

**CHRISTINE  
BREINER:**

Welcome back to recitation. In this video, what I'd like you to do is use least squares to fit a line to the following data, which includes three points: the point (0, 1), the point (2, 1), and the point (3, 4). And I've drawn a rough picture where these points are on a graph, and I'll be talking a little bit about that after you try this problem. So why don't you try to solve the problem based on the technique you learned in class, and then bring the video back up. I'll show you again what you're actually doing, and then I will also solve the problem and you can see, you can compare your answer to mine.

OK, welcome back. So the first thing I want to do when we're talking about least squares to fit a line to these data, what I'd like to do is I'm going to draw a line on here, and we're going to talk about what the least squares actually is trying to do. And so I'm going to draw this line, say, something like this. That's my first guess on what might be the actual least squares line for these data.

And if you recall, what you're actually trying to do is you're trying to minimize a certain quantity, and the quantity you're trying to minimize is the difference between the actual value you get and the expected value you get, the square of that. So this is one thing we're going to square. So even pictorially, we could say we have something that contributes some area, OK, and then this is another thing that's the difference between the expected value in  $y$  and the actual value you get. So let me try and make that squared, because we square that value. And then this last one is pretty small. The difference between the expected value and the actual value is very small over in this case from the picture I drew.

And so what you're trying to do is you're trying to find the line that minimizes the area of the sum of these three, in this case, three squares. So if I, you know, if I tip the line down further here, but I kept it fixed there, the area of this square would get bigger, because this distance would get bigger. But, of course, the area of this square would get smaller. And so you're trying to figure out a way to optimize, in this case, minimize, the area of these-- sorry, the sum of the areas of these squares by figuring out where to put the line.

Now, I mentioned if I could fix this here and move it up and down, then that would change the slope but fix the intercept. But I could also move the intercept up and down, and I could slide the line up and down, and that would fix the slope but vary the intercept, or I could do both. And ultimately, that's what you're doing. You're trying to find the best  $y$ -intercept and the best

slope at the same time, and that's why you learned this process in multivariable calculus, because you're trying to optimize something that has two quantities that you can change. You can change slope and you can change the y-intercept, so that's why you learned it now instead of in single-variable calculus.

Now, if we come over here, I wrote down ahead of time what the equations are you get when you optimize, when you're trying to optimize the least squares line. So if you write down, you know, the function as a function of a and b, and you take the derivative with respect to a and you set it equal to zero, you get the first equation. When you take the derivative with respect to b and you set it equal to zero, you get the second equation. And so I'm just going to fill in the details and then tell you what the solution is, and then I'm going to mention a little more sophisticated thing we can do, or I guess maybe the next level of least squares we can do, that's not a linear problem. So I'll do that last.

So what I want to do is I put all the values I need over here on the right. So all the values we need are on the right. We're going to need the sum of all the  $x_i$ 's and that's 5. The sum of all the  $y_i$ 's is 6. The sum of the  $x_i$  squareds is 13, and the sum of the  $x_i y_i$  product is 14. Now, where are these coming from? Remember, what is  $x_i$  and what is  $y_i$ ? If we come back over here, this is  $(x_1, y_1)$ ,  $(x_2, y_2)$ ,  $(x_3, y_3)$ , OK? So that's sort of how we get-- if we switch the whole pairs around, it doesn't matter. Obviously, if we switched x- and y-values together, it does matter, and that you see actually from the equation does matter. Right here you have a product of them.

So let me write this out. And again, we're solving for a and b, so in this case, a is the slope and b is the intercept, right? So let me write everything in. So we actually get that we want to solve the following system:  $13a + 5b$  is equal to 14. Make sure I put in everything correctly. And then the second one is  $5a + 3b$  is equal to 6. OK, this is a system of equations. It has two equations and two unknowns, and if you solve this-- I should look at my notes, because I didn't memorize it and I'm not going to do all the algebra. If you go to solve this, what you actually get is that a is equal to  $6/7$  and b is equal to  $4/7$ . So you get that the line-- let me write it here. You get a is equal to  $6/7$  and b is equal to  $4/7$ , OK?

And so what that tells you is if you come back over here, the line we actually want-- I'll write the solution right here-- the line we actually want is the line  $y$  equals  $6/7 x$  plus  $4/7$ . That is our least squares line that is our solution. OK? So that is actually-- solving for a line to fit these

data, you get the following solution. But I want to point out, and what we're going to do, after this video there will be a challenge problem, to have you actually finish this. I want to point out that you can actually also fit a higher-degree polynomial to these data. For instance, you could try and use the technique of least squares to fit a parabola to these data. And let me point out what the function would look like in this case. So this is for the challenge problem.

For the challenge problem, it now will be a function of three variables, so it will look something like this. We'll say  $a$  comma  $b$  comma  $c$ , and the point that we'll be wanting to minimize, or the function we'll be wanting to minimize, we're summing over  $i$ , again from 1 to 3-- I didn't write that down above, but it was fairly obvious that that was what we were summing over-- the following thing. We have  $y_i$  minus this big quantity:  $a x_i$  squared plus  $b x_i$  plus  $c$  and we square the whole thing.

And so what this is actually giving you is, just as in the picture before, this is giving you the difference between the expected value and the actual value. So this is your actual  $y$ -value, and based on the parabola you're looking for, what's in parentheses is your expected  $y$ -value, right? So this would actually be my parabola. When I evaluate it at  $x_i$ , I get a number. I get a value, and that value will be on the parabola. This will be the  $y$ -value associated to the  $x$ -value from the data. So this is the actual value. This is the expected value. We take the difference and square it. And because we want to minimize how we're going to solve this problem and this is what you'll have to do is, just like with the line situation, you're going to take the derivative of this function with respect to each of these three variables separately. So you'll take  $f$  sub  $a$ , and then you'll set that equal to zero. And then you'll take  $f$  sub  $b$ , you'll set that equal to zero. And then you'll take  $f$  sub  $c$ , and you'll set that equal to zero. Again, we set them all equal to zero, because this is really an optimization problem. We want to minimize something. We want to minimize this quantity.

So you take each of those three derivatives, partial derivatives, set them equal to zero, and you have a system of three equations with three variables. And you know how to solve such a system, by using matrices, for instance. That would be one way to solve such a system. And so then you can actually come up with a formula for the sort of parabola of best fit, you could think of it as. And so I'm going to stop there, because I think from here, you guys will do it. But I just want to point out this is where you're going to be starting this next sort of challenge problem to find the quadratic of best fit for these three data points. So that's where I'll end it.