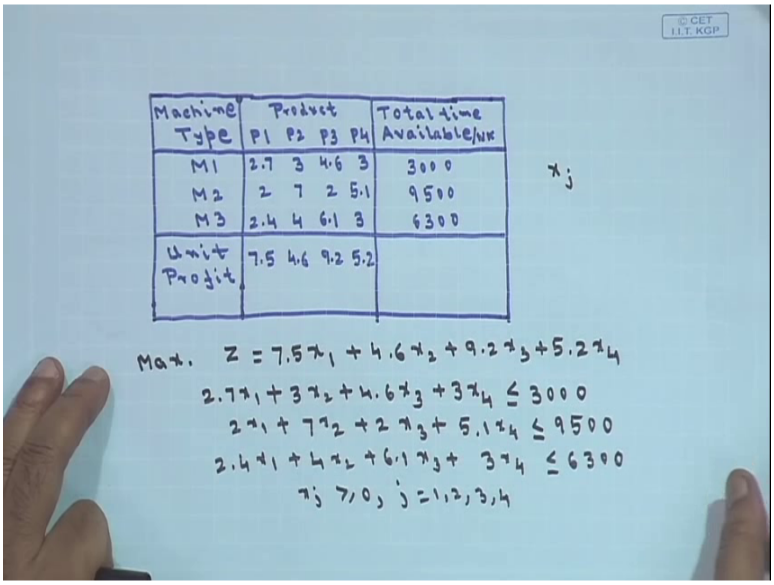
Constrained and Unconstrained Optimization
Prof. Adrijit Goswami
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture - 01
Introduction to Optimization

Today, we will start the constrained and unconstrained optimization. In the first lecture, we will discuss little bit about the preliminaries of optimization.

(Refer Slide Time: 00:35)



If you see what is optimization? Optimization usually we tell as a mathematical process, optimization is a mathematical discipline that concerns with the minimum finding the minimum or maximum value of an objective function. Where the objective function consisting of one or more than one variables subject to certain constraints or restrictions we say.

So, basically the optimization means we have a function of several variables and certain constraints are there. We want to optimize it that is we want to minimize or maximize the function. Basically this optimization technique or operations research initiated in 1940s during second world war. So, let us see one particular problem. If you see this particular problem, we have machines 3 machines we are having M1, M2 and M3. This machines can produce 4 products P1, P2, P3 and P4. Each machine can produce how much quantity of each product is written. On this column if you see the last column total time available per week say in terms of minutes, 3000 minutes, 9500 minutes, 6300 minutes and the per unit profit is also given on the last row.

So, if I want to develop an optimization problem from this, where I have some machines. Some machines are producing some products what is the production rate that is given to us. What is the time available that is given to us and what is the unit profit that also is given to us? So, from here if I want to optimize this particular problem. I can assume that x j is the number of units that the product j produces, x j is the number of units that the product j is been producing per week say. So, our point will be a we can develop a function z equals what will be the this one our aim is I am producing x 1 quantity of product P1 x 2 quantity of product P2 like this and per unit profit is this one. So, therefore, 7.5 is the unit profit for product x 1. So, 7.5 into x 1 plus 4.6 into x 2 plus 9.2 into x 3 plus 5.2 into x 4.

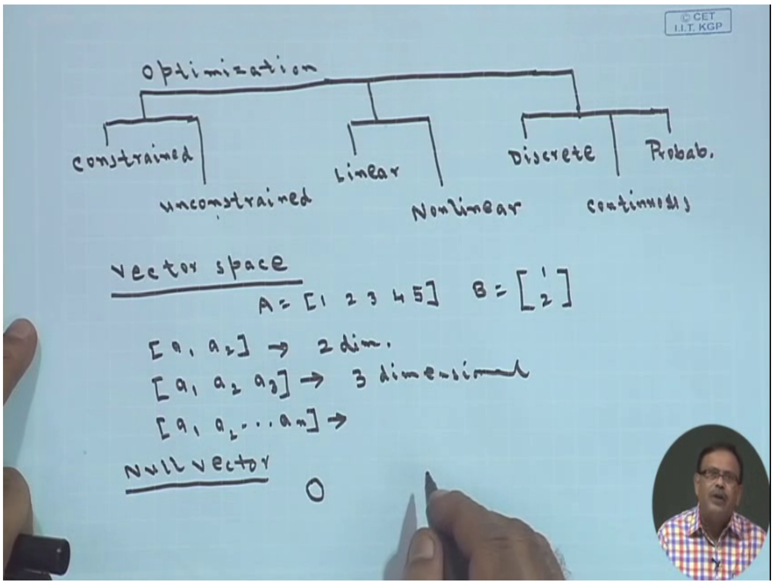
So, this is the actual profit, if I am producing x 1 quantity of product P1 x 2 quantity of product P2 x 3 quantity of product P3 like this and their unit profit is this I can find out one objective function like this. So, this is a function z which is a function of 4 variables, but this is only not the thing there are certain constraints, because each machine can produce a maximum quantity of each product and the time constraint is also given; that means, if you see these first row for the machine M1. It can produce maximum of 2 point 7 of product P1 that is 2.7 x 1 it can produce plus for product P2 it is 3. So, it is 3 x 2 plus 4.6 into x 3. Here it is 4.6 and for the product 4 it is 3. So, it is 3 x 4.

So, this is these machine can produce these 4 products within this time, but we had the available time limit is 3000; that means, I have to produce this total product this one within this time. So, it should be less than equals 3000; that means, total time cannot exceed the 3000 units of time. Similarly, I can create another one for the second row that is for machine M2 that is 2 x 1 plus 7 x 2 plus 2 x 3 plus 5.1 into x 4 which should be less than equals 9500. So, this is the second one and for the third machine it is 2 point 4 into x 1 plus 4 x 2 plus 6.1 into x 3 plus 3 x 4 which is less than equals 6300.

So, if you see this is the objective function and these are the constraints or the restrictions. Where; obviously, I will tell that x j is greater than equals 0, where z is one 2 3 and 4 since these are the products. So, cannot be they cannot be negative. So, like this way I have formulate for this particular table one optimization problem. If you see here this is the objective function and this is the constants. So, optimization is a technique by which we can find out the values of the where decision variables say here x 1 x 2 x 3 and x 4, which will satisfy this constraint. And since it is a objective function and this is a profit function so; obviously, we will try to maximum this function always we will try to maximize. So, maximize z equals this subject to this one.

So, this is the basic idea of the optimization. Now your optimization if you see various type of optimizations are there.

(Refer Slide Time: 06:58)



One I can tell as constrained, constrained optimization another one we can tell as unconstrained optimization. The meaning of constrained and unconstrained optimization means unconstrained means only you have the objective function and you want to find out the optimum value of this objective function without using these constraints. So, this is if these particular constraints are not there. If I have only say max z equals x 1 plus x 2. So, this is one unconstrained optimization problem whereas, if I have the constraints then we can tell that the this is a constrained optimization. So, this second one is the constrained optimization whatever we have discussed.

So, this is one part the other part is, I can put it as it can be linear it can be non-linear also. Linear means whatever objective functions you are considering here. The objective functions should be linear in nature. So, as you know one particular function may be linear function liner function means you are just like x 1, x 2, x 3, x 4 is there. So, this is linear and the constraints are also linear. And other type of function I can make that is say maximize z which is equal say x square plus y square. So, it becomes non-linear. So, in optimization we try to solve both linear and non-linear functions in both cases, it may be linear it may be non-linear.

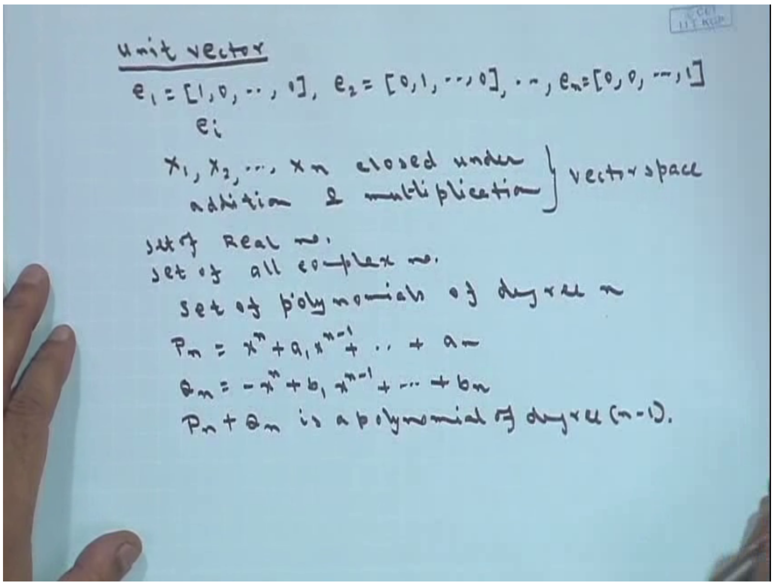
The other type which we can talk about that is about the variables that is it should be discrete it can be continuous or it can be probabilistic. So, whenever you are talking about the decision variables that are for this particular problem x 1, x 2, x 3, x 4 are decision variables. What type of variables are they? So, they may be discrete values they can take they can take continuous values that is any value on the real line or the values may be probabilistic. If it is probabilistic then we have to convert it in to the corresponding continuous cases. So, the optimization I can derive like this it can be unconstrained or constrained optimization. The problem may be linear it can be non-linear. The decision variables which we are associating they may be discrete they may be continuous or they make be probabilistic.

Now, let us go little bit about little mathematics which is required for the solution of the initial problem because after this we will start the linear programming problem at first. So, before going through the linear programming problem. Let us see some basic definitions which you may be knowing. So, we will just brush up all these things. One is vector space whenever you are writing a matrix, you sometimes you write a matrix like this 1, 2, 3, 4, 5. Or sometimes you write a matrix that is one 2 also. So, the first one is a row vector we call it the second one is a column vector.

So, the vectors can be given you know may be knowing this is a row vector then it is a column vector. And this row vector or vector can be given geometric representations also if you think this a 1 a 2 this is a vector. So, this represents a point in 2 dimensional space. Please note this one this represents a point in 2 dimensional space. Similarly, if I take a point if I take a vector like this a 1 a 2 a 3 this will represent a point in 3 dimensional space. And if I take a vector of n elements a 1 a 2 a n, then also if I imagine it then it will be one point in n dimensional vectors. And you have to remember that the row vector and the column vector both are equivalent.

Ah next one is null vector. Null vector is denoted by this o where all the elements are 0 of that vector that is it will look like this. So, this is a null vector. Similarly, you are having the unit vector.

(Refer Slide Time: 12:30).



Unit vector we define like this. I can tell e one equals this thing e 2 will be equals to 0 one and all other elements are 0. Similarly, e n will be like this. So, basically the unit vector you can denote by e i, where the vector with unity as a value of the ith component or in other sense for e i the value of the ith component will be one and the value of the other elements will be 0, which we have represented like this. I think it is clear that unit vector is the vector whose ith component is one and all other elements are zeros.

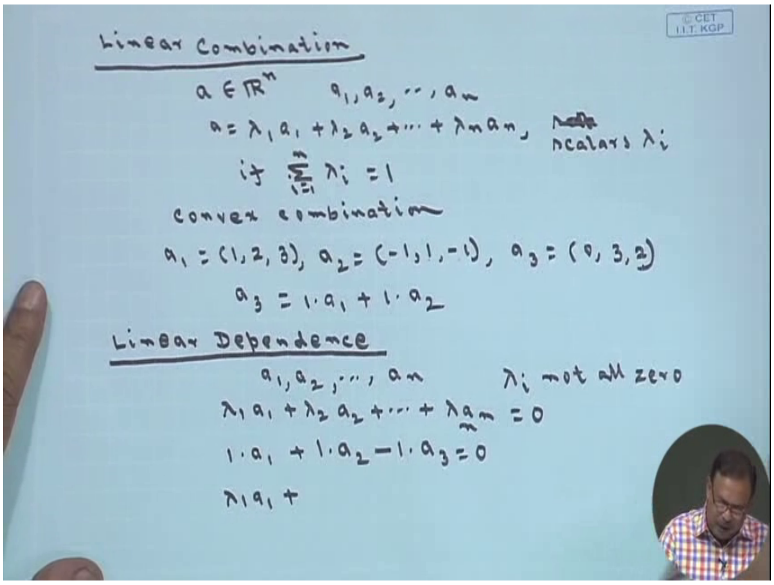
Now, suppose you have a set of vectors x 1 x 2 x n. So, you have set of n vectors which is say closed which is closed under addition and multiplication which is closed under addition and this 2 factors multiplication, then we call that this set of n vectors will form a vector space. So, basically what is a vector space if I have a set of n vectors x 1 x 2 x n which is closed under addition and multiplication. Then this set of n vectors and will form a vector space. So, quite naturally the question will come what is what do you mean by closed. Under addition closed under addition means if we take the sum of any 2 vectors then the resultant vector will also be a member of the set. So, if x 1 x 2 x n is there, if I take addition of any the any of this 2 then the resultant or the sum of any 2 vectors will also be a member of the set then we say that it is closed under addition.

On the same way multiplication closed under multiplication means whenever I have 2 vectors, if I take any 2 vectors from set of n vectors x 1 x 2 x n. And if I multiply then resultant vector should also be a member of this x 1, x 2, x n. So, basically closed under addition and multiplication means whenever you will perform this 2 arithmetic operations. The resultant vector also belongs to the set itself as an example if you see set of all real numbers or set of all complex numbers. If you take they will form a vector space. So, set of all real numbers we are saying set of all real numbers similarly, say set of all complex numbers set of all complex numbers they will form vector space.

Suppose I take a polynomials set of polynomials of degree n what happens set of polynomials of degree n. Will it form a vector space or not? If I think I can take 2 any 2 polynomials like this of degree n x n plus a 1 x n minus 1 like this way plus a n, I am taking another one q n minus x n plus b1 x n minus 1 dot like this, b n if I form the Pn plus q n. And if you see Pn plus q n this x n vanishes therefore, the sum Pn plus q n is a polynomial that is true it is a polynomial of degree how much the degree will be not n because the nth term to the power n th term vanishes. So, the degree is n minus 1.

So, from here we can conclude that; that means, addition after addition whatever result you are getting that does not belongs to the set of polynomials of degree n. So, in other sense you can say that the set of polynomials of degree n will not form a vector space. So, I hope it is clear. Let us go to the next one that is linear combination which is very important.

(Refer Slide Time: 17:54)



This parts which we will be using after words, the you have a vector a, I am just writing in simple form belongs to R n say this will be a linear combination of the vectors a 1 a 2 a n. If we can form like this a equals lambda 1 a 1 plus lambda 2 a 2 like this way plus lambda n a n for some scalars lambda, i for some scalars lambda i. So, if I represent a in terms of some n vectors a 1, a 2, a n like this way a equals lambda 1 a plus lambda 2 a 2 plus lambda n a n where lambda i is are scalars.

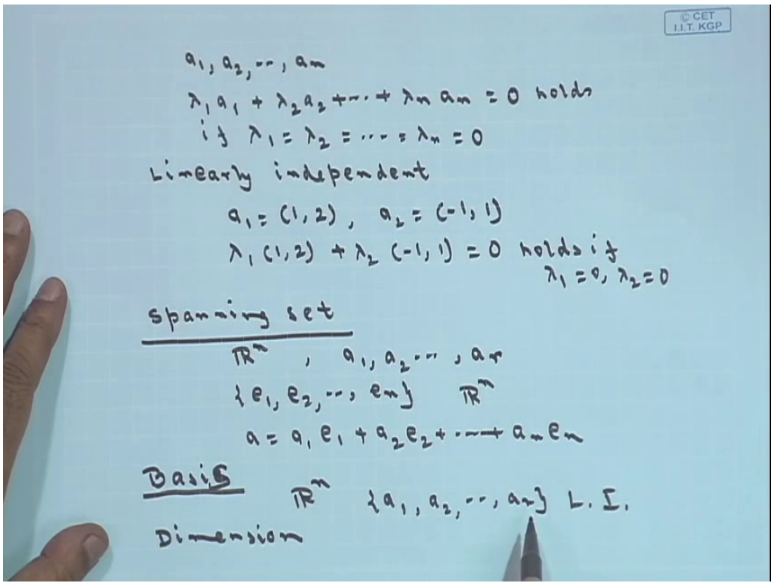
So, then we say a is a linear combination of a 1 a 2 a n in addition to this if summation over i equals 1 to n lambda I this is equals to one; that means, summation of these all scalars is equals to 1. Then this linear combination we call it as convex combination we call it as convex combination. So, if I can represent a vector a like this in terms of n vectors a 1 a 2 a n like this a equals lambda 1 a 1 plus lambda 2 a 2 plus n a n and if summation i equals 1 to n lambda i equals 1 then the combination, we call it as convex combination.

If you consider 3 vectors a 1 equals this, if you have the other vector a 2 equals minus 1 1 and minus 1 say and if I take another one say 0 3 2, can I write down a 3 equals how much a 3, is a 1 1 into a 1 plus 1 into i 2. So, one minus 1 0 2 plus 1 3 and 3 minus 1 2. So, basically I have 3 vectors a 1 a 2 a 3, I am writing a 3 as a linear combination of a 1 and a 2 where lambda 1 is 1 and lambda 2 is 1. So, this is a combination linear combination of this, 2 from this linear combination itself we come to the other one that is linear dependence you may have studied it in matrix notations linear dependence.

So, you have a set of n vectors a 1 a 2 a n, you have the set of n vectors a 1 a 2 a n and I will say that this set of n vectors are linearly dependent, if there exist some lambda I not all 0 some lambda i is I can find out where all the lambda is will not be vanishing and this lambda 1 a 1 plus lambda 2 a 2 like this way plus lambda n a n which is equals to 0. Then we say that a 1 a 2 a n this n vectors are linearly dependent. So, please note that here all the lambda is will not be 0. In the earlier example, if you take this example I think one into a 1 plus 1 into a 2 minus 1 in to a 3 this is equals 0 1 into a 1 plus a 2 this is 0 2 plus 1 3 minus 3 0 similarly 3 minus 1 2 minus 2. So, this is 0.

Therefore, we can say that this vectors a 1 a 2 a 3 this vectors are linearly dependent. On the other hand, if you see if I can find out the vectors such that if lambda 1 a 1 plus let me write it in the next page that would be better if you have n vectors.

(Refer Slide Time: 23:04).



A 1 a 2 a n and likely in your dependency I am getting lambda 2 a 2 plus lambda n a n this is equals 0, say this holds this one holds if all the lambda is vanishes that is lambda 1 equals lambda 2 equals lambda n equals 0. Then we say that these vectors a 1 a 2 a n are linearly independent this, we will use frequently.

So, please note the definition of linearly independent set of vectors a 1 a 2 a n will be linearly independent, if for some scalars lambda is we can get one equation like this lambda 1 a 1 plus lambda 2 a 2 plus lambda n a n equals 0 this holds. If all these scalars vanishes than only we say that this vectors are linearly independent. For an example if I take a 1 equals 1 2 say a 2 equals minus 1 1 and if I take lambda 1 into 1 2 plus sorry into plus lambda 2 into minus 1 and one can you find out some value of lambda 1 nonzero values of lambda 1 and lambda 2 for which this will vanish, but if we calculate you will find, I am getting any value if I solve this will hold if your lambda 1 is 0 and lambda 2 is 0 only.

So, we say that this lambda 1 equals 0 and lambda 2 equals 0. So, this a 1 and a 2 are linearly independently. So, basically if 3 vectors are 2 vectors are linearly dependent geometrically means they lie on the same line and whenever some vectors are linearly independent; that means, one vector can be represented as linear combination of the others. So, that is one thing next is spanning set you had the set of all vectors in R n from here if, I can get a set of vectors a 1 a 2 a R say we say that this set of R vectors a 1 a 2 a R will span or generate R n means if I take any linear combination of this that will be a vector of R n or in the sense other way I can tell that in from R n if you take any vector that will be a linear combination of this a 1 a 2 a r.

So, therefore, if in R n I can find out a set of vectors a 1 a 2 a R which can generate or span all the vectors of this R n, then we say this set as a spanning set this consider this one, e 1 e 2 e n. This set spans R n I think means using this set linear combination of this set will be giving us any vector of R n, because I can write down something like this a equals a 1 e 1 plus a 2 e 2 plus a n e n for certain values of a 1, a 2, a n, I can obtain this value.

So, this e 1 e 2 e n spans all the vectors of R n. Similarly, there is another concept which we call as basic not basis, but it is basis. So, a basis for this R n is a linearly independent subset of vectors where the it is spans which spans the entire space. So, basically basis is nothing, but the linearly independent. Please note this one linearly independent set of vectors these are linearly independent set of n vectors a 1 a to a n which spans the entire vector space and this we call as the basis. And similarly your dimension is the number of linearly independent vectors in the spanning set is known as the dimension of the basis. So, in one-word basis is nothing, but the a set of linearly independent n vectors which spans the entire R n and the dimension is the elements number of elements or number of linearly independent columns of that set. So, next we will start in the next class.