Data Analysis and Decision Making-3

Professor Raghu Nandan Sengupta
Department of Industrial and Management Engineering
Indian Institute of Technology, Kanpur

Lecture 15

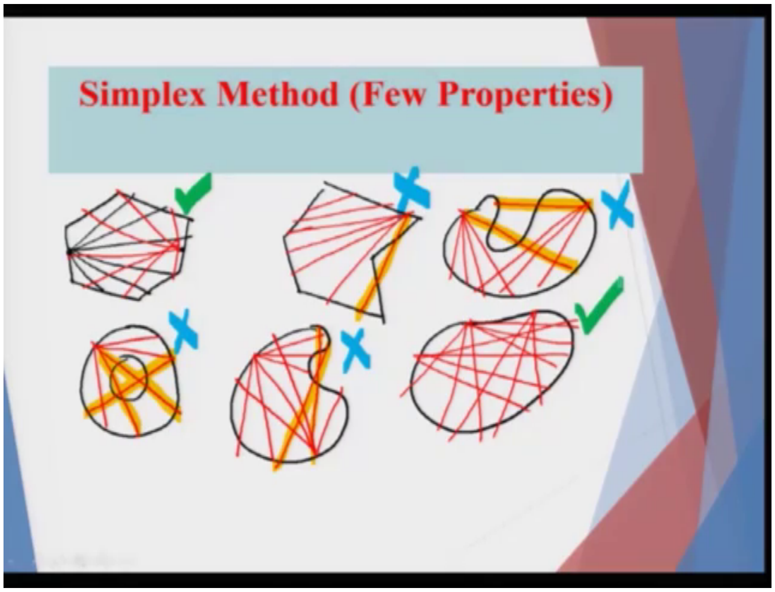
A welcome back my dear friends. A very good morning, good afternoon, good evening, wherever you are in this part of this globe. And this is the DADM-3, which is Data Analysis and Decision Making-3 lecture series, under NPTEL mock. And, as you know, I keep repeating that but just to make you understand and also to set the tone of the class, rather than immediately going into the concepts.

So, this total course duration is for 12 weeks, which one converted into number of contact hours is 30. And when converted into number of lectures is 60 because each lecture is for half an hour and as you know that each week we 5 lectures. After 5 lectures you have then assignment and you have already taken two assignments. So, if you see the lecture which is lecture number 15, which is in the third week. We are going to wrap up with this class, the third week and immediately after that you will basically be taking the third assignment.

So, as you know my name is Raghu Nandan Sengupta from the IME department at IIT, Kanpur. So what we were discussing, if you remember. We were discussing about the convexity and the concavity. Obviously, concavity would not be discussed here, I just thought I should mention that the concept of strictly convex and similarly the counterpart of strictly concave, concave would be there from the functional point of view. And then I told that the, those corner points on holes would not be there.

So I was drawing the diagram, marking how if you join any two points between in that overall feasible region, they should always lie in that region. But in for few examples they were not. So I will basically continue. I have already done three examples. I will do another three and then go into the actual concept of how you solve the problem of optimization using the simplex method.

(Refer Slide Time: 02:02)

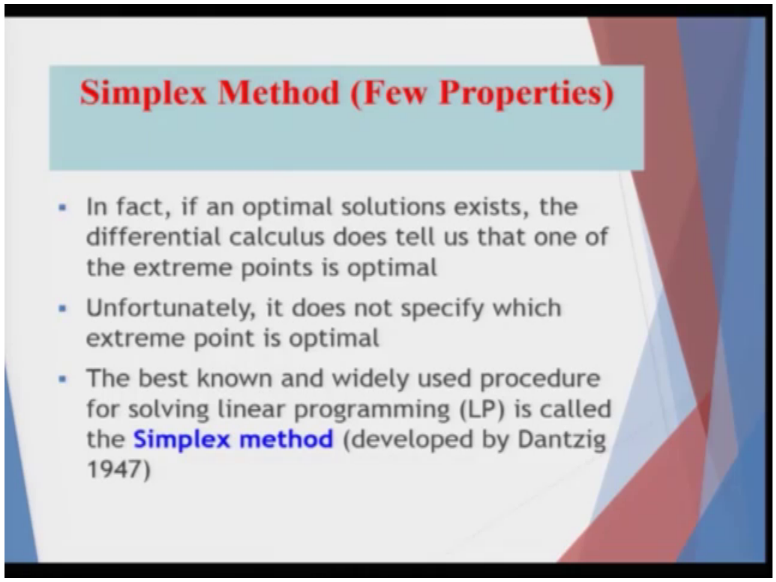


Okay, I mean okay, so these are the diagrams. The first one is possible, second was not, third was not, fourth I have drawn it but I will basically mark it as a little bit different way. Let me mark it. So let us go one by one. Possible, possible, not possible, not possible, possible, not possible. So there are different combinations, so what are the not possible one, I again mark with the color. So this is not possible, this is not possible, this is not possible. So there are other combinations also. So whether its possible, this overall region is convex? The answer is no.

Let us come to the fourth one, possible, possible, possible, possible, possible, possible, possible, (posib) not possible see, possible, possible. This one, if I take here, sorry, if I take here, not possible. So there are I have only for the diagram which I have drawn, there are no not two possibles. So obviously I mark them with the color as I am doing. So this is not done and this is not done. So, the final answer is not possible.

For this region, I am not been able to draw the last diagram nicely, but I will try to highlight it as it is possible, as it is applicable. So let us consider possible, possible, possible, possible, possible, possible, possible, possible, possible, possible any combination for this diagram is possible. So does it meet the criteria of convexity? This is yes. So these diagrams can be extended to the higher dimension but I just thought I will draw the diagrams in 2-dimension to make you understand.

(Refer Slide Time: 04:41)



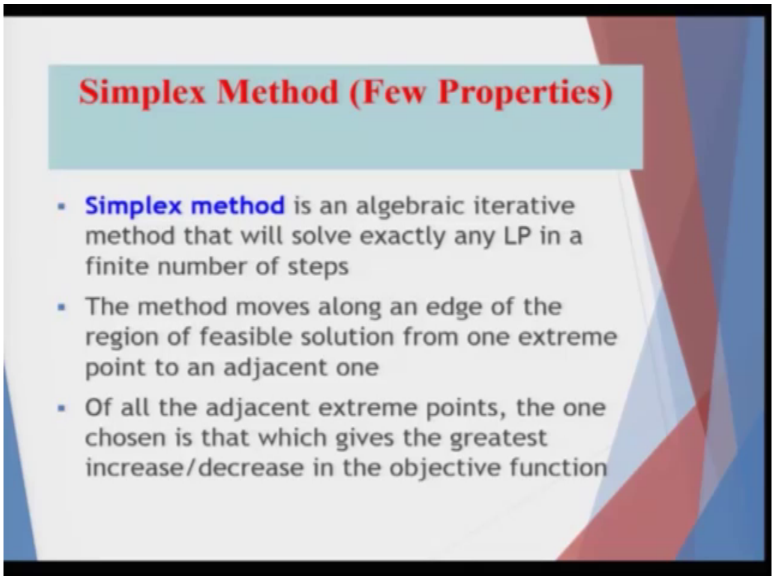
So continuing with the simplex properties, few (property) methods Few Properties. In fact, if an optimum solutions exists, set of optimal solutions exists, it can be set also considering that for different combinations it always leads the same answer or optimum solution only one exists. The differential calculus which is dy by dx concept which we use, when we want to basically optimize, find out the maxima, dy by dx is equal to 0 or minima, dy dx is equal to 0.

And depending on the maxima and the minima, you try to take it, the second derivative or if it is a matrix you take the Haitian, and whether the second derivative is less than 0 or greater than 0, we will basically comment the concept of maxima minima accordingly. And if you remember some simple calculus we have done if the second derivative is also 0, we say it is a point of inflection, where the line is changing its rate of change of the function from dy dx, the maxima to the minima, the minima to maxima where the change is happening, dy dx concept is changing from positive to negative, negative to positive.

So if you use a differential calculus, is does really tell us that one of the extreme point is optima, optimal but which point we do not know. Unfortunately it does not specify which extreme point is optima based on which you want to find the solution. So that is what is the essence of the problem, (you) we need to find out that which set of decision variables, the (X1), X1, X1, X2, X3 X4 which one is, will give you the optimum solution whether the maxima or the minima.

The best known and widely used procedure for solving linear programming problem or LP is called the simplex method and it was proposed or developed by Dantzig in 1947. And will try to utilize that concept very simply.

(Refer Slide Time: 06:32)



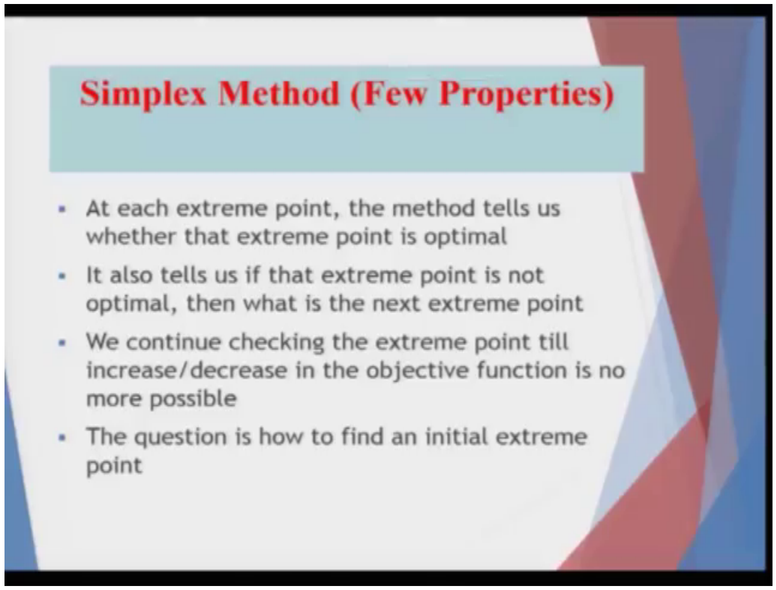
The simplest method is an algebraic iterative method that will solve exactly any linear programming in a finite number of steps. Now the question is that why it is a finite number of steps? Because as we remember I did mention about Karmakar's Interior Point algorithm, the concept that all actually the feasible region can be infinite. But your searches would be only restricted to the corner point because from there you will get your problem solution, the set of decision variables which will give you the optimum solution as well as the value of Z, which is maxima or minima.

The method moves along an edge of the region of feasible region. So this is what, what I am saying. So if you are at the corner point you will try to basically, if you remember the first diagram or the problem, which I solved. I did not solve it in order to give you the exact solution methodology or the final answer. But I said that if you started D, D was the origin. You will basically take a direction of move depending on whether the maxima or minima and try to find out the rate of change of the function which is happening at the highest rate or in the minimum rate depending on whether you are trying on to go for a objective function which is maximization and minimization.

So, this is exactly what is says, that you should basically move along an edge from any point, so let me read it. The method moves along an edge of the region of the feasible solution from one extreme point to an adjacent one. Of all the adjacent extreme points, the one chosen is that which gives the greatest increase or decrease in the objective function based on which you are trying to basically achieve the, the maxima or the minimum. So if you want to obtain the maxima, so you will move in that direction where the rate of change is the maximum. So you are trying to reach that point as fast as possible. And if it is a minimization (prom) you will basically find out the rate of change of the increase is happening in the, in the negative direction and the opposite directions, such that you reach the minimum as soon as possible.

That means the slopes, if slope that means, if you are going up, up the hill. So you will take the steepest one such that the climbing amount which you want to take, per unit jump which you want to do, would be such that you will be reaching the objective function maximum ASAP. In the case, when you are going to go downhill, you will take that path, where the jump should be happening in such a way that the steepest negative direction movement would be such that you will reach the minimum point as soon as possible.

(Refer Slide Time: 09:05)



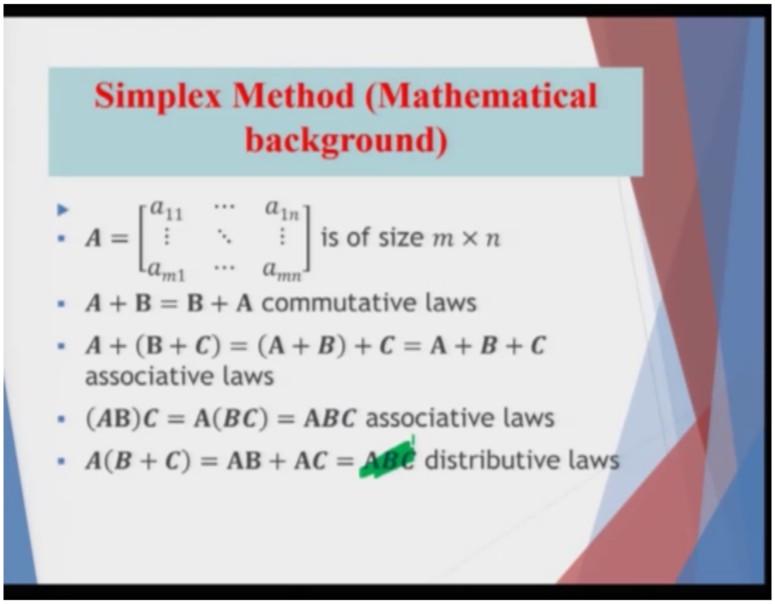
Now at each extreme point, the method tells us whether the extreme point is optimal. Because, reason is that say for example you reached the optimal point and we will see that how it is done mathematically. Now you try to move. Now when you move, the rate of change the function whether in a negative direction or positive direction, if it is not adding up to the objective function value, that means in the maximum problem, it is not increasing the objective function or in the minimum, minimization problem is not decreasing the objective function, obviously it would mean the optimum point or the extreme point from where you are starting or from where you have just reached and you want to start off again, is the optimum point.

So, obviously that can would come out from solution technique also, very simple. So it also tells us that if that extreme point is not optimum, then what is the next extreme points and in which direction which you should move? As I told you the problem of trying to basically make the maximum jump when you are going positive or make the minimum jump when you are going to the negative direction. We will continue checking the extreme point till the increase and decrease in the objective function is no more possible. So if I consider the first point which I mentioned at each extreme point, the method tells us whether the extreme point is optimum. Because any extra change in, in the this values of the X1 and X2, because the corner points are what?

The corner points are the feasible points based on some of the constraints being, the overall constant set, feasible region is met, but the that corner points will be given by solving some of the constraints which actually passed through this, that point, the corner point. And once you reach there, if you try to move in the, in the other direction and you find out there is no increase in the objective function or no decrease in the objective function considering the maximization and minimization problem, you will basically stop there.

So that is what the third point says, we continue checking the extreme point till increase and decrease in the objective function is no more possible. The question is, how do we find an initial extreme point based on which we will start the solution. Now if you remember in the problem which you have solved where I have drawn the diagram, we started at origin. So was there any logic why we started in the origin? I will try to basically show it algebraically and when we solve the problem, will try to go into details for that also.

(Refer Slide Time: 11:33)

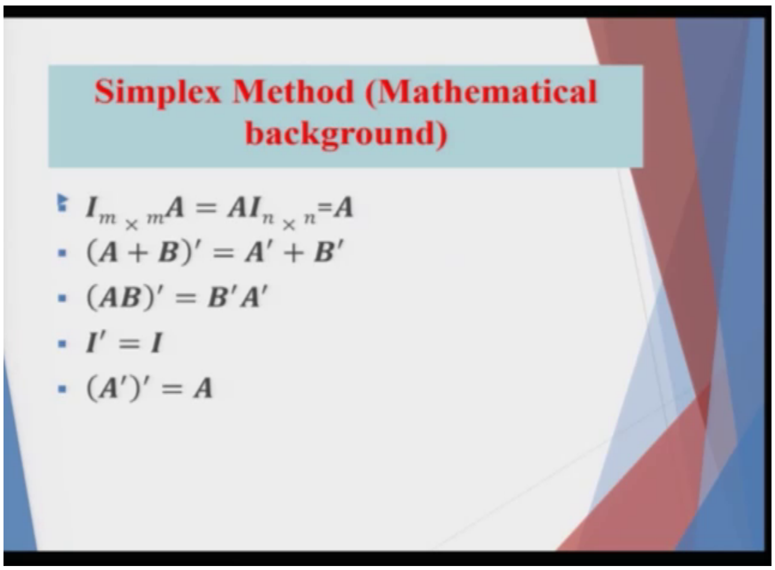


Now very simple properties of matrix, so consider the matrix is there, where the size of the matrix is m into n. So you basically have m number of rows and n number of columns. I am not going to mention anything about m and n, which is m is more or n is more. The commutative law for addition means adding A plus B also gives you the same property of B plus A and these would be utilized later on. So please bear with me. The associative law of addition means that if you add up B and C first and then add up to A, it will give you same answer if you add up A and B first and then basically add up C or you basically add up one at a time, that means A plus B plus C would give you the same answer.

When I go into the associative property and obviously remember, this case of A plus B and is equal to B plus A or this third property of associative property for positive one, would always mean the size of the matrices are possible. So that means m and m, what I am mentioning they would be applicable for all the other 4 bullet points, which is the second, third, forth, fifth which I am just discussing. So, let us come to the fourth bullet point. It is the associative law for multiplication. So if you multiply A and B first and then multiply with C, would give you the same answer if you multiply basically B and C first, and then multiply with A. But remember the, the size or the dimension of the matrix should hold true. So, if you multiplying A and B, so the size of A, if it is m and n, B should definitely be n cross something, because n and n should match.

Similarly, when I go into the distributive law of multiplication, along with or for multiplication, it means that if I first add up B and C and then multiply A with that, would give me the same answer if I basically add up the concept. So this is, this should not be there sorry, my mistake. This should not be here. So basically if I add up first multiply A with B and then A with C and add them up accordingly. So I will try to remove this.

(Refer Slide Time: 13:46)

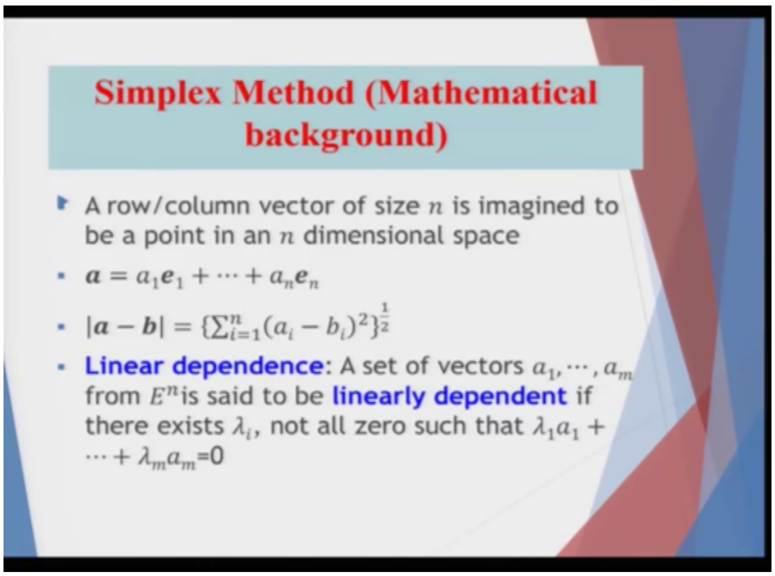


Coming back to the identity, this is important because this identity points concept would be utilized. Coming into the identity, so if I multiply matrix, identity 1, pre-multiply it with before A and A is a size m cross n and if I basically multiply, post-multiply A with identity matrix of size n cross n. Considering the matrix multiplication concept should be hold true, considering the sizes, we should get the matrix A. That means, pre-multiplying a number with 1, post-multiplying a number with 1, positive 1, would always lead to the same number. That means 1 into 3 is equal to 3 into 1 is equal to 3.

Now in the case if A is of size m cross n, m cross m sorry, both the rows and columns are same then the left identity and the right identity would be the, of the same size and you will basically left later on will see left inverse and right inverse would be basically will be giving us the answer such that you pre-multiply or post-multiply, the answer would always be the same.

If you take the transpose, so A plus B transpose is equal to A transpose plus B transpose. If I multiply, so AB transpose would be B transpose into A transpose. So if you have three matrices it will be just the reverse order of the transposes. I transpose, identity matrix would basically be the identity matrix itself and transposes or transpose would obviously give you back the same matrix.

(Refer Slide Time: 15:27)



Now, we will consider vectors and why the vectors? That will become clear when we consider the concept of simultaneous equation. So we will consider either a row vector or column vector based on how we are trying to basically do the nomenclature of the direction where this vector is. Row or column vector of size n, that means n elements are there. It is imagine to be a point in an n dimension. So if you have a vector (2) 2, 1. So obviously it means that I move 2 units in X direction and 1 unit in Y direction considering it is 2-dimension one.

If I have a matrix of 0, 0, 2, it means I move 0 direction, zero unit in the X direction, 0 direction in the Y direction and 2 units in basically, it should be 0 units in the X direction, 0 units in the Y direction and 2 units in the Z direction, such that I reached that point which is given by the matrix 0, 0, 2.

If I have a vector of size 2, 3, on the vector which is 2, 3, 4 that means I would move, 2 units in the X then again I move 3 units in the Y and then again I move 4 units in the Z direction and basically I reach that point such that originally joining that point would give me the vector as I suppose. Now whenever I am going to consider a any-any-any row or a column vector, I will basically take that as a combination of units of movement I am going to take in the orthogonal directions which is E1, E2, till En.

And I will technically consider that E1 with respect to E2, with respect E3, till with respect to En are all orthogonal, that means 90 degrees, when you are trying to solve the problem. And any component of the vectors which is there in any direction of E1 to En, if you basically rotate it and basically try to find out the component on the vector in other direction, considering it is orthogonal, it will be 0.

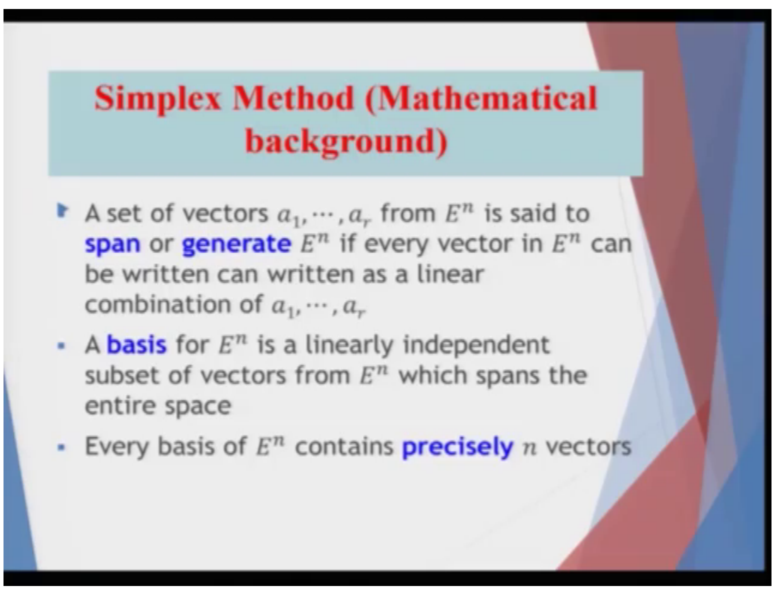
So any vector A would be comprised by taking units of A1 to An being multiplied by the concept of, of the unit vectors which is there in the direction E1 to En. And what is E1? E1 would basically be, if it is of dimension n, the first element will be 1, rest from the second to the nth one are all 0. If I consider E2, the first one is 0, second one is 1 and third to nth one is all 0. Similarly if I go to the last one, which is En, from the first to the nth minus 1, are all zero and nth one is only 1.

If I take the concept of distances, you will consider the distances as a square one or L2 norm. So the L2 norm distance would be given by the square of difference between each A and B, add them up for all the n different dimensions and basically find out the square root. Now having said that we will basically go slowly to understand the concept of linear dependence and the concept as it is required. And why, why linear dependence is important you will immediately find out as we slowly go into solving the problems.

So linear dependence would mean a set of vectors a1 to a m and here a1, a2, a3 would basically be defined as, as using the combinations of unit vectors of E1 and Em or En. From the space of En, now remember the dimensions of m and n and m and n are different, the size of m and n are different. So that will come up later on. So a set of vectors a1 to a m from the space E, of N dimension is said to be linearly dependent if there exists some lambda, some lambda is such that and lambda is basically I is basically 1 to m, such that taking any convex combination of lambda 1 into a1 plus lambda 2 into a2 till lambda m, lambda m into a m would lead to 0 such that they, any one of the vectors actually means, it can be explained by the convex combination of this m minus 1 vector.

That means, if there are linearly independent, all of them are independent and if you solve them, the number of solutions would be unique and only one point. But if they are linearly dependent such that one or more of the simultaneous equation values can be thrown away such that the number of solution which you have for the rest of x1 to x n can be infinite in number technically, because you will have many solutions in that case. So, no unique solution would be there.

(Refer Slide Time: 20:34)



A set of vectors a1 to a r, the r values would basically change accordingly and obviously it will be less than n. Maximum value of r can be n. So a set of vector a1 to a n in the space En is said to be a span that means it is able to span the whole space. It is able to cover the whole space or generate the whole space. If every vector and remember that the size of r maximum can be n or less than n. So if their dependent structure is not there, now obviously in that case r and n would be equal if the dependent structure is there, the sum of the lambda values are not 0. Still you can basically get a convex combination of value 0, which actually mean some of the vectors or some of the vector which are in some plane can be explained as a convex combination or lesser number of vectors, you are going to take.

So if each, every vector in En can be written, it can be written as a linear combination of a1 to a n, such that as I mentioned now, before, just before that the dimension of that overall space can be reduced in order to compare or give the expression of any other set of vectors. Now having said the span or the generation, generation means that you can generate the whole space, the whole set of vectors which are there. So, these are the minimum tools, based on which, you can generate all the vectors.

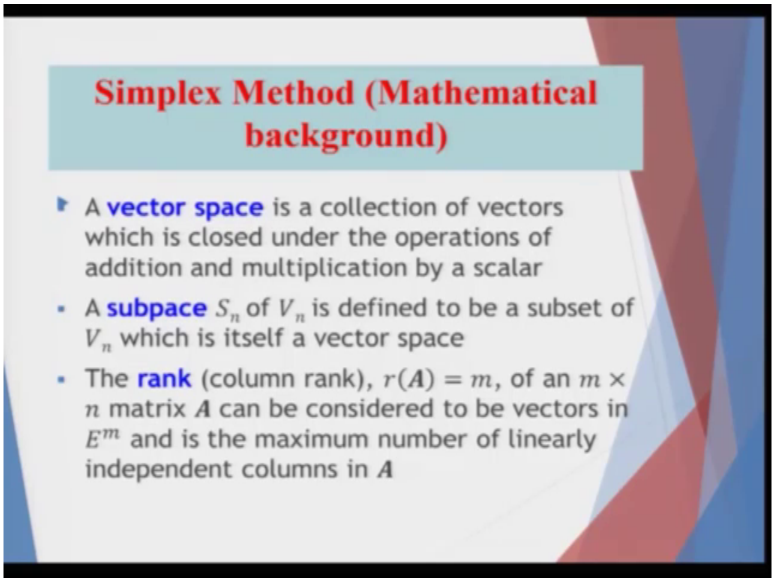
A basis for of dimension En is a linearly independent subset of vectors from En which spans the entire space. So basically the vector which you are going to take, it is able to span all the spaces, all the vectors from all, overall space of En basically that would be known as the basis. The simple the set, unitary set or unitary element, building blocks which can be utilized to build up all the other set of vectors, that is the minimum one.

So say for example, if I am going to build a house and if I consider the minimum one is 1brick. So based on how many such bricks I utilize, I can build the whole house. So this are the minimum block which I mean. Or say for example, if I am going to build up a bridge and considering the minimum number, not the minimum number, minimum size of the steel rod or the steel truss, which I will have or the linkages which I will have is basically 1, that means trying to multiply 1 or trying to take infinite such or huge amount number of such, such rods of steel or iron I can build the whole bridge.

So, this is the minimum building block, which you will have and multiplying it many number of times, I am basically able to span or generate the whole area. So there can be more than one basis like building blocks can be more than one type. So say for example, I use building a house, I use a brick of size, say for example of weight, of size of 6 inch by 3 inch by 4 inch. In another case say for example, I use it 5 inch by 4 inch by 5 inch. So these can be two different basis, based on which I can build the house what I need to do.

Every basis in En contains precisely n vectors, based on which you can generate the whole span which I required to do. If they are independent? If dependent structure is there, then it will be less.

(Refer Slide Time: 24:06)



A vector space is a collection of vectors which is closed under the operations of addition and multiplication by the scalar. Remember that, why I am saying that? When we come into the concept of basis, in the basic solution which we have, based on which we will start the solution, would basically give you, give us the answer. So a vector space, so there are many vectors which basically makes the space is a collection of the vectors which is closed, that means any addition would always lead into the same space.

If you remember I drew the diagram, six diagrams. So they were in from, in few of them, the when joining 2 points they were going outside the region. So it means that they are closed on the operations (addition), operations of addition and multiplication by a skill that means, the vectors which we will be get would be inside the overall space. They would not be going out, in very simplistic sense, which I am trying to say.

A subspace of the vector space Vn is defined to be a subset of Vn which is itself a vector space. So if I keep shrinking or keep going down and down like taking a microscope and going more into a inner depth. So if I am able to find out the, the smallest subset of Vn, obviously that would be the subspace based on which I can generate and go and basically generated the outer sphere also. So consider that, that I have a building block or 3 building blocks I may able to build a house.

So if I am able to build an apartment after apartment, so the smallest building block which I take would be a subspace based on which I can build up the whole apartment or a compartment complex. Now whenever we consider a matrix, we know this is important. We know that what is a rank. Rank of a matrix in simple sense, what I have already considered, I intuitively, I have already discussed rank. Rank means the size of the number of rows or number of columns based on which you can basically explain and find out the rest of the rows or the columns, as a convex combination.

But, if you are trying to mention a set of simultaneous equations such that you have a unique solution and none of the equations can be explained as a convex combination of the others. So the ranks of that, of that matrix would be exactly equal to the number of rows or the columns and it will be exactly equal to the number of variables you want to find out. Say for example, I have a equation 2 X1 plus 4 X2 is equal to 5 and 10 X1 plus 20 X2 is equal to 20.

So if you solve that, you will get unique solutions. That means none of them can be expressed as a convex combination of the other. But if I have X1 plus X2 is equal to 1 and 2 X1 plus 2 X2 is equal to 2, which means the second can be (combi) explained as a convex combination of the first by multiplying by 2, that means trying to find out the values of X1 and X2 would not give you the unique solution.

So if I am able to basically give such examples where you have more than two variables and you are able express them as a vector, as a matrix then it would mean that the rank would basically be for the later case, would be less than the number of rows and columns (which) we have. Because one of them is being explained as a convex combination of the other but for the other example which I gave where none could be expressed as a convex combination of the other. If I had more than 3 variables such that we have more than 3 equations and if I am able to write them as a matrix, then the rank of the matrix would be exactly equal to the (rank), the row or the column of that matrix.

So with this I will come and explain the last point. The rank, here I am going to consider the column rank also and obviously it can express as the row rank also. But for our examples later on we will consider the row and the column rank to be equal. Equal because the size of the matrix, what we will consider. Technically initial is m cross n but will add those slacks and the surpluses variable, which I will come to that later on, such that the size of the matrix or the dimension of the matrix would be m cross n or n cross n depending on whichever is larger.

The rank, where rank is m, of an m cross n matrix can be considered to be the vectors, in the dimension Em and is the maximum number of linearly independent columns of A such that any other combination (you) using that you be able to explain and find out the rest of the vectors, which are there. So you are only considering the column. Similarly if you consider the rows again it can be mentioned as rank of A, in the row wise it will be n such that trying to find out the combinations of any other rows would always be given by the convex combination of this n number of rows.

So n or m is the minimum number of columns or rows which we will be able to utilize to explain the rest of the combination of the rows or the columns, columns or rows which you have. So with this, I will end the 15 lecture and continue discussing about that in the later classes and then explain (that) how matrix multiplication will be coming in solving the equations. Have a nice day and thank you very much.