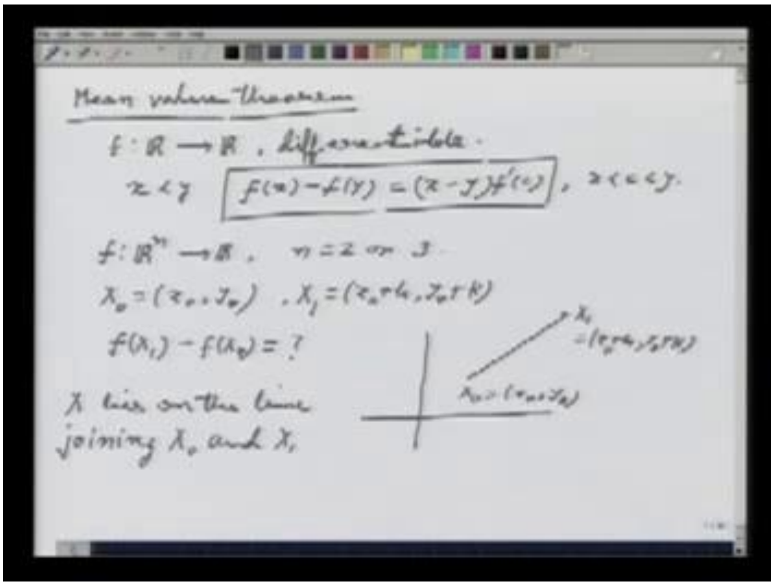
Mathematics-I
Prof. S.K. Ray
Indian Institute of Technology Kanpur

Lecture - 25
Mean Value Theorem

(Refer Slide Time: 00:29)

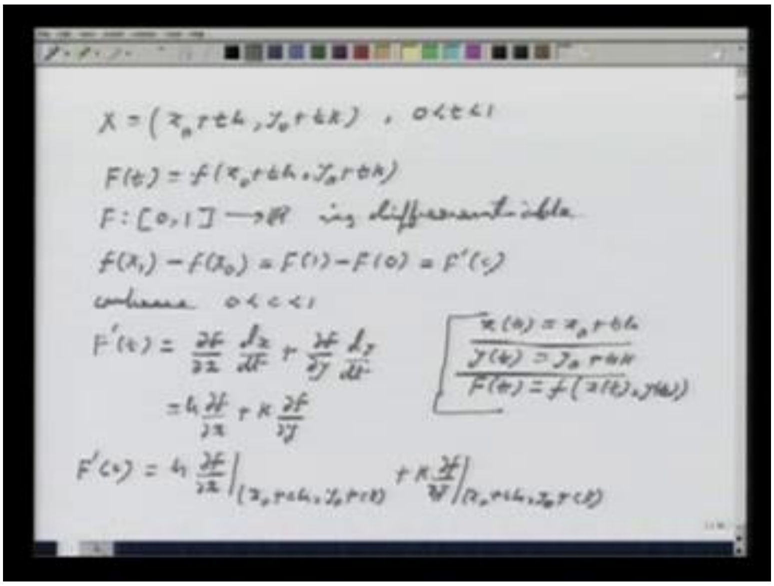


In today's lecture, we are going to talk about one important result regarding derivatives. That is the Mean Value Theorem. We know what mean value theorem says in one variable. So, suppose f is a function from R to R. So, it is one variable function and f is differentiable. If I choose two points say x and y, suppose x is less than y. Then I know that f x minus f y is equal to x minus y times f prime c, where c lies between x and y. This is the standard mean value theorem. Now, suppose I consider functions, which are functions on R 2 or R 3 let us say. So, let us say f is a function from R n to R.

As usual I will assume that n equals to 2 or 3. Now, differentiability of f we have talked about, so f differentiable makes sense. Can we make sense of the mean value theorem. So, that is let us say I am concentrating on two dimension. I take a point x naught, which is x naught, y naught. I take any other point call it x 1 that is x naught plus h, y naught plus k. Then I look at the quantity f of x 1 minus f of x naught.

Question is can I say anything regarding this quantity using the derivative of f. So, let us look at the situation pictorially it means, suppose this is my point x naught which is x naught y naught. I have taken another point here that is x 1. That is x naught plus h y naught plus k. And then, I look at f x 1 minus f x naught. And I want to get something like x minus y times f prime c. So, I look at this line joining x naught and x 1. If mean value theorem is at all true, it should be let with some point which lies between x naught and x 1. So now, how do I represent this line. Well, if I take any point call it x, suppose x lies on the line joining x naught and x 1.

(Refer Slide Time: 04:02)



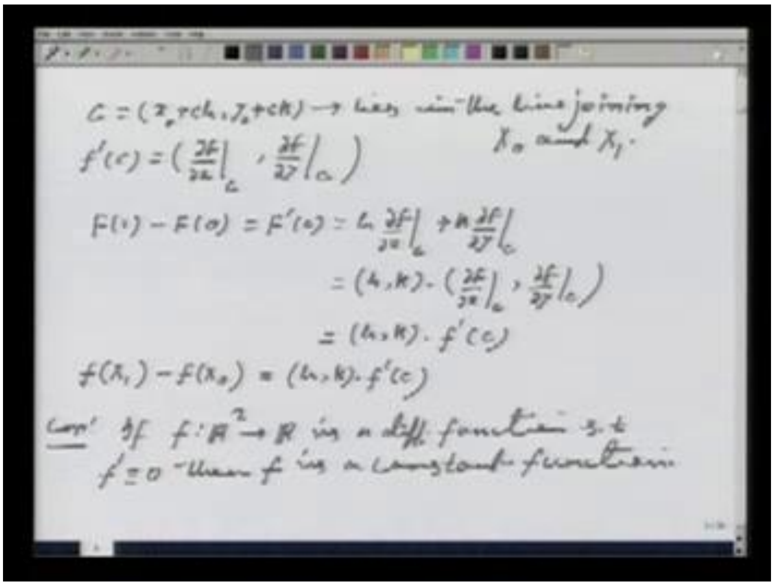
Then, I can easily see that x must have a coordinate x naught plus t h and y naught plus t k, where 0 strictly less t strictly less 1. If I take t equal to 1 I get the point x 1, if I take t equal to 0 I get the point x naught. Now, this suggests, that I look at the function capital F of t to be equal to f of x naught plus t h comma y 0 plus t k. Notice, then that f is a function from the closed interval 0, 1 to R. And since little f is a differentiable function, I get that capital F is differentiable.

And notice that, little f of x 1 minus f of x 0, this is capital F of 1 minus capital F of 0. Then certainly, I can use the 1 variable mean value theorem. That is this quantity then is f prime little c, where 0 less c less 1. Now, I want an expression of f prime c. Well in general, what is f prime t to calculate that, I can certainly use the chain rule, because capital F is given by composition of 2 functions. So, f prime t, then is del f del x times d x d t plus del f del y times d y d t.

So, this is d x d t is easy to calculate, what is my x t. Let me note it here, x of t is x 0 plus t h, y t is equal to y 0 plus 2 k. And capital F t is nothing but little f of x t y t. So, this is the perfect form for applying the chain rule. Now, then what is d x d t, it is a simple calculation to check. That d x d t is just t h, because x t equal to x 0 plus t h. So, if I differentiate it with respect to t, what I get is h.

Similarly since y t equals to y 0 plus t k, if I differentiate it with respect to t all I get is k. So, what I get is h del f del x plus k del f del y. Then, what is f prime at c, now I have to put the points. This is h times del f del x at the point x 0 plus c h y 0 plus c k plus k del f del y at the point x 0 plus c h y 0 plus c k.

(Refer Slide Time: 08:20)



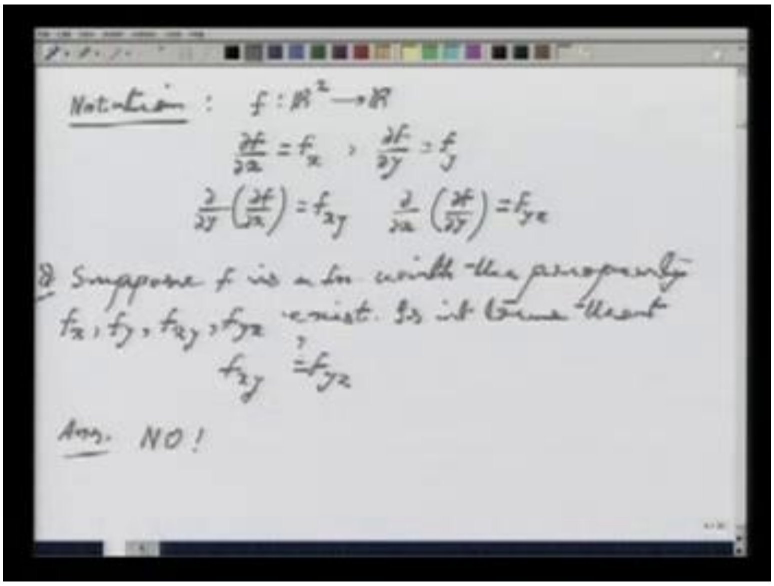
In a different notation, let capital C denotes the point x 0 plus c h and y 0 plus c k. And then what is f prime at c. Since f is differentiable, the derivative is given by the partial derivatives. So, this is then del f del x at c del f del y at c mind you, capital C here stands for the point f not plus c h y 0 plus c k. So, in this slide, then f 1 minus f 0 is f prime at little c. Which is nothing but h del f del x at c plus k del f del y at c. That means the dot product of the vectors h k with del f del x at c del f del y at c.

But, then this is h k dot f prime at c notice here the point c, which I got here. If you look at this formula, since little c lies between 0 and 1, this point capital C lies in the line joining x 0 and x 1. Then, if I put the value of f 1, which is nothing but f at capital X 1 minus little f at capital X 0, which is equal to h k dot f prime at c, where c lies in the line joining x 0 and x 1.

This is an exact analog of the mean value theorem, which we are looking for. Then obvious corollary of this is like one variable case. That if f from r 2 to r is a differentiable function. Such that, f prime is identically equal to 0, then f is a constant function. This is simple, because whatever x 1 I choose f prime c is always 0. That means, f x 1 is equal to f x 0 for all x 1, that precisely means, that f is a constant function.

So, this is a typical use of mean value theorem, which we have seen in one dimensional case. It appears in r 2 also. Now, with the usual definition of derivative in r 3, the mean value theorem, the proof of the mean value theorem actually works, even on r n. And you get the exact result. That f x 1 minus f x 0, in case of three dimension, you will get h k t dot f prime c. In higher variables there will be more variables coming. So, this is what is mean value theorem?

(Refer Slide Time: 12:21)

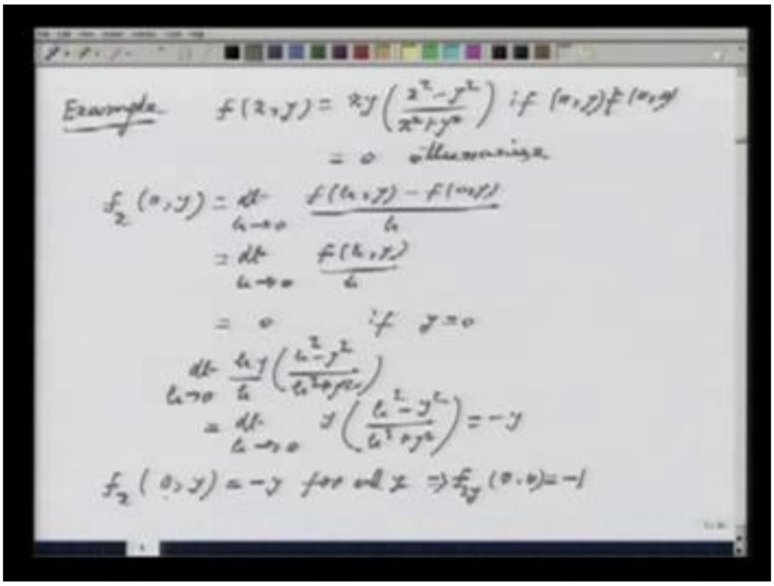


Now, before proceeding further, we ask a curious question, which we will need for that first I need a notation. So, suppose f is a function from r 2 to r. Then first I know, what is del f del x, that is the partial derivative whose notation I say is f x. Then del f del y whose notation I say is f y. Then, I can also look at del del y of del f del x. Because del f del x, again is a function of two variables. So, I can look at the partial derivative now with respect to y, this I denote by f x y.

Similarly, del del x of del f del y this I denote by f y x. So, if x comes first from the left hand side. That means, the first partial derivative is with respect to x. If the second 1 is y that means, I am looking at the second partial derivative with respect to y. Now, suppose f is a function with the property f x, f y, f x y, f y x exist. So, this is a question we are going to see.

So, suppose f is a function with the property, that the partial derivatives exists. F x, f y, f x y and f y x, is it true that f x y is always equal to f y x. These are called mixed partial derivatives. So, first I we look at the partial derivatives with respect to x, then I look at the partial derivative with respect to y. On the other hand now, I first look at the partial derivative with respect to y and then with respect to x. Do I get the same result, well the answer is in general, no.

(Refer Slide Time: 15:01)

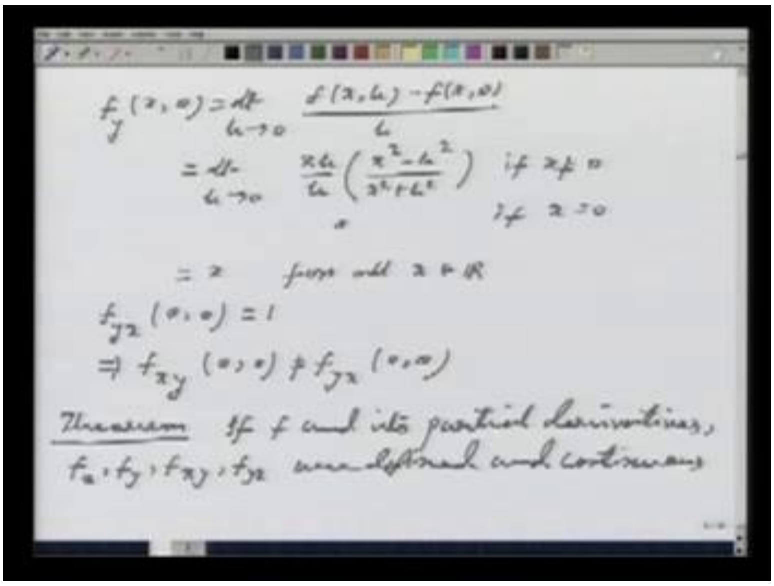


Let us see an example which illustrates that. Let us look at the function f x y equal to x y times x square minus y square divided by x square plus y square. If x y is not equal to 0, 0 equal to 0 otherwise. That means the function is 0 on the origin, it is 0 on the x axis, 0 on the y axis. And in other cases it is x y times x square minus y square by x square plus y square. Now, let us try to calculate what is the partial derivatives f x for 0 y. So, I just write down the definition this is limit h going to 0. F of h y minus f of 0 y divided by h, but f of 0 y is 0 by definition of the function.

So, this is limit h going to 0, f of h y divided by h. Now, this is certainly 0 if y is equal to 0. So, this is equal to 0 if y equal to 0. If y is non zero, then I write down the definition. It is h y times h square minus y square divided by h square plus y square, then divided by h limit h going to 0. What I get is limit h going to 0, h cancels it is y times h square minus y square by h square plus y square. What is this limit as h goes to 0 what I do is, I divide the numerator and denominator by h square.

This is simply if h goes to 0, then it is minus y square by y square. So, I get minus 1, so answer is minus y. So, in any case it follows that f x at 0 y is equal to minus y for all y. Because, for y equal to 0 I already got that it is 0. This then implies that now if I look at the partial derivative with respect to y what do I get. So, f x y at 0, 0 is certainly equal to minus 1. Because, the function is minus y if I look at it is derivative, does not matter where I always get minus 1. So, f x y at 0, 0 is minus 1.

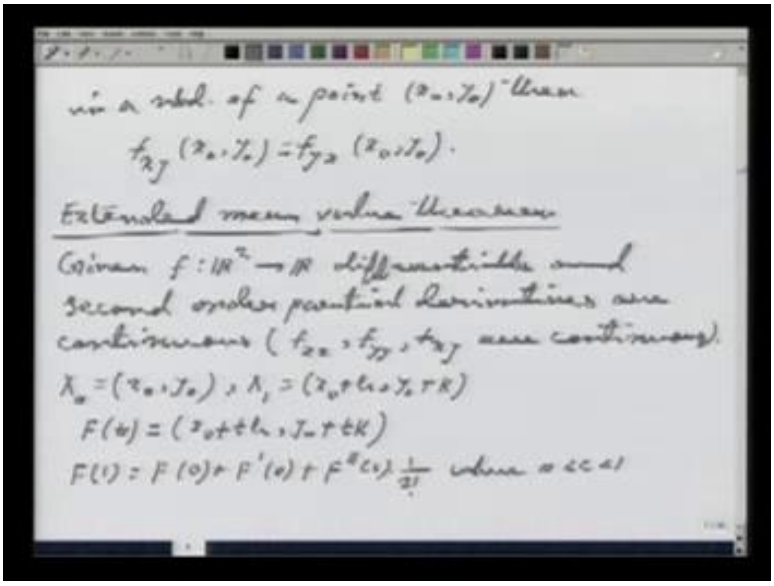
(Refer Slide Time: 18:13)



Now, let us look at what is f y at x 0 by definition. Then this is limit h going to 0 f of x h minus f of x 0 divided by h. This is then limit h going to 0 assuming that x is non zero which is x h divided by h into x square minus h square by x square plus h square. This is if x is not equal to 0, this is 0, if x equal to 0. So, if x is not equal to 0 then in this case the above limit clearly h cancels is each other x remains. And in the bracket, I can just put h to be equal to 0 what I get is 1.

So, this is turning out to be x, for all x in R. As a result now I look at the derivative with respect to x which is f y x that at 0, 0 turns out to be equal to 1. This implies f x y at 0, 0 is not equal to f y x at 0, 0. So, in general it is not true that if I have a function whose derivatives exist this mixed partial derivatives also exist. Then f x y equal to f y x, it is not generally true. But, what we have is that in most cases which consensus this would be true by the following theorem. So, it says that if f and it is partial derivatives. That is f x, f y, f x y, f y x are defined, and continuous in a neighborhood of a point x naught, y naught.

(Refer Slide Time: 20:52)

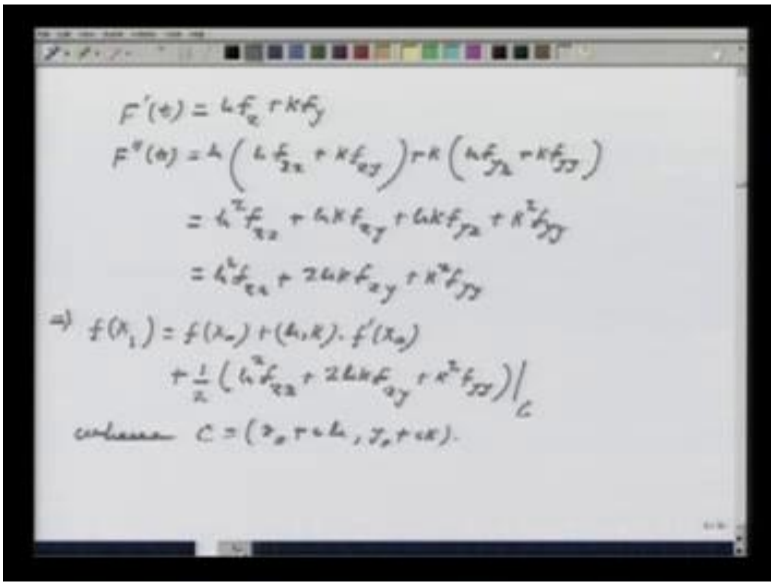


Then f x y at x naught y naught is same as f y x at x naught y naught. So, only thing which can go wrong is continuity of the partial derivatives. If you lose that, then the mix partial derivatives may not be same, but if all the partial derivatives are continuous. Then the mixed partial derivatives are also same.

So, here more general, I mean almost all the time going to work with the cases, when all these partial derivatives are continuous. So, with this in mind now we are going to extended mean value theorem. So, what kind of extension we are going to talk about. Well we are just going to use the Taylor's theorem of order 2. Remember the function which we have defined, so given let us say f from R 2 to R differentiable. And second order partial derivatives that is, suppose some such function is given to us.

I choose the point x 0 equal to x 0, y 0, x 1 equal to x naught plus h. Exactly as we have done in the mean value theorem. And I define f of t, I have k here I look at this function. Then I know this is also true that F 1 is equal to F 0 plus F prime 0 plus F double prime c into 1 by factorial 2. Where 0 less c less 1, this is the second order Taylor's theorem. Now, what does this imply in terms of the function little f from the capital F, That is what we want to see now. So, the expression of f prime 0 we anyway know. All I need to know now is what is F double prime?

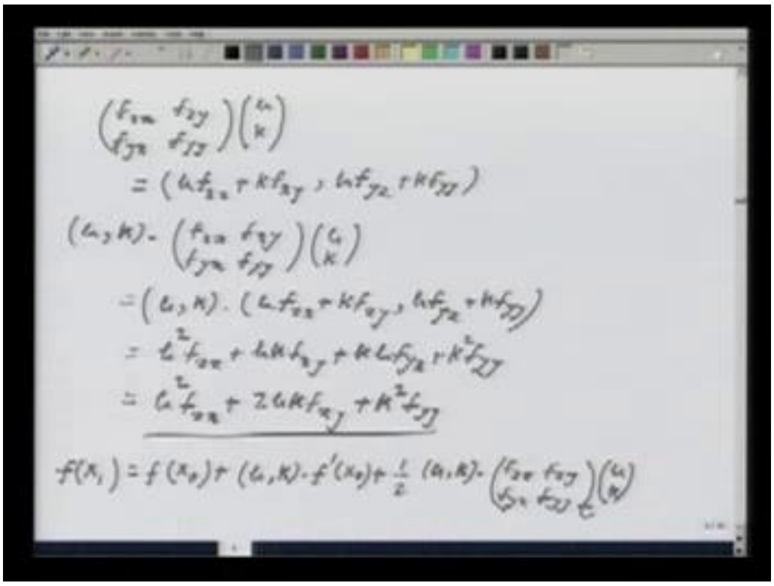
(Refer Slide Time: 25:04)



So, let me start with this. So, what was F prime t it was h f x plus k f y. So, F double prime then is now I have to differentiate again f x and f y. So, it is h f times I will write it is h f x x plus k f x y plus. Now, k times, now it is the derivative of f y again by chain rule. That is h f y x plus k f y y. So, I can again write this as h square f x x plus h k f x y plus h k f y x plus k square f y y. Now, since I have assumed that all the partial derivatives are continuous.

In fact, the second order partial derivatives are also continuous. Then by the previous theorem I know that f x y is same as f y x. Then I can write this as h square f x x plus twice h k f x y plus k square f y y. Now, then this implies that little f at x naught that means, little f at x 1 that is capital F at 1, which is equals to f of x naught plus h k dot f prime at x naught plus now comes F double prime t.

(Refer Slide Time: 28:38)



So, it is half I can write it as h square f x x plus twice h k f x y plus k square f y y at a point capital C, where capital C is equal to x naught plus c h comma y naught plus c k. Now, to write it in a different form I again use some notation what I do is. I look at this 2 by 2 matrix f x x, f x y, f y x, f y y. I apply it on the vector h k what do I get, just think of this as numbers. Because at the point c if I am evaluating this f x x, f x y, f y x and f y y are indeed numbers.

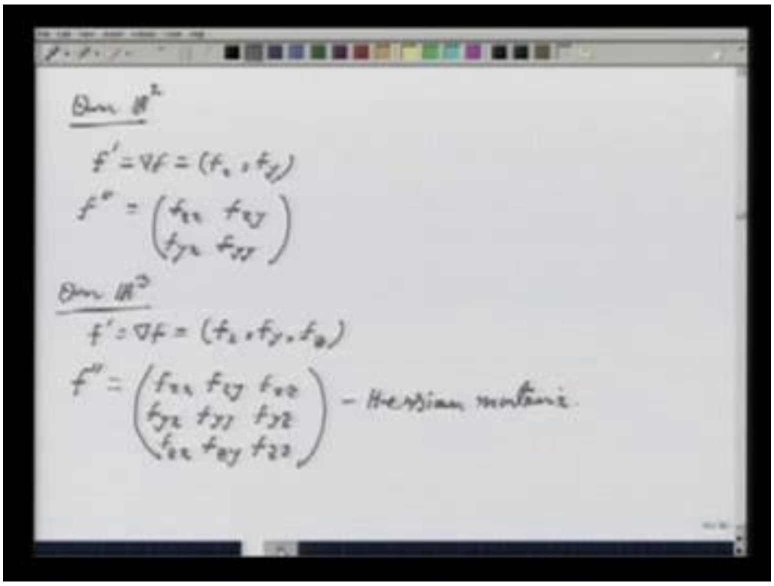
So, what I get is h f x x plus k f x y comma h f y x plus k f y y. Suppose now I take the dot product of this with h k. So, h k dot f x x, f x y, f y x, f y y of h k, what I get is h k dot h f x x plus k f x y. What I get is h square f x x plus h k f x y plus k h f y x plus k square f y y which is h square f x x plus twice h k f x y plus k square f y y. Notice that this quantity is exactly what I got here.

So now, I can write it in the form. Finally, what we got is that f of x 1, this is the extended mean value theorem; is equal to f of x naught plus h k dot f prime at x naught, which is a vector. Plus half h k dot the 2 by 2 matrix applied on the vector h k. So, where is the role of the point c, well the matrix has to be evaluated at the point capital C. So, when you calculate this f x x, f x y, f y x and f y y evaluate them at the point capital c, which is given by x naught plus c h comma y naught plus c k. So, this is the extended mean value theorem.

I am going to use this to talk about maxima minima of functions. So, this is the reminiscent of the Taylor theorem just up to order 2. If you try to go to the Taylor's theorem of order n, then the only complication which arises is we have to calculate capital F n of t. That means, the n th derivative of capital F, in the second case I got, so much of complications. If you go for higher ends, of course it can be calculated.

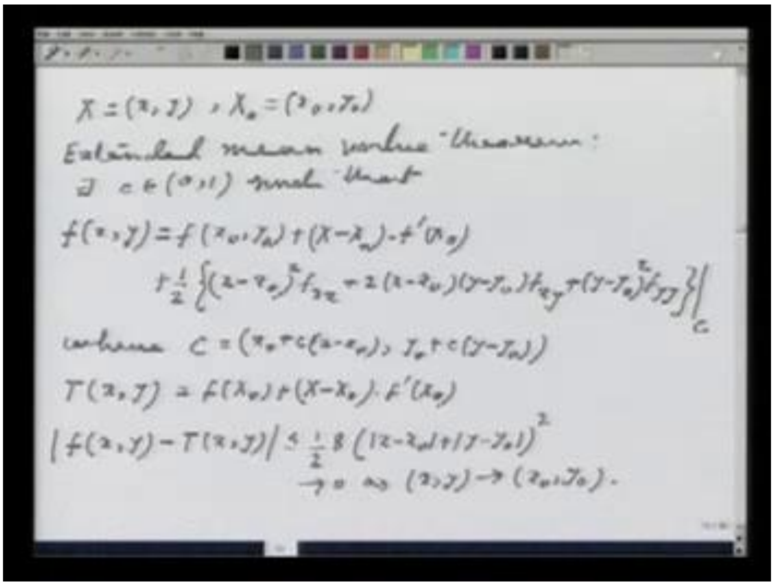
But, that form exactly we do not need and you know conceptually what it means. That you look at the function capital F, look at the n th derivative of that. And that put them in the one variable Taylor's formula, what you get is the two dimensional Taylor's theorem.

(Refer Slide Time: 33:00)



Now, notice that suppose I am on R 2. Then analog of f prime, in fact equal to f prime that is grad f, that is just f x, f y. And then, what is f double prime we got a matrix write that matrix. This is f x x, f x y, f y x, f y y this is what we got. On R 3 what happens it is analogous that f prime again is given by grad f, that is f x, f y, f z. And analog of f double prime here in this case it turns out to be 3 by 3 matrix which you can easily calculate. It is f x x, f x y, f x z, then f y x, f y y, f y z. Then f z x, f z y, f z z this is called the hessian matrix. Now, let us look back at the extended mean value theorem what did it say.

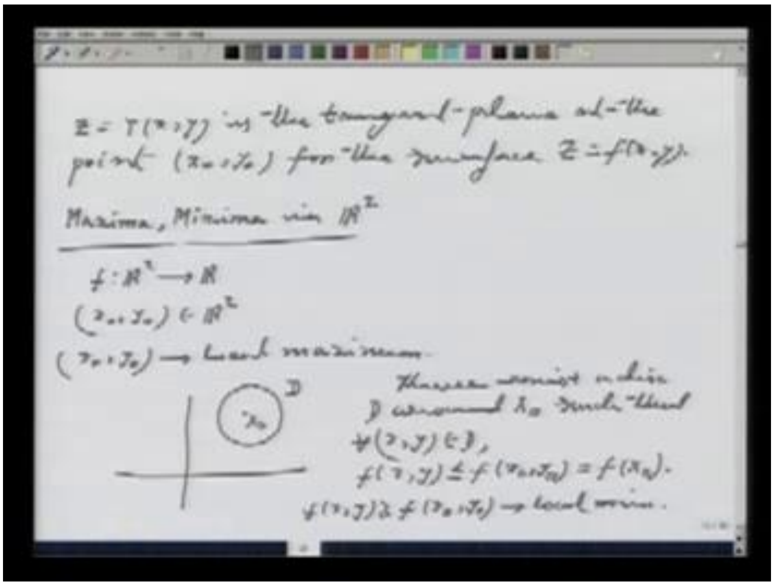
(Refer Slide Time: 34:53)



Now, instead of h k what I will do is I take a point x, which is just x y. And I take a point x naught let us say it is x naught, y naught. Then extended mean value theorem says what, it says there exist a little c between 0 and 1. Such that, f of x y equals to f of x naught y naught plus x minus x naught dot f prime at x naught plus half. I write this portion here little x minus x naught square f x x plus 2 times x minus x naught into y minus y naught into f x y plus y minus y naught square times f y y. This whole portion is at capital C, where capital C is equal to x naught plus where C equal to x naught plus little c into x minus x naught comma y naught plus little c y minus y naught. That is the partial derivatives which are involved in the last expression. They have to be evaluated at the point capital C. Now, let me define T of x y to be the first line. That is this is equal to f of capital X naught plus x minus x naught dot f prime at x 0. Once I do this then let me see what is f of x y minus T x y.

So, this then is less than or equal to half, because I am looking at a neighborhood of x y. Where all these partial derivatives they are continuous, so they are obviously bounded. So, it is lesser equal to some constant B times mod x minus x naught plus mod y minus y naught whole square, which certainly goes to 0 as x y goes to x naught y naught.

(Refer Slide Time: 38:34)



Now, notice that z equal to T x y is just the equation of the tangent plane is the tangent plane. At the point x naught y naught for the surface z equals to f x y. So, the extended mean value theorem actually tells me that if I look at the surface z equal to f x y. Then the plane z equal to T x y which is defined by this formula actually best approximates. The surface at the point x naught y naught, that is why this plane was called tangent plane.

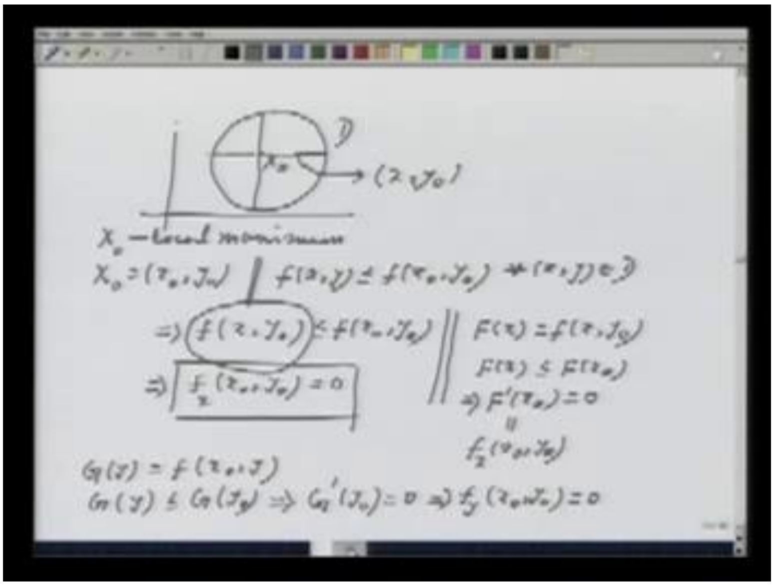
So, in the previous lecture what we have defined. This is another way of justifying that why this is called tangent plane. Now, equip with all this now I want to go to the concept of maxima minima in several variable. So, now let us look at maxima minima in two variables. So, what do you mean by maxima minima in R 2. So, let us say f is a function from R 2 to R. And let us choose a point x naught y naught belonging to R 2. Now, I want to say when is x naught y naught a local maxima what does that mean.

It means, if I draw the picture let us say this is the point x naught y naught. Let me denote it by capital X 0. Then there exist a disc around this point x naught. Such that, if I call that disc D. So, there exist a disc D around x naught such that, for all x y in this D, f of x y is less or equal to f of x naught y naught that is f of capital X naught. In one variable we said that a point x naught is called local maximum. If there exist an interval around that point, such that for all x, y in that interval f of x is less than or equal to f of x naught.

In two dimension in R 2, instead of the interval we say disc the rest of the concept is the same. That if you concentrate on the disc there for all the other points except the point x naught y naught. The value of the function is less than or equal to f of x naught y naught. So, now you obviously understand what do we mean by saying that something is local minimum. Local minimum means f of x y is bigger than or equal to f of x naught y naught. This is the criteria for local minimum.

Now, the point is how to get hold of these points using derivatives, because here we have to deal with either partial derivative or the two dimensional derivative. One dimensional derivative will not work. But, we will see that one dimensional derivatives are not completely useless in this situation either.

(Refer Slide Time: 43:08)



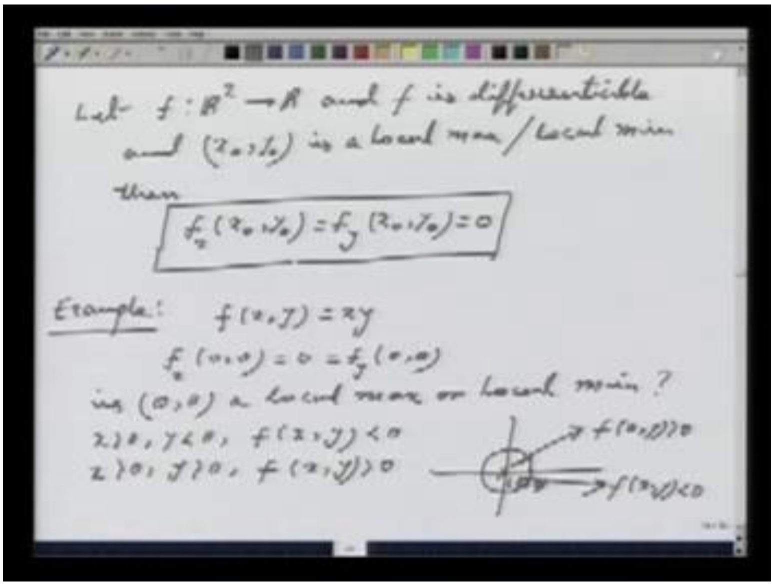
So, what we do is, let us say this point x naught is let us say local maximum. So, let us say x naught is local maximum, so x naught is given by, let us say x naught y naught. So, for all points on the disc. Now, I know that f of x y, if I call this disc D. That f of x y is less than or equal to f of x naught y naught for all x y in D. In particular, this implies that f of x y naught is lesser equal to f of x naught y naught. That means, that if I look at this line on this line all the points are actually look like any point on this line look like x y naught, the height is y naught.

Now, if you notice the left hand side that is here. This is a function of one variable and for that 1 variable function f of x naught, y naught. Now, y naught is fixed is a local maximum. Because, now I got an interval that implies by the one variable, result that the partial derivative f x at x naught y naught must be equal to 0. So, what we are saying is now define a function capital F x equal to f of x y naught. Because y naught is fixed, so only thing which varying is little x.

So, it is a function of one variable and I see that f of x is less or equal to f of x naught in an interval. This implies then that f prime at x naught must be equal to 0. But since I know f x capital F x is little f x y naught capital F prime x naught is nothing but, f x at x naught y naught. That is why I got the condition that f x at x naught y naught is 0. Now, the same r argument applies on this line also. This is where the x coordinate is constant only y is varying. So, I define now let us say g y that is f of x naught times y.

Then, I know, that g of y is lesser equal to g of y naught. Because, for every point x y on the disc, f of x y is less than or equal to f of x naught y naught. So, this g y naught g achieves it is maximum at y naught, so g y is lesser equal to g y naught. So, again the one variable, result implies that g prime at y naught is 0. But then this would imply, since I know that the form of g y is just fixing the first coordinate of f. This would then imply that f y at x naught y naught is 0.

(Refer Slide Time: 47:04)



So, what we have achieved here is as follows. Let f is from r 2 to r and f is differentiable. So, that would automatically imply that the partial derivatives at least exist. If this is the case and x naught y naught is a local maximum. Or it will equally work with local minimum. Then f x at x naught y naught equals to f y at x naught y naught equal to 0.

So, this is analogous to the case that to find the extremum. You always look at the solutions of f prime x equal to 0. In two dimension, what you have to do is you have to look at the partial derivatives f x and f y. And look at their common 0. But, this is a necessary condition, it is not sufficient exactly like the one variable case. That is if f x and f y both vanish at a point is, it a maximum or a minimum.

So, we look at an example for that, let us look at the function f of x y equal to x y. Then I look at what is f x at 0 0, that is 0, because f x is just y. So, at 0 0, it is 0 which is same as f y at 0 0. So, question is is 0 0, a local max or a local min the answer is it is neither a local maximum nor a local minimum. Because, if I take x bigger than 0 y less than 0. Then, f of x y is strictly less than 0, and if I take both of them to be bigger than 0. Then f of x y is strictly bigger than 0, but now again this is the point 0 0. If I take any disc around 0 0, it will contain x y where both coordinates are positive. Because, I can take points from the first coordinate or it will contain points, where x is positive y is negative. That means, I can choose points from this region. So, if I choose points from this region. Then f of x y is negative, if I choose points from this region, then f of x y is positive. So, at this region f of x y is bigger than 0.

At this region f of x y is less than 0. So, whatever ball I choose around 0 0, I can always get points. Where the value of the function is bigger than 0 0 and the value of the function is less than 0 0. So, the point 0 0, the origin is neither maximum nor minimum. So, if you look at the partial derivatives f x and f y and collects its zeros. That does not mean, all zeros are either maximum or minimum.

But, what we can guarantee is that, wherever is maximum or minimum, they actually counts in that set. That is maximum and minimum, if that is given at some x naught, y naught. Let us say then the partial derivatives vanish there. But, just vanishing of partial derivatives is not enough, so then the question appears that. Then what are the extra things, we need to characterize the maximum and minimum of functions of two variables.

So, that is the question, we are going to deal with in our next lecture. So, that would be the analogue of the double derivative test as we do in the one variable case. But, that is for the next lecture.