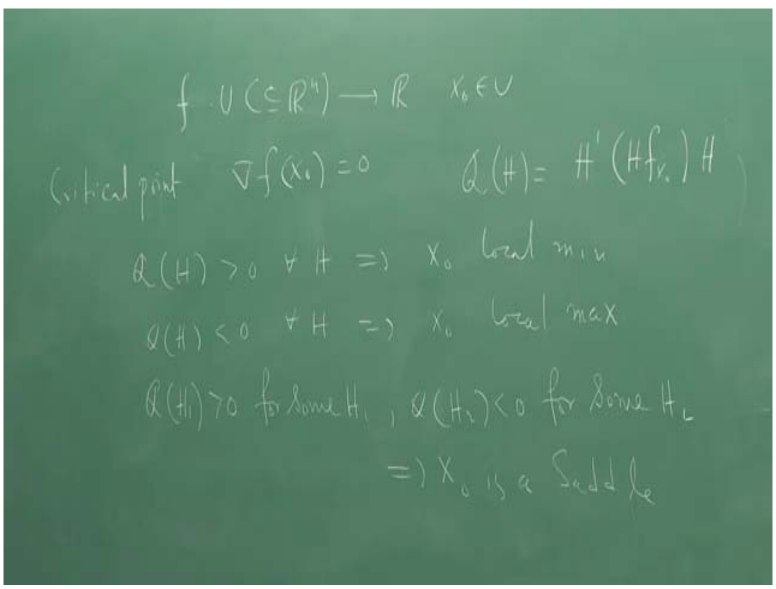
Differential Calculus of Several Variables
Professor Sudipta Dutta
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur
Lecture Number 14
Practical Test based on Hessian Matrix

Okay, in the last lecture we have derived the 'second derivative test' in terms of this, Q(H), that is, second term in the, Taylor's series expansion. So what we have done is this thing, that's for f on an open set R, and it's not in U, I have a critical point, that is grad f of x0, 0. Then we've taken Q of H, with half or without half, H prime, Hessian at x0 H. Okay. And what are the test that Q(H) greater than 0 for all H implies x0 local minimum, Q(H) less than 0 for all H, x0 local maximum, and Q(H1) greater than 0, for some H1, and Q(H2) less than 0, for some H2 implies x0 is a saddle.

(Refer Slide Time: 02:09)



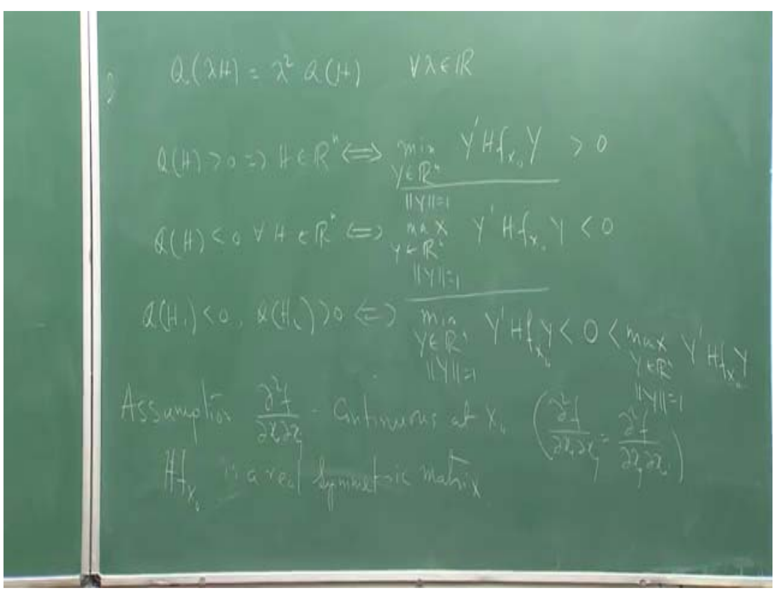
That's what the test was, 'second derivative test'. Now, we also had noticed, that in the proof also, that we have used mainly the scaling property. So because of the scaling property, any H can be brought down to norm 1. That has no, nothing wrong in assuming instead of for all H, for all H norm 1, for all H, for all H norm 1, and, H1 norm 1, H2 norm 1, right? So what we can do here, is that, that, we can write here, again, Q(H) is greater than 0 for all H in Rn. This is to same as saying minimum, of let's say, let me write in this form now, instead of H.

Y, in Rn, norm y equal to 1, yHf(x0), this is greater than 0. This is if and only if. Similarly, Q(H) is less than 0 for all H in Rn, this is if and only if maximum y in Rn, norm y equal to 1, y prime Hf(x0) y is less than 0. And, Q(H1) less than 0, and Q(H2), the third statement there, greater than 0, if and only if minimum of y in Rn, norm y equal to 1, of y prime Hy is less (than) Hf(x0)y, is less than 0, less than maximum y in Rn, norm y equal to 1, y prime Hf(x0)y.

Right? Because this is because of scaling property you can deduce that. So look at these two board cleaned. At it's end point you will see, that what you've written is, is just because of the Q(lambda, H) in lambda square Q(H). But why I have written like this? Well you see, every time I have to check, I have to compute this Q(H) and see, if I want to use this board for checking minima, maxima or saddle point, I have to kind of check for every edge.

And if I write it in this way, well I have reduced the problem to norm of, if you look at this form at norm of y only 1, (you) that's the reduction, but it's not a great reduction. But if you, now I apply some results from linear algebra, then this will give us a great advantage in actually checking minima, maxima or saddle point. And, what is that? Under the assumption we made, that del square f, del xi xj, they are continuous at x0, this Hf(x0), under the assumption we have proved that, mixed partial derivatives are equal.

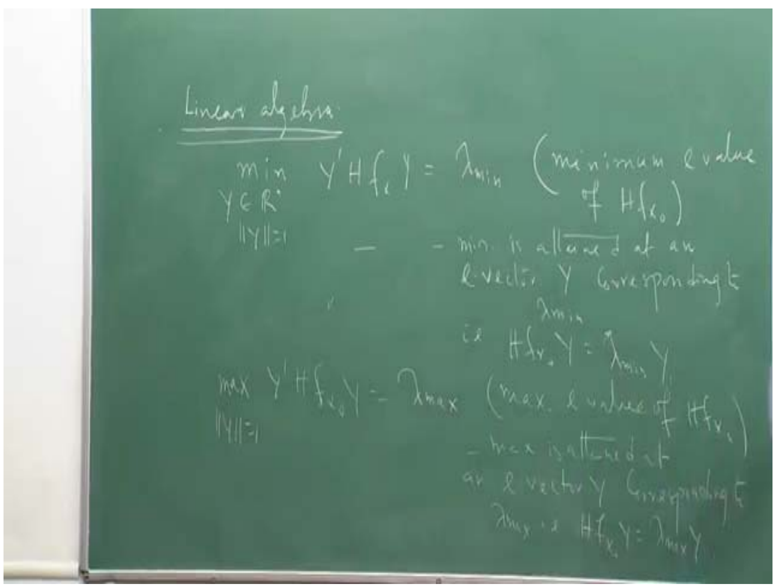
(Refer Slide Time: 06:42)



So Hf(x0) is a 'real symmetric matrix'. And what do you know by 'real symmetric matrix'? That they have, real i-n values and, there i-n vectors can be chosen to be real as well. And, this quantity has a special meaning here, this quantity has special meaning in terms of i-n values. So if you use your linear algebra, the first course in linear algebra, minimum y in Rn, norm y equal to 1, y prime Hf(x0)y equal to lambda min, the minimum i-n value of Hf(x0), and, the minimum is attained at an i-n vector y, corresponding to lambda min, that is, Hf(x0)y equal to lambda min y.

Similarly, maximum y prime Hf(x0)y, norm y equal to 1, this is maximum of the lambda max, that is maximum i-n value of H of Hf(x0), and the maximum is attained at an i-n vector y, corresponding to lambda max, that is a, sorry, Hf(x0)y equal to lambda max y. This is a simple result from linear algebra. You prove it. This can be verified with calculus as well. And we'll be able to do it after the next week, when we do it 'Lagrange Multiplier Technique', may be I'll do it or may be I'll, , give it as an assignment.

(Refer Slide Time: 09:18)



But for time being, let us recall linear algebra and this result. So what he says now here, that, if the minimum i-n value is greater than 0, and all i-n values are greater than 0, then it's a local minima. If maximum i-n value is less than 0, then all i-n values has to be less than 0, then it's a local maximum. And, if there's some i-n values less than 0, and, some i-n values bigger than 0, then it's a saddle. So in terms of i-n values, this is a test, of, (secon, secon) second derivative test reduces to this (pa), this thing, that, (())(09:52) looking for the i-n values.

But in general what happens, that i-n values if n is large, n equal to 2, no problem, n equal to 3, okay. Not take too much of problem. But n equal to say 100, then, calculating the i-n values may be a difficult task, in general. So, from this observation, and this linear algebra result, we'll now write down a practical test for actually checking minima maxima and saddle points, in case of non degenerate critical point, that is when Hessian has non zero determinant. So instead of checking Q(H) for every H, we just go through this check and that makes life little easier.

Now if you can calculate i-n values, nothing better. Okay, so let's write it as a theorem. So, again the same setup. f is from open set to R, x0 in U and grad f(x0) equal to 0, that is x0 is a critical point. Instead of avoiding writing grad f(x0) every time, let me put a notation for it. A equal to, because I am looking at one x0, so let me write it as A, just say n cross in matrix A.

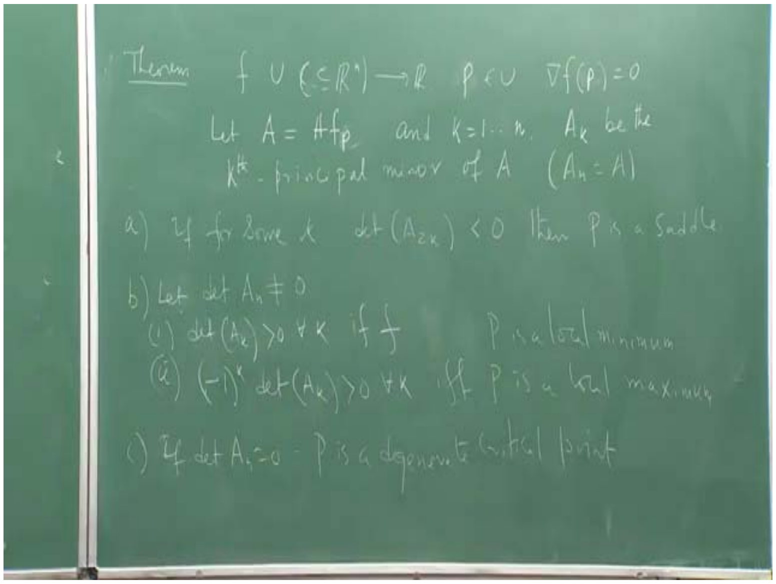
Okay. So from that, this same board whatever we have, I will write down the statement. , okay. Don't (())(11:46), I just want to make it P, because I have to use this notation x in the proof. Okay? So don't worry. In the x0, I will use P, instead of (y, y) point x0, I will use some point P. Okay. So let H is equal to Hessian, and, K equal to 1 to n, Ak be the Kth principal (min), Kth principal minor of A.

What does it mean? So A is this n cross n matrix. You look at the one by one, first one by one matrix. This is A1. The first row, first, entry of the first row first column. Then, in the first row, you get two elements, and then, second row two elements. So I think you all know, what is principal minor. So, this is, Ak is the A matrix where K plus 1 to n and K plus 1 to n entries of rows and columns are deleted. Okay?

So conclusion is like this. If for some K, determinant of A2k is negative, okay? Not A Ak, but, so K has to be between 1 to n, then, P is a saddle. b, if determinant of An, An is A, so in; is not equal to 0, then, so let determinant of A is not equal to 0, the (determinant), the Hessian is non zero determinant. If for all K, determinant of Ak is greater than 0, then P is a local minimum. Okay?

In fact, I can write it in this form. Determinant of Ak is greater than 0 for all K, if and only if P is a local minimum. Okay? Minus 1 power K, determinant of Ak is greater than 0 for all K, greater than 0 or less than 0? Greater than 0 for all K, if and only if P is a local maximum. And finally, if determinant of An is 0, P is a, we have already called it, degenerate critical point, and in this case, the test fails.

(Refer Slide Time: 15:47)



Okay. Before I prove it, let me give an example, of how this test can be applied. So let's take, (there are) many examples here, let's take one. Okay. x square y, y square z, z square z minus 2x. f(x), example, (x,y,z) x square y, y square z, z square x minus 2x. So first search for critical point. Correct? Grad f(P) equal to 0. What is the solution?

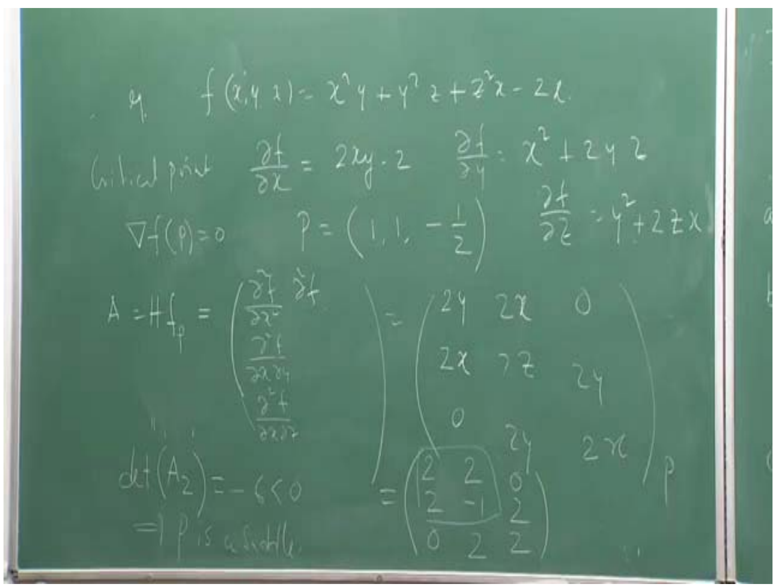
You'll see, the only solution is, if you solve this, the only solution is, P equal to 1, 1, and, 1, 1, and, minus 1. I have written something wrong, right? Yeah. Sorry. Del f del y is, x square plus 2yz, and del f del z is, I have missed this part, y square plus 2zx. 2z, 2x 2x. No 2zx. Thike? Del f del z is, y square plus 2zx, x square plus 2yz, y square plus 2zx yeah. That will give us this thing.

So let us calculate the Hessian, Hf at P, which is A in that, which is equal to, del square f del x square, del square f del x del y, del square f del x del z, so on. You calculate how much is that. First column, del square f del x x square, that is gives you 2y, del square f del x del y, del x del y, that gives you 2x, and this gives you del f square del (de) del z, that is 0. Now, del square, so I know what will be here, this will be 2x here, because symmetric matrix and here it will be 0 again, okay?

So del square f del y square, that is, 2z, del square f del y square 2z, and, del square f del z, del square f del z del y is, just (del) del del square f del y del z, so this quarter is del square f del y del z, which is equal to, how much? 2y, so 2y here, it will be, and del square f del z square is equal to simply, 2, is that okay? That at the point P, so which is equal to the matrix 2, 2, 0; 2, minus 1, 2; 0, 2, 2. Okay?

And you see, determinant of A2, which is this part, minus 2 minus 6, this is minus 6 less than 0, so P is a saddle, okay? So I don't have to check for Q and all those things. Check the calculation carefully but okay. (Theorem)(())(21:27) okay. So for today, I will start the proof, because time is running out, and I will complete the proof in the next lecture. Proof is very interesting.

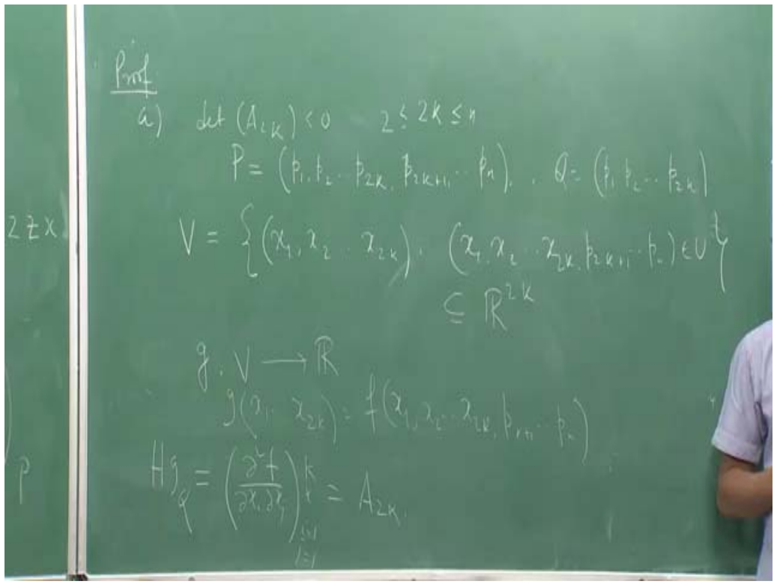
(Refer Slide Time: 21:21)



So first a part. So let me try, if I can finish the proof of a part, today. Okay. What I have assumption? So determinant of A2k is less than 0 for some 2k, less than equal to n, less than equal to 2. Okay. So, let me have this point P, critical point. Let me write it as P1, P2, P2k, P2k plus 1 Pn. Okay. Let's take this set V. This is set of all x1, x2, x2k, this set, such that, x1, x2, x2k, P2k plus 1 Pn belongs to U, that is I take the first 2k section of U, with the last 2k plus 1 to n coordinate fixed. Okay?

And I define, so this is half set in R2k. I define g from this V, this is an open set in R2k to R by g of x1, x2k equal to f of x1, x2, x2k, Pk plus 1 Pn. Okay? Then, what I want you to do; so, P is this, let Q is this, you calculate Hg at Q. You find out, since Pk plus 1 Pns are fixed, this is del square f del xi del xj, i equal to 1 to K, j equal to 1 to K, which is A2k. So, now what do I do? I know determinant of A2k is given to be negative.

(Refer Slide Time: 24:37)



But what is determinant of A2k. A2k is a real symmetric matrix. So let beta 1, beta 2, beta 2k are i-n values of A2k. And then all of you know this result, determinant of a matrix symmetric (())(24:56) have the i-n values, this is product of the i-n values. 2k is a even number. Even number of real (number), real terms, their product is less than 0. What does it mean? All of them cannot be negative, because even then it'll be greater than 0. All of them cannot be positive.

(Refer Slide Time: 25:43)

Even they will be positive. So there exist beta i, such that, and beta j, such that beta i is positive, and beta j is negative. Let me complete it here, and continue from this point. So from this point, we’ll continue in the next lecture.