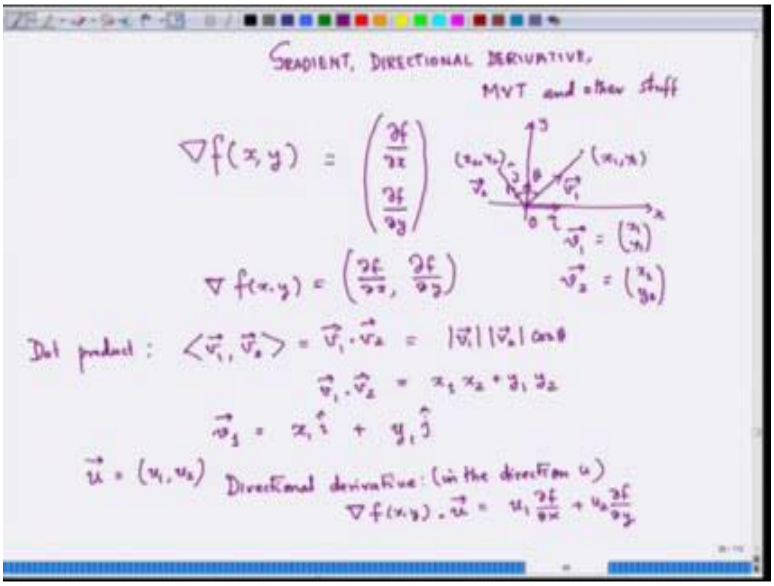
Basic Calculus for Engineers, Scientists and Economists
Prof. Joydeep Dutta
Department of Humanities and Social Sciences
Indian Institute of Technology, Kanpur

Lecture – 22
Differentiation of Functions of Two Variables – 2

Now, we are coming to the second part of our story about the derivative associated with functions of two variables. Again, I would like to assure you that you really have to give a little more concentration; here we are going up with differently.

(Refer Slide Time: 00:40)



Note that we have been writing a gradient directional derivative MVT; it is not MTV, by the way it is MVT means Mean Value Theorem and other stuff. So, just have to be very attentive here. Gradient of a function of two variables at any given point x, y or x naught y naught does not matter whatever you want to say is written like this.

And this is nothing but a vector consisting of. Now you might question me what about vectors, you have not spoken about vectors, it is too difficult to speak about everything in the same course. But I assume that you know what is a vector, so vector in a two-dimensional geometric framework is object which has both the magnitude and a direction. If you take a point in a two-dimensional plane and joint it with a straight line with the origin o, then these are the particular direction and of course, it has a particular magnitude. If you change the point to some other point here, for example, so it has another direction and it has a magnitude so these quantities are called vectors.

But it can be easily seen that this point is this point which we had drawn first can be given a coordinate x 1, y 1 because this is in the x y plane; and the second one is given x 2, y 2. So, x 1, y 1 has a different magnitude and a direction; x 2, y 2 has a different magnitude and a direction. So I can instead if I call this vector as v 1, and if I call this vector as v 2, it does not make any problem, if I say v 1 is x 1, y 1, or you can write it just x 1 y 1 in this same rho type fashion.

This is, I am writing it like this because this is one type of the standard ways how we are always writing it. You may not write it like that also you may just you can write like this as you know in informal rho that is if you write like this I think mathematician preferred this physic is preferred though I am writing now. It does not matter whatever you want to do.

So you have here that I am essentially expressing a vector in terms of the its coordinate. So, here is basically the coordinates are point whose x coordinate is x del f del x y coordinate is del f del y. This is a vector because these are real numbers right. So if you put some x particular value of x and y, this will give me one real number this will give me another real number. There is a notion of dot product of two vectors; so it is either denoted like this or denoted like this, this is called the dot product.

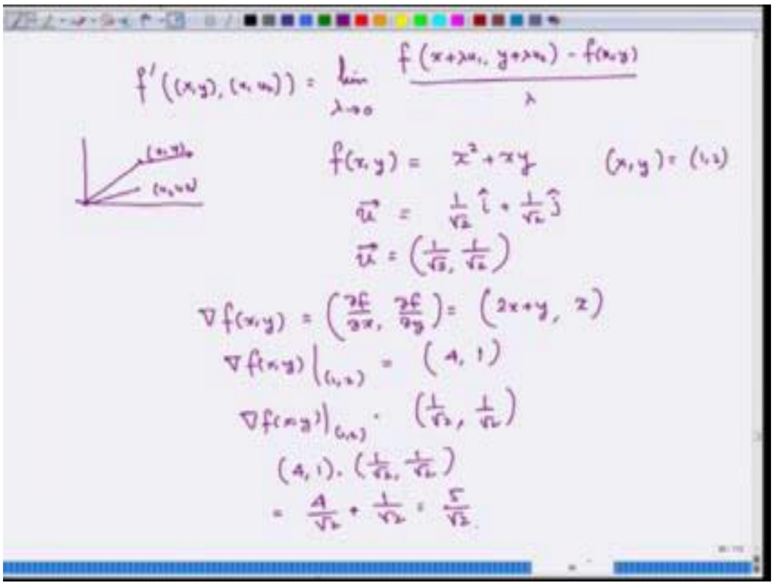
And how do I compute, because just (Refer Time: 03:54) v 1 dot v 2, it does not mean anything; it essentially means the product of the it essentially means that the product of the magnitude this absolute value of v 1 actually means the magnitude of v 1 into the magnitude of v 2 into cosine of the angle between them theta - cos theta. This can be much more easily computed when you have represented in terms of coordinates. So in terms of coordinates, v 1 dot v 2 is written as x 1 into x 2, if I have y 1 into y 2 and this is exactly same as this.

The interesting part is that all of you have learnt that I can think of unit vector i along this direction, and I am thinking of a unit vector j along the y-direction then any vector v 1, this is possibly known to can be expressed in the following way x 1 into i vectors are usually given with the hat to show their specialty. Those who know some linear algebra will immediately understand this we are talking about the basics or vector space. I cannot resist telling these terms, because I am a mathematician, but from a physicist point of view, these are simple unit vectors and you can express any vector as a skill of as a addition of two unit vectors skilled up by the amount given by the coordinates of the vector. So it is written as y 1 j vector.

Now, once you do that, we can now think of something called directional derivative that is finding the derivative of the function of two variables in a certain given direction. So given direction would be another vector, see it is a u vector. And u 1, it is given in terms of two coordinate's u 1 and u 2. And I am telling, I am asking what is the meaning of derivative or what is the meaning of the directional derivative of the function f in the direction u. This is given as follows.

In the direction u, it is given as the inner dot product of the gradient vector into the u vector, so it is u 1 del f del x plus u 2 del f del y is as simple as that. Of course, there is a much more involved definition which many of you might not find very comfortable, but we will just provide it, it is in your book also.

(Refer Slide Time: 07:23)



It is says that it is ok. If you are talking about the directional derivative of the function x, y at the point x, y, in the direction u 1 and u 2, what you do is move from x, y in the direction u 1 and u 2, so this simply means that if this is my x, y point this is my vector x, y, and this is my u vector with u 1 and u 2 as coordinates, so I have to move from x, y in the direction of u 1 u 2 that is in the direction parallel to the u 1 u 2 vector, it is just parallel to the u 1 u 2 direction. And in that direction if I move what is the derivative of this function. It essentially means to find limit lambda going to 0, f of x plus lambda u 1, y plus lambda u 2 minus f of x y and that divided by lambda. There are lots of issues here; I am just not going to go through.

This is the simple definition, and in this book also there is a say f (x, y) is equal to x square plus x y; f (x, y) is equal to x square plus x y and they ask you guys using this first principle find the limit. And for the directional derivative in the direction u 1 u 2 or rather I would say that they have given in the direction of a vector u, which is given as 1 by root 2 I vector plus 1 by root 2 j vector.

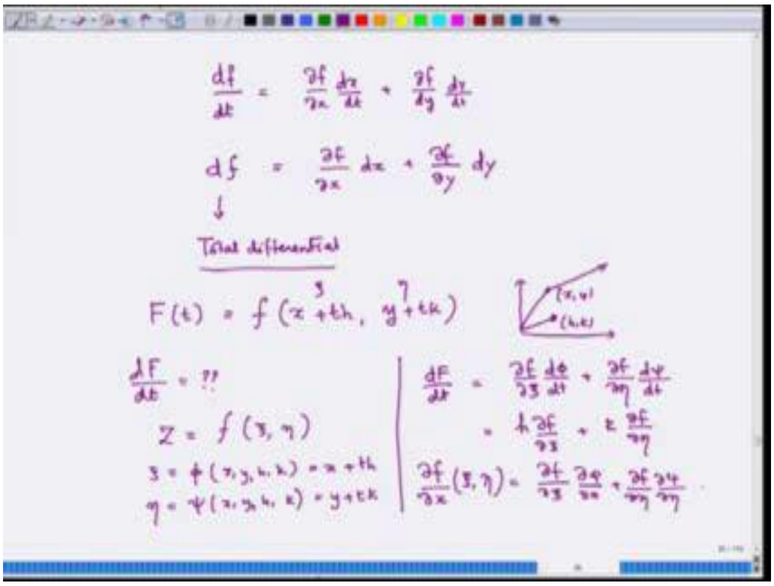
Now, you will immediately realize the following fact that this is u or u hat is nothing but the vector in the Cartesian, if we once again in the Cartesian coordinate this is representing the coordinates of 1 by root 2, 1 by root 2, and I have going to find the directional derivative. If you go with the first principle, the answer given in the book is 5 by root 2, but you will see that this is much easier to find in the way I have shown you. So, first I will find the gradient vector which is del f del x del f del y and that is given here as 2 x plus y. And of course, 5 by root 2 answer cannot come unless you are looking at a certain point, which I have missed sorry.

So, you are you are telling that you have to find it in the directional derivative in the direction u, at the point (x, y) which they have given as (1, 2). So I will now compute it add x is equal to (1, 2). So, what is del f del y here del f del y is just x. Now, I have to now compute the grad f (x, y) vector at the point (1, 2) and that will give me I will put x is equal to 1 and y equal to 2, so that will give me 4 and this will give me 1, (4, 1). Now my directional derivative is basically taking the dot product of f (x, y) at (1, 2) with the vectors u, which is 1 by root 2, 1 by root 2. I would have (4, 1) dot 1 by root 2, 1 by root 2, so this is nothing but 4 by root 2 plus 1 by root 2 that is 5 by root 2, which is same as the answer given in the book, which they have calculated by doing this limit. So, you see this is exactly the fact.

Now, once directional derivative have some uses, they are lot hugely used in optimization which is a subject where you talk about maximization and minimization of function, which we will not get into unless we come into the next thing. There are certain issues like what do you mean by is there a mean value theorem for functions of two variables that is a big question. So, we have going to now talk of about the mean value theorem for functions of two variables; and once we do that, you will have a interesting idea about how things can be bought from the higher dimension to the lower dimension, so there is something called a total derivative.

What we have spoken is a partial derivative. So what is called a total derivative, so or the total differential. So, you have already I will just come, but let me just mention this concept of (Refer Time: 12:47) full differentiability slightly difficult at this stage to give you which is that is why I am not going to talk about the concept of full differentiability. But what I am going to talk about is that I can talk about what is called differential of the function, and how to use that idea to speak about the mean value theorem.

(Refer Slide Time: 13:14)



Now, for me do this, see we have already spoken about this chain rule. The total differential of a function of two variables is actually defined as follows, del f del x into the differential of x plus del f del y into the differential of y, so it is nothing but the extension of the whole idea what you had in function of one variable. Function of one variable if I refer the function of one variable then you essentially have to close this part, then you do not you can actually forget this part.

But because now you have the function of two variables, you essentially have to bring in the y component, so you have this part so this is total differential of a function of two variables. So I call this has a total differential. Now, this total differential has some meaning; the meaning is as follows; it can be usefully you can use two approximate functions. This leads us to what is called the mean value theorem.

Of course, I can use the function of one variable to handle all this, but instead of going into the details, for example, let me tell you one thing. For example, if you take a function f of t, function of one variable t, take a fixed x and fixed y, and a fixed h and a fixed k; so h and k, these are these are fixed vector which is giving you direction, and x and y are fixed points - reference points. So, we are moving from that so basically from (x, y), you are moving in the direction of h and k, and h and k is here. Now this is the origin and this is the full stop. The similar things like that directional derivative.

Now, if I take the derivative of this f dash t, d F dt, how will I do it d f d t is nothing but taking the chain rule what we have learnt, because this now I have made these variables as a function of t. So, we want to compute d F dt; I just wrote del x d t and I have to move ahead, but that I thought that and just do it and ask how to compute this.

So just for a second observe the following that consider here the function f where (x, y), (h, k) are the independent variable, t is for the moment of constant. Then I can write this Z as f of psi eta, where psi is a function say phi of x, y, h, k and that is given by x plus t h. And eta is a function of function say psi of x, y h, k given as y plus t k. Now t is the constant for the time being.

Now if I go by the chain rule, now psi for the moment, if I fix up x and h, (x, y), (h, k) fix up then psi and I start varying t then psi becomes the function of t, and eta becomes the function of t. Then by that rule my d F dt what I have learnt earlier should be del f del psi, because this is my psi and this is my eta del f del psi del psi del eta which is same as writing as del phi del eta or d psi d t or d phi d t is the same thing plus del f del eta d psi d t.

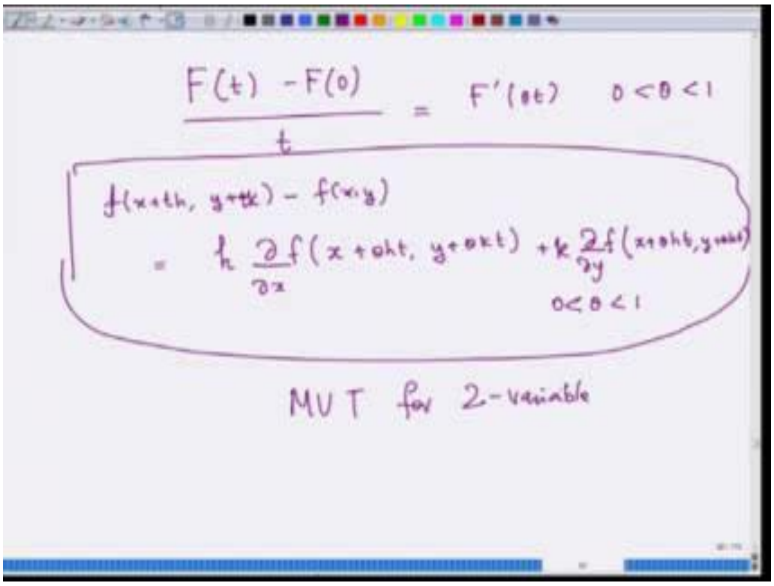
So if you look at so now, what is now del del f del psi at the point psi eta basically. If you look at this so it becomes del psi del t is just h so it is h into del f del psi plus k into del f del eta. Now, once I have done this, I need to know how to calculate del f del psi in terms of del f del x, because psi is a function of x, h y, k. And if I want to calculate del f del x on this function, because I can write this as phi of x, y, h, k this as phi of h k del f del x then what should I do then I should go by the chain rule.

So, del f x, now basically I am computing del f del x at psi eta, this is; and then if I go by the chain rule what I should do I should first calculate del f del psi into del psi into I should have now del f del psi into del phi of or del psi del x or del phi of del x del f del eta into del psi of del eta; similarly, for the others h and k.

Of course, it does not mind, if I for the moment even fix up my h and k, it does not matter. I can even forget the h and k and say h and k are just constants and then go on so. It should be del (Refer Time: 20:07), so if you write now del f y del h so del f, so you have to write del f with respect to del h here, basically then again then you have to write del phi by so you have to write again here psi and eta del phi del x del phi, so you have to again come and write del f del psi basically you with respect to psi, here you have got del phi del x. So this is the chain rule, so del phi del x here is just one remaining goes.

And del psi del x here is just 0, because there is no x, so it becomes del phi del psi at into del phi del x which is 1. So what is happening so del f del x at psi eta is same as del f del psi at psi eta. Now I can then write this d F dt as h into del f del x that is k times del f del y at psi eta so these are all evaluated at the point psi eta, whatever be or given the psi eta does not matter. So, psi is a f of x plus t h for some given t, so that is exactly the calculation. Now this will immediately lead us to a mean value theorem, very simple.

(Refer Slide Time: 21:58)



So, what I will do, take the function F (t) minus F (0) divided by t, this is nothing with the function F is differentiable which I am seeing that it is differentiable I can computed out is some for some c which is lying between t and 0. So, any point lying between t and 0 is a fraction of t, where theta is something lying between 1 and 0. So, once this is done, you can immediately write down the whole thing. So, you know what is F(t) and you know what is F of 0.

Immediately once you do it, because if you, so it will become f of x plus t h y plus t k minus f of x y is equal to what is this, f dash theta t. So, what is f dash theta t it is nothing but h into del f del x computed at x plus theta h t y plus theta k t plus k into del f del y computed at x plus theta h t and y plus theta k t, where theta is between 1 and 0.

Just by using the function of mean a value theorem for one variable we have got the mean value theorem for two variables; MVT for so del f del y for two variables. So, here again we have calculated at del f del x at x is equal to h. when instead of t, I have put theta t; so instead of h t, I have h of theta t instead of k t, I have theta of k of theta t, the same thing here. So, you see how much the chain rules are involved and how much calculations we did. So, you really need to go through this calculation once yourself to convince yourself.

With this, we end the session, and we come back to talk about Maxima and Minima of Two Variables in the next class.