Least squares

Sample Subtitle

Juan V. Vía

August 11, 2021

Example

Showing why least squares

We have a variable y. We know that it's dependent of another variable x in some way. But we don't know how, exactly.
So we go to the field and measure certain points. Those that we can reach. Six of them.

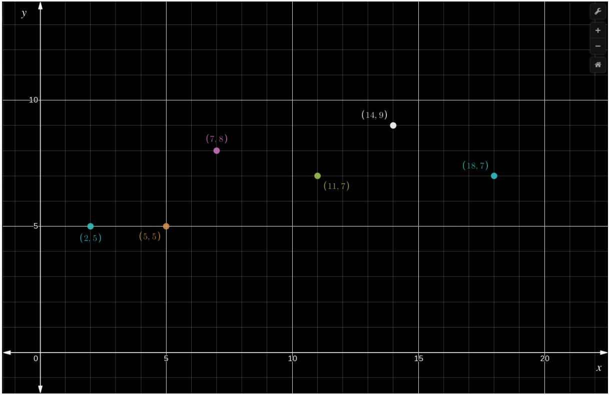
$$(2, 5), (5, 5), (7, 8), (11, 7), (14, 9), (18, 7)$$

That is: at $x = 2$ we measure $y = 5$, at $x = 5$ we measure $y = 5$ again, but at $x = 7$ we got $y = 8$, and so on.

Example

Showing why least squares

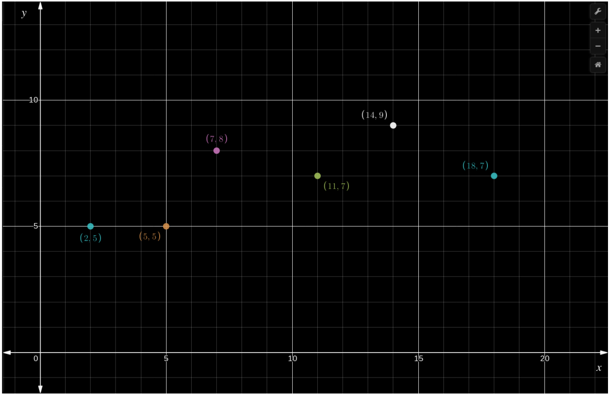
Next step, obviously, is to plot these points.



Example

Showing why least squares

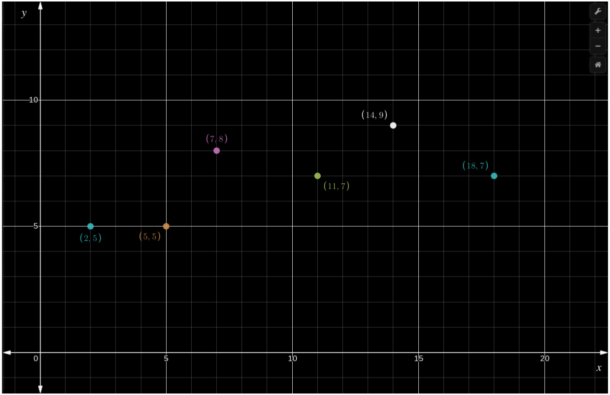
Let's start by tracing a line wich "best fit" that data



Example

Showing why least squares

What line? How to choose that line?



Example

Showing why least squares

A very common way to express a line is using the equation (a polinomial function)

$$y = mx + b$$

Using vector notation:

$$y = \begin{bmatrix} x & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix}$$

Or, switching terms:

$$\begin{bmatrix} x & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = y$$

Example

Showing why least squares

Ok. Fine. That's it. This form

$$\begin{bmatrix} x & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = y$$

will be a useful one because we know x and y in six points. For example take the first point $(2, 5)$

$$\begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = 5$$

Example

Showing why least squares

And yes, your guess is true. We can incorporate the second point $(5, 5)$ and get

$$\begin{bmatrix} 2 & 1 \\ 5 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$

...and the third point...and the fourth...

Example

Showing why least squares

Reaching the end this is the result

$$\begin{bmatrix} 2 & 1 \\ 5 & 1 \\ 7 & 1 \\ 11 & 1 \\ 14 & 1 \\ 18 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ 8 \\ 7 \\ 9 \\ 7 \end{bmatrix}$$

Example

Showing why least squares

Time to name things.

Example

Showing why least squares

Call A to the matrix

$$A = \begin{bmatrix} 2 & 1 \\ 5 & 1 \\ 7 & 1 \\ 11 & 1 \\ 14 & 1 \\ 18 & 1 \end{bmatrix}$$

Example

Showing why least squares

Call x to the column vector

$$x = \begin{bmatrix} m \\ b \end{bmatrix}$$

Example

Showing why least squares

Call b to the column vector

$$b = \begin{bmatrix} 5 \\ 5 \\ 8 \\ 7 \\ 9 \\ 7 \end{bmatrix}$$

Example

Showing why least squares

Thus, the equation

$$\begin{bmatrix} 2 & 1 \\ 5 & 1 \\ 7 & 1 \\ 11 & 1 \\ 14 & 1 \\ 18 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ 8 \\ 7 \\ 9 \\ 7 \end{bmatrix}$$

becomes

$$Ax = b$$

Example

Showing why least squares

In other words: if you want to know the values m and b in the equation $y = mx + b$ of the line we are searching find the x such that $Ax = b$, but...

...that equation have no solution!...

...because the matrix A is tall, so $Ax = b$ is over-determined, there are more equations (m) than variables to choose. There is no x such that the norm of the difference between Ax and b equals zero

$$\|Ax - b\| \neq 0$$

We need the least squares approximate solution

Example

Showing why least squares

In the least squares approximate solution of $Ax = b$ we, provided that $\|Ax - b\| \neq 0$, select x that the *norm* of the *residual* $r = Ax - b$ be minimal.
It's all here: http://vmls-book.stanford.edu/ and in many other sources. The least squares approximate solution is very known. Let's see it.

Example

Showing why least squares

All start with our equation $Ax = b$.

Example

Showing why least squares

Take A,

$$A = \begin{bmatrix} 2 & 1 \\ 5 & 1 \\ 7 & 1 \\ 11 & 1 \\ 14 & 1 \\ 18 & 1 \end{bmatrix}$$

transpose it.

$$A^\intercal = \begin{bmatrix} 2 & 5 & 7 & 11 & 14 & 18 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Example

Showing why least squares

Multiply A^{T} by A

$$A^{\mathsf{T}}A = \begin{bmatrix} 2 & 5 & 7 & 11 & 14 & 18 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 5 & 1 \\ 7 & 1 \\ 11 & 1 \\ 14 & 1 \\ 18 & 1 \end{bmatrix} = \begin{bmatrix} 719 & 57 \\ 57 & 6 \end{bmatrix}$$

Example

Showing why least squares

Invert the product

$$(A^\intercal A)^{-1} = \begin{bmatrix} 719 & 57 \\ 57 & 6 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{2}{355} & -\frac{19}{355} \\ -\frac{19}{355} & \frac{719}{1065} \end{bmatrix}$$

Example

Showing why least squares

Multiply inverse and transpose

$$(A^\intercal A)^{-1}A^\intercal = \begin{bmatrix} \frac{2}{355} & -\frac{19}{355} \\ -\frac{19}{355} & \frac{719}{1065} \end{bmatrix} \begin{bmatrix} 2 & 5 & 7 & 11 & 14 & 18 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{3}{71} & -\frac{9}{355} & -\frac{1}{71} & \frac{3}{355} & \frac{9}{355} & \frac{17}{355} \\ \frac{121}{213} & \frac{434}{1065} & \frac{64}{213} & \frac{92}{1065} & -\frac{79}{1065} & -\frac{307}{1065} \end{bmatrix}$$

Let's call this product $(A^\intercal A)^{-1}A^\intercal$ the *pseudo-inverse*

Example

Showing why least squares

Finally multiply the pseudo-inverse by b

$$x = \begin{bmatrix} -\frac{3}{71} & -\frac{9}{355} & -\frac{1}{71} & \frac{3}{355} & \frac{9}{355} & \frac{17}{355} \\ \frac{121}{213} & \frac{434}{1065} & \frac{64}{213} & \frac{92}{1065} & -\frac{79}{1065} & -\frac{307}{1065} \end{bmatrix} \begin{bmatrix} 5 \\ 5 \\ 8 \\ 7 \\ 9 \\ 7 \end{bmatrix} = \begin{bmatrix} \frac{61}{355} \\ \frac{5539}{1065} \end{bmatrix}$$

Example

Showing why least squares

That's it

$$x = \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} \frac{61}{355} \\ \frac{5539}{1065} \end{bmatrix} \text{ or } \begin{bmatrix} 0.171831 \\ 5.20094 \end{bmatrix}$$

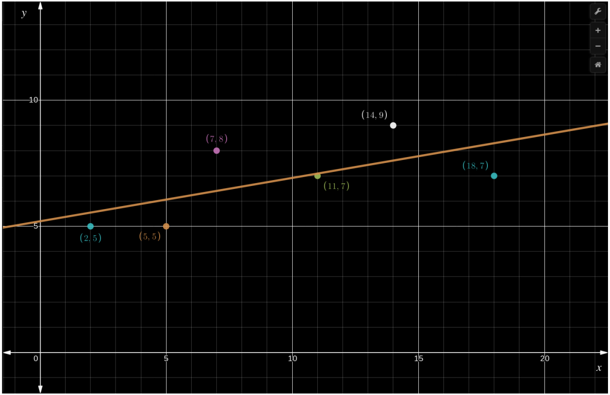
and $m = 0.171831$ and $b = 5.20094$ and the line is

$$y = 0.171831x + 5.20094$$

Example

Showing why least squares

And the best fit is



Least squares

Recapitulation

You have a bunch of data points

You decide derive a polinomial function $y = Ax + B$ from this points.

You build a matrix A and a vector b from the polinomial and the points.

The polinomial coefficients are a vector x so $Ax = b$

Then $x = (A^\intercal A)^{-1} A^\intercal b$

Least squares

The polinomial function

In the example we picked a line for regression.
A line in a plane is a polinomial function where y (the dependent variable, a real number) is a function of x (the independent variable, a real number)

$$y\colon \mathbb{R} \to \mathbb{R} = Ax + B$$