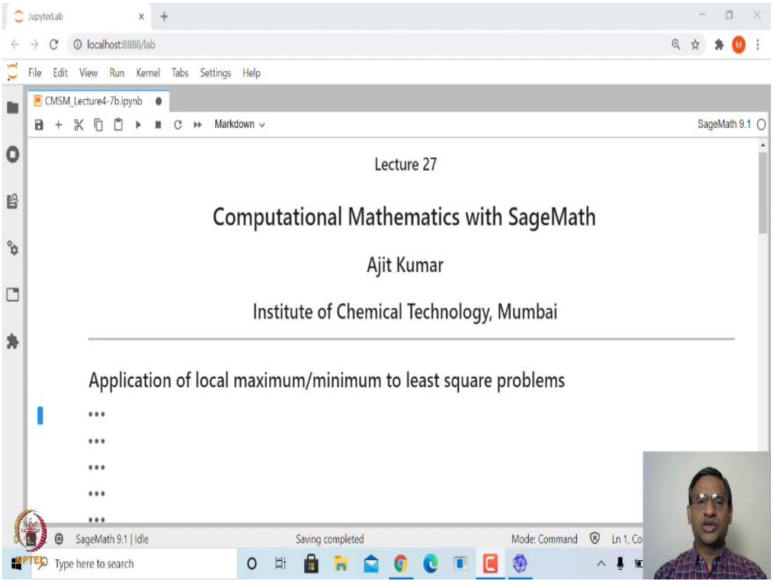
Computational Mathematics with SageMath
Prof. Ajit Kumar
Department of Mathematics
Institute of Chemical Technology, Mumbai

Lecture – 27
Application of local maximum and local minimum

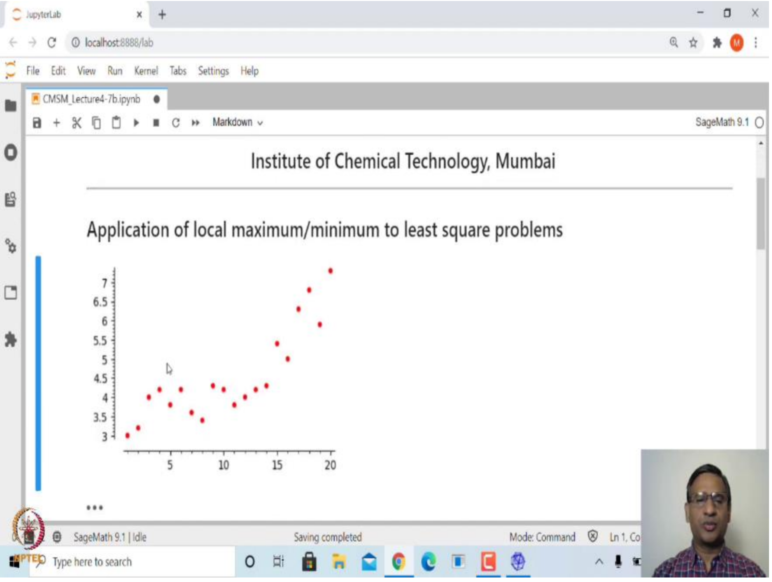
(Refer Slide Time: 00:15)



Welcome to the 27th lecture on Computational Mathematics with SageMath. In this lecture, we will look at an application of local maximum, local minimum to least-square problems. In the last lecture, we saw how to find the local maximum and local minimum of a function and in this lecture, we will look at one of its application.

It has many applications from which we will look at one such. This problem is called the regression analysis problem in statistics or this is also known as the best fit problem. Let us look at what it is.

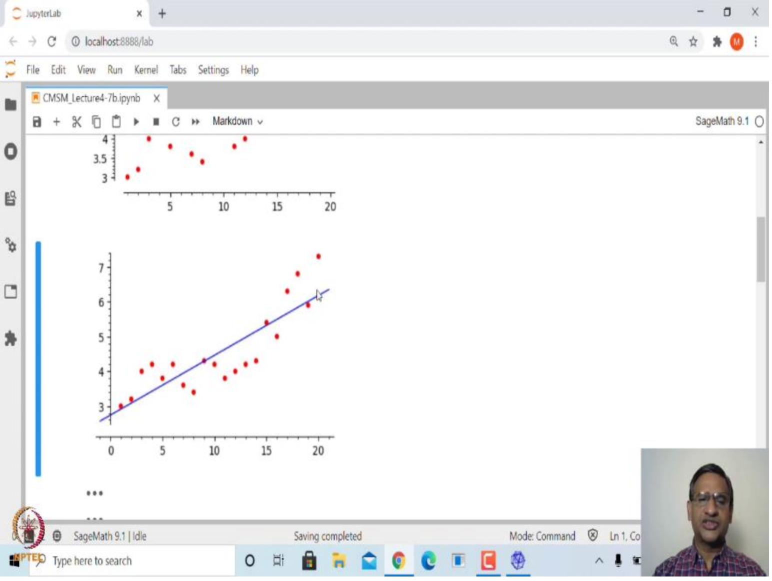
(Refer Slide Time: 00:59)



Firstly, suppose you have a set of points in the plane and you want to fit the best fit straight line to the set of points. For example, you can think of this as a real-life problem, suppose these are a set of villages in some area and the government wants to build a railway track.

Now, the railway track will be straight and also for this, these villagers are happy, so, they will have to find out what is the best possible way in which this railway track can be built. Now, first of all, what do you mean by building the best-fit railway track or finding the best fit straight line using this given set of points?

(Refer Slide Time: 01:47)



Suppose this is the best fit straight line. Let us say its equation is $y = mx + c$, where m is the slope and c is the intercept. Then what is the meaning of this best fit straight line? One way of defining is that to look at its distance from all of these points.

Let us suppose this is a point (x_i, y_i) and look at the vertical distance of this point from this straight line. If we look at this point (x_i, y_i), somewhere it will be x_i and somewhere y-coordinate of this will be $mx_i + c$.

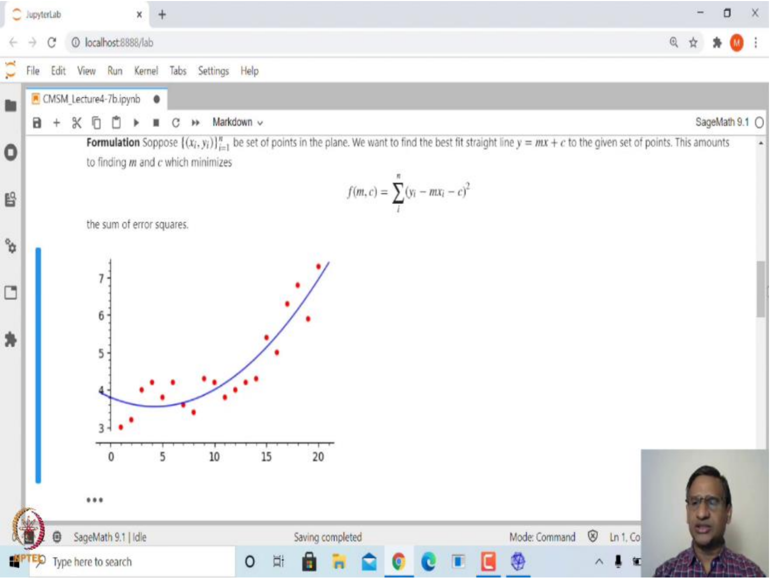
Look at the difference between the y-coordinate that is $y_i - mx_i - c$, this is the distance. But this could be a negative. For example, if I look at the point which is below the straight line, $y_i - mx_i - c$ will be negative. But the distance is positive.

You can look at the absolute value of this distance, i.e., the vertical distance of (x_i, y_i) and then you want to minimize this distance for each of this point. Now, this modulus is not a differentiable function. So, you can take the square of this distance that will be $(y_i - mx_i - c)^2$.

And this you want to minimize for every i ; that means, you can take the sum of the squares of the distance and then minimize that. This distance you can think of as an error term because this is the best possible straight line would have been when all these points are on this straight line.

That is the ideal situation, but that is not the case. The vertical distance will be the error. So, what you need to do is to find a, b or let us say m and c, the slope and the intercept for which the sum of the squares of the distances of this straight line from all these points is minimum.

(Refer Slide Time: 04:01)



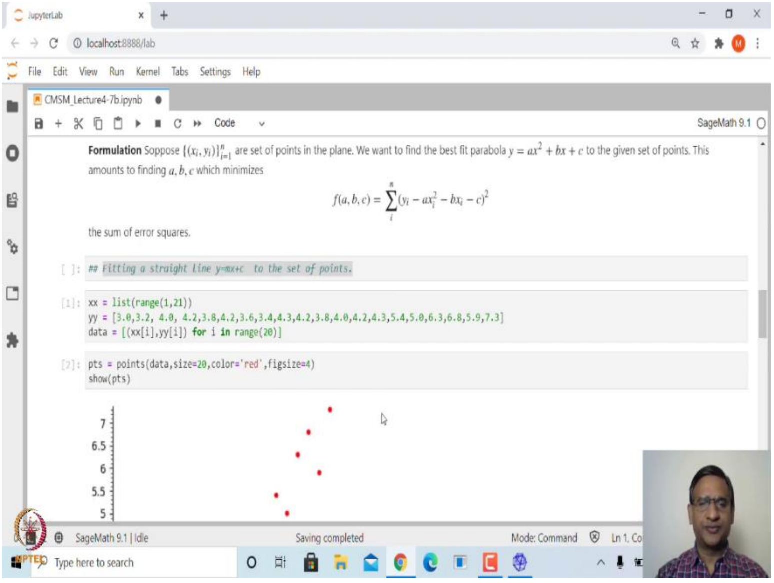
This is the same as saying, you can formulate this problem as follows-

Suppose you are given $\{(x_i, y_i)\}$, $1 \leq i \leq n$, the n set of points in the plane. We want to find the best fit straight line $y = mx + c$ to the given set of points. This amounts to finding m and c which minimizes this function $f(m, c) = \sum_{i=1}^{n} \quad (y_i - mx_i - c)^2$, this is sum of the error square. This is what you want to minimize. That is the best fit problem or linear least square problem.

You can also do the same thing instead of the straight line you wanted to fit a parabola say $y = ax^2 + bx + c$.

In that case, again you can find the square of the vertical distance, take the sum of the squares of the vertical distance and that sum of the squares of the vertical distance will be now a function of three variables a, b and c.

(Refer Slide Time: 05:15)

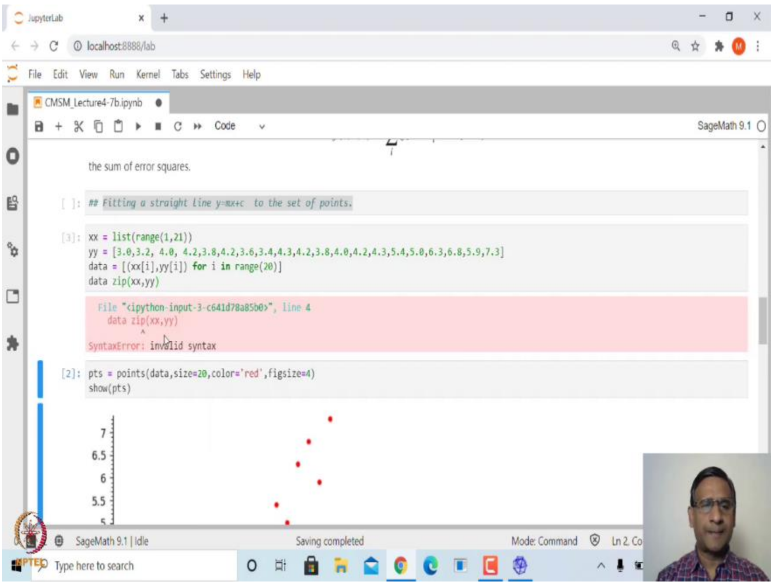


So, how do we formulate this? In this case, the formulation will be like you are given a set of n points and you want to find the best fit parabola $ax^2 + bx + c$. The vertical distance of the point i will be $(y_i - ax_i^2 - bx_i - c)$ and then take the square of this and then sum. That is the function of a, b, c, i.e., $f(a, b, c) = \sum_{i=1}^{n} \quad (y_i - ax_i^2 - bx_i - c)$^2. You need to find a, b and c which minimizes this function.

So, that is a minimization problem and as this is quadratic, it will have just only one minimum which will be the global minimum.

Let us take an example. Suppose at first you want to fit the best fit straight line $y = mx + c$ to a given set of points same as what we saw earlier.

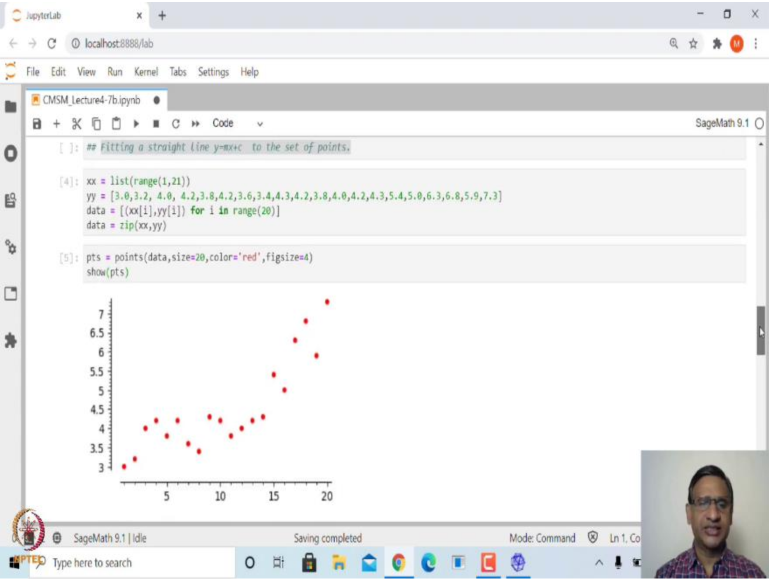
(Refer Slide Time: 06:14)



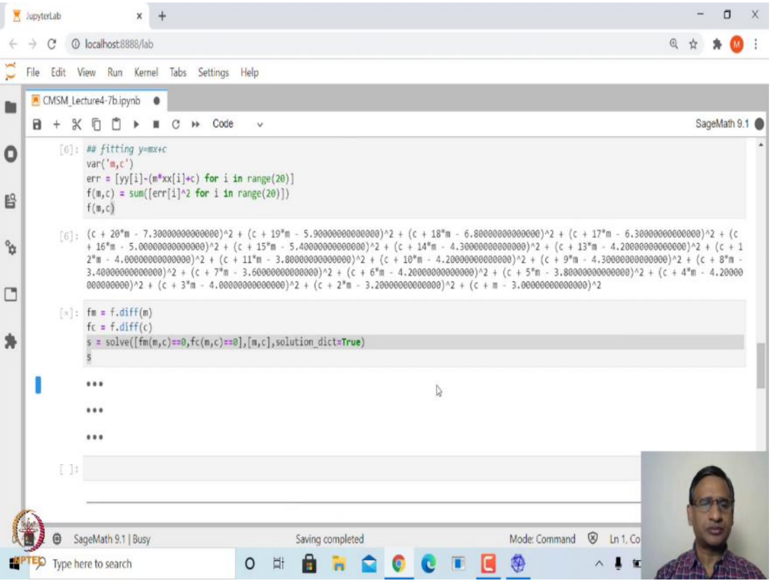
How we are defining? These are some random points I have taken. Here the x-coordinate I have put in this list xx that is, i.e., the range 1 to 21 will give you $1, 2, 3, \ldots, 20$. At 1, the y value is 3, at 2, y value is 3.2, at 3, y value is 4, at 5 y, value is 4.2 and so on.

That is stored in a list yy. Then you are looking at this data which is a tuple $(xx(i), yy(i))$, i going from 0 to 20 because there are 20 points. You can also define this as then you can just plot this set of points, this is another way of doing it.

(Refer Slide Time: 07:12)



(Refer Slide Time: 07:24)

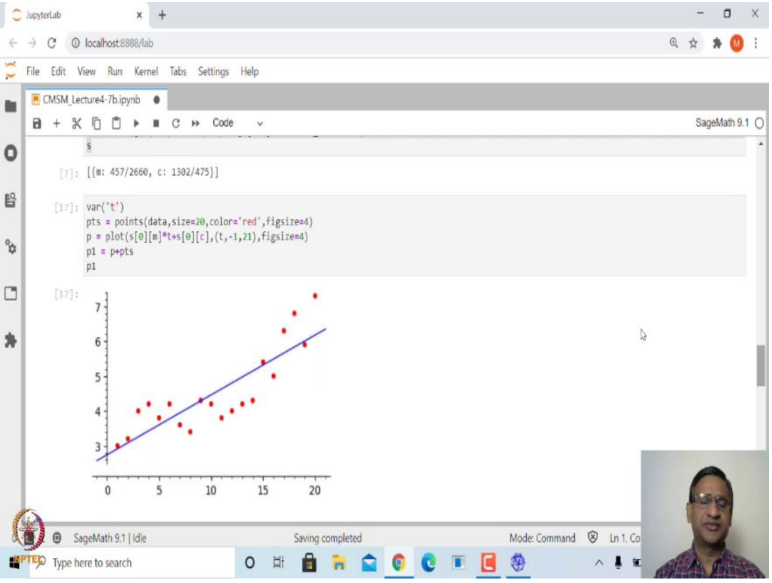


Now, let us define some of the error terms. Here, err is the error which is the vertical distance of $yy[i]$ from $(mxx[i] + c)$, the point on a straight line and then define the sum of the error squares.

Then the function looks quite big because it has written all the sum of the squares of the distance of all these 20 points. Now, what we need to do? In order to minimize this function, we need to find the critical points, which is the same as saying find the partial derivative of this function f with respect to m and c then solve these two equations simultaneously for m and c.

That is what is done here, the first-order partial derivative of f is carried out with respect to m and then with respect to c, and then solved this with a solution dictionary equal to true enabled. In this case, let us see what solution we get. We should get only one solution because this is a quadratic function. We can see it has only one solution m.

(Refer Slide Time: 08:33)

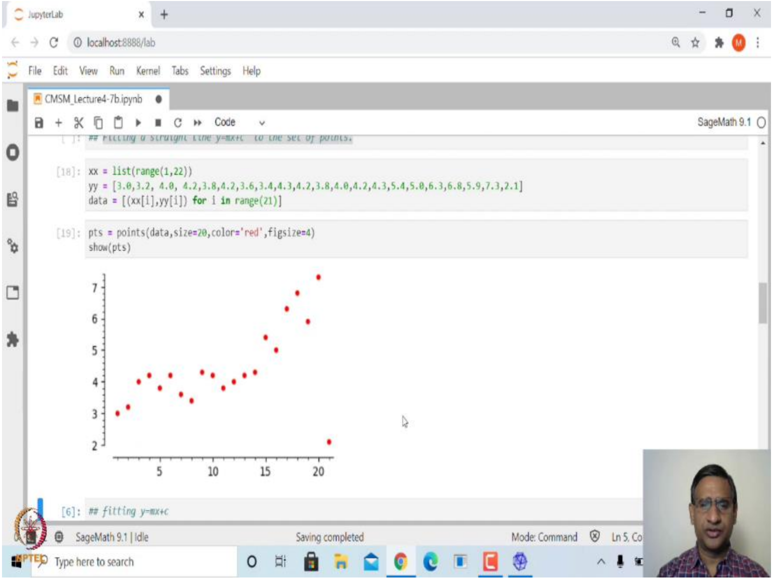


Let us plot the best fit straight line along with the set of points. This is already we have done. Here p is the straight line that is $y = mx + c$.

This is the value of m which is extracted from this solution dictionary and this is the value of c which is again extracted from the solution dictionary and then plot these two together. When we execute this, we get the best fit straight line that fits into this set of points.

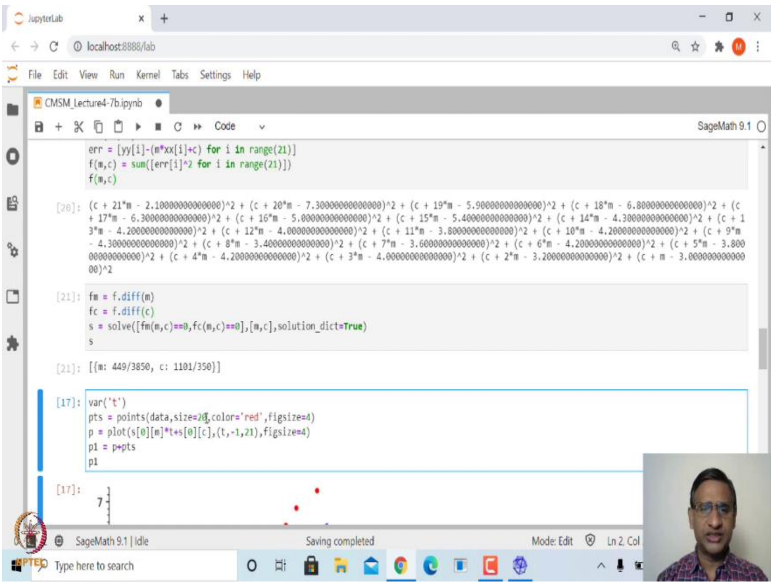
Now, if you add some more points in this and then try to find out. For example, let us just add some more points instead of 21, one more point that is 22 and in this case in yy let us put 2.1 and the range would be 21 data. The set of points with one extra point is plotted.

(Refer Slide Time: 09:23)



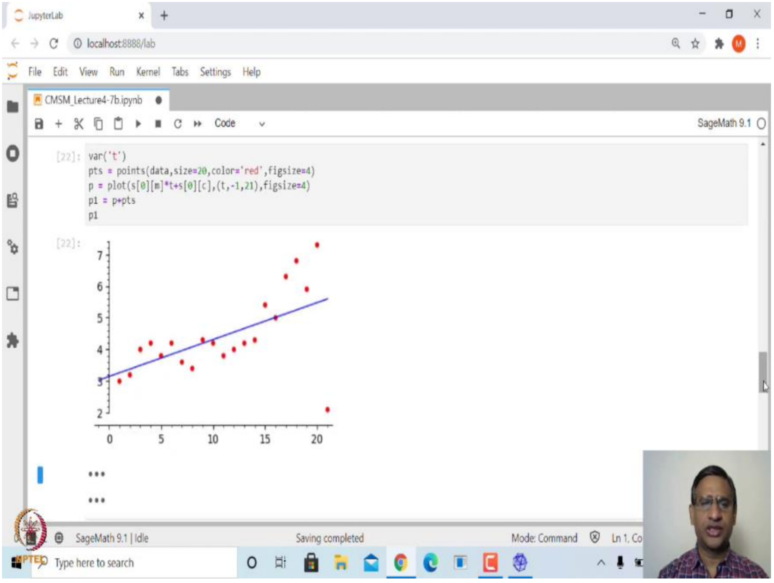
So, for example, let me just add some more points. So, let us say instead of 21 let us make it say let us just add one just one more point let us say this is 22 and in this case 7, I will put, for example, let us say 2.1. This is 2.1 and in this case, it will be 21. So, that the set of points so, one extra point which is plotted is the extra point that is given.

(Refer Slide Time: 10:10)

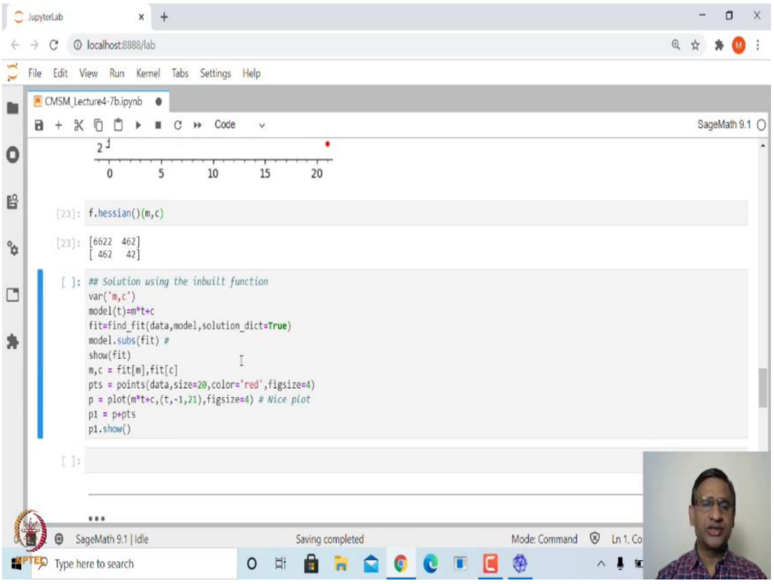


Next, let us solve this function value you can see here it has changed the solution. So, you can have as many points as you want.

(Refer Slide Time: 10:32)



(Refer Slide Time: 10:42)



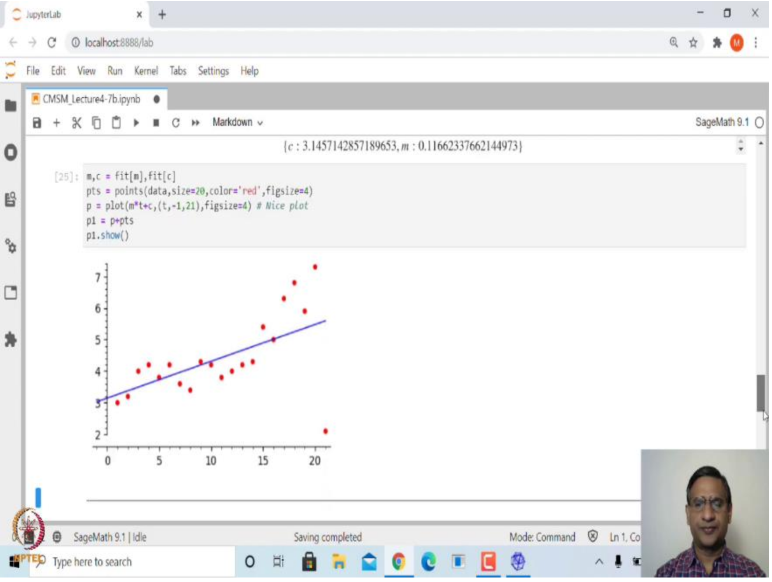
But you can also find hessian of f at m and c. This is independent of m and c because this is quadratic. Its hessian will be a constant, all the second-order partial derivative will be constant. You can see here this is positive definite.

You can check whether this is positive definite by using an inbuilt function or this is the determinant that is positive and this first term here is also positive. Therefore, point m and c which we have obtained is a local minimum. In fact, this is a global minimum because the function a quadratic, only one critical point is there.

Sage also provides you with an inbuilt function to do this fitting. Let us see how do we do that? Again, let us declare m and c as variables and declare a model which you want to fit. Here $mt + c$ we could have written $mx + c$.

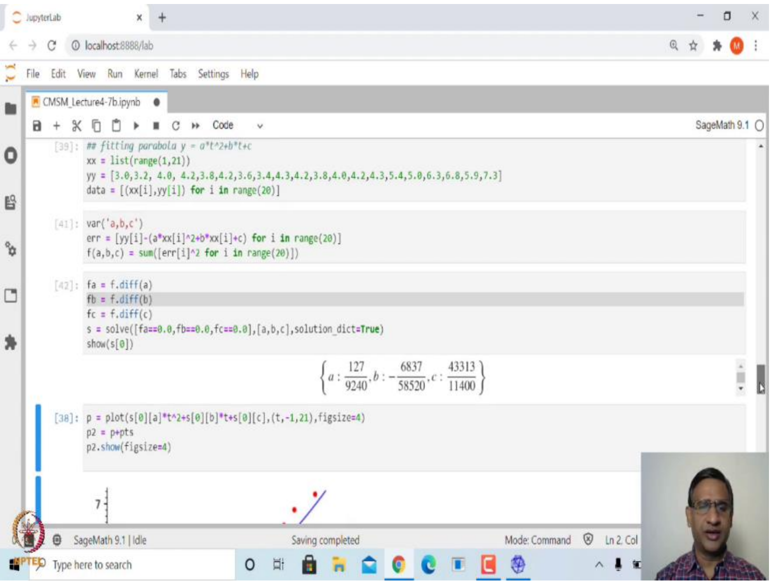
Then used a function called find underscore fit and give the data and then give the model which you want to fit. This is the solution dictionary equal to true. The solution will be printed inside the dictionary and then this is to show what are these.

(Refer Slide Time: 12:03)



This is the solution m is equal to this value and c is equal to this value. Now, we can plot all these things. Store the first value in m and the second value that is in c and then this is a set of points and then plotting the line and then plot the entire thing. That is how again we have found the best fit line using the inbuilt function. So, again you can even look at this function find the best fit taken the help on this and then see how it has been defined, you can look at the source code.

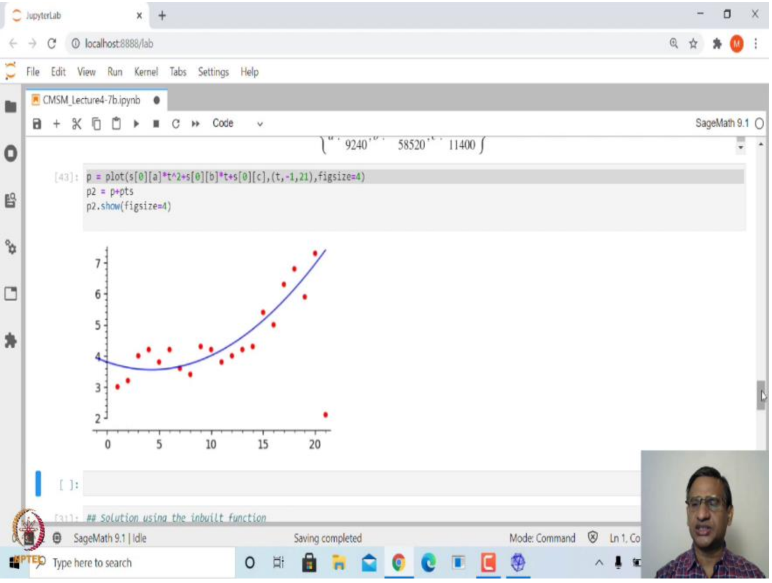
(Refer Slide Time: 12:41)



Now let us look at how to fit a parabola $y = at^2 + bt + c$ to the given set of points as we have done earlier. Again, let us just take the same set of points, just taking the 20 points, not adding the last point. Now, let us declare a, b, c as variables and define the sum of the error square as a function of a, b and c.

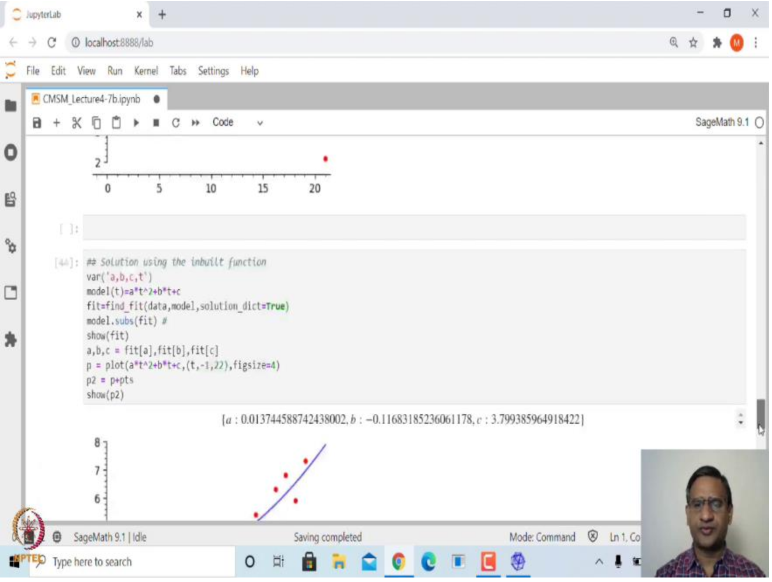
Then let us find the first-order partial derivatives of this and then solve for a, b, c. This is the solution you get equal to this value, b equal to this value and c equal to this value and then let us try to plot the set of points along with the fitted parabola.

(Refer Slide Time: 13:46)



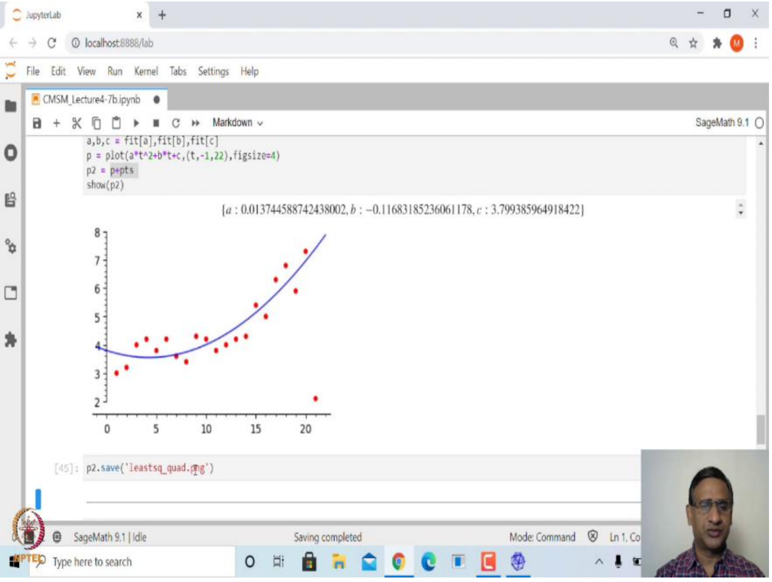
This is the value of a which we have obtained from this solution dictionary, this is the value of b, this is the value of c and then we plot the set of points along with the fitted parabola. So, this is how it looks like.

(Refer Slide Time: 13:53)



Again, you can see here in both these cases the values are the same and again you can use an inbuilt function. For example, in this case, you only need to change the model fit and you can just repeat the same code what we have done earlier. So, $model(t) = at^2 + bt + c$ and then find the fit of this model to this given data and then plot all these things together.

(Refer Slide Time: 14:16)

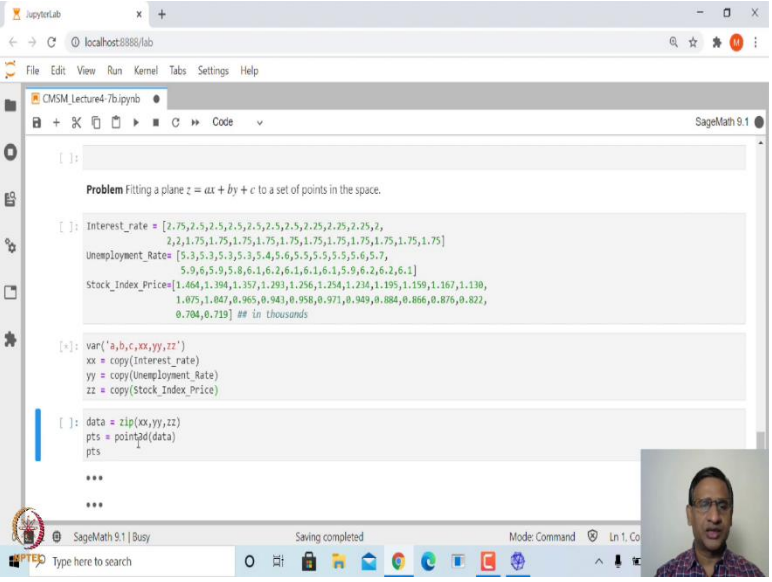


This is what you get. Again, you can see that these values will be the same as what we obtained using the calculus, i.e., using local maxima, local minima.

Also, if you want to save this figure, one way is you can right click and save it and other one is you can use the save command $p2.save$. So, $p2$ is the variable in which we have saved and stored this plot.

Then using $p2.save$ in the $leastsq\_quad.png$, you can save it in png format. Also, you can save it in pdf format then it will save in your default working directory you can find out.

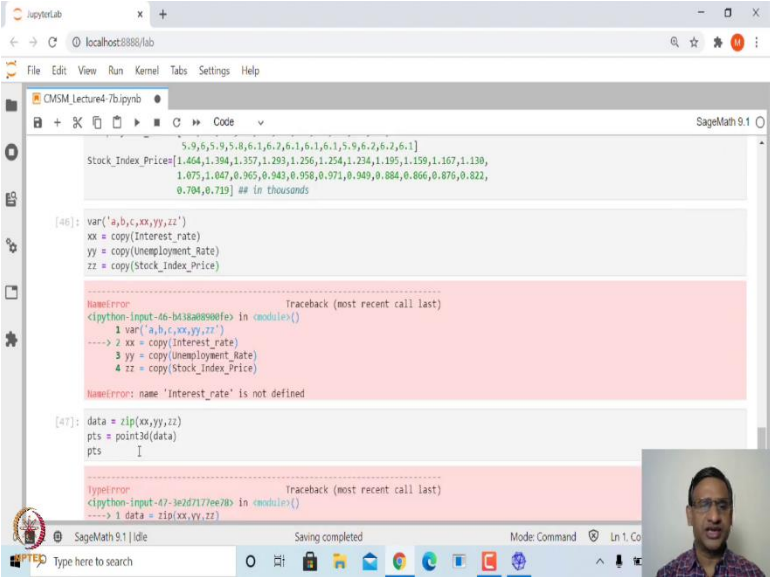
(Refer Slide Time: 15:02)



Let us look at another problem. This is a problem of fitting a plane say $z = ax + by + c$ that is a plane to the set of points in the space. As an example, let us look at a problem. Suppose you have an interest rate as a list, this is the unemployment rates and then the stock price in thousands. Let us say the stock price depends linearly on the interest rate and unemployment rate. We want to estimate what kind of relation it will have i.e., we want to estimate what should be a value of a, b, c which will be the best fit to the given set of data points.

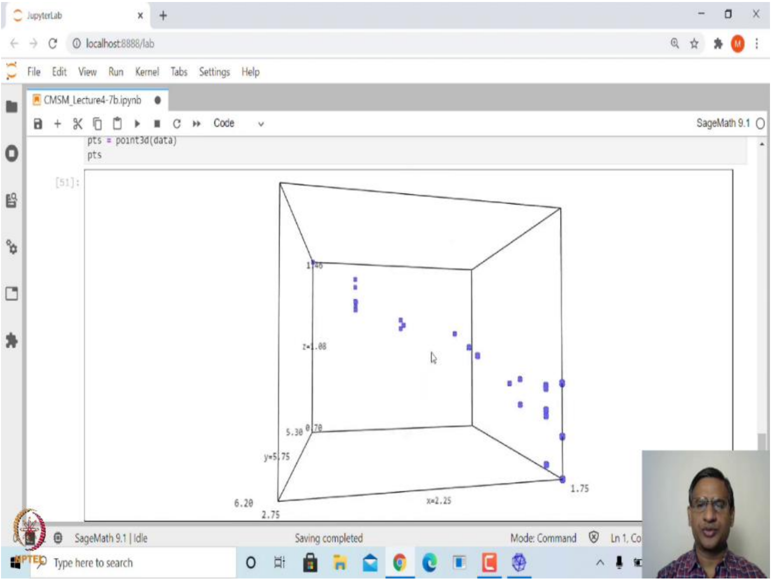
First, let us store these values interest rate in xx, the unemployment rate in yy and stock prices in zz. We have taken a copy of this, it is already explained what is meaning of copy in Python and what is its advantage over just writing $xx = Interest\_rate$.

(Refer Slide Time: 16:23)



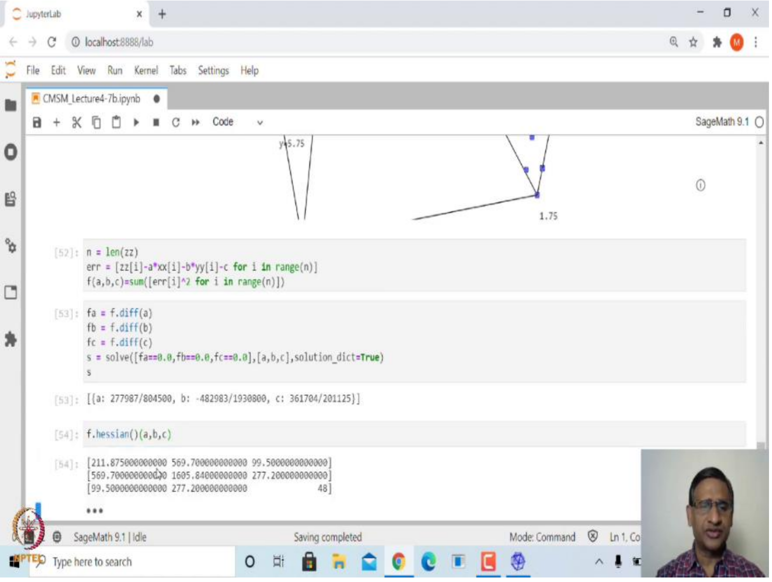
So, let us store this and now we can create this data using zip. Just a second, I think we did not run this. Let me now run this and then create the data and then plot this data.

(Refer Slide Time: 16:34)



This is the set of points in the space and to this set of points, you want to fit the best-fit plane. Earlier in two variables, we were looking at the best fit straight line and this can be generalized to the best-fit plane in three variables.

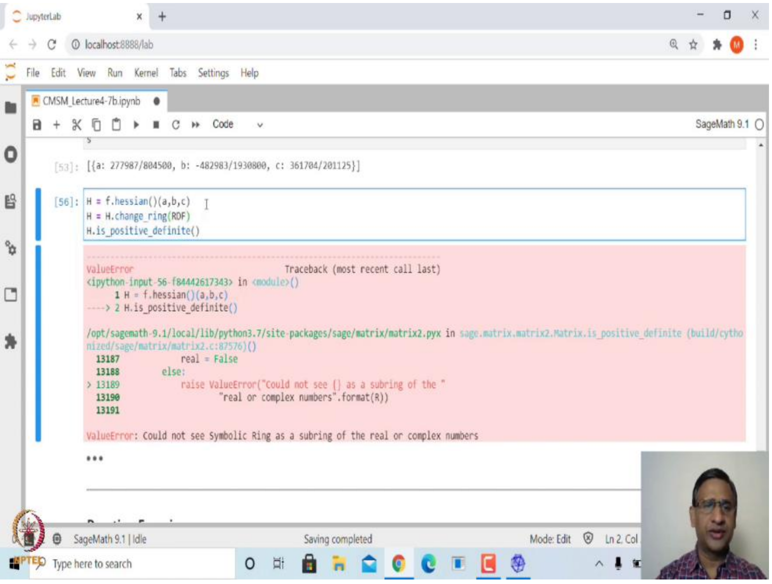
(Refer Slide Time: 16:52)



Now what we do? Again, we define the vertical distance i.e., the vertical distance of the point (x_i, y_i, z_i) from the plane $ax + by + cz$ and then take the sum of the error square that will be a function of three variables a, b and c.

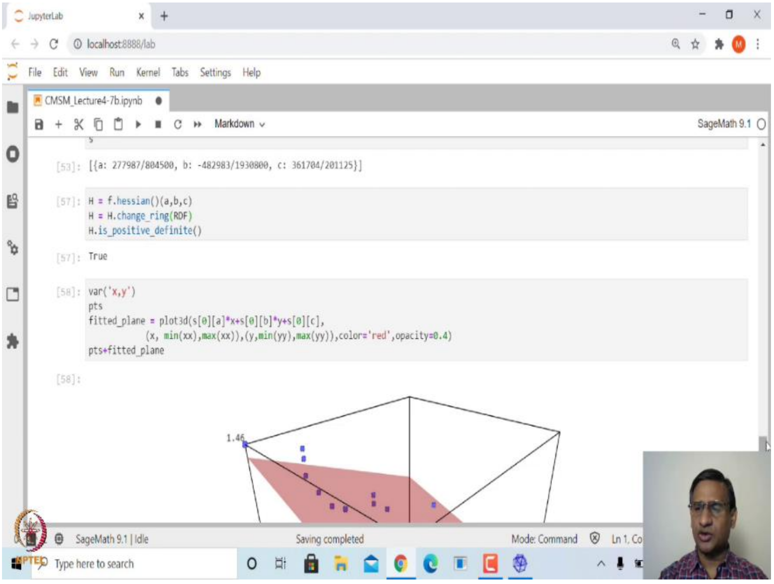
Once you have done this then find the partial derivative with respect to a, b and c and solve these three for a, b, c. These are the values you are getting and once you can find out hessian, it will be a constant again.

(Refer Slide Time: 17:38)



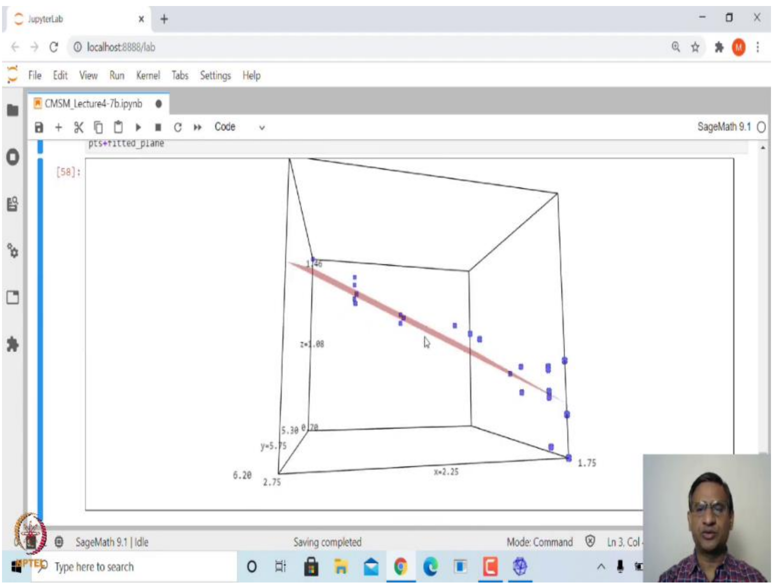
Also, you can check that it is a positive definite. For example, let us store this in H and check whether $H.is\_positive\_definite$. For this, you have to change the ring as this says that it is in a symbolic ring. So, first, we will define $H = H.change\_ring(RDF)$ and then we are getting this is $true$.

(Refer Slide Time: 18:11)



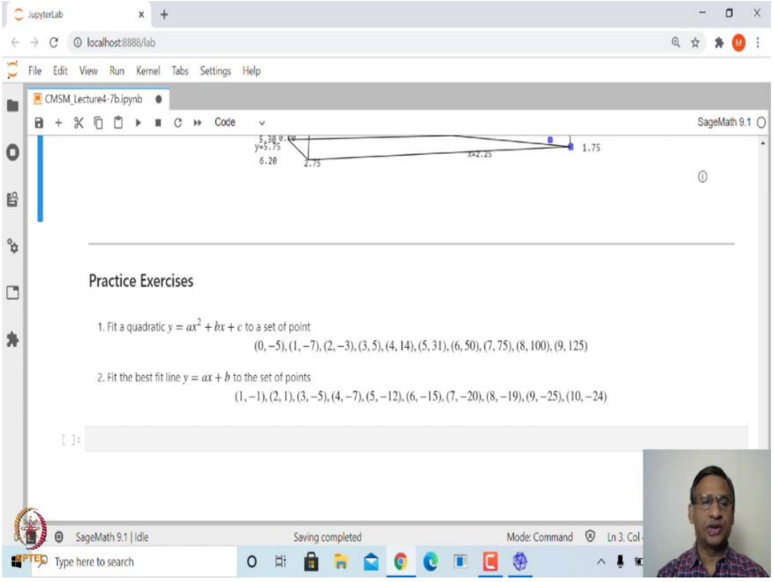
And, again you can plot the set of points along with the fitted plane.

(Refer Slide Time: 18:27)



So, the plane fitted is the best-fit plane to this given set of points. You could extend this to points defined in higher dimension also.

(Refer Slide Time: 18:40)



Let me just leave you with some practice exercises that are fairly straightforward. So, you want to fit a quadratic to this given set of points and best fit a straight line to this given set of points.

Thank you very much.