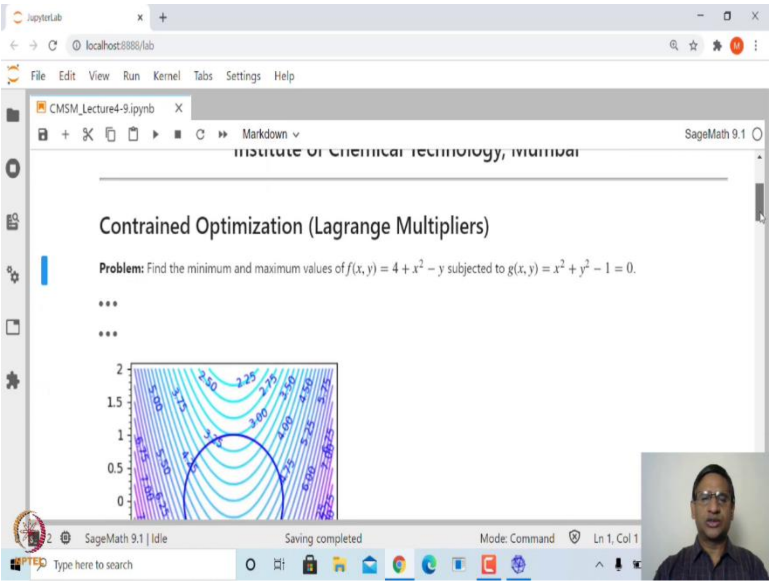
Computational Mathematics with SageMath
Prof. Ajit Kumar
Department of Mathematics
Institute of Chemical Technology, Mumbai

Contrained Optimization (Lagrange Multipliers)
Lecture – 28
Constrained optimization using Lagrange multipliers

Welcome to the 28th lecture on Computational Mathematics with SageMath. In this lecture, we shall look at finding minimum, and maximum value of a function under certain Constraints. This kind of problems are very frequent in our real life, and also in industry. Most often, the minimization/maximization problem that we encounter in our real-life are subjected to certain constraints.

(Refer Slide Time: 00:49)

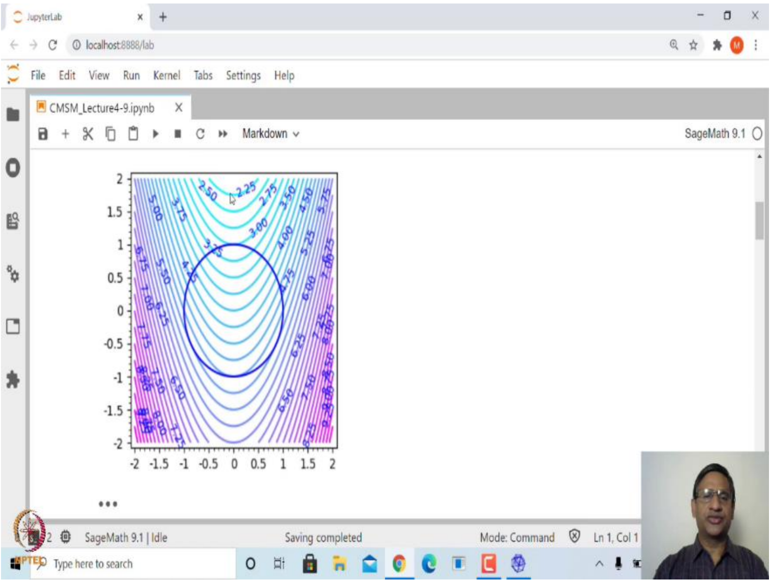


So, let us start with an example. Suppose we want to find minimum, and maximum values of f(x,y) is equal to 4 plus x square minus y, with constraints g(x,y) is equal to x square plus y square minus 1 is equal to 0. So, this problem is of finding the minimum or maximum value of this function on the unit circle; g(x,y) equal to 0 is nothing, but unit circle, right?

If you look at now the value of the function f(x,y); let us say suppose f(x,y) is equal to 2; that means, you are looking at a curve x square minus y plus 4 is equal to 2, that is same as saying x square minus y plus 2 is equal to 0. So, it's same as saying y is equal to x square plus 2. So, that is the parabola, right? So, for various values of f, you have set of parabolas. So, you want to look at what happens to this parabola on this circle.

Now, let us look at geometrically what it means. So, let us try to plot various graphs; a graph of this function f(x,y) for various values of x, various values of f, that is same as saying let us plot contours of f, and let us plot also the circle, and then see what happens.

(Refer Slide Time: 02:27)

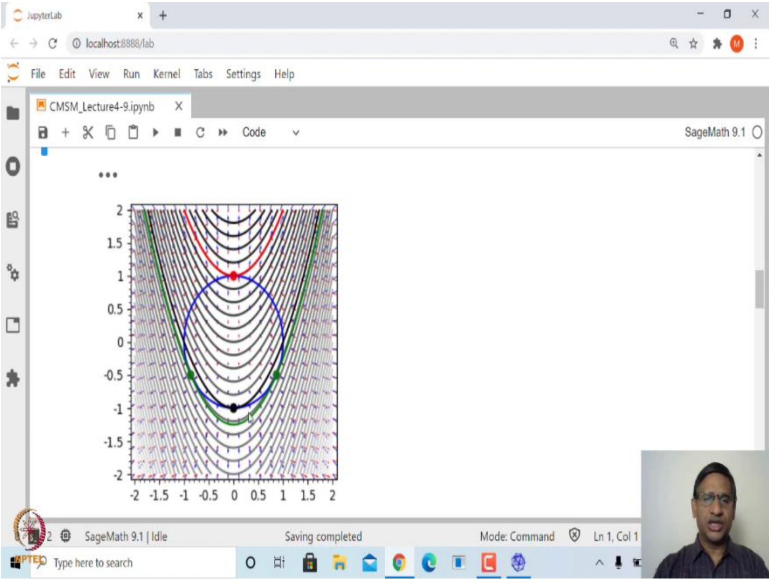


So, let us look at this, this contour plot. So, this blue curve is the circle, and this, these are all contours for f. You can see here, this is level curve at 2.5. This is at 2.50, this is 2.75, 2.2, 3.0, and so on. So, as you come down, the value of the function is increasing right now. So, what we want to find? We want to find the maximum value or the, and the minimum value.

So, if you can see here, maximum value of the function on this unit circle. So, you can see here for example, if I look at this curve for f(x,y) is equal to 3 at this point, value is 3. As we go little down, this function value will increase, right? So, this seems to be a minimum value.

And similarly if you, as you come down, as this, this parabola leaves this unit circle, for example, there is one here; this is touching this point at this stage, and this, I think it seems to be at 5.25, and again there is something below. For example, this, this, this curve, this also is, after that this will just leave this unit circle, and that seems to be at 5.25. So, 5.25 seems to be the maximum whereas, 3.0 seems to be the minimum, right?

(Refer Slide Time: 04:15)



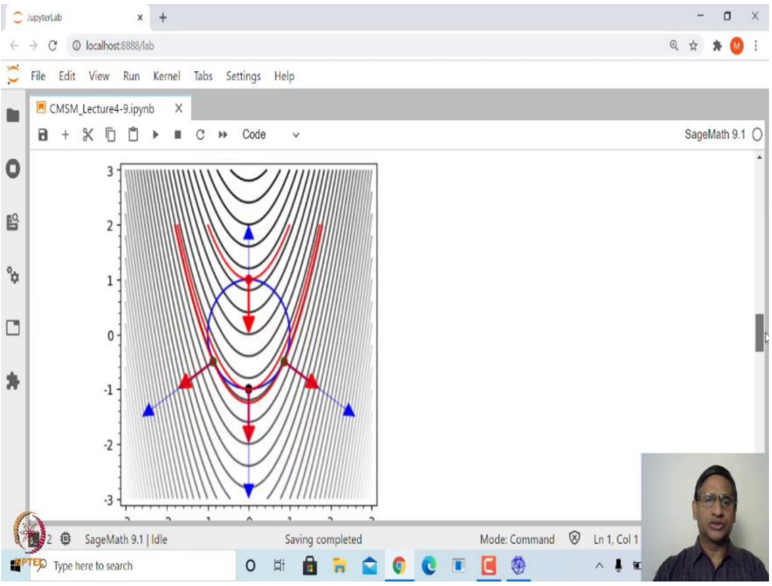
So, let us look at; so, this, this seems to be the 4 points at which the function, this contour of f seems to be touching this unit circle, and at this point when, when it touches these two, will have common tangent, right?

So, this seems to be the one perspective candidate for minimum, this seems, this seems to be one candidate for maximum, and at this point, the f and g have common tangent; is same as saying the gradient of f and g at this particular point, and at this particular point, they will be parallel to each, ok?

Just once second, they will be parallel to each other. Of course, we, we need to consider the condition because this g you want to be circle, you do not want something like some kind of wedge kind of curve. So, in particular you need to assume that the gradient of g, at least in the neighborhood of this point, should be non-zero. If it is non-zero everywhere, that is what is called regular curve, and in more, in higher dimension, it will be regular surface, right?

So, if you try to plot the gradient of f and g at this stage, you can see here, here, this gradient will be almost parallel to each other. Let me, let me plot in a, in a bigger; let me just plot gradient separately. Tthis is plotted using vector field plot.

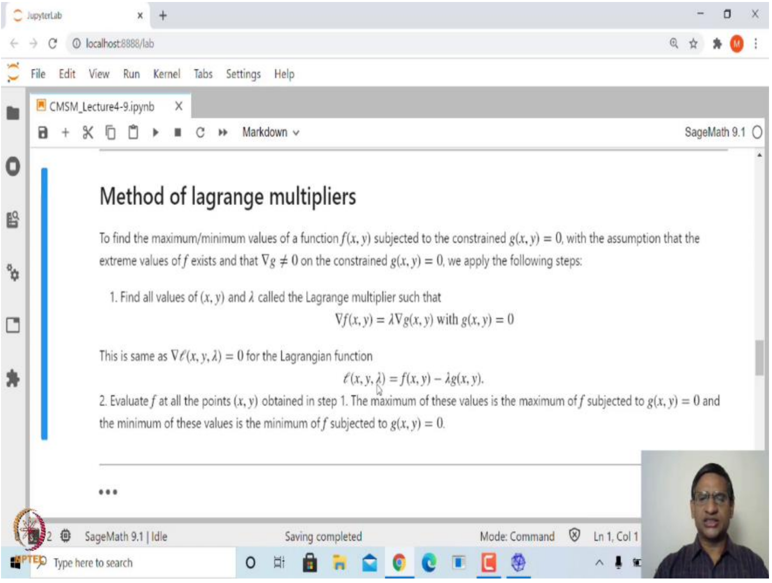
(Refer Slide Time: 05:56)



So, if, if you look at here, so, at this, at this point, you can see here, this is the blue one is gradient of g, and the red one is gradient of f, and gradient of f and g at these 4 points are parallel to each other. So, actually that is the geometric meaning. So, in particular, we are looking for a point on this unit circle that is g(x,y) equal to 0, at which gradient of f and gradient of g are parallel to each other. Of course, provided this, we assume that the condition that gradient of g at that point is non-zero, ok?

So, that actually gives a necessary condition for a point f x for a point (x,y) to be minimum or maximum value of f subjected to g(x,y) is equal to 0. In case such a point exists, then we can, we can show that the gradient of f and gradient of g will be parallel to each other, ok?

(Refer Slide Time: 07:04)



So, let us just formulate that. So, this method of solving such problem is known as Lagrange multiplier. So, let us see what it is. So, in order to find maximum or minimum value of a function subjected to g(x,y) equal to 0. In this case, we are looking at only the equality constraints.

Inequality constraints can also be converted into equality constraints. So, in principle, one can take care of such problems as well using this. However, the size of the problem will increase, and therefore, there are certain methods to, to handle inequality constraint separately, ok? So, you are looking at finding maximum and minimum value of f(x,y), subjected to g(x,y) equal to 0. Geometrically, we saw what it means, and let us assume that g, gradient of g is non-zero, ok?

So, the, the set of points at which g(x,y) equal to 0 is a curve; such curve is called a regular curve. So, you are looking at a point (x,y) on this regular curve such that, and so, in case a point (x,y) is the point of minimum or maximum, then we can find a lambda, a real number which is known as Lagrange multiplier, such that, gradient of f(x,y) is equal to lambda times gradient of g at (x,y), and of course, that point is on, on this g(x,y) equal to 0. So, this, this is the equations.

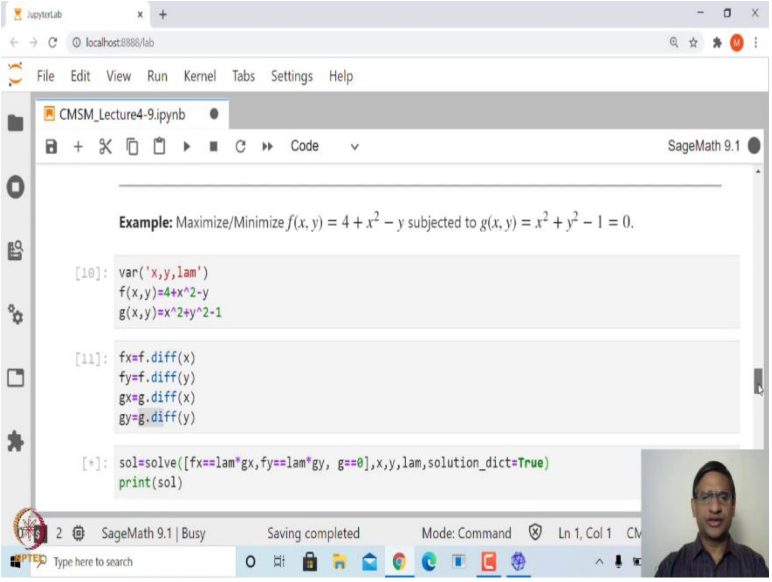
So, if you look at the first one, grad f is equal to lambda times grad g, that will give you two equations; one with partial derivative of, with respect to x, and one partial derivative with respect to y, and this is the third equation. So, you are looking for solving these three equations in three variables x, y, and lambda right, ok?

Now, this you can also rewrite as follows. One can think of this equation as gradient of l(x,y,lambda) equal to 0, for a function l which is l(x,y,lambda) is equal to f(x,y) minus lambda times g(x,y). This is what is called Lagrangian function. So, you can, this, this particular equation, these three equations is same as, or equivalent to gradient of l is equal to 0 for Lagrangian function f(x,y) minus lambda times g(x,y). So, that is another way of looking at this problem, right?

So, once you get these values of x, y, and lambda, what we are concerned with? Only the, the lambda, sorry (x,y) value. So, you, it is possible that you may get more than one points. So, at those points, evaluate the function value, and then choose one which gives you the maximum, and one which gives you the minimum, if you want to find maximum and minimum value respectively, right? So, that is the procedure. This is the, these are the steps, ok?

And of course, there are second order necessary and sufficient condition to check once you get a point (x,y) which is prospective candidate whether that is maximum or minimum. There are necessary and sufficient conditions for checking those critical points, but we will not be going into this. In case you are interested, you can look at some book on optimization technique, and that will give you this entire theory of Lagrange multipliers, right?

(Refer Slide Time: 10:40)



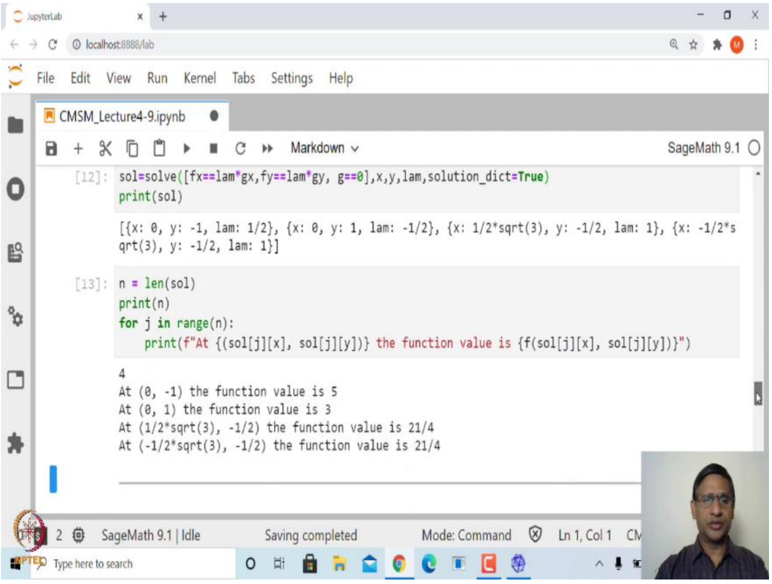
So, let us work out the same example which we were looking at, geometrically. So, this is the problem. Maximize or minimize; there is a spelling mistake, maximize or minimize f(x,y) is equal to this, subjected to g(x,y) equal to x square plus y square minus 1 is equal to 0. So, here when I say maximize or minimize, it's same as saying find a point (x,y) at which f(x,y) is maximum or minimum, right?

So, that is what you, or if you want you can reformulate this as find the, the maximum or the minimum value of f(x,y) subjected to g(x,y) equal to 0, ok? So, first we will, we have to actually solve this set of three equations; that means, we need to define a variable x, y, and lambda, and f(x,y) is equal to 4 plus x square minus y; g(x,y) is equal to x square plus y square minus 1.

So, once we have defined these variables, and then declared these functions, now we need to find these gradient of l is equal to 0. So, we can define Lagrangian function, and then define the gradient, or we can separately write this equation, equate the first-order partial derivative of, of l with respect to 0, that will give me first-order partial derivative of f with respect to x, minus lambda times first-order partial derivative of g with respect to x.

Similarly we can do with partial derivative with respect to y, right? So, let us find partial derivative of f with respect to x, partial derivative of f with respect to y, partial derivative of g with respect to x, and partial derivative of g with respect to y, and store them in fx, fy, gx, gy, right?

So, once we have obtained this, then now we need to solve the equations fx is equal to lambda times gx, fy is equal to lambda times gy, and g(x,y) is equal to 0 for x, y, and lambda. And let us enable this solution dictionary is equal to true. So, we will get solution as a list of as a list of dictionary. So, that is the, let us run this, we will; once we run this, then it is still running. You can see here, this star means it is running, yeah? (Refer Slide Time: 13:05)



It has found the solution, and you can see here, it has four solutions. One is this where x is equal to 0, y is equal to minus 1, x is equal to 0, y equal to 1, that is what we saw it seems to be minimum of the function. This is x equal to 3 by square root 3 by 2, y is equal to minus half, and last one is minus half square root 3 minus half, ok?

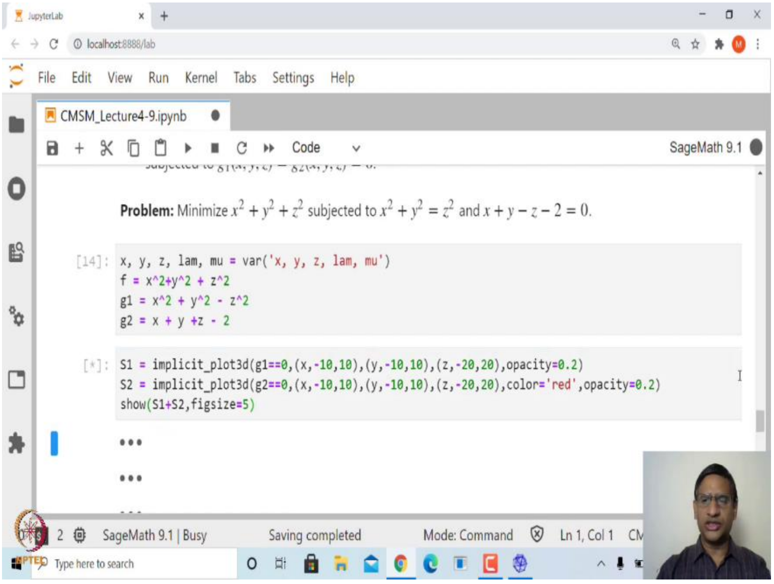
So, now let us evaluate the function value at these points. So, first of all let me declare that n is the number of solutions of this, and then run a loop on each of these solutions. So, for j in range 4 in this case, take the, the point the x-value, and y-value of the solution, and then find the value of the function at this point, and let us print all of these things.

So, first, this is the number of solutions, and at 1 minus 1, the function value is 5. At 1, 0, the function value is 3; at half square root 3 comma minus half, the function value is 21 point 21 by 4 which is 5.24, and at this point also, this function value is 5.24. So, this, this is a point of minimum, and this will be a point of maximum, and the minimum value is 3, and maximum value is 5.24. That is what we saw geometrically if you go back, and see this, this contour plots.

So, this was the, the minimum, and this is the, the maximum value. This is, that this point is the point of maximum, and at these points, the gradients are parallel to each other, ok? So, that is the, that is how we can solve minimization and maximization problem of two variables in, subjected to single equality constraints.

Now, let us, and of course, you can extend these two function of three variables, and more variables, the procedure will be same. In case if we have function of three variables, this, this grad f is equal to lambda times grad g will give you three equations, and then g(x,y) equal to 0 is the fourth equation. If it is n variables, this will give you n plus 1 equations in n plus 1 variables: x1, x2,…, xn, and lambda, right?

(Refer Slide Time: 15:50)



Now, let us look at, suppose you have two constraints, we want to find minimum and maximum value of the function subjected to two constraints. So, let us look at this problem. You want to minimize x square plus y square plus z square subjected to x square plus y square is equal to z square, and x plus y minus z minus 2 is equal to 0.

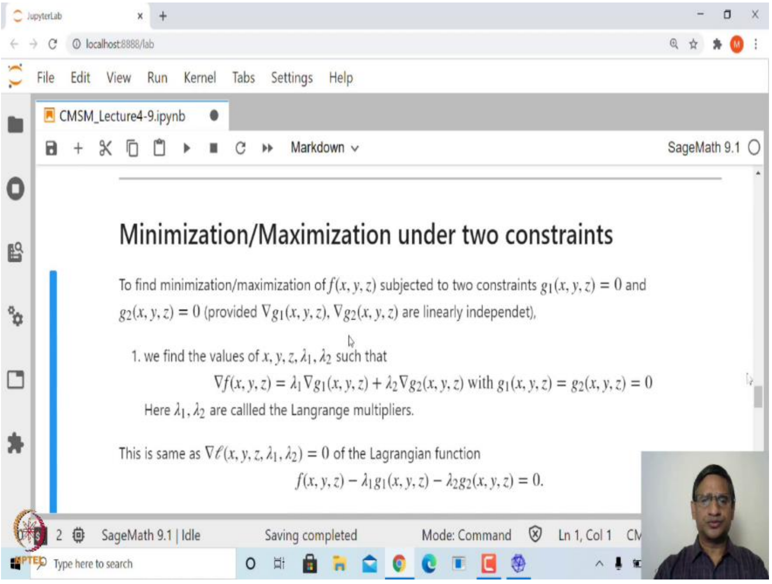
So, if you look at this constraint, this first constraint actually gives you a cone, and this is a plane. So, you are looking the set of points which satisfy these two equations, is nothing but the intersection of this cone with this, the plane. So, what you are looking at? You are looking at minimum/maximum value of, in this case minimum value of this function.

If you look at, this is nothing, but square of the distance of any point (x, y, z) from the origin. So, you want to minimize, or find the point at which this distance is minimum, and that lies on this intersection of these two surfaces right; that is the geometric meaning of this.

So, if you try to plot the, the level curves of these, these two constraints, let us see what we get. So, first let us declare x, y, z, and lambda, and mu; so, instead as two variables.

Before that let me, let me also explain how, what are the steps involved in solving this kind of problems.

(Refer Slide Time: 17:19)



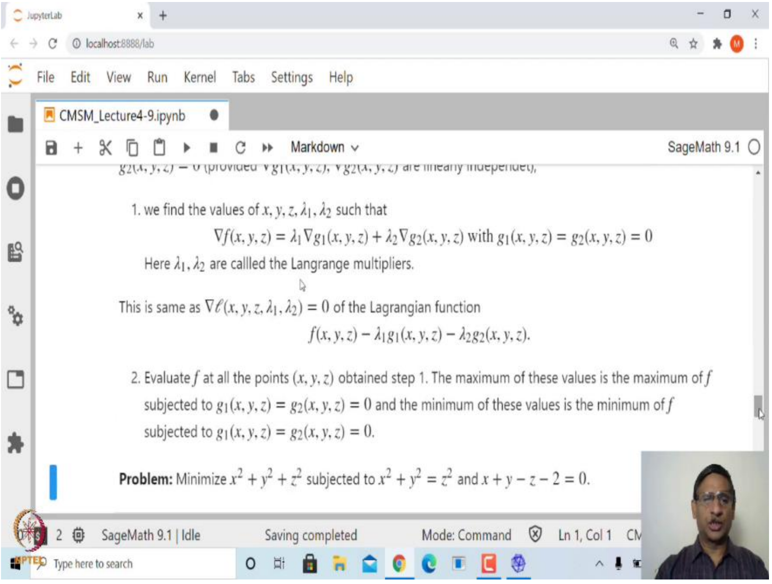
So, steps involved in solving this kind of problem are as follows. In order to find minimum or maximum value of f(x,y) subjected to two constraints g1(x,y) is equal g1(x, y, z) is equal to 0, and g2(x, y, z) is equal to 0. In this case again, we will be assuming that grad g1, and grad g2 are linearly independent.

This should be non-zero, and that linearly independent. Of course, the the first step is again to if x, y, z is point at which the minimum or maximum value occurs, then one can show that there exist two real numbers; lambda 1, and lambda 2 such that grad f at (x, y, z) is equal to lambda times grad g 1 at (x,y,z) plus lambda 2 times grad g 2 at (x,y,z).

And of course, g 1 g 2 should be equal to 0 that is the constraint, right? So, here lambda 1, and lambda 2 again are called Lagrange multipliers. And so, that is the the first order necessary condition for a point (x,y,z) to be minimum or maximum of f subjected to g1 equal to 0 g2 is equal to 0.

So, and again this this particular set of equations can be thought of as grad l (x, y, z, lambda 1, lambda 2) is equal to 0 where l is f(x,y,z) minus lambda 1 g 1 minus lambda 2 g 2 which is called Lagrangian. So, this should not be equal to 0 here. This is the the Lagrangian, ok?

(Refer Slide Time: 19:05)



So, and once you find solution of this set of equations so, so how many equations you have? You have 3 equations here, and you have 2 equations here. So, 5 equations, and 5 unknowns x, y, z, and lambda 1, and lambda 2. Once you obtained this solution, then evaluate the function value at these solutions.

And then take one which gives you the minimum, and one which gives you the maximum. So, these are the steps involved. So, only in the previous one this, this term was missing.

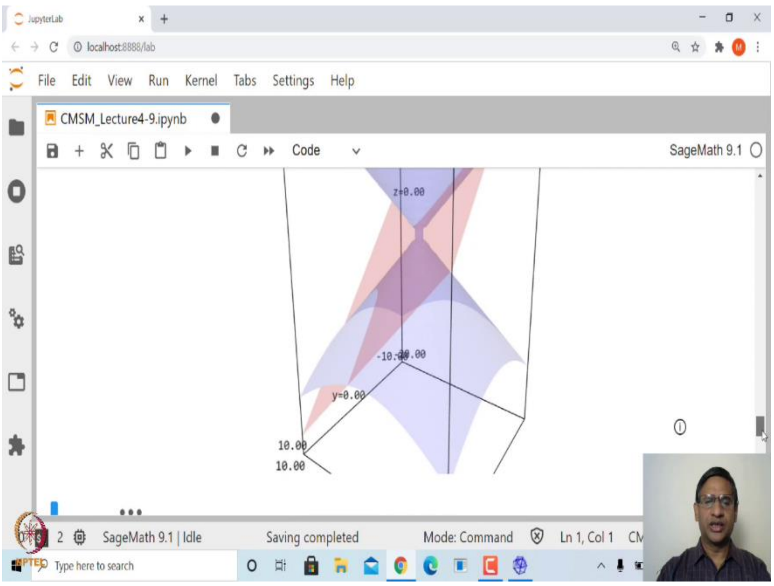
So, you can now see that in case you want to generalize with more constraints, you will have so many Lagrange multipliers. That is the generalization.

And again in this case also you one has second-order necessary, and sufficient condition for a point to be maximum or minimum subjected to the, these set of constraints. So, we will not get into the nsecond-order conditions, but we will simply find a set of points which satisfy these equations, and check, and then take one which gives me the maximum value or minimum value.

Now, if you look at this, this problem, finding the, the minimum distance of x, minimum distance from the origin on the curve of intersection of these two surfaces. In this case, actually this, this cone, and this surface is in has infinite extension so, you may not expect this, these to have the point of maximization. That may not be the point which will give me maximum value, it could go to infinity.

So, let us plot graph of these constraints first. So, first, declare the variables x, y, z, w lambda, and mu. Instead of lambda 1 and lambda 2, I am just calling it as lambda lam, and mu. So, these are the, the declares. Next, let us plot these two surfaces which satisfy this g1, and g2 is equal to 0.

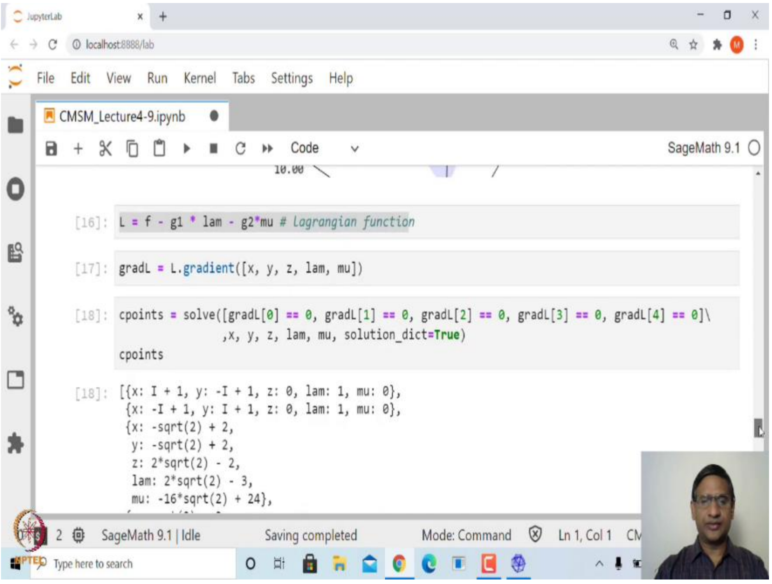
(Refer Slide Time: 21:08)



So, when you plot these surfaces, this is what you get. So, let me just rotate this, and you can see here, this curve of intersection of this cone with this plane; that is the, the curve of intersection; that is the curve of intersection. So, you want to find this is somewhere, here it is in the origin.

So, from origin, you want to look at the distance of, of point on the curve of intersection, and you want to find the minimum value of the function, right?

(Refer Slide Time: 21:42)



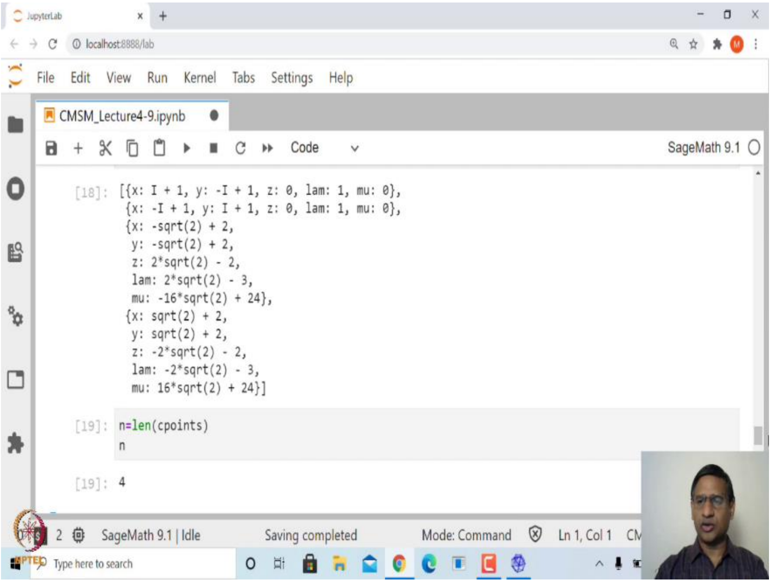
So, let us look at the steps. So, first we will declare this Lagrangian, that is L is equal to f minus gl, g1 times lambda, and minus g2 times lambda times mu. So, that is the, the Lagrangian function.

Now we will find the gradient of L, the gradient of L with respect to x, y, z, and lambda, and mu. So, when we, when we find this, this is the Lagrangian. I am calling this in gradL, L dot grad, gradient at x, y, z comma lambda comma mu. So, now let us solve this grad g. gradL is equal to 0 is same as saying each coordinate equate to, to 0.

So, gradL 0th coordinate that will be partial derivative of l with respect to x equal to 0. The second coordinate equal to 0 is same as saying partial derivative of l with respect to y equal to 0. This is partial derivative of l with respect to 0, and partial derivative of this with respect to lambda will give you, actually g1 is equal to 0. Partial derivative of l with respect to mu equal to 0 will give you g2 is equal to 0.

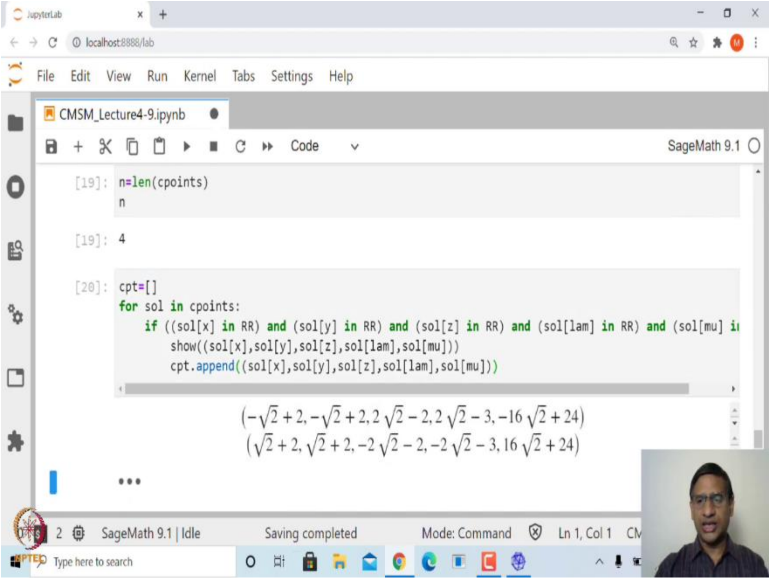
And then solve this for x, y, z, lambda, mu, and declare the solution dictionary equal to True. So, these are the, the critical points.

(Refer Slide Time: 23:12)



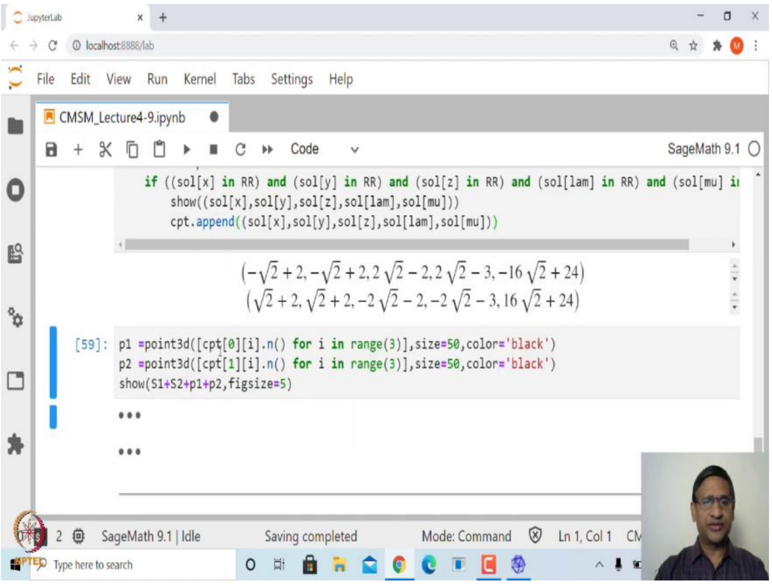
And if you go through this, there are some points which are not real. So, for example, x value is 1 plus i, which is not real. So, we will not consider such points, we will not consider such points. So, we will consider only when this solution is real. So, first, let us find out how many solutions we, we have We have 4 solutions in this case, but not all of them are real. So, we will extract only the real solutions.

(Refer Slide Time: 23:43)



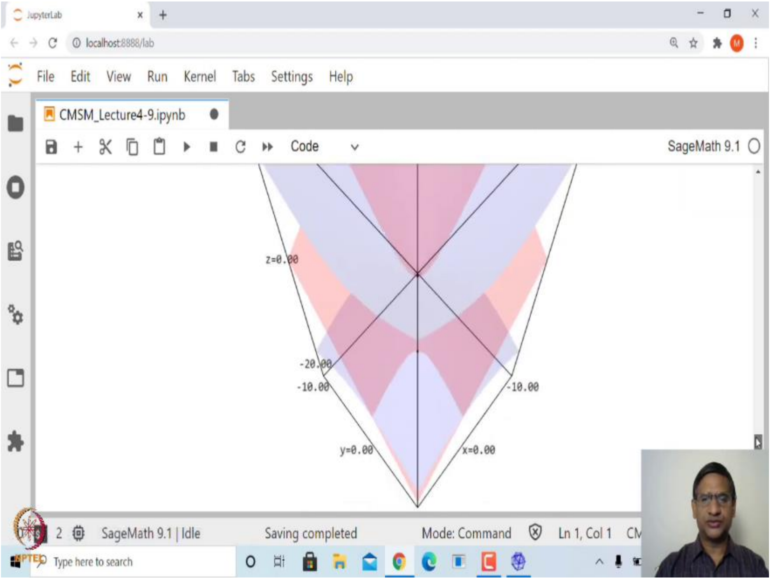
So, this is a small Sage code to extract only the real solution. So, what we are doing? We will extract these real solutions, and put it in, in a list cpt. So, first, we have declared this as empty list, and then solution in this set of critical four critical points as a list of dictionary, and then look at whether the, this x value is real, y value is real, z value is real. If all of them are real, then you show that x, y, z, lambda and mu, and then you append that in cpt, and so, these are the, the two real solutions; only two real solutions. And these two real solution if you look at graphically, it should be, one should be somewhere here, and other one should be somewhere here, ok?

(Refer Slide Time: 24:40)



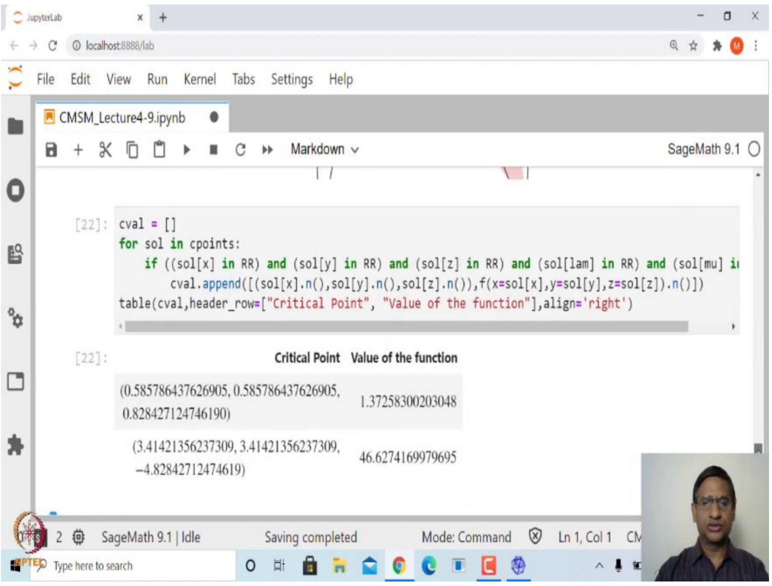
Now, if we try to find the function value at these points; if we try to find function value at these points so, before that let us plot these two points along with the constraint surfaces.

(Refer Slide Time: 24:49)



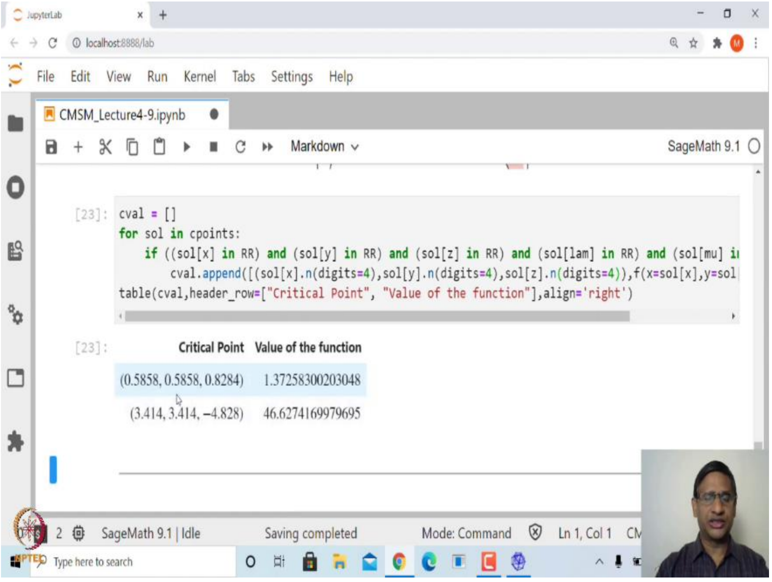
So, when you plot this, and let me rotate. So, if these are the two points, the perspective points, and I mean geometrically it is quite clear. This will be point at which function value will have smallest, and this is the bigger one, right?

(Refer Slide Time: 25:08)



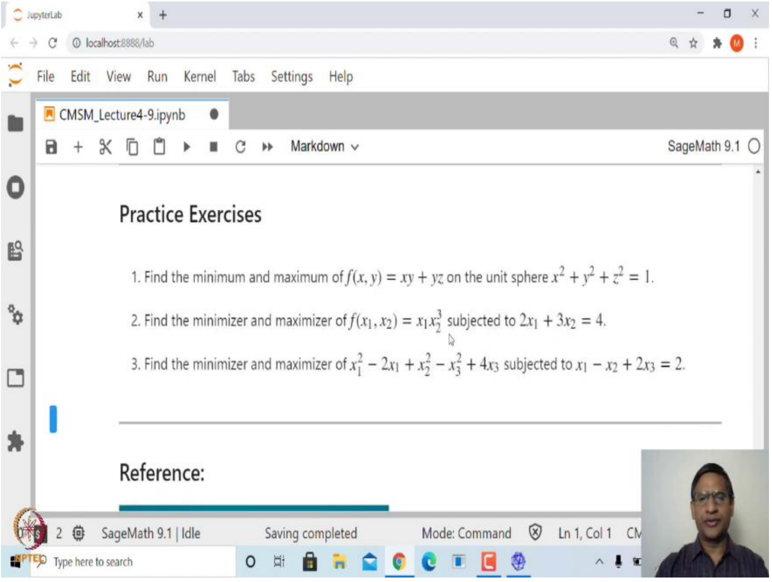
So, let us tabulate the critical points along with the function value. So, this, these are the, the points in the values in the decimal.

(Refer Slide Time: 25:32)



If you want, I can reduce the number of decimal. So, how do I reduce? I will say here digits is equal to let us say only 4 digits; let us show only 4 digits, here also 4 digits, here also 4 digits, and yeah so, this is, these are the two, two points, and this point is a point at which the function has minimum value. Of course as we saw geometrically, this may not have any maximum value, right? So, you can try to solve more problems on more number of constraints. You can extend this to function of n variables with as many constraints as you want.

(Refer Slide Time: 26:02)

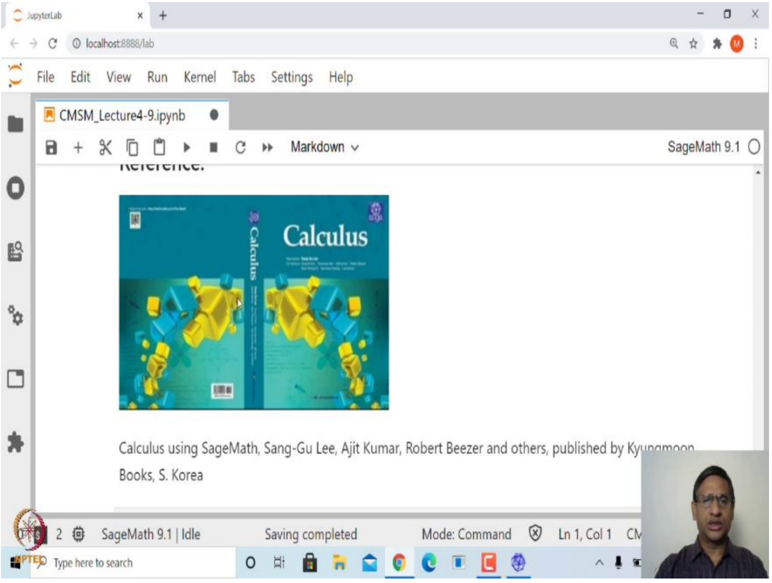


Let me leave you with some practice exercises. These are the quite straightforward ones. So, you are expected to solve these problems using SageMath.

So, 1st is to find the minimum, and maximum value of this, subjected to this unit sphere, and the 2nd one is to find minimum and maximum; minimizer or maximizer of f, which is x 1 into x 2 square subjected to this constraint, and the 3rd one is in function in three variables, right? So, these are the straightforward practice exercises. Try to solve them, and of course, also try to plot the graph to get geometric idea.

So, these are things we wanted to cover in calculus of 1 and 2 variables. Of course, there are other concepts like integral of multivariate functions. Similarly, the vector calculus etcetera; all these things can be done in, in, using SageMath. However, because this is a short course, I thought basic introduction will be fine, and remaining things you all can explore on your own.

(Refer Slide Time: 27:26)



If you want to look at some reference you can look at this particular book on Calculus. This uses SageMath, and this is a book by Sang-Gu Lee, myself, and Robert Beezer, and there are few other authors, and this has been published by Kyung Moon publication in South Korea, and this is available from Amazon.

So, you can go through this book, and it gives you all the Sage codes, starting from very basic one variable calculus to multivariable calculus including vector calculus, ok? So, from next class, we will start learning or exploring concepts in linear algebra.

Thank you very much.