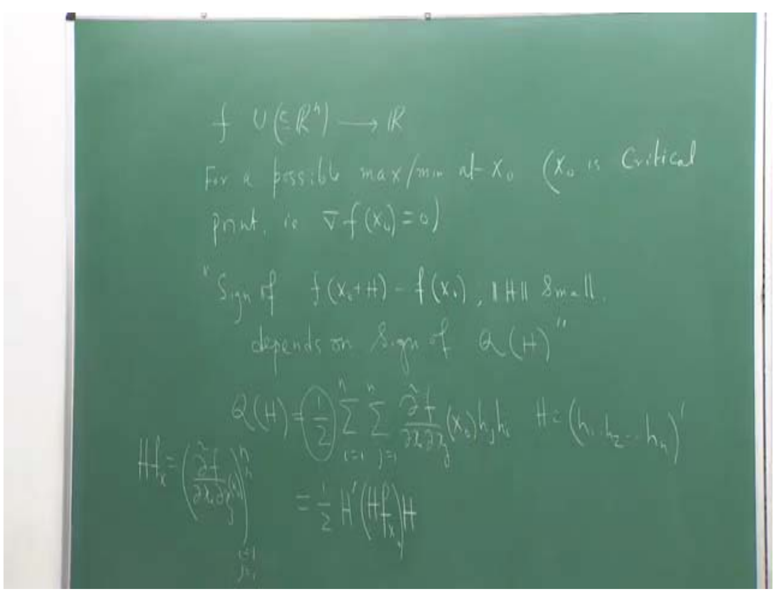
Differential Calculus of Several Variables
Professor Sudipta Dutta
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur
Lecture Number 12
Second Derivative test for Maximum, Minimum and Saddle point

So today's lecture, we'll continue with the discussion on maxima minima, and we'll formalize a statement we made, last time. So, so for a, what we did last time if you recall, for a possible maximum, or minimum, at x0, where x0 is a stationary point, a critical point, we have decided to call it critical point, some call it stationary point. Critical point, that is grad f at x0 is 0. We wanted to do the second derivative test and, what we did? We said something like, sign of f of x0 plus H, for H small, minus f(x0), H small, norm of H small, depends on sign of Q(H).

That was informally, what we did, and what was Q(H)? Q(H) we defined half of the second term in the Taylor's series, summation i equal to 1 to n, j equal to 1 to n, del square f, del xi, del xj, hj, hi, where h is the factor (h1, h2…. hn) prime. Now as far the sign is concerned, this fellow is doing nothing. Right? This is simply a constant, so this is we have also written as H prime. So everything evaluated at x0, I keep on forgetting that. So this is H prime, Hessian f at x0 H.

This is Hessian, I mean, one should not confuse with this H. Hf at x is del f del square f, del xi, del xj at x0, 1 to n, 1 to n. Okay. So, we makes two assumption here, okay? And we'll see,

(Refer Slide Time: 03:31)



what assumption is necessary that in our, so we're actually discuss about it, last time. We have also assumed it. Assume that second order partial derivatives are all continuous. Continuous at x0. And also assume, that determinant of the Hessian at f(x0) is non zero.

I will come back to this assumption, later, why we need that. May be this lecture, in the next lecture, or next lecture. So if determinant of H of f(x0) is zero, we call x0, a 'degenerate critical point'. And in this case, the second derivative case, test will fail. It will not give you anything. As a proof we'll suggest, you'll see if you see the proof, you will see that (if) this determinant is zero, the second derivative test doesn't work.

It's simply like that if you do the function of one variable, if f double prime at that critical point, (())(05:21) function from interval to R, f prime (())(05:25) x0 is 0, and if double prime x0 is also 0, then second derivative test for one variable also fails. Similar situations also occur here, and, in this case, degenerate case, we have to do for that test, either by, observation or by going to higher derivatives. Anyways, so, since the, we have only (conce) concerned about sign, we forget about this half part, and take our Q(H), since you are (conce), it half will not change any sign, so we will take Q(H) (())(06:05) to be this thing. If you want to keep half, no problem, but this time I have to write a half, so let us forget about the half part.

Okay. The first observation you made here, that if you have any lambda in real, then Q of lambda, H, for any H. So this is lambda will be multiplied with each of the component here. So lambda each of the component with (())(06:43) , it will come out to be, this. So if I multiply with any lambda, sign of Q(H) doesn't change, it remains same. Sign of Q lambda H is sign of Q(H), for any lambda in R.

This is needed for scaling, as you will see. So, let us suppose, that, first case one. Q of H is greater than zero. Suppose, for all H in Rn, okay? Then, write this way, f of x0 plus H minus f of x0 is equal to, okay? I have (())(07:46) taken out half part, so may be, I write half Q(H), why? I'm writing Taylor's series expansion, remember, that it had to have grad f at x0 dot H, but grad f(x0) this is 0, so it will actually start at Q(H), correct?

Plus, I will have error term, so I will do it only for second order term. Second order Taylor's series expansion. Okay. Now, look at this fellow. Infimum of Q(H), infimum sorry, infimum of

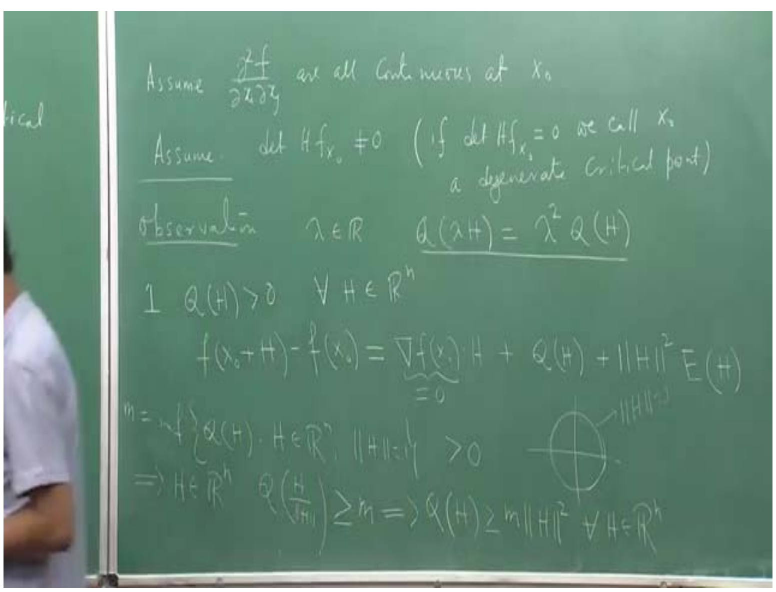
Q(H), H in Rn, and norm H equal to 1. Let's call it m. My claim is, this m is bigger than 0. Why? Well, Q(H) is always bigger than 0, for all H in Rn. Norm H equal to 1, this is a circle, right?

Now in circle, in the very (first) second lecture, or first (lect) second lecture of this course, we have shown, that, if I have a continuous function, then it attains it's maxima and minima over a compact set. The circle S1 is a closed and bounded (ci) circle. This is a set of all H, such that norm H equal to 1. This is a compact set, so on that compact set, infimum will be attained. Because Q is a continuous function of H, from definition.

And infimum is greater than 0, because Q(H) is greater than 0, for all H, and on the compact, if (())(10:13) continuous function is greater than 0, on a compact set, it's infimum is also greater than 0. That's very easy to prove. Okay. So what does it say? That if I take any H in Rn, then if I look at Q(H) by norm H, H by norm H is on the circle, because H by norm H has norm 1. So this will be greater than equal to m, because m in the infimum.

Which implies, from this equality, that for any H, Q(H) is greater than equal to m into norm of H square, for all H, in Rn. Very good. So what about f(x0) plus H minus of f(x0), from there, this is greater than equal to, m into norm of H square plus norm of H square E(H). Now E(H) is

(Refer Slide Time: 11:09)

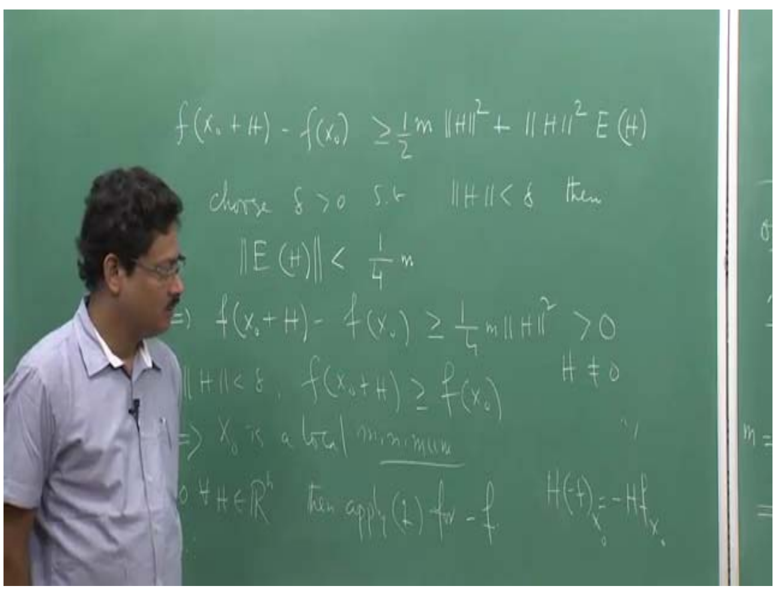


the error term. So can choose a delta greater than 0, such that, if norm of H is less than delta, then, E of H, m is a fixed number, which is infimum, this is less than, let's say, I should have put a half here na, if I have (())(12:12). Doesn't matter. So just say anything. 1, 4 by m.

So that should it show, that f of x0 plus H minus of f(x0) is bigger than equal to, so it is a half here, 1, 4, m H square, which is always bigger than 0, as long as H is non zero. And H equal to 0, this is f(x0). That shows, for norm of H less than delta, f of x0 plus H is always bigger than f of x0. Correct? So that's for, for all y, in a neighbourhood of x0, this sign is bigger than, so x0 is a local minimum. So if Q(H) is greater than 0 for all H in Rn, then x0 is a local minimum.

Now maximum part is easy. Next, if I have Q(H) is less than 0, for all H in Rn. Hmm? Then, apply, one, this part for minus f. Okay? How does it help? Because, you will see that H minus f at x0 is equal to minus Hf at x0. So, and so a local minima for minus f is a local maxima for f. So, if this is so, I will get (())(14:26) same consideration, x0 is a local maxima. Apply everything, this analysis for minus f. So that solves the question for (minu) , f as, x0 is a local minima or maxima, when Q(H) is either greater 0, or less than 0.

(Refer Slide Time: 14:16)



So third (possibility) there exist H1 in Rn such that Q of H1 is positive, and, H2 in Rn, such that Q of H2 is negative, that is H attains both positive and negative sign. In that case, what happens? In that case my claim is, that x0 will be a saddle. How? Well. So again, the same situation. Grad of x0 is 0, so for a fixed H, and for any lambda in R, you look at this fellow. f of x0 plus lambda H minus f of x0.

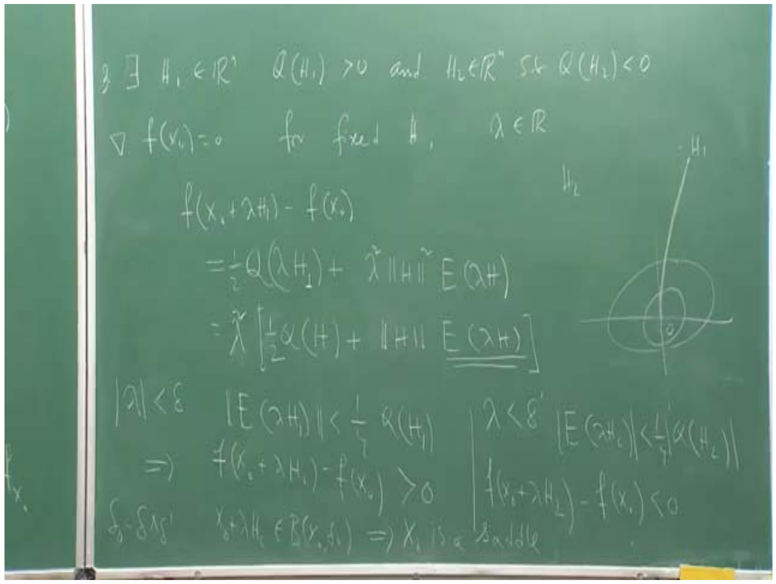
See now, two, there is U, okay, here is my U, here is some x0, H1 is somewhere here, and, H2 may be somewhere here, I don't care. Actually, without loss of generality, by translating, I can say that x0 is 0. So this is translation, so this is 0. Without loss of generality, you just look at U minus x0, it will have same kind of property. Okay what is that? Follow geometry. This is equal to Q, half okay, lambda H plus lambda square H square E(lambda H).

Which is equal to half of lambda square Q(H) plus lambda square. So let's take the lambda square part out. Now gain the same idea. If I choose, , lambda, let's say, mod lambda less than some delta, here is a lambda square there, so E(lambda H) is the error term, so I can have E(lambda H), this is less than one fourth of Q(H), right? So in particular, if I apply to H1, so I will have H1 here, H1 here, everything H1 here, so sign of this fellow, in this case f(x0) plus lambda H1 minus f of x0, Q of H1 is positive, so this will be positive.

Similarly, choosing lambda some less than some delta prime, I can make E(lambda, H2) less than one fourth of Q(H2), and in that case I will have f of x0 plus lambda H2 minus f of x0, less than 0. So what is happening? Now lambda is any arbitrary number less than delta. H is somewhere. Assume x0 is 0. Now, multiplying by lambda I can always bring it to a neighbourhood around 0.

H, H(one), whatever we H1, multiplied by suitable lambda, I can always (bri) bring E(lambda, H1), sorry lambda H1, near a neighbourhood around 0. So that means I can make, so if I take delta0 to be minimum of delta and delta1, I can bring x0 plus lambda Hi in B(x0, delta0). So I have a point, x0 plus lambda H1 minus f(x0) is greater than 0, so if this is greater than 0, (())(19:57) x0 plus lambda H1 is greater than f(x0), and f(x0 plus lambda H2) is less than x0. And that is precisely the definition for x0 is a saddle point.

(Refer Slide Time: 20:20)

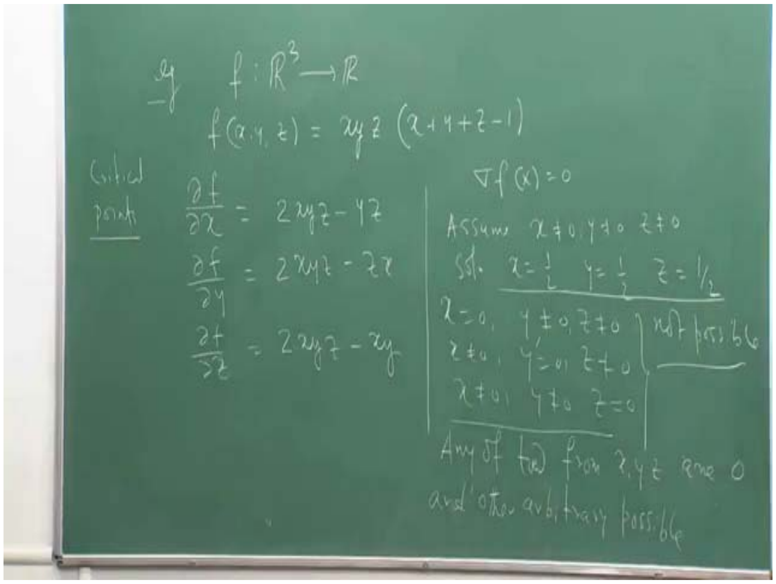


So this is the so called, second derivative test for function of n variable. Okay? So go through the proof again. You have to atleast go, you have to, you (have) have to go, through it, atleast twice, then you understand what is going on here. Very easy, but still go through it once again. And I'll do an example for you to end today's class. Let me write down the example. I have already written down for it. f from say R3 to R, f of (x,y,z) equal to xyz, x plus y plus z minus y.

So I have to look for critical point first. There may be many, and in this case there are uncountably many actually. As you see, this will be equal to 2xyz minus, 2xyz minus yz, right? Here. Del f del y, this will be 2xyz minus zx, and del f del z will be 2xyz minus xy. Okay. So if I put grad f at x equal to 0, then, if I assume x0 equal to 0, y0 equal to 0, z0 equal to 0, then the solution is, only solution is you can find x equal to half, y equal to half, z equal to half.

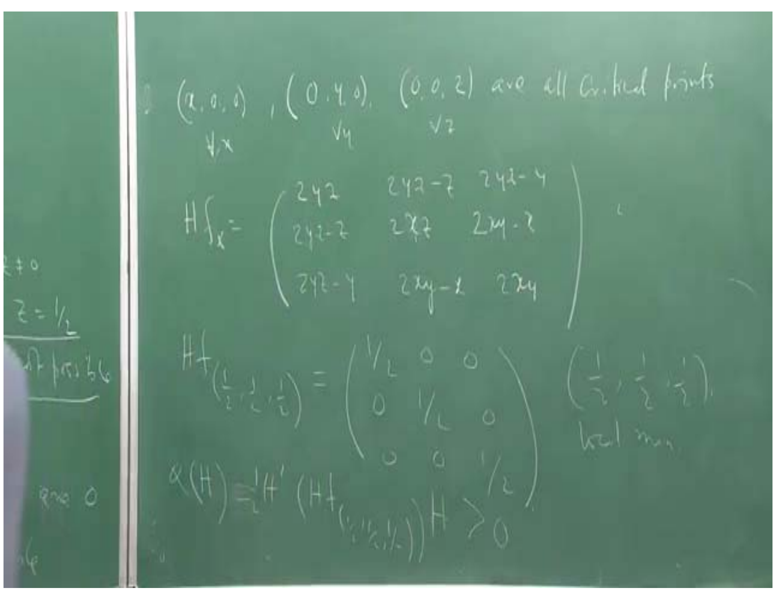
x equal to 0, y not equal to 0, z not equal to 0. x equal to 0, y not equal to 0, z not equal to 0, but this is equal to 0, so this is not possible. Similarly x not equal to 0, y equal to 0, z not equal to 0; x not equal to 0, y not equal to 0, and z equal to 0, they are all not possible. Okay? So one solution, but, if you see, any of two from x, y, z, is 0, are 0, and other anything, other arbitrary, possible. That is, are all critical points, for any z. For all x, for all y, all z.

(Refer Slide Time: 23:35)



If you calculate the Hessian, f at x, that comes out to be 2yz, 2yz minus z, 2yz minus y, 2yz minus z, 2 , just a minute, Hfx fx x del f del y 2xz, sorry, 2xz, and, 2yz minus y, 2xy minus x, 2xy minus x, 2xy; I think I made a mistake here, del fx, , del fx del x x 2yz minus z, 2yz, , del f del y del x, minus z, del f del z x 2yz minus y. Thike. Yeah. So, Hf at half, half, half is equal to half 0, 0, 0 half 0, 0, 0, half. So Q(H) is always equal to, how much? You can check, Q(H) is always Q(H), half of H prime Hf at half, half, half, Horlicks, this is always greater than 0. Okay?

(Refer Slide Time: 26:26)



So, this is, local minima, whereas, Hf at, let's say, (x,0,0), if you check determinant of this, it's same as determinant of Hf at (0,y,0), same as determinant of Hf at (0,0,z) equal to 0. So these critical points are all degenerate, and we have not, so far, deduced any test for degenerate critical points. So, we'll talk about this test in some other language, something from linear algebra is happening in the (nex) next lecture.