Menu
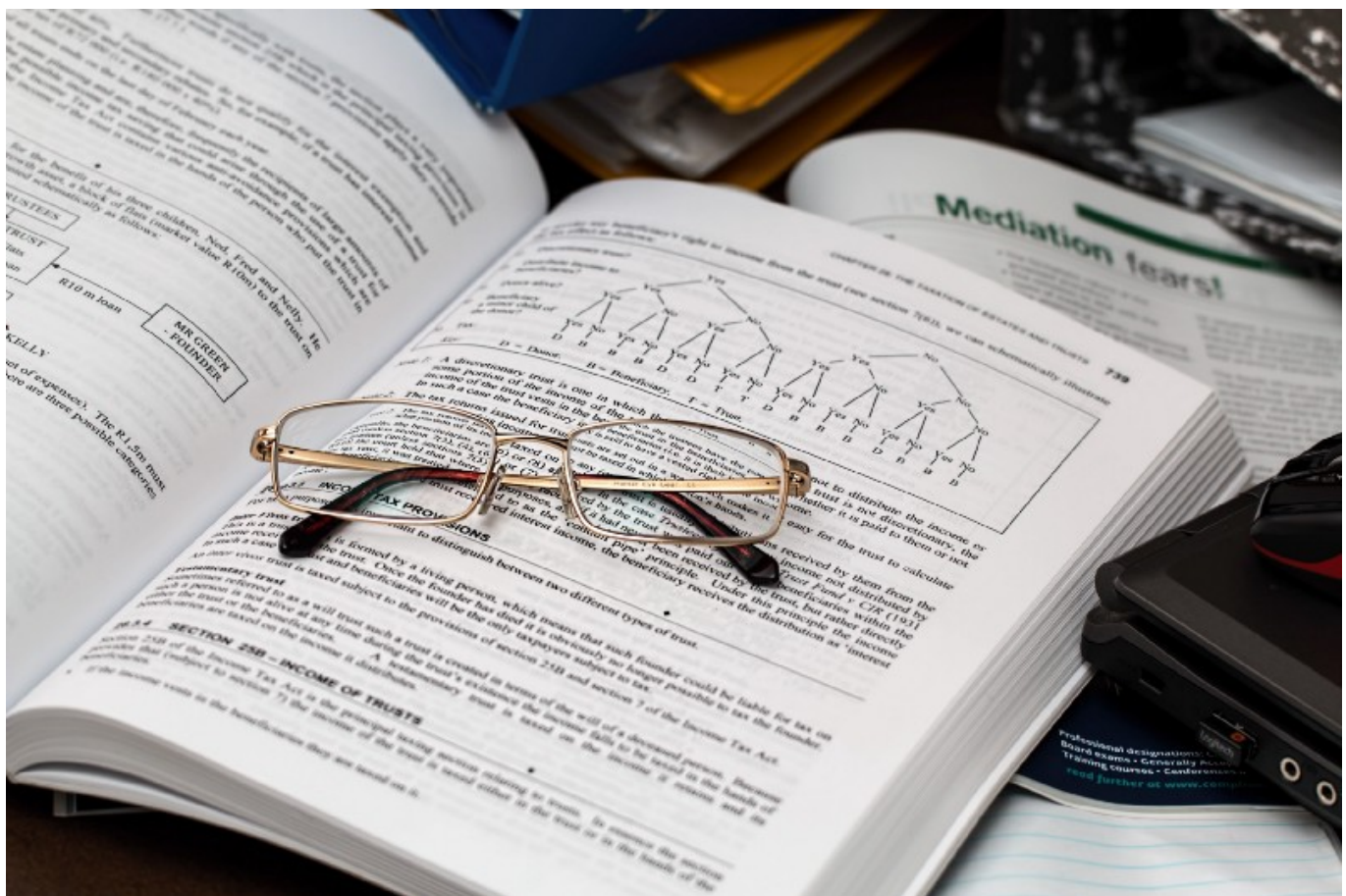
Machine learning: an introduction to mean squared error and regression lines

by Moshe Binieli

Introduction

This article will deal with the statistical method **mean squared error**, and I'll describe the relationship of this method to the **regression line**.

Menu

The example consists of points on the Cartesian axis. We will define a mathematical function that will give us the straight line that passes best between all points on the Cartesian axis.

And in this way, we will learn the connection between these two methods, and how the result of their connection looks together.

General explanation

This is the definition from Wikipedia:

> In statistics, the mean squared error (MSE) of an estimator (of a procedure for estimating an unobserved quantity) measures the average of the squares of the errors — that is, the average squared difference between the estimated values and what is estimated. MSE is a risk function, corresponding to the expected value of the squared error loss. The fact that MSE is almost always strictly positive (and not zero) is because of randomness or because the estimator does not account for information that could produce a more accurate estimate.

The structure of the article

- Get a feel for the idea, graph visualization, mean squared error equation.

- The mathematical part which contains algebraic manipulations and a derivative of two-variable functions for finding a minimum. This section is **for those who want to understand how** we get the mathematical formulas later, you can skip it if that doesn't interest you.

- An explanation of the mathematical formulae we received and the role of each variable in the formula.

Menu

Get a feel for the idea

Let's say we have seven points, and our goal is to find a line that **minimizes** the squared distances to these different points.

Let's try to understand that.

I will take an example and I will draw a line between the points. Of course, my drawing isn't the best, but it's just for demonstration purposes.
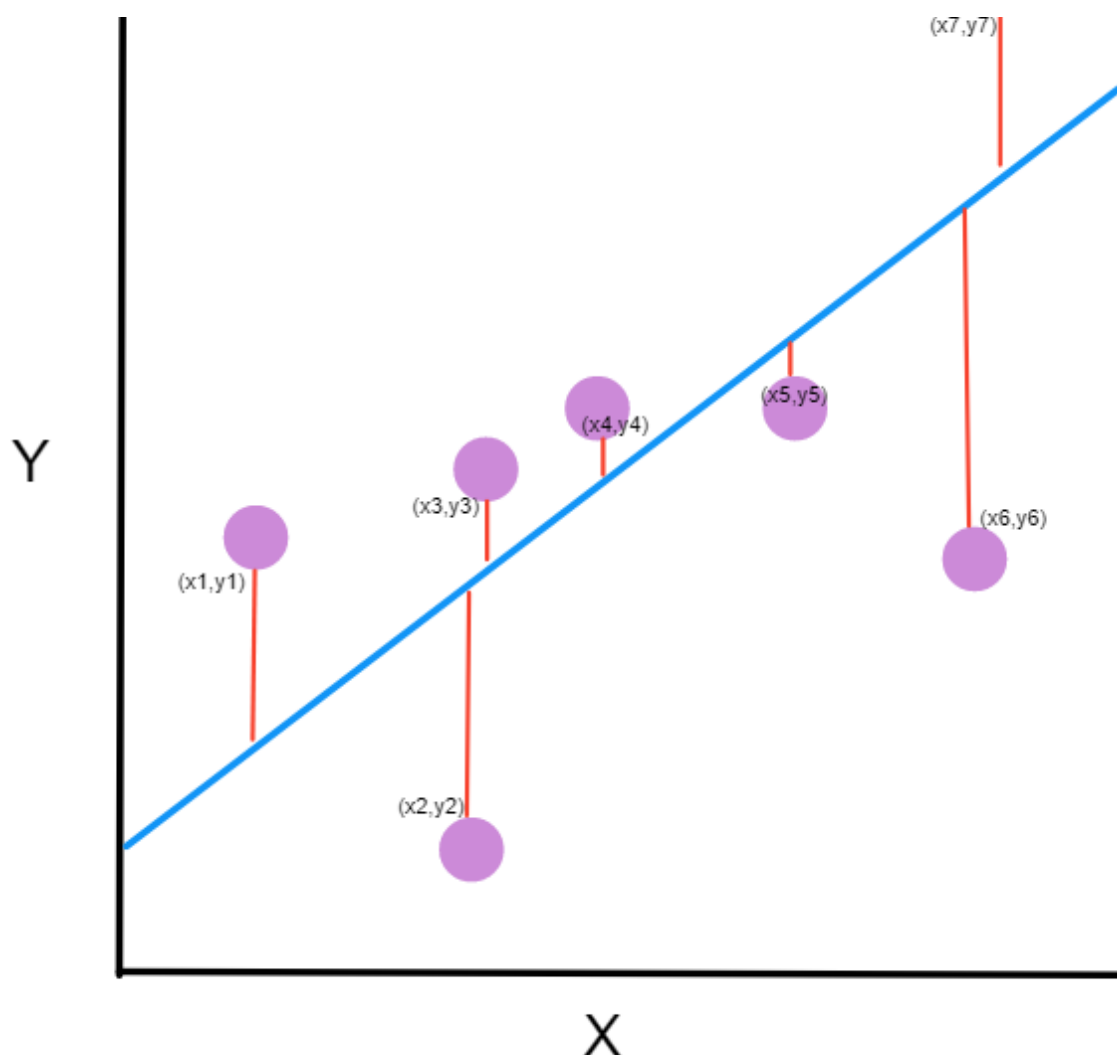
Points on a simple graph.

You might be asking yourself, what is this graph?

- the **purple dots** are the points on the graph. Each point has an x-coordinate and a y-coordinate.

- The **blue line** is our prediction line. This is a line that passes through all the points and fits them in the best way. This line contains the predicted points.

- The **red line** between each purple point and the prediction line are the **errors.** Each error is the distance from the point to its predicted point.

where **M** is the <u>slope</u> of the line and **B** is <u>y-intercept</u> of the line.

We want to find M (<u>slope</u>) and B (<u>y-intercept</u>) that **minimizes** the squared error!

Let's define a mathematical equation that will give us the mean squared error for all our points.

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

General formula for mean squared error.

Let's analyze what this equation actually means.

- In mathematics, the character that looks like weird E is called summation (Greek sigma). It is the sum of a sequence of numbers, from i=1 to n. Let's imagine this like an array of points, where we go through all the points, from the first (i=1) to the last (i=n).

- For each point, we take the y-coordinate of the point, and the y'-coordinate. The y-coordinate is our purple dot. The y' point sits on the line we created. We subtract the y-coordinate value from the y'-coordinate value, and calculate the square of the result.

divide it by n, which will give the mean.

Our goal is to minimize this mean, which will provide us with the best line that goes through all the points.

From concept to mathematical equations

This part is **for people who want to understand how we got to the mathematical equations**. You can skip to the next part if you want.

As you know, the line equation is y=mx+b, where m is the <u>slope</u> and b is the <u>y-intercept</u>.

Let's take each point on the graph, and we'll do our calculation (y-y')². But what is y', and how do we calculate it? We do not have it as part of the data.

But we do know that, in order to calculate y', we need to use our line equation, y=mx+b, and put the x in the equation.

From here we get the following equation:

$$MSE = (y_1 - (mx_1 + b))^2 + (y_2 - (mx_2 + b))^2 + \ldots + (y_n - (mx_n + b))^2$$

Let's rewrite this expression to simplify it.

$$MSE = y_1^2 - 2y_1(mx_1 + b) + (mx_1 + b)^2 + \ldots + y_n^2 - 2y_n(mx_n + b) + (mx_n + b)^2$$

Let's begin by opening all the brackets in the equation. I colored the difference between the equations to make it easier to understand.

$$MSE = y_1^2 - 2y_1mx_1 - 2y_1b + m^2x_1^2 + 2mx_1b + b^2 + \ldots + y_n^2 - 2y_nmx_n - 2y_nb + m^2x_n^2 + 2mx_nb + b^2$$

Now, let's apply another manipulation. We will take each part and put it together. We will take all the y, and (-2ymx) and etc, and we will put them all side-by-side.

$$MSE = (y_1^2 + \ldots + y_n^2) - 2m(x_1y_1 + \ldots + x_ny_n) - 2b(y_1 + \ldots + y_n) + m^2(x_1^2 + \ldots + x_n^2) + 2mb(x_1 + \ldots + x_n) + nb^2$$

At this point we're starting to be messy, so let's take the mean of all squared values for y, xy, x, x².

Let's define, for each one, a new character which will represent the mean of all the squared values.

Let's see an example, let's take all the y values, and divide them by n since it's the mean, and call it y(HeadLine).

$$\frac{y_1^2 + y_2^2 + \ldots + y_n^2}{n} = \overline{y^2}$$

$$y_1^2 + y_2^2 + \ldots + y_n^2 = \overline{y^2}n$$

Which will lead us to the following equation:

$$MSE = n\overline{y^2} - 2mn\overline{xy} - 2bn\overline{y} + m^2n\overline{x^2} + 2mbn\overline{x} + nb^2$$

If we look at what we got, we can see that we have a 3D surface. It looks like a glass, which rises sharply upwards.

We want to find M and B that minimize the function. We will make a partial derivative with respect to M and a partial derivative with respect to B.

Since we are looking for a minimum point, we will take the partial derivatives and compare to 0.

$$\frac{\partial MSE}{\partial m} = \frac{\partial MSE}{\partial b} = 0$$

Partial derivatives formula

Menu

$$\frac{}{\partial m} \Rightarrow -2nxy + 2nx^-m + 2onx \Rightarrow -xy + x^-m + 0x = 0$$

$$\frac{\partial MSE}{\partial b} \Rightarrow -2n\bar{y} + 2mn\bar{x} + 2bn \Rightarrow -\bar{y} + m\bar{x} + b = 0$$

Partial derivatives

Let's take the two equations we received, isolating the variable b from both, and then subtracting the upper equation from the bottom equation.

$$First \Rightarrow m\frac{\overline{x^2}}{\bar{x}} + b = \frac{\overline{xy}}{\bar{x}}$$

$$Second \Rightarrow m\bar{x} + b = \bar{y}$$

Different writing of the equations after the derivation by parts

Let's subtract the first equation from the second equation

$$m(\bar{x} - \frac{\overline{x^2}}{\bar{x}}) = \bar{y} - \frac{\overline{xy}}{\bar{x}} \Rightarrow m = \frac{\bar{y} - \frac{\overline{xy}}{\bar{x}}}{\bar{x} - \frac{\overline{x^2}}{\bar{x}}}$$

Merge two equations together

$$m = \cfrac{\overline{y} - \cfrac{\overline{xy}}{\overline{x}}}{\overline{x} - \cfrac{\overline{x^2}}{\overline{x}}} \cdot \frac{(\overline{x})}{(\overline{x})} = \frac{\overline{x}\,\overline{y} - \overline{xy}}{(\overline{x})^2 - \overline{x^2}}$$

Final equation to find M.

And there we go, this is the equation to find M, let's take this and write down B equation.

$$m\overline{x} + b = \overline{y} \Rightarrow b = \overline{y} - m\overline{x}$$

Final equation to find B.

Equations for slope and y-intercept

Let's provide the mathematical equations that will help us find the required slope and y-intercept.

$$m = \frac{\overline{xy} - \overline{x}\,\overline{y}}{\overline{x^2} - (\overline{x})^2} \qquad b = \overline{y} - m\overline{x}$$

Slope and y-intercept equations

equations?

They are actually simple to understand, so let's talk about them a little
bit.

$$\frac{x_1 + x_2 + \ldots + x_n}{n} = \overline{x}$$

Sum of x divided by n

$$\frac{x_1^2 + x_2^2 + \ldots + x_n^2}{n} = \overline{x^2}$$

Sum of x² divided by n

$$\frac{x_1 y_1 + x_2 y_2 + \ldots + x_n y_n}{n} = \overline{xy}$$

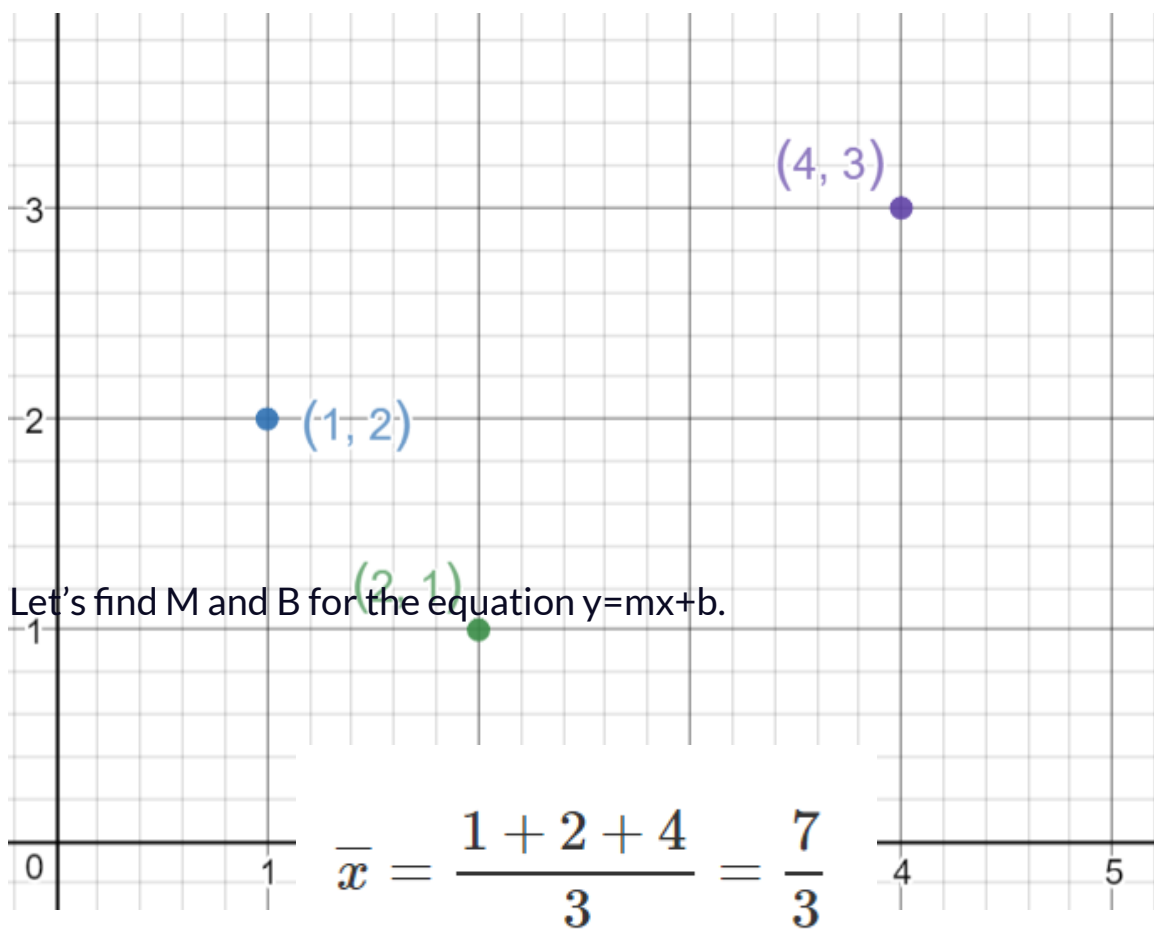$$\frac{y_1 + y_2 + \ldots + y_n}{n} = \overline{y}$$

Sum of y divided by n

Now that we understand our equations it's time to get all things together and show some examples.

Examples

A big thank you to <u>Khan Academy</u> for the examples.

Example #1

Let's take 3 points, (1,2), (2,1), (4,3).

Menu

(4, 3)

(1, 2)

Let's find M and B for the equation y=mx+b.

(2, 1)

$$\overline{x} = \frac{1+2+4}{3} = \frac{7}{3}$$

Sum the x values and divide by n

$$\overline{y} = \frac{2+1+3}{3} = 2$$

Sum the y values and divide by n

$$\overline{xy} = \frac{1 \cdot 2 + 2 \cdot 1 + 4 \cdot 3}{3} = \frac{16}{3}$$

Sum the xy values and divide by n

$$\overline{x^2} = \frac{1^2 + 2^2 + 4^2}{3} = \frac{21}{3} = 7$$

Sum the x² values and divide by n

After we've calculated the relevant parts for our M equation and B equation, let's put those values inside the equations and get the slope and y-intercept.

$$m = \frac{\frac{7}{3} \cdot 2 - \frac{16}{3}}{\left(\frac{7}{3}\right)^2 - 7} = \frac{\frac{14}{3} - \frac{16}{3}}{\frac{49}{9} - 7} = \frac{-\frac{2}{3}}{-\frac{14}{9}} = \frac{3}{7}$$
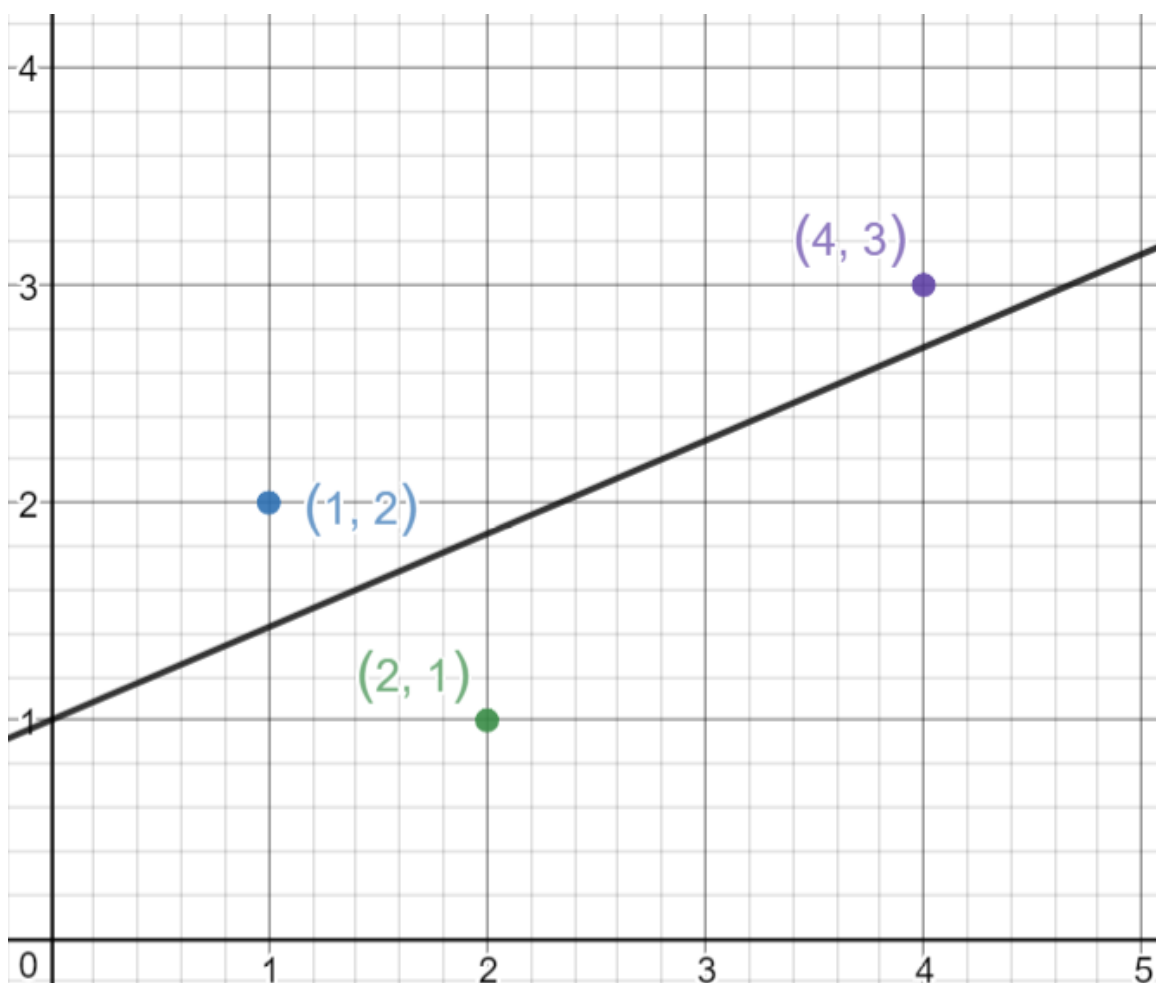
Slope calculation

$$b = 2 - \frac{3}{7} \cdot \frac{7}{3} = 2 - 1 = 1$$
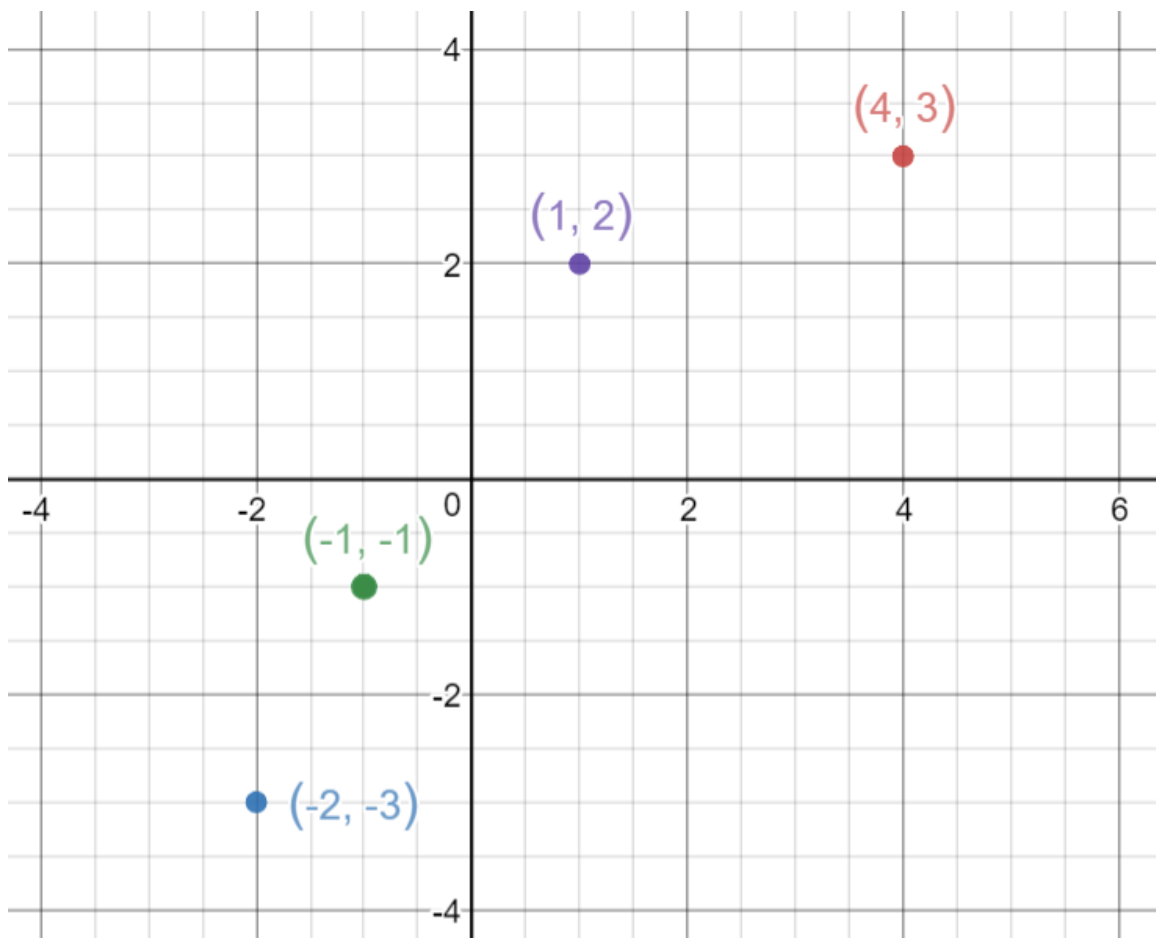
y-intercept calculation

$$y = \frac{3}{7}x + 1$$

Now let's draw the line and see how the line passes through the lines in such a way that it minimizes the squared distances.

Regression line that minimizes the MSE.

Example #2

Points on graph.

Let's find M and B for the equation y=mx+b.

$$\bar{x} = \frac{(-2) + (-1) + 1 + 4}{4} = \frac{1}{2}$$

Sum the x values and divide by n

Menu

$$\bar{y} = \frac{(-3) + (-1) + 2 + 3}{4} = \frac{1}{4}$$

Sum the y values and divide by n

$$\overline{xy} = \frac{(-2) \cdot (-3) + (-1) \cdot (-1) + 1 \cdot 2 + 3 \cdot 4}{4} = \frac{6 + 1 + 2 + 12}{4} = \frac{21}{4}$$

Sum the xy values and divide by n

$$\overline{x^2} = \frac{4 + 1 + 1 + 16}{4} = \frac{11}{2}$$

Sum the x² values and divide by n

Same as before, let's put those values inside our equations to find M and B.

$$m = \frac{\frac{21}{4} - \frac{1}{2} \cdot \frac{1}{4}}{\frac{11}{2} - \left(\frac{1}{2}\right)^2} = \frac{\frac{21}{4} - \frac{1}{8}}{\frac{11}{2} - \frac{1}{4}} = \frac{41}{42}$$
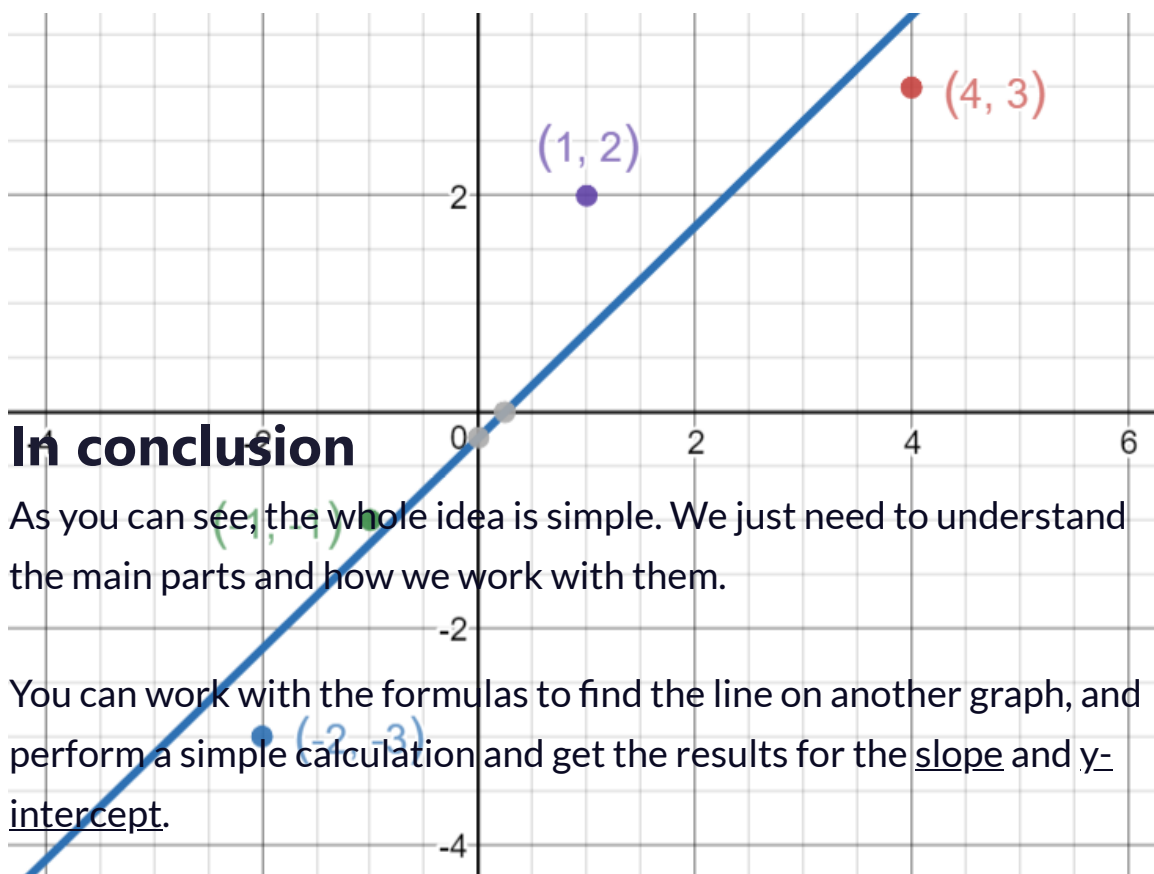
Slope calculation

$$b = \frac{1}{4} - \frac{11}{42} \cdot \frac{1}{2} = -\frac{5}{21}$$

y-intercept calculation

Let's take those results and set them inside line equation y=mx+b.

$$y = \frac{41}{42}x - \frac{5}{21}$$

Now let's draw the line and see how the line passes through the lines in such a way that it minimizes the squared distances.

Menu

(4, 3)

(1, 2)

(-1, -1)

(-2, -3)

Regression line that minimizes the MSE

In conclusion

As you can see, the whole idea is simple. We just need to understand the main parts and how we work with them.

You can work with the formulas to find the line on another graph, and perform a simple calculation and get the results for the slope and y-intercept.

That's all, simple eh?

Every comment and all feedback is welcome — if it's necessary, I will fix the article.

Feel free to contact me directly at LinkedIn — Click Here.

If this article was helpful, | tweet it. |

Learn to code for free. freeCodeCamp's open source curriculum has helped more than 40,000 people get jobs as developers.

| Get started |

Menu

freeCodeCamp is a donor-supported tax-exempt 501(c)(3) nonprofit organization (United States Federal Tax Identification Number: 82-0779546)

Our mission: to help people learn to code for free. We accomplish this by creating thousands of videos, articles, and interactive coding lessons - all freely available to the public. We also have thousands of freeCodeCamp study groups around the world.

Donations to freeCodeCamp go toward our education initiatives, and help pay for servers, services, and staff.

You can make a tax-deductible donation here.

Our Nonprofit

About

Alumni Network

Open Source

Shop

Support

Sponsors

Academic Honesty

Code of Conduct

Privacy Policy

Terms of Service

Copyright Policy

Trending Guides

2019 Web Developer Roadmap

Python Tutorial

CSS Flexbox Guide

JavaScript Tutorial

Python Example

HTML Tutorial

Linux Command Line Guide

JavaScript Example

Git Tutorial

React Tutorial

Java Tutorial

Linux Tutorial

CSS Tutorial

jQuery Example

SQL Tutorial

CSS Example

React Example

Menu

Bootstrap Example

How to Set Up SSH Keys

WordPress Tutorial

PHP Example

Menu