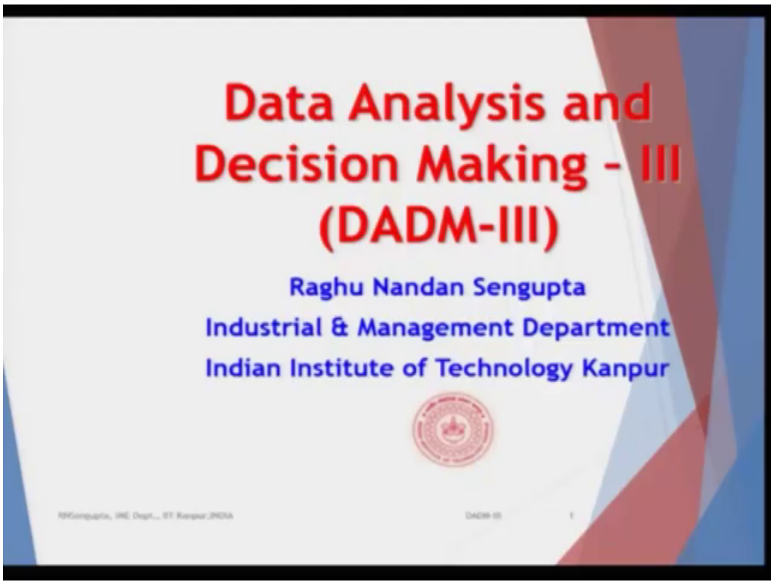
Data Analysis and Decision Making - III
Professor. Raghu Nandan Sengupta
Department of Industrial and Management Engineering
Indian Institute of Technology Kanpur
Lecture No. 51

(Refer Slide Time: 00:27)



Welcome back my dear friends, a very good morning, good afternoon, good evening to all of you wherever you are in this part of the globe, and this is the DADM 3 which is Data Analysis and decision making 3 course on the NPTEL MOOC series. And as you know this course total duration is for 12 weeks which is the total contact hours is 30 and the number of lectures is 60 considering that each lecture is for half an hour and we are going to start the eleventh week that is the eleventh and a twelfth week are left and each week we have 5 lectures of half an hour each, you have already completed 10 lectures, which means you have already taken 10 assignments.

We will complete the eleventh one you will take the eleventh assignment then you continue the twelfth one and we will take the twelfth assignment and they would be a final examination after the end of the course. And my good name is Raghu Nandan Sengupta from the IME department at IIT, Kanpur. So, if you remember in the last almost 5 lectures or 4 lectures we basically considered 2 detailed examples of branch and bound and with certain flavours in one of then the branch and bound was the problem where there are integer solutions only for the decision variables.

And in the other case the decision variables was 0 1 programming, so and we have already done linear programming and all this things we will consider other models later on also but today we will go into the little bit of quadratic programming then go into the concept of reliability based optimization then later on go into robust optimization.

And then again come back with multi objectives few formulations and we have already done the genetic algorithm, (())(02:08) handling, consider all this things we will just brush then and do not go into the depth but just brush then accordingly and go and consider few of the important topics such that we are able to handle those type of problems accordingly. As, you see in the slide we are in the eleventh week, we are going to start the 51st lecture which is the first lecture in the eleventh week. So, we will basically go through the concept of quadratic programming, without any theory consider a concept of quadratic programming and what are the basic (())(02:47) based on which quadratic programming can be solved.

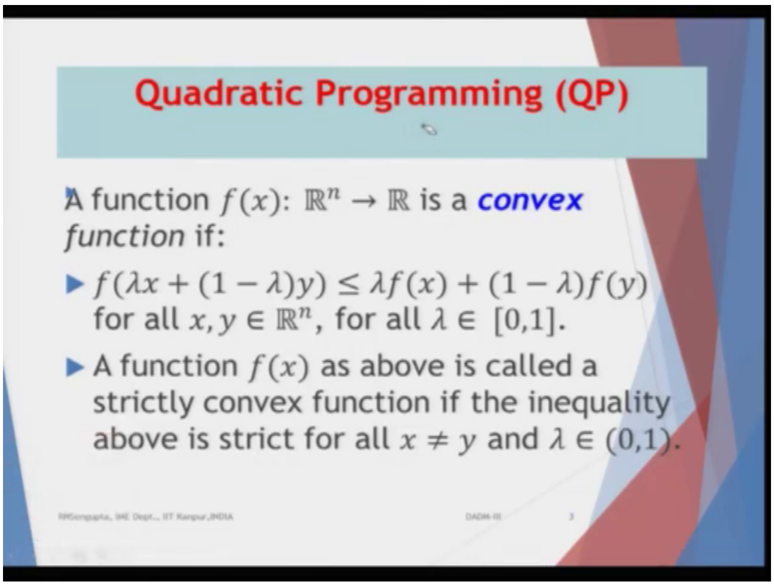
And the quadratic programming concept we have already discussed in one of the problem formulations where you are required to basically you are actual objective formulations would be half xQx where xQ and x all 3 are vectors corresponding like x is a vector and the x transpose is a vector also because when I am writing xQx obviously one of them would be a transpose that x1. And Q is basically a matrix of corresponding size and then that is the quadratic part and if you remember I had discussed that how that quadratic programming can be formulated from the point view of finance where you have a set of assets you want to formulate a portfolio try to minimise the variance of the portfolio where the variance concept is basically something to do with a square term which from where we can understand that why I am saying that is it can be formulate as a quadratic programming.

And there is a square term and there is a linear term also depending on, on the fact that in the quadratic programming you can have c into x where x is also vector c is also vector is that basically s e can be a cost structure and this can be considered as the thus the overall expected value of portfolio where now in this problem when you are talking about the finance thus axis are basically the weights w and each w is between 0 and 1. Or if they can be between wi minimum or wi maximum where i is basically the stock or script number and their total, in totality there are capital N number of scripts. So, coming back to the slides, so consider by we will just go through some basic very basic definitions all of you are aware but I will still repeat it because they would be utilized later on in the concept of quadratic programming and further on, when we are discussing the concept of reliability and robust optimization and how the problems can be solved.

So, consider you have a function which you are mapping from the N space to the real line and N space means there are N number of variables so consider in the same way if you have N numbers of stocks you want to basically find out the combined weights of each and every stock which will give us the overall portfolio expected value. So, you have basically trying to find out the functional form from the N space to the real line so that is why it means is a convex function if an only if we have this. so, what it would mean that the convex combination of the variables and their functional form would always be less than equal to the convex combinations of the functions if taken separately. And this would be utilized say for example I give an example were they can be utilized not about convex of this convex function but about the convex combination of this functions.

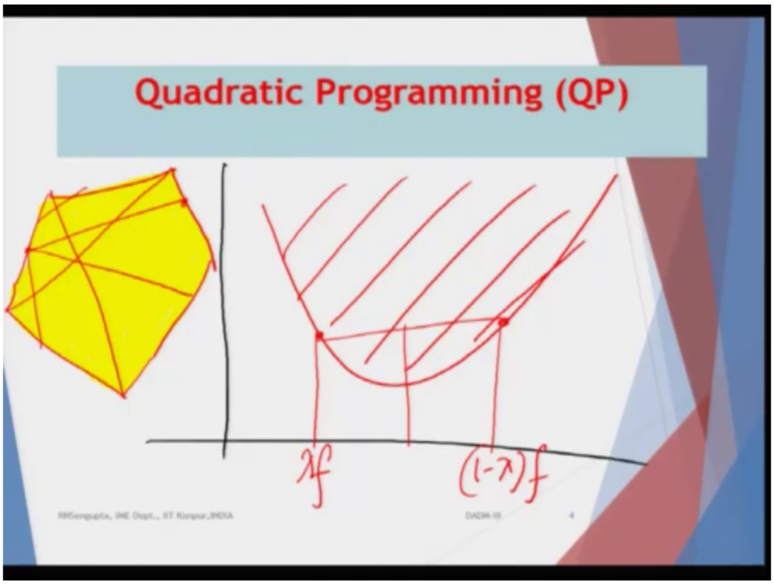
We have already discuss some of this problem in a very nice way in the multiple linear regression step up when we were considering the balance loss function. With the balance loss function had the concept where the precision or estimation and the goodness of fit were being considered, precision on estimation was related to trying to find out how good or how close the values of beta hat are with respect to beta and obviously we will use take the recourse of trying to find out whether the properties of unbiasedness consistency holds for the beta hat values. And here I am saying that beta hat values can be either a vector or scaler but if it is a linear multiple regression simple one you have only one x so one beta, so beta is scaler if you have multiple linear regression you have more than one beta so obviously in that case beta would be considered as a vector.

(Refer Slide Time: 07:12)



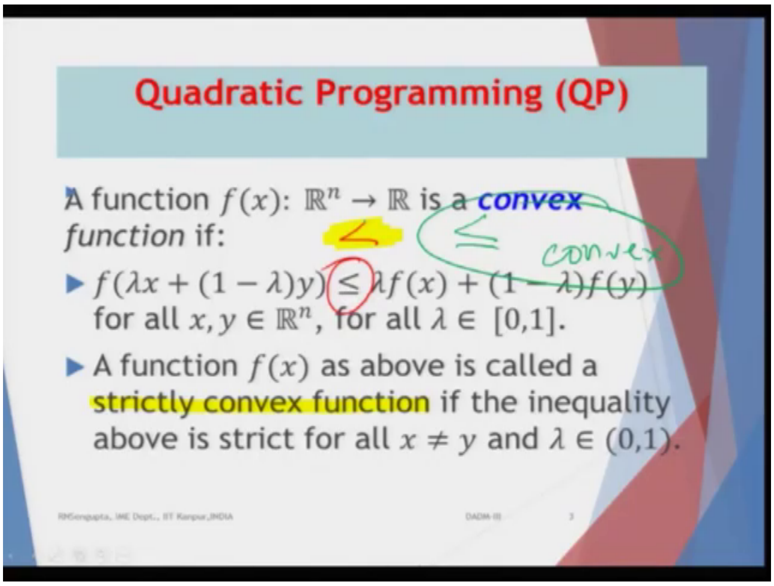
So, functional f of x as I said is a convex function, if the convex combination of the variables and their functional form is less than equal to the convex combination of the functional itself. So, is basically means that if I have, so the (con) concept of convex and concave functions if you consider the spaces would be like this, so if I have.

(Refer Slide Time: 07:57)



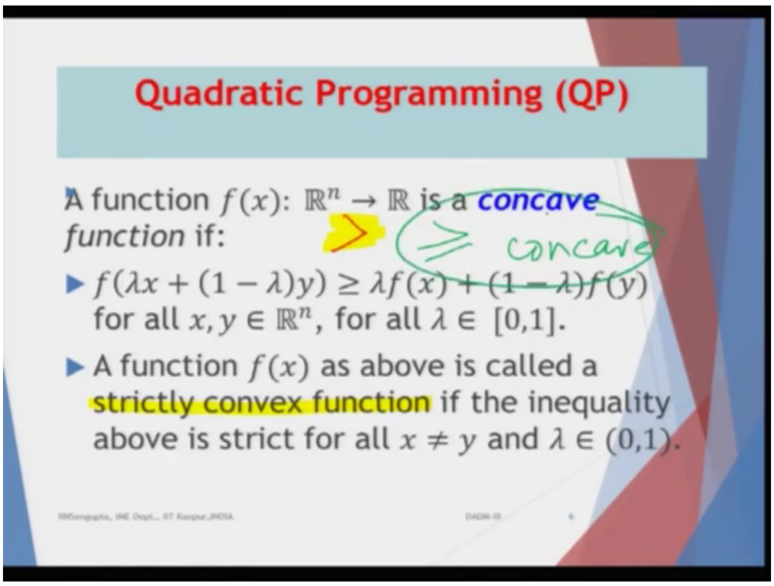
Let, me utilize the separate slide so it will easier for all of us to consider. So, here I will consider you have a function, now what you have is basically you want to join them, so you are taking lambda of f 1 minus lambda of f the function form, which ever lambda is half half so you are taking 50 percent 50 percent of them and case if we say. So, obviously all the points would be inside, so if we consider this space this one, so all the once are inside. So, any combinations are inside, so I am just drawing all the sets of points which are inside. So I just colour it in order to make it much more prominent because a red in the yellow background would look good. So, if I am taking 2 points any orbit points join it, it is always inside, join it is inside, join it is inside, join them it is inside, join them it is inside, so obviously this would be convex.

(Refer Slide Time: 10:05)



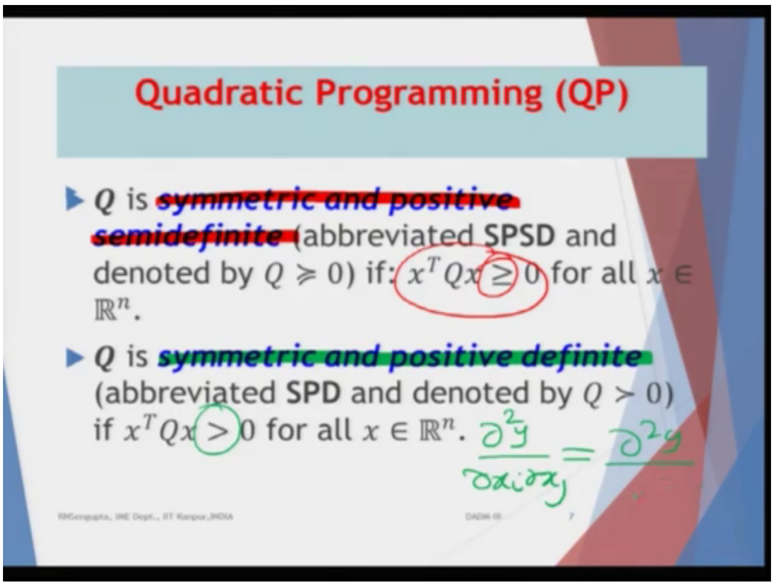
So, in case if the less then equal to sign is replaced by less than sign it will be a strictly convex and it is less than equal to it will be convex function and in the case obviously lambda would always be between 0 1. So, a function as above is called a strictly convex is the inequality sign is of strict type, that means you have this, so in that case it would be a strictly convex function. In the case less than type oh sorry it should, less than type it will be utilized as a convex function of the not strictly, strictly word will be use for the less then type only, not the less then equal to type.

(Refer Slide Time: 11:13)



Now, I go to k function, again you are mapping from the N dimensional space to the real line one dimension and in this case if the convex combination of the variables that function is greater than equal to the convex combination of the function by itself, then it is called an an obviously lambda is between 0 and 1 and x and y whatever the decision variable of the variables are in the end dimensional space. So, this would be a concave type and if the greater then equal to sign is replaced by the greater then sign only it will be a strictly concave type. So, again let me mark it, so greater than type I have the strictly convex function and in case I have the greater then type, it should be convex, concave sorry my mistake concave this will be concave function. So that I can change it. So, this would be concave function. So, a function fx is called a strictly concave function if the inequality is strict of an for all x not equal to y and lambda is between 0 and 1.

(Refer Slide Time: 12:53)

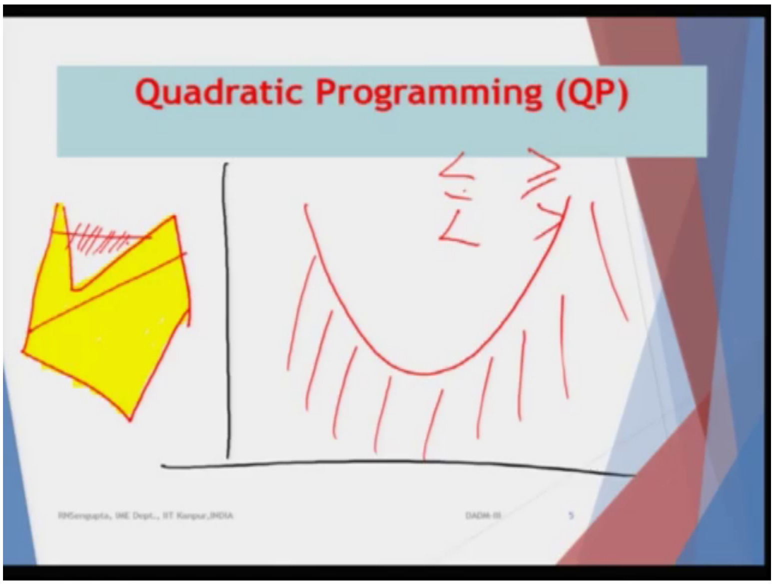


Now, key Q which is the symmetric, which use the matrix which in the case for the finance one was basically variance covariance matrix it is symmetric obviously because the principle diagonals are all the covariance of the stock with itself which is the variance, which is the first element 1 comma 1 is sigma suffix 11, 2 comma 2 is sigma suffix 22 and the last element in the N comma N is sigma N square N comma N. And if we consider the off the diagonal element they are symmetric because sigma ig is equal to ji, so Q is a symmetric and positive semidefinite matrix and in this case it will be semidefinite and definite depending on the concept if the greater then equal to or greater than sign basically happens.

So, in this case, if the x transpose Qx if it is a greater then type greater than equal to type it is semidefinite and if it is greater than type it is definite. So, I should basically use a different colour, chose green, so it is symmetric and positive, positive definite obviously in that can if it is the signs are reversed or it is less than equal to and less then it will be negative, symmetric and negative and symmetric and semidefinite and positive and semidefinite. So, in this case I have symmetric and positive semidefinite so it greater than equal to sign hold in this greater than sign holds. So, Q you are basically taking in any and some of the cases it will be taken as the hessian matrix which is the different double differentiation which is d2y dx2.

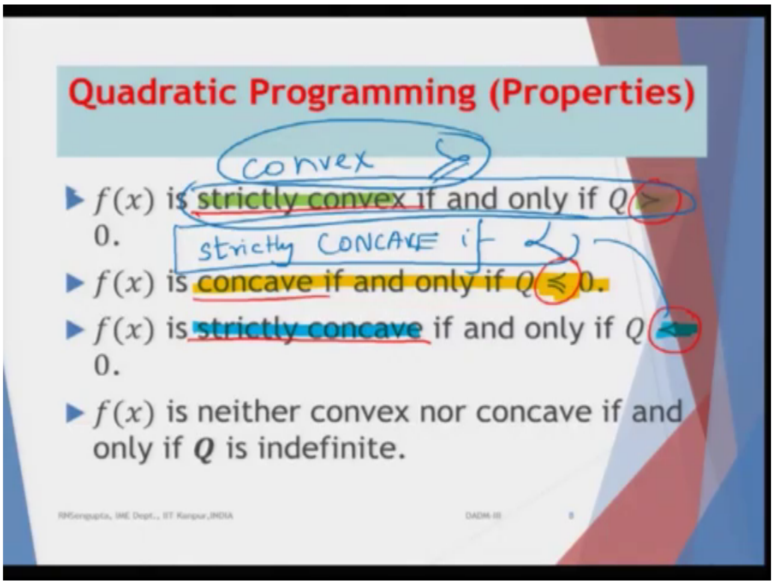
So, in the first element 1 comma 1 you will have del 2y del x1 del x2 considering it is basically of an n dimension, so for the first element it would be del 2y del x1 whole square then the 2 comma 2 element will be del 2y del x2 whole square. So, this 2 which I am mentioning is basically the suffix and the n comma n element will be del 2y del xn whole square and the off the diagonal element we will considered the symmetric that means del 2y del xi del xj is equal to del 2y by del xj and del xi which means technically I would have for symmetricity del 2y del xi del xj is equal to del 2y del xj del xi, so this should hold.

(Refer Slide Time: 15:58)



So, in the case of the concave one let me draw the diagram which I did not do, so in the case of the concave this is and the points which you have, so if I have, I have drawn just drawn it, very simply in order to make you understand, so this is, so again colouring out yellow. So, in case I points so this seems to be convex the moment I have this, this whole portion outside, so this is not as a concave area. And similarly you can basically draw as you have drawn this one as a convex the opposite portion in this area you can have basically, we are looking from the other side it would be concave one. And again less than type less than greater than type greater would be in this case it would be strict and then this would be just less than type or concave or convex strictly concave, strictly concave strictly convex and another case would be just concave convex.

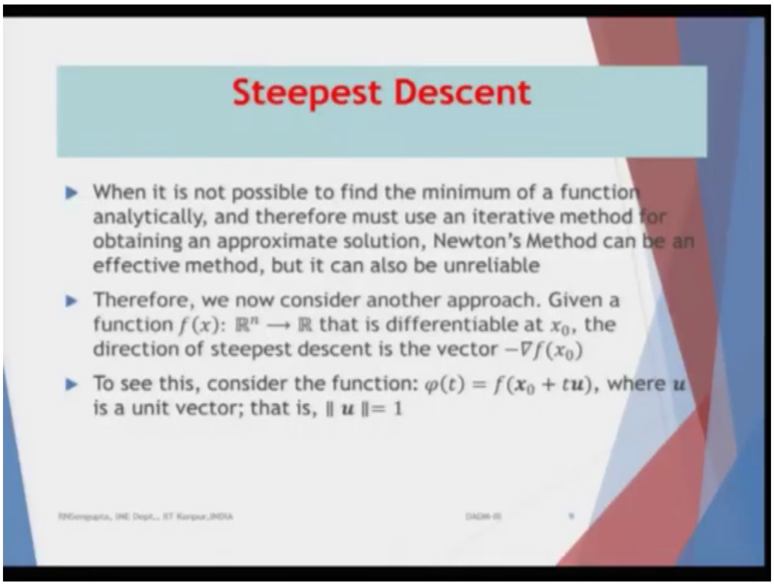
(Refer Slide Time: 17:55)



So, in n case if the hessian matrix this cube factors for the semidefinite and definite positive are true, then you will have fx is strictly convex if an only if Q is definitely greater than 0. It will be concave if an only if this is less than equal to 0. And it will be strictly concave and convex properties would be hold true, if so in the case if it is greater than equal to it is concave. If it is greater than only it is strictly convex it is less than it is concave this is less than equal to and this is less than it is strictly concave. So, I will highlight, in the first case concave for greater then equal to, concave less than equal to, so strictly convex and concave are true dramatic, strictly convex greater then and finally strictly concave less then.

So, again I will repeat concave for greater then equal to, strictly concave only, so this is strictly concave I should use the, so first if we go to the concave if an only if this holds, in case it is strictly concave which I have not included sorry for that, this holds then it is which is here and convex only convex if this greater than holds and strictly convex is greater than equal to holds and it strictly convex is greater than holds. fx is neither convex now convex if concave if an only if Q is indefine N definite and nothing can be say by greater than or equal to is the same thing like the point of inflection.

(Refer Slide Time: 20:52)



When it is not possible to find the minimum of the function analytically and therefore must use an iterative method for obtaining approximate solution, many of the method which we know like the Newton Raphson method, Runge-Kutta method all this iterative methods can be utilized. But it can also be unreliable you may not get the actual solution, because in the Newton Raphson method or Runge-Kutta method in the one dimensional you start at one point and basically move according to where in this direction you want to find or divide dx and move in that direction where it is the maximum.

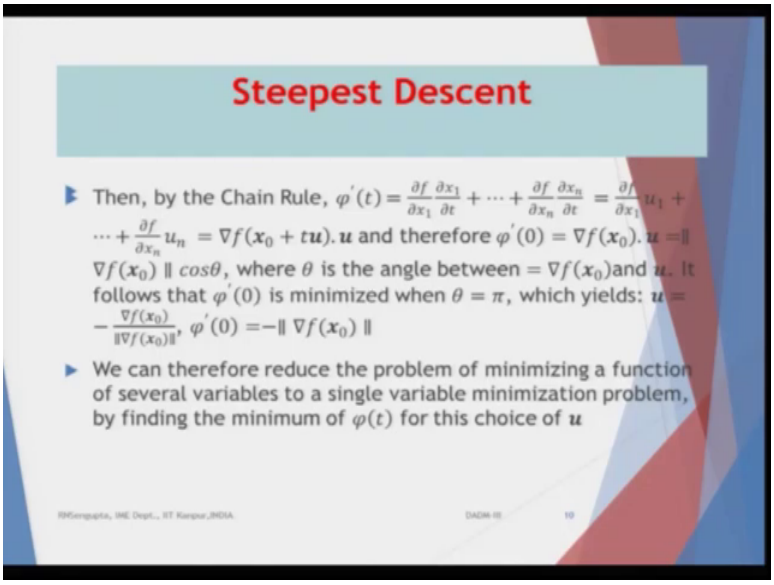
And in the second derivative obviously it is greater than equal to 0 or less than 0 depending on maximum or minimum if it is equal to 0 it is basically the point of inflection. When in the case when you have a multidimensional case you have to basically find out that the rate of change of that functional form which is d2y dx2 or dy dx where it would be the highest in which direction. When we are considering the concept of del it should be remembered that if we are trying to differentiate the functional form y with respect to any one of the x either x1 x2 x3, then in that case obviously the others axis value should be considered as not effecting the rate of change of the function that means they are constant.

So, therefore we now consider another approach which we will discuss now which is steepest descent and also the steeped as on depending on which direction you are moving. And we have a function such that it maps from the n dimensional space to the real line and it is differentiable at x naught, that means you can find other rate of change of the function. The direction of the steepest descent would be the vector where we will take the rate of change of that functional form delt delta of functional form at x naught and then take the decision whether it is a maxima or the minima.

To, see this we will consider the functional form of as fx naught plus functional form of tu and here we will consider that u is basically of a unit vector the dimension, such that we are considering that we are at a point x naught and we will basically take a decision in which direction to move and it was in end dimensional one. And each of this t's would basically give you the quantum of jumps or quantum of movement we are going to make in the u direction.

So u direction can be n in number also, where u is a unit vector such that the mod of the u is 1, so is basically like having a vector and this vector is defined by unit vectors e1 e2 e3 till en and each of this unit vectors e1 to en are multiplied by the corresponding factors. Such that it gives you in which direction that in e1 direction what is the total quantum of distance you should move and that would basically if you find out the total vector it will give you that in as the vector is broken down into the orthogonal parts, so which are e1 to en the quantum of multiplication of e1 to en would give what distance you are trying to move in the orthogonal directions starting from e1 to en.

(Refer Slide Time: 24:26)

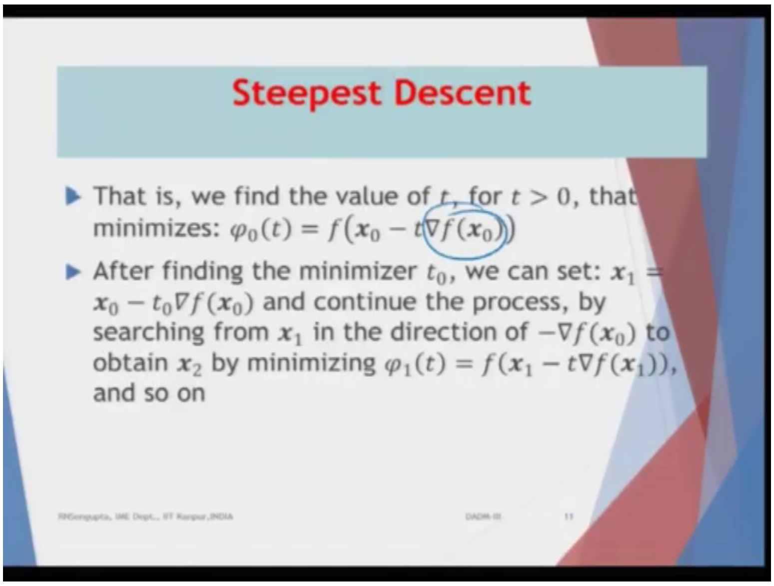


Then by the chain rule you will have that you want to find out the rate of the function, so you want to find out the del phi del t, del phi del t would be basically trying to take del f del x1 that means we are trying to find out the rate of change of the function with the variable x as well as find trying to find out the rate of change of the function variable x with respect to the the rate of change of t.

So, what we will be taking is that, we will try to basically find out the rate of change of the function with respect to each and every quantum of the variables x1 multiplied by the corresponding the values of that u1 vectors u1 u2 u3 till un, where u1 to un would be the corresponding rate of change of that decision variables in the direction of the unit vector which you are going to take. And such that we will basically have that at the point when you want to find out the functional format x naught it would basically give you that vector form into the cross of the theta that means if you have in 2 dimensional format what you do is that you break up that vector into f cos theta and f sin theta that is what we are going to do.

Where theta is the angle between del of x naught and u, u is basically the unit vectors which you have is vector remember and it will follow that phi of 0 is minimized depending on the fact that we have been able to find out the maximum rate of change of that function or the minimum rate of change of the function depending on where whether we are going to find out the maximum on the minimum, maximum descent fall on the maximum ascent depending on whether you are trying to basically take a maximization on the minimization problem. We, can therefore reduce the problem minimization a function to several variables to a single variable minimization problem or the maximization problem by finding the minimum of phi t for this choice of u, u is basically the unit vectors which you have.

(Refer Slide Time: 26:43)



That is, we find out the value of t, so there are different values of t depending on how you find out the rate of change of the function where you will find out the phi of t's based on the fact that you are starting a iteration process as x naught and you will try to find out that what is the quantum form x naught you will move in each and every directions or the rate of change such that the value of del of f of x would give you the rate of change of that function in each and every direction and you will multiply by the quantum of t to find out in which direction the rate of change of function would be maximum.

So, after finding the minimization of t naught we can set x naught again so once you find out the next position you basically move to x1 where you have found at that the steepest descent again try to find out in which direction you are going where the maximum descent is there or steepest descent is there and basically move accordingly. So, in a 3 dimension 1 considering that (xy x) z is vertically up, x is on to the right and y is coming towards me, so you are at a point with where the coordinates are x naught y naught and z naught which is basically x naught in this problem this is a vector and you will try to basically find in which direction is moving the maximum, so in that direction it will move and that each step once you have move into the point that next point would be x1 which will now have a new coordinate of x y z in this 3D figure.

And they will keep moving in that direction where the maximum rate of change of the function is and in that case you will be able to reach the minimization or the maximization depending on how the problem formulation has been done. So, with this I will close the 51st lecture which is the first lecture in the eleventh week and continue discussion about this, its steepest descent in further lectures in the 52nd and further on have a nice day and thank you very much.