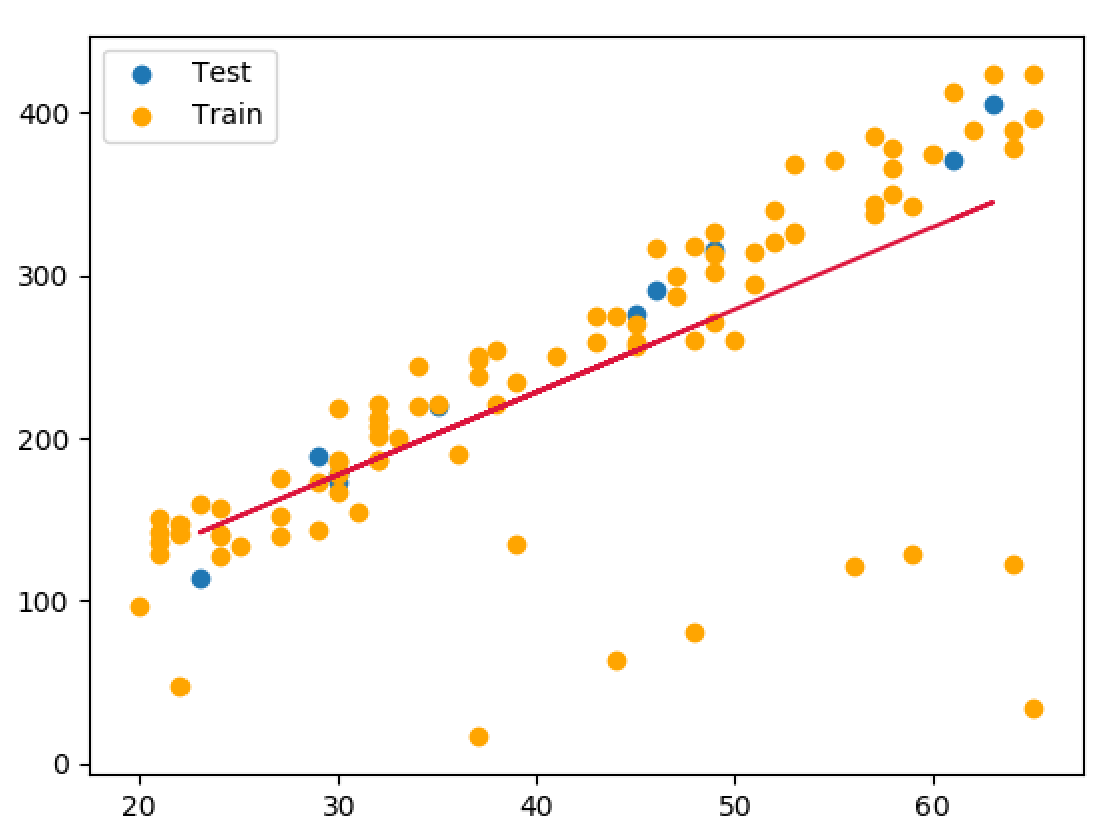
Linearity Assumption (Biggest Problem)



The main problems of linear regression

Sensitive to Outliers

Because MSE squares the error, a single extreme point can damage the model badly.
Example:
1 wrong measurement = huge squared error = model is distorted.



x	y
2.4	1.5
9.8	2.7
5.4	2.1
7.8	1.8
8.9	2.4
4.1	1.5

Find a linear fit for these data using the two methods seen in class.