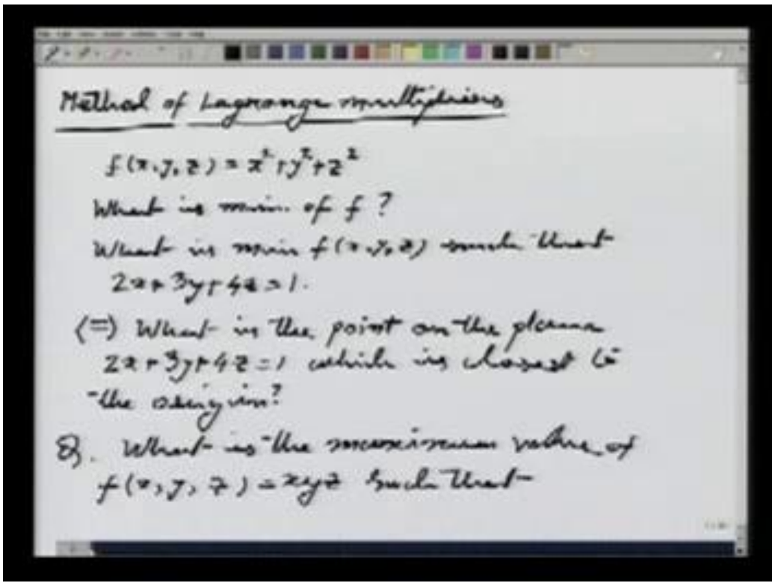
Mathematics - I
Prof. S.K Ray
Indian Institute of Technology, Kanpur

Lecture - 27
Method of Lagrange Multipliers

(Refer Slide Time: 00:32)



In today's lecture, we are going to talk about something related to maximum, minimum; it is called the Method of Lagrange Multipliers. So, what are the kind of problems, we are going to deal within this is as follows. Let us look at the following function defined on r 3 given by f of x, y, z. This is very simple function, x square plus y square plus z square. Now, suppose, I ask you, that what is, minimum of f? Suppose, I ask this question, then obviously, the answer is 0. Because, x square plus y square plus z square is always bigger than or equal to 0. The minimum possible value then is 0 and the minimum is achieved at 0, 0 and 0. So, that is very simple.

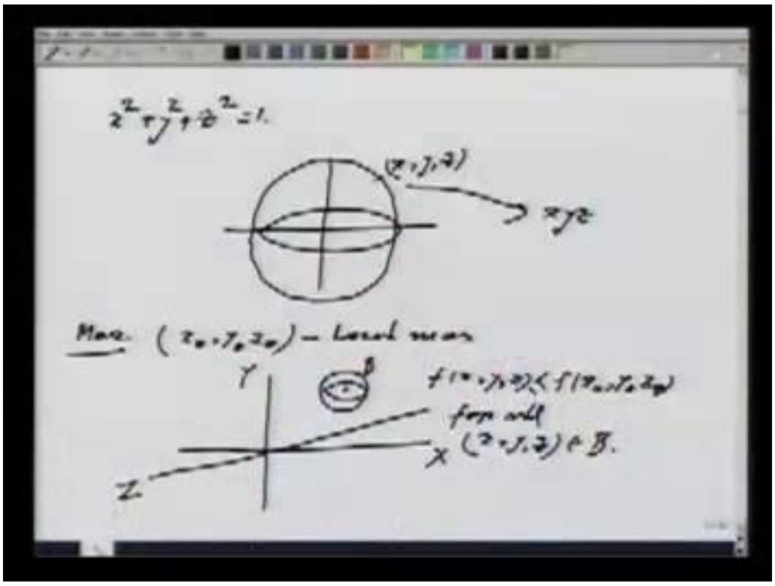
But, suppose I ask you, that what is, minimum of f x, y, z, such that, 2 x plus 3 y, plus 4 z equal to 1. See, now the question is very different. Because, the global minimum of the function is at the 0, 0, 0, but that point, does not satisfy the given equation, 2 x plus 3 y plus 4 z equals to 1. So, the method of Lagrange multipliers, essentially deals with this kind of problems. That is you want to find extremum of a function, subject to the condition, that x, y, z is on some surface.

In this case, my surface is a plane. So, geometrically speaking, what I am asking is, that this is same as asking. That what is the point, the point on the plane 2 x plus 3 y plus 4 z equal to 1, which is closest to the origin. Usually, what you would do is, take any point x y z on the plane 2 x plus 3 y plus 3 z equal to 1, then see what is the Euclidian distance of that point with the origin, which is square root of x square plus y square plus z square. It is enough to minimize the square of this function.

That is, x square plus y square plus z square. And you want to find the minimum value of x, y, z; such that this function is minimum. So, this is a typical example of a constraint extremum problem. Let us look at another example. Let us say, what is the maximum value of the function f of x y z? Let us say, equal to x, y, z. Now, a moment thought would tell you, that this maximum of the function must be infinity.

Because, it is certainly an unbounded function, because once I fix a value of y and z, x I can choose as large as, I like. So, you cannot certainly find out a real number, which is the maximum value of this function. But, at the same time I can add a constraint on it. That is, what is the maximum value of this function such that, x square, plus y square, plus z square equals to 1.

(Refer Slide Time: 04:39)

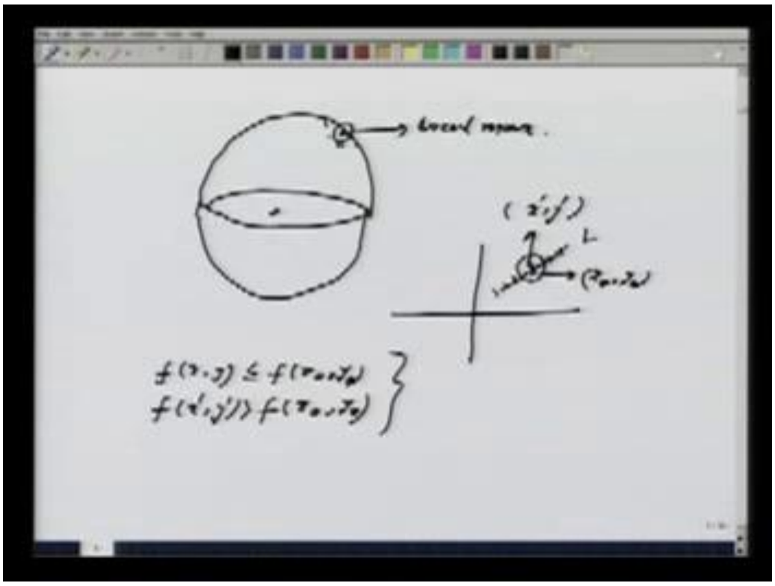


That is, what I am saying is now that look at the sphere in three dimension. Now, a function f, is defined on this, given by, if x, y, z is a point on the sphere, then it is going to x, y, z. I am asking, what is the maximum of this function. Now, in all these questions, there is something very involved here. In three dimension, what exactly we mean by, let us say maximum or a local maximum as we say, x naught, y naught, z naught. This point is the local maximum.

What does this means? That means around this point? So, if I draw the picture in three dimension. Let us say, that these are the three axis x, y, z. Suppose this is my point, x naught, y naught, z naught, this is the point x naught, y naught z naught. Then, this point is a local maximum, means what it means there exist, a ball, which looks like this. Around the point x naught, y naught, z naught, such that for all points in that ball, f of x y z is strictly less than or equal to, let us say f of x naught, y naught, z naught.

If I call this ball b; for all x y z in b, this is the minimum meaning of local maximum. That a point is a local maximum means; there exist a ball around that point. Such that, for each point in that ball, the value of the function is less than or equal to the value of the point, where I am saying the function as the local maximum.

(Refer Slide Time: 06:56)



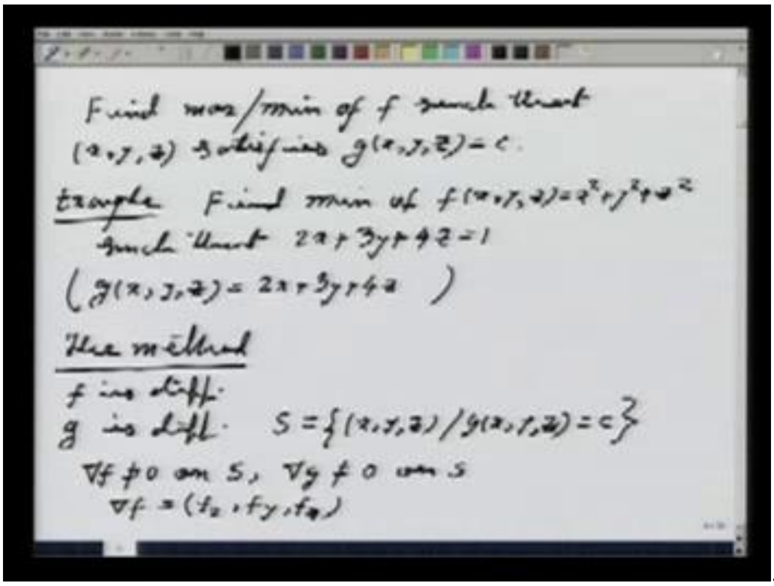
Now, if I look at the constraint situation. That suppose, I am on the big sphere, then what do I mean by, f has maximum at this point. That means, there exist some portion of the sphere, which is given by the dotted line. Where, the value of the function is less than this point. Now, notice that, this has nothing to do with the usual three dimensional balls. For example, if the function has a local maximum at this point, let us say. Then, among all the neighboring points, the value of the function at this point is bigger than the other points.

So, this point will then suffice as the local maximum on the sphere, no problem about that. But, it might happen, that the function has a minimum here and this point, which I am looking at is not at all a minimum of the function. But, when I am looking at the constraint situation, in that situation, the value of the function at this point is less than the value of the function at the neighboring points. I can very well see a two dimensional situation, that look at this line.

Look at this point, suppose the value of the function at all the other points, it is smaller than these points. If I call this point x naught, y naught, then if I call this line L, then f of x y, let us say is lesser equal to f of x naught, y naught, this can happen. But, whenever, I look at any point very close to this. That means, if I look at a disc around that point, and then whatever disc I choose there exist a point here, inside the disc.

There exists a point here, inside the disc, such that f of x prime, y prime. If I call this point x prime, y prime, here it is bigger than f of x naught, y naught. This is precisely, what happens in three dimensions, it can happen, in the case of constraint maximum or minimum. So, the abstract of the whole discussion is, that our usual method of finding maximum minimum is not going to work in the case of constraint maximum, minimum. That means, I am searching for extremum of a function. Where, the points where the function is, the points where I am concentrating I am, on, they are actually varying on some surface.

(Refer Slide Time: 09:47)



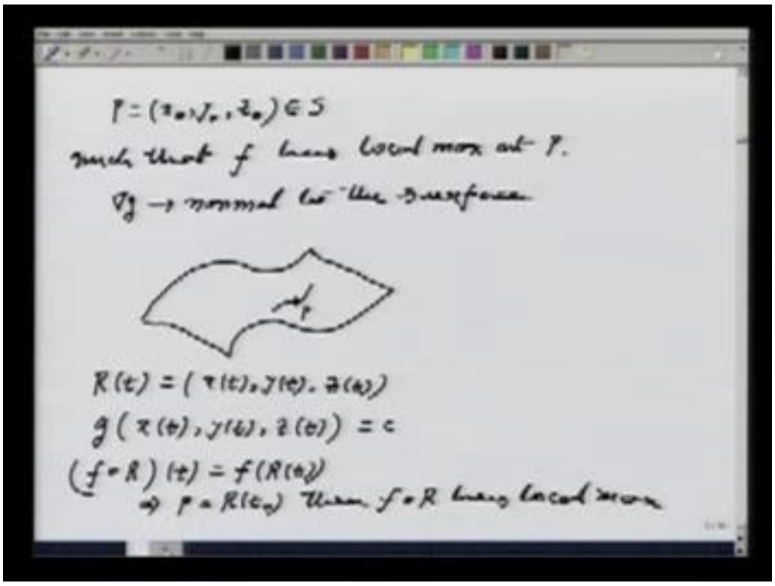
.

So, a typical statement, of a constraint extremum problem is something like, find maximum or minimum of f; such that, x, y, z satisfies g x, y, z equal to some constant. So, a typical example, which I have mentioned, that find minimum of f x, y, z equal to x square, plus y square, plus z square, such that 2 x plus 3 y plus 4 z equal to 1. So, in this case, g x, y, z is given by 2 x plus 3 y plus 4 z. So, g x, y, z equals to 1 gives me the surface, which is a plane and I want to minimize the function, on these planes.

So, the question is, how does one go about it, to have to develop a total new method. Now, the point is, first to get hold of, what is the collection of points, where the extremum can lie. And that is actually, what is suggested by the method of Lagrange multipliers. And once, you get those collections of points, among those points, now you have to search, examining the function to see whether you can locate the maximum or the minimum.

So, let me now, describe the method. So, I assume f is differentiable, g is also differentiable. I define a surface s, which is x, y, z, such that g x, y, z equal to c. I will also assume that grad f not equal to 0 on s and grad g not equal to 0 on s, where we have the obvious meaning of grad f, that is the partial derivatives f x, f y, f z. So, that means, typically, I am dealing with real valued functions, here f and g. Now, first, I have to get hold of the extreme points.

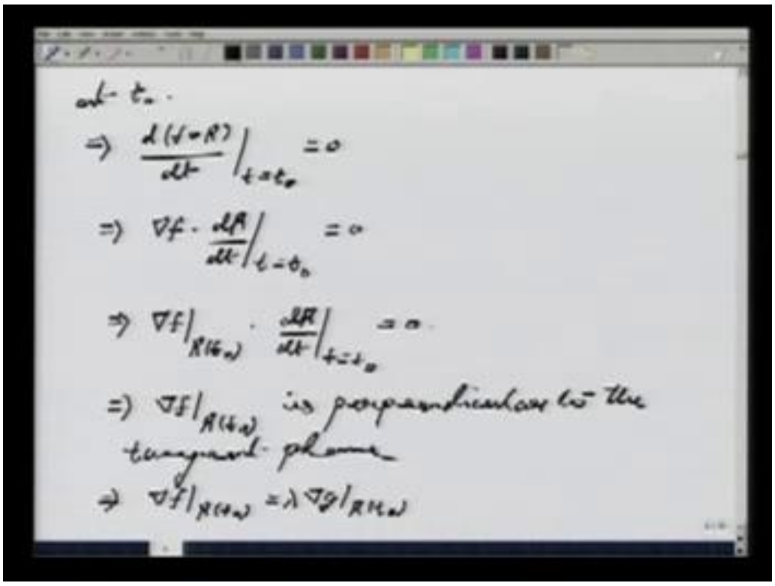
(Refer Slide Time: 12:50)



So, let us say p equal to x naught, y naught, z naught. Let us say, this is a point on s, such that f has local max, let me work with local max only. Similarly, you can work with local minimum, at p. Now, the first thing about s, what I know is, that grad g, if I look at any point on the surface, this is the normal, normal to the surface. Now, suppose, p is a local maximum at p. Look at the situation here, I will just draw some surface for you, it looks something like this, let us say.

And I have a point p here. I look at any curve passing through p. Then, let us say, r t is the curve. So, the curve r t looks like x t, y t, z t; this is a curve and since this curve is on the surface, I also know that g of x t, y t, z t is a constant c. Now, look at the function, f compose r, by that I mean, f of r t, look at this function. I know at the point p, f has a local maximum at p. That means, if I look at this is my point p. If I look at the neighboring points on the curve, the value of the function at p is always bigger than the value of the function at the neighboring points, because p is a local maximum.

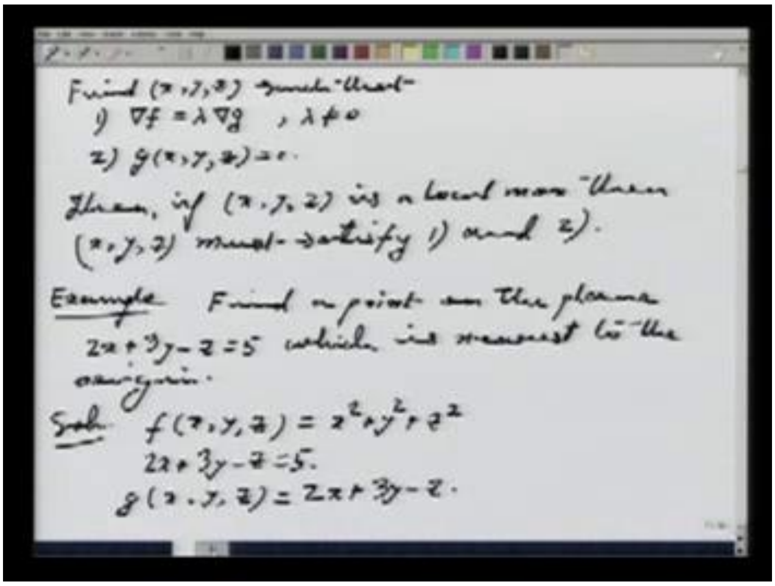
(Refer Slide Time: 15:35)



So, this then implies that, if p is given by r t naught. Then, f compose r, has local max, at t naught. And then this function is a function on real line, going to real line, our usual one variable results of calculus works, which implies then. That d of f compose r d t, at t equal to t naught is 0. But to calculate this derivative, I can use chain rule. So, this implies, grad f, dot d r, d t at t equal to t naught, sorry, it is t equal to t naught here, is equal to 0. This, then implies, that grad f at r t naught, dot d r, d t at t equal to t naught is 0.

Now, notice, this is happening, for all possible curves. Because, it has a local maximum, that is, if I go back to the picture, instead of this, I could have taken this curve, I could have taken this curve. All possible curves passing through p, if I take, then this relation is true. But, d r, d t at t equal to t naught, that gives us the tangent plane. So, grad f at r t naught is a vector, whose dot product with any tangent vector is 0. That means, grad f at r t naught is perpendicular to the tangent plane. But, this then implies, since the normal is already perpendicular to the tangent plane. That means, grad f at r t naught must be a scalar multiple of the normal vector.

(Refer Slide Time: 18:24)



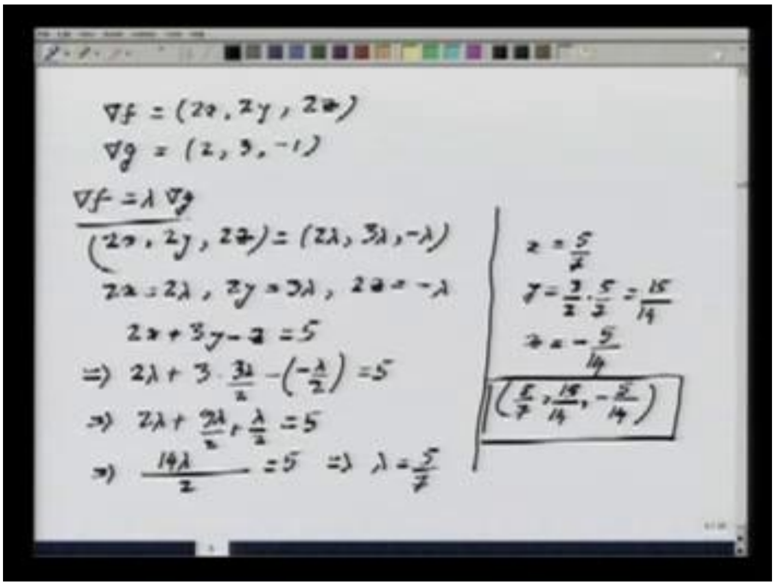
So, this gives me one idea of, how to get the extremum points. That is, what I do is, I look at the equation, grad f equal to lambda grad g. That means, I find x, y, z; such that, grad f equal to lambda grad g, that is number 1, of course lambda is not equal to 0. And number 2, g of x, y, z equal to c; that means, I am looking at points on the surface. Then, if x, y, z is a local extrema, that is, if it is a local max, then x, y, z must satisfy, 1 and 2.

So, this actually, gives me the collection of extremum. That is the collection of critical points, among them; I have to find the local maximum and minimum. But, that will depend, on the nature of the function. So, let us see the method once again, what it says is, that if I want to find the local extremum of a function f, on a surface s, which is given by a function g. All I do is, first I calculate grad f, then I calculate grad g, I look at the equation grad f equal to lambda grad g.

I do not know, what is this lambda? I have to find out x, y, z which satisfies this. And x, y, z is a point on the surface, I get a collection of points, there might be too many, I do not know, that is the collection of critical points. Among those points, I will try to find, what is the maximum or minimum of the function. So, all these things will be very much clear, if I look at the following example.

Let us, look at the following function. Let us look at the following problem, find a point on the plane 2 x plus 3 y minus z equal to 5, which is nearest to the origin. So, which function I want to minimize here, well the function is obviously the distance function? I look at a square of the Euclidian distance. That is x square, plus y square, plus z square. This is the function, I want to find minimum of this, on 2 x plus 3 y, minus z equal to 5. So, what is my function g here? So, g x y z is 2 x plus 3 y minus z.

(Refer Slide Time: 21:50)



So, let me try to calculate then what is grad f, grad f is 2 x, 2 y, 2 z. Now, I have to calculate, what is grad g, grad g can be easily computed, it is 2, 3 minus 1. So, what does it mean to say, grad f equal to lambda grad g. If I look at this equation, that means 2 x, 2 y, 2 z that is lambda times grad g, that is 2 lambda, 3 lambda, minus lambda. That means, 2 x equal to 2 lambda, 2 y equal to 3 lambda and 2 z equal to minus lambda.

Now, I have another equation, that is given by g. So, the equation is 2 x plus 3 y minus z equal to 5. So, that means 2 lambda because, 2 x is 2 lambda. 2 y means what, so it is plus y is 3 lambda by 2, so 3 into 3 lambda by 2 minus lambda by 2, that is my z, that is equal to 5. So, this implies 2 lambda, 9 lambda by 2 plus lambda by 2; that is equal to 5. So, this implies I have 2 here, 14 lambda here in the top, that is equal to 5. So, this implies, lambda equal to 5 by 7, which is a simple calculation.

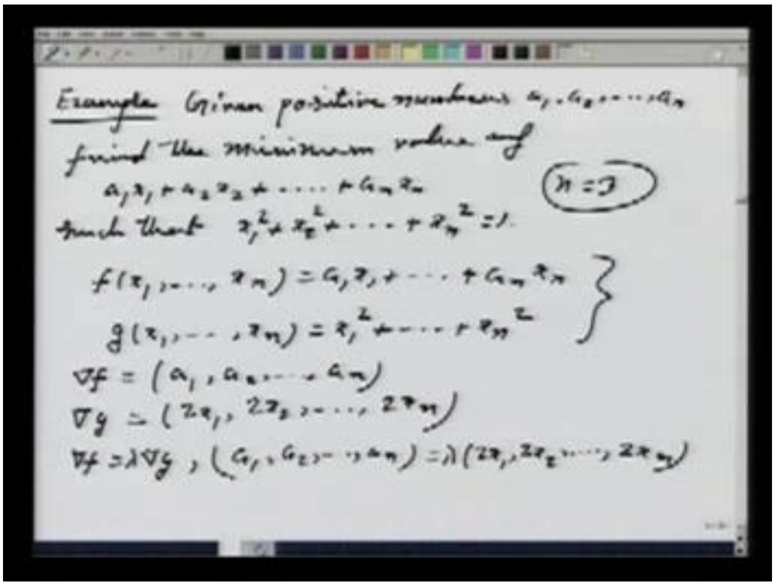
Now, once I know lambda equal to 5 by 7. Then, I certainly know, what, are the points. So, let us, calculate the points here, that means, x equal to 5 by 7, because x equal to lambda, y equal to 3 by 2 into 5 by 7, that means, 15 by 14, Z equals to minus lambda by 2. But, lambda is 5 by 7, that means, 5 by 14. That means what, so I found a point precisely, 5 by 7, 15 by 14 and minus 5 by 14.

Now, I always know that, there is certainly a point on the plane, which is closest to the origin. That means, the function, which I am looking at it does have a minimum. That is known from elementary geometry. But here, I see, that I got a point, which is candidate for local extremum and that is the only point. There is no other point coming, that certainly means, that the minimum is achieved at this point only.

See, I am using something here, which is not included in the method. Because, method just gives me the candidates for critical points. I know that, there exist a minimum and luckily, I got only one critical point. That means, this must be the minimum. So, this is the extra thing, which is not in the method. So, what does method give me, method give me the critical points. But, which one of the critical points is the candidate, that has to be found out by some other considerations, which I am doing it here, from geometry.

Because, I know there exist a point on the plane, which is closest to the origin. Given any plane, you can always find a point on the plane, which is closest to the origin. So, you know it exist, but because of this method, I could exactly find it out, that what exactly the point is...

(Refer Slide Time: 26:18)

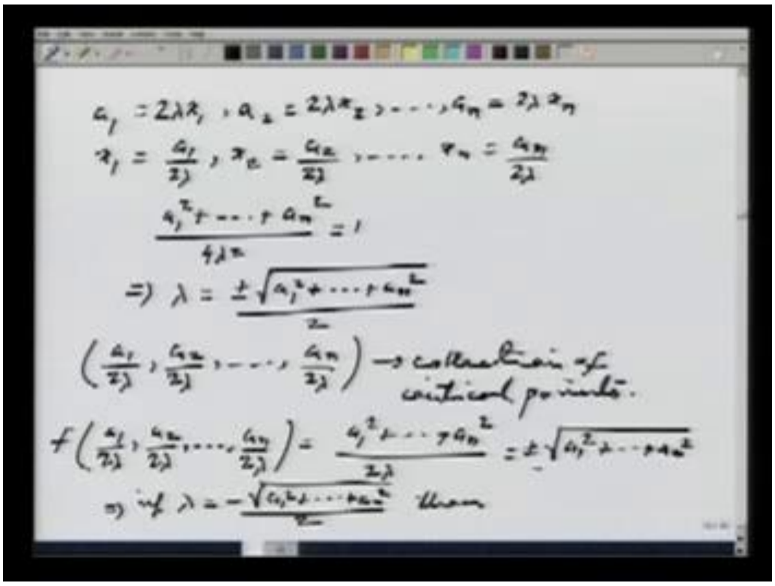


Now, let us look at another example, given positive numbers a 1, a 2. Let us say up to a n. Find the minimum value of a 1, x 1 plus a 2, x 2, a n, x n, such that. Now, what does it say, it says, that the unit sphere is given to me? The given equation x 1 square plus x 2 square plus x n square equals to 1 gives m e the unit sphere. The given problem is to minimize the function, a 1, x 1 plus a 2, x 2 plus an x n on the sphere. What is the guarantee that some such minimum exist, well that is given from advanced mathematics.

That if I have any closed and bounded region and I have a function, which is continuous on that, then the function has a maximum and a minimum. Now, the function f of x 1, x 2, x n equal to, so if I look at the function f of x 1 up to x n equal to a 1, x 1 plus a n, x n. The function is certainly continuous function, in fact, it is differentiable. So, it will have a maximum and minimum on the unit sphere, because it is closed and bounded.

So, in this case, I want to apply the method of Lagrange multipliers again. Because, it is a constraint extrema problem, that the given function f is the function of course, which I want to minimize and the function g is given by is x 1 square up to x n square. Well, I describe the method of Lagrange multipliers, only for three dimensions. But, it works, equally well, if I am looking at higher dimension. So, if you want to convince yourself with that, you can assume n equals to 3, no problem with that, but it works more generally. Now, fine I again start with grad f. Grad f, in this case turns out to be a 1, a 2 up to a n. And grad g, that is twice x 1, twice x 2 up to twice x n. So, grad f equal to lambda grad g, if I want to solve, I have to look at the equations, a 1, a 2 up to a n. That is equals to lambda times, 2 x 1 up to twice x n.

(Refer Slide Time: 30:00)



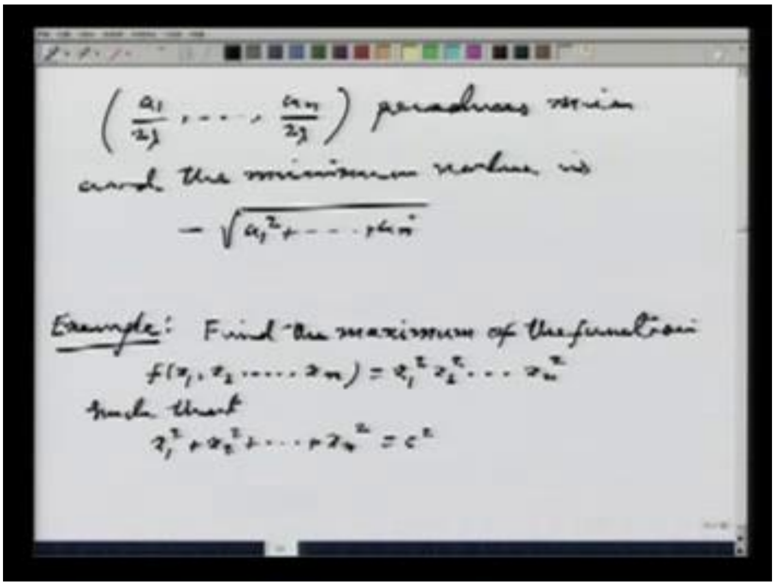
If I equate, I get a 1 equals to twice lambda x 1; a 2, twice lambda x 2 up to a n equals to twice lambda x n. Once I have this, I know then that x 1 equals to a 1 by 2 lambda; x 2 equal to a 2 by 2 lambda up to x n equals to a n by 2 lambda. Now, since summation x I square equals to 1, that is x 1, x 2, x n lies on the sphere. I get, that a 1 square plus up to a n square divided by 4 lambda square, that is equals to 1. I am just squaring and adding. So, this gives me the value of lambda, I get that lambda is equals to plus minus.

Now, if I put these values of lambda, then I get actually two choices of critical points correct. Because, then the points actually, I have are then a 1 by 2 lambda, these are my critical points, a 2 by 2 lambda up to a n by 2 lambda. But lambda has got two values, that we can see, because I am get a plus minus. So, this is actually the collection of critical points. That means, one of them will work as my local extremum and I am searching here for the minimum value.

Notice that, the maximum value is has also achieved. Because, the function is continuous function, so it will have a maximum and a minimum. So, for one point, I will have a maximum, for one point, I will have a minimum. So, how does one do that, just I try to put the value of the points, on the function and see what happens. So, if I do that, so what is f of a 1 by 2 lambda, a 2 by 2 lambda, a n by 2 lambda. Then I have a 1 x 1, that means, a 1 square plus up to a n square divided by twice lambda.

That means, it is plus minus, square root of. Now, since maximum has to be bigger than the minimum, I will certainly choose the value minus. Because, one is a positive value, other is a negative value.

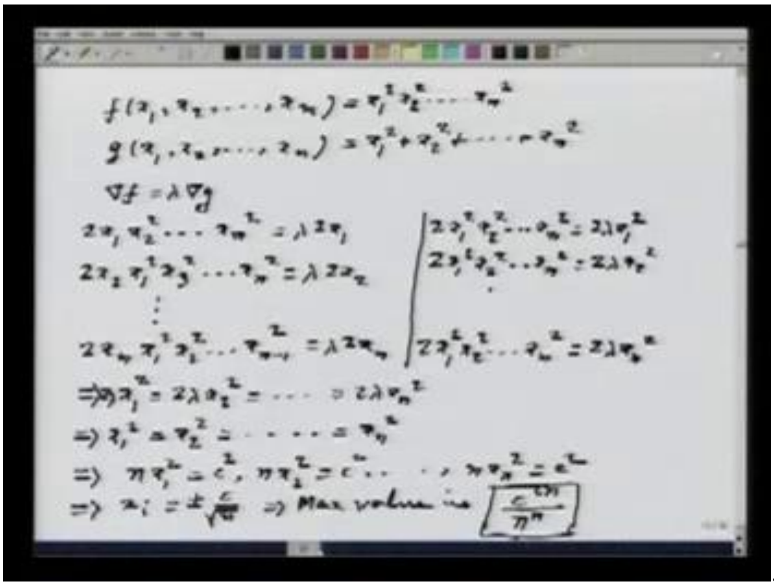
(Refer Slide Time: 33:57)



So, I will choose the negative value. That means, if lambda equal to minus of square root of, then a 1 by 2 lambda up to a n by 2 lambda produces minimum. And, certainly, the minimum value of the function, that I can see from here. It is minus of and minimum value is minus square root of, so I think, it is more or less clear. Now, that what exactly one has to do, I will finish this with one more example, which will provide, an interesting proof of a very well known inequality, which all of us know.

So, let us, move to the next example. So, here the problem is, find the maximum of the function f x 1, x 2 up to x n given by x 1 square, x 2 square up to x n square. Such that x 1 square plus x 2 square, plus x n square equal to c square. Look at the given function f, which is given by x 1 square plus x 2, x 1 square into x 2 square up to x n square. And we want to maximize the function over the sphere, x 1 square plus x 2 square plus x n square equal to c square. That means, it is a sphere of radius c.

(Refer Slide Time: 36:07)



So, what I do is, I proceed with the given one, that f of x 1, x 2, x n. That is, x 1 square x 2 square up to x n square and the function g is given by x 1 square plus x 2 square up to x n square. So, we have to start with, again the method of Lagrange multipliers. So, we start with grad f equals to lambda grad g, so what do we get, if I calculate. So, I get, twice x 1 into x 2 square up to x n square, that is lambda times twice x 1, that is my first equation.

So, what I am doing is, I am differentiating the first equation. That is x 1 square x 2 square x n square with respect to x 1, that gives me the left hand side. Lambda times, then the partial derivative of g with respect to x 1, that is twice x 1. Because, rest are constant, I get lambda times twice x, I am going to repeat this procedure with x 2, x 3 up to x n. So, what I get then is, twice x 2, times x 1 square, then x 3 square, up to x n square.

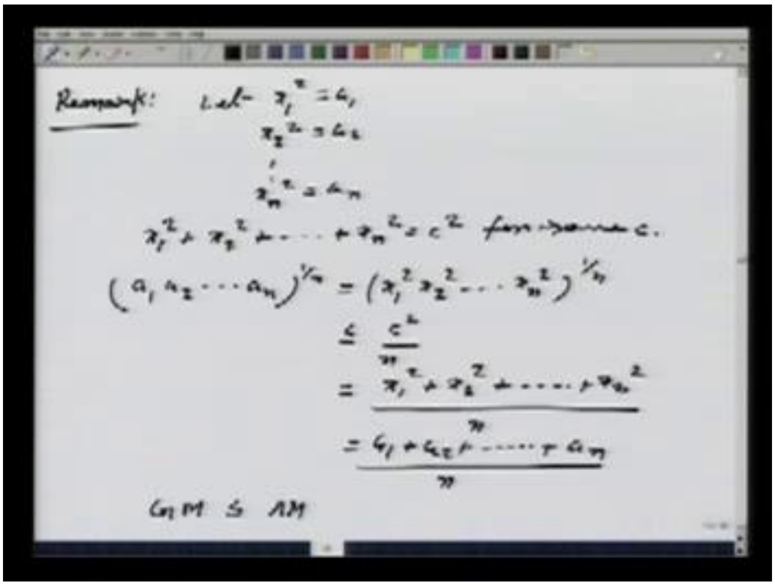
And the right hand side, then is lambda times, twice x 2. That is the partial derivative of g with respect to x 2. I go on doing this at the last stage, I get 2 times x n into x 1 square, x 2 square up to x n minus 1 square, which is equals to lambda times, twice x n. Now, after this, what I do is, I multiply all these equations, the first equation, I multiply with x 1, the second equation, I multiply with x 2 and the third last equation, I multiply with x n. Once, I do this, this implies, that x 1 square, 2 lambda x 1 square equals to 2 lambda, x 2 square up to 2 lambda x n square.

Because, if I multiply all these equations, the first equation, I multiply with x 1. What I get, I get 2 times x 1 square, x 2 square up to x n square equals to 2 lambda x 1 square. Second equation, I multiply with x 2 on both sides, what I get is, 2 x 1 square, x 2 square, up to x n square, that is 2 lambda, x 2 square. I go on doing this, the last equation, I multiply by x n on both side, what I get is, twice x 1 square, x 2 square up to, x n square. That is equals to twice lambda, x n square.

Look at the left hand side, all the quantities are same. That means, right hand side quantities should also be same. That means, twice lambda x 1 square, is twice lambda x 2 square up to twice lambda x n, square as I said. But, since grad f and grad g both are non 0, what I get is, I can cancel the lambdas also. This implies then x 1 square equal to x 2 square up to x n square. But x 1, x 2, x n lie on the sphere of radius c, that then implies, that n times x 1 square is equal to c, n times x 2 square is also equal to c and so on, that means n times x n square is equal to c.

So, this then implies, that x i equal to, sorry, these are equal to c squares, because that is the given equation. So, I get x i, equal to plus minus c by root n. Again, I got two values, as the extreme value, but whichever value I take, it does not really matter, because the function is given by, x 1 square x 2 square x n square. Because, I am going to take squares of all the xi's, so even if I take minus, it is going to turn out to be plus. So, the maximum value is the product square product of all these. That means, c square n times; that means c to the power 2 n, divided by n to the power n. So, this is the answer to the problem, that I wanted to maximize, I have maximized.

(Refer Slide Time: 42:00)

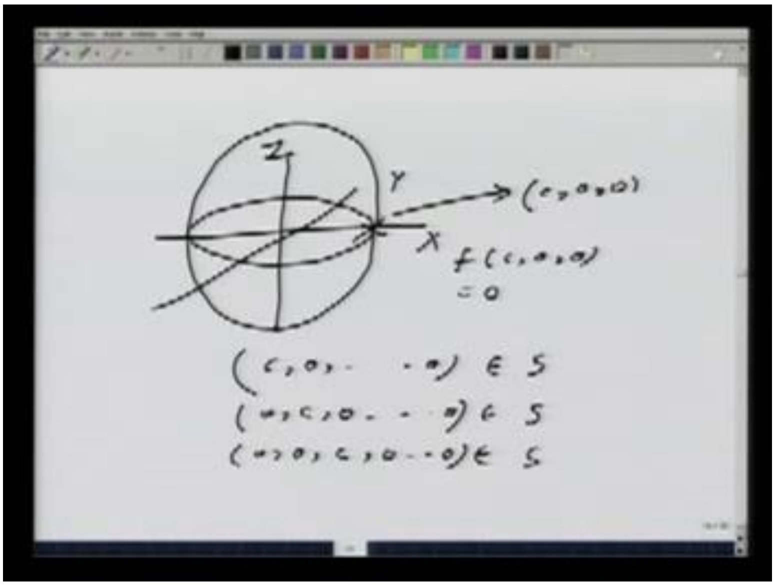


Now, some more comments in order. So, I will put it as a remark, what I do here is, let x 1 square, I call it a 1; x 2 square call it a 2 up to x n square, let us say equal to a n and x 1 square plus x 2 square plus up to x n square, that is equal to c square, for some c. Then, I look at the quantity a 1, a 2 up to a n, whole to the power 1 by n. That means, I am looking at, x 1 square, x 2 square, x n square to the power 1 by n.

Now, I know the maximum value of x 1 square, x 2 square, x n square as I have done in the previous case. It is c to the power 2 n by n to the power n, but I am taking a power 1 by n. So, it is lesser equal to c square by n. But then c square is x 1 square, plus x 2 square, plus x n square divided by n. But, x 1 square is a 1, plus a 2, plus a n divided by n. What, I finally landed up with is a familiar inequality, which you all know, that is, geometric mean is less than or equal to arithmetic mean. So, using the method of Lagrange multipliers, one can actually prove that a m is bigger than or equal to g m for n quantities.

Now, let us look back at the problem again. The given problem was find the maximum of these functions, the maximum is something, which we were trying to look at what happens, if we ask about the minimum of this function.

(Refer Slide Time: 44:32)



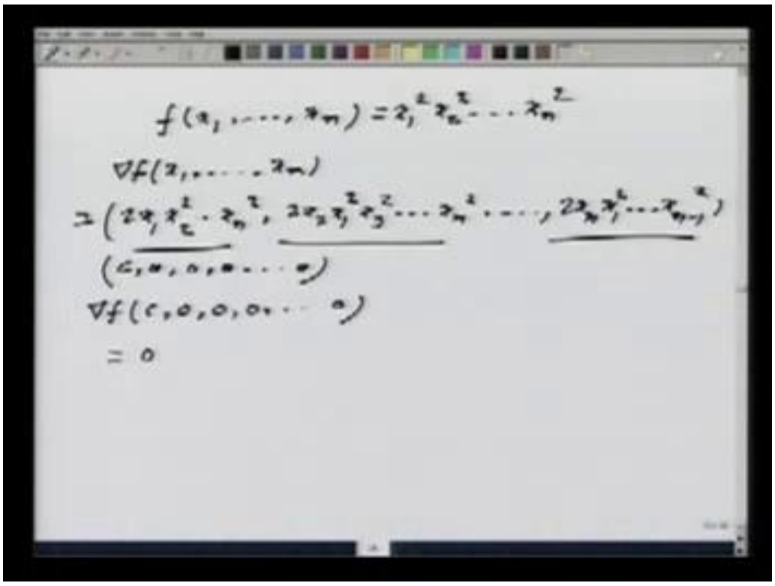
Well, a moment's thought would tell you that perhaps the minimum is 0. Why, so because when I look at the sphere, look at the three dimensional case. It is a sphere of radius c and then what is the coordinate of this point. The coordinate of this point, if this is x axis, this is y axis, this is z axis. The coordinate of this point, is c 0, 0, what is the value of the function at this point. What is f of c 0, 0, it is product of the squares of the component.

But, there is 0 also appearing in the components. That means, this is 0 and since the function is always non negative. Because, it is product of the squares, this is certainly the minimum value of the function. That means, this point c 0, 0 is a critical point of the function, correct. So, in the n dimensional case also, it is going to happen. Whenever, the sphere intersects the axis, then except 1. All the other coordinates at 0, that means, if I look at the points of the form c, then rest at 0 or 0 c rest at 0 and 0, 0 then c rest at 0.

Notice that, all these points belong to the sphere s of radius c. Why, this points, which are certainly minimum of the function and not appearing, as critical point in the method of Lagrange multiplier. That is the question; I look back at my calculations again, what are the points, which served as the critical points here. That means, the xi's satisfies this, the points 0 does not appear here, the points c 0, 0.

That is, what I mean, the intersection of the sphere, with the axis, those points are not appearing here, as the points of extremum. Why is that happening, well the reason is very simple.

(Refer Slide Time: 46:41)



The given function notice, f of x 1 up to x n is, x 1 square, x 2 square up to x n square. Now, what is grad f, then at any point x 1, x 2, x n, it can be easily calculated, it is twice x 1 times x 2 square up to x n square. Then, twice x 2, x 1 square, x 3 square up to x n square and it goes on like that. The last term here is, twice x n into x 1 square up to x n minus 1 square, I take a point c 0, 0, that means, a point on the axis. If this is the point, then I can see, that grad f, at c 0, 0, 0 up to 0, this is equal to 0.

Because, all the points x 1, x 2, x n appear in all the terms, which I am writing here, x 1 here, does not appear as a square, but 2 x 1 appears. I look at the second term, x 2 square does not appear, certainly, but twice x 2 occurs. So, all the terms are actually present and one of the term and some of the terms are always 0, when I look at points like c 0, 0, 0, 0. You know where it intersects, so the gradient of f, at that point becomes 0.

But, look back at the method of Lagrange multiplier. I have always assumed that the points of local extremum, which I am going to find out, they satisfy two conditions. That grad f, is not equal to 0 at that point, grad g is not equal to 0 at that point. If I look at grad of g, then at c 0, 0, 0, 0, this point which I am looking at grad g is non 0. But, grad f is certainly equal to 0 here, and that is why, the method of Lagrange multiplier is failing here.

So, this is not the case, that although the function is continuous, it does not have any minimum, it has a minimum certainly. We could easily find out, what is the minimum. But, that point escapes, the method of Lagrange multiplier, simply because, gradient of f was 0 at that point. So, that finishes our discussion on the method of Lagrange multipliers. From the next lectures onward, we are going to deal with another important thing of calculus. That is integration. But, this time, we are going to do with functions of two or more than two variables and we are going to talk about Riemann integration there.