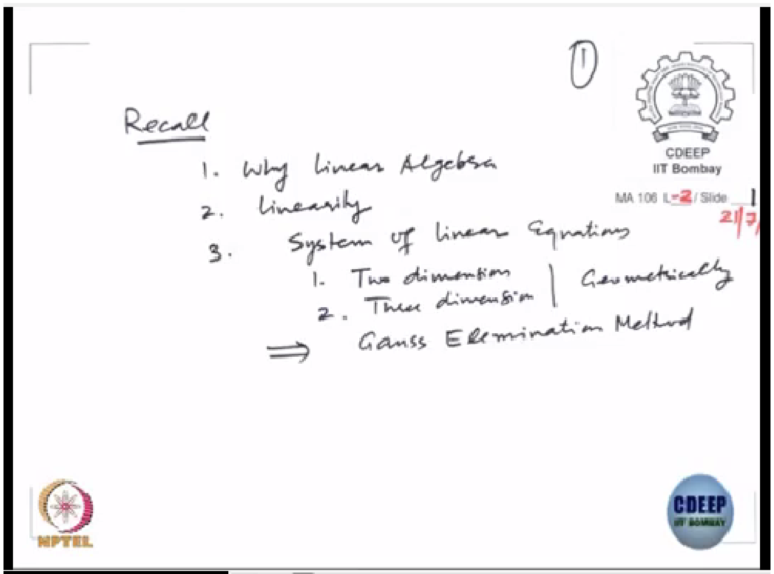
Basic Linear Algebra
Prof. Inder K. Rana
Department of Mathematics
Indian Institute of Technology - Bombay

Lecture – 04
Systems of Linear Equations I

Okay, so let us begin with today's lecture. In the previous lecture, we had started looking at why linear algebra is important to study.

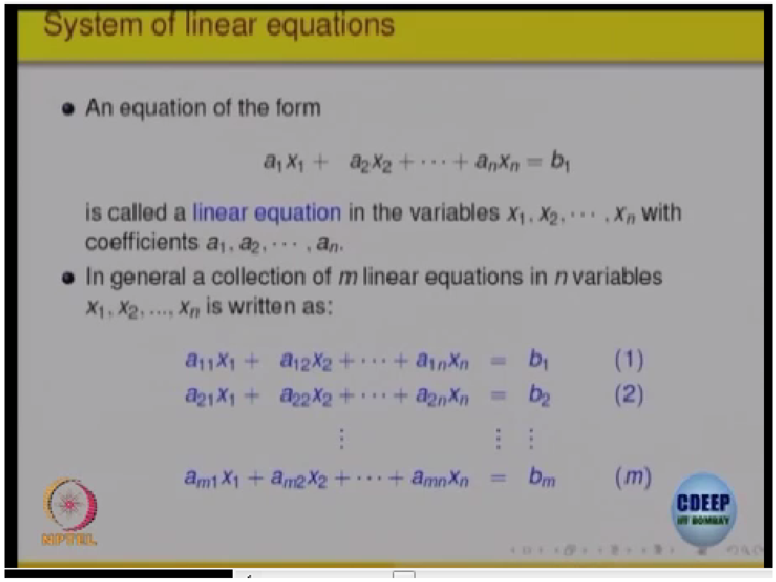
(Refer Slide Time: 00:35)



Then we started looking at how does linearity arise in setups. Basically there are 2 ways linearity can arise. One is because of linear equations and the other is studying geometry in the algebraic setup. We started looking at the system of linear equations. We looked at system of linear equations in 2 variables and then in 3 variables.

And both dimensions 2 and dimension 3, that is 2 variables and 3 variables, we analyze geometrically how the solutions can be obtained and that led us to algebraic method of solving those equations namely trying to eliminate variables one at a time and trying to see whether the system has a solution or not. So we will start with looking at general system of linear equations and generalize Gauss elimination method for a system of m linear equations in n variables.

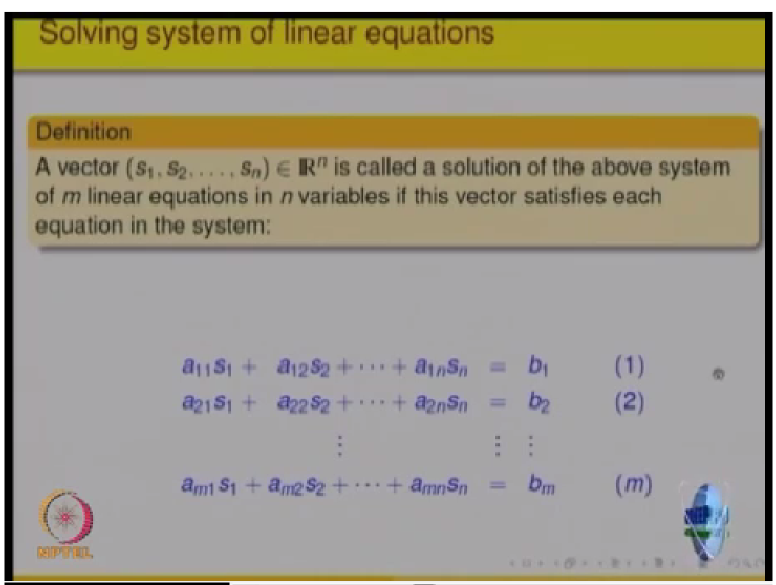
(Refer Slide Time: 01:43)



So let us look at, so we recall that the equation of the type a1x1+a2x2 and so on anxn=b1. We call it as a linear equation because the variables x1 x2 x3 and xn, all have power 1. And you say this is linear equation as the coefficients a1 a2 an. In general, we may have a collection of m linear equations in n variables and we will write it as a11x1+a12x2 and so on.

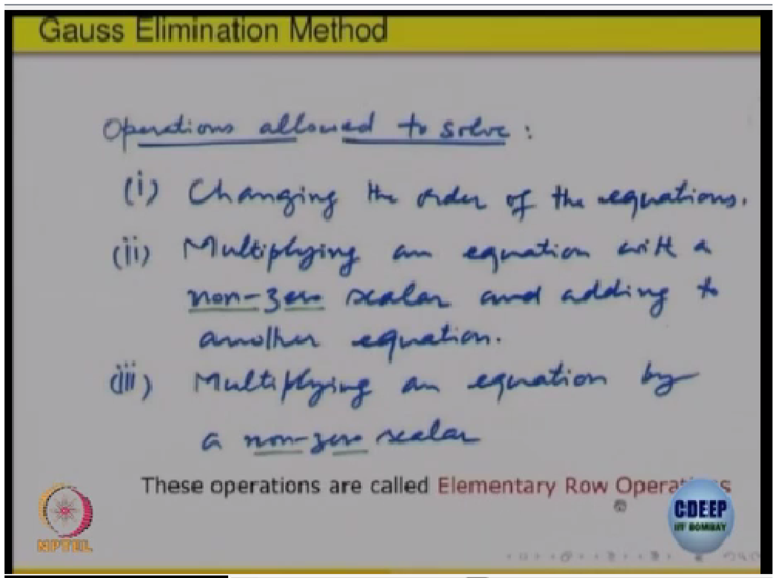
So the first subscript in the coefficients indicates the equation. So this is the a11 a12 and so on, that is the first equation. Second equation will have coefficients a21 and so on. And the last equation that is mth equation will have coefficients with subscripts am1 am2 and amn. So this is a system of m equations in n variables.

(Refer Slide Time: 02:49)



And we call a vector in Rn s1 s2 sn a solution of the system if it satisfies all the equations, namely if we replace x1 by s1, x2 by s2 and xn by sn in all the equations, then we should get left hand side is equal to the right hand side. So in such a vector, s1 s2 sn is called a solution of the system.

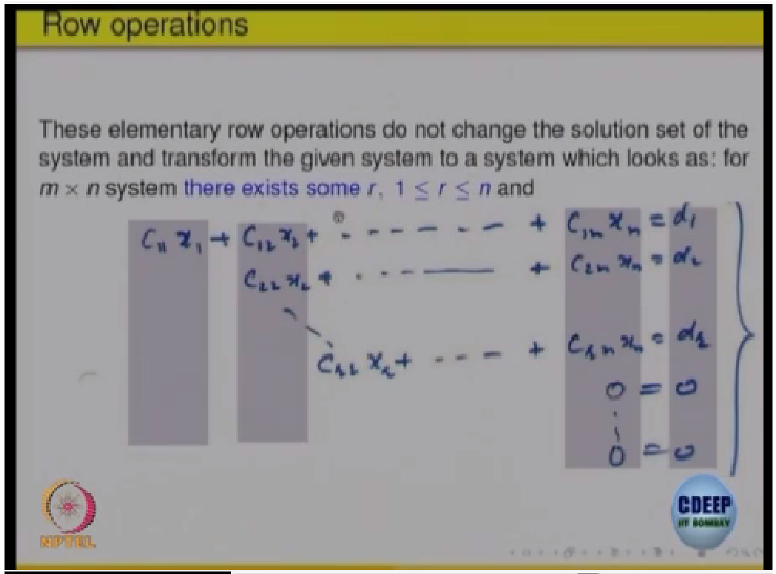
(Refer Slide Time: 03:19)



And we saw that the solution of the system does not change if we change the order of the equations. That is just renumbering the equations. Instead of first with the second. The solution does not change. Multiplying an equation with a non-0 scalar and adding it to another. So we will get a new equation in place of the old one. But still the system has the same solution. Whichever was the solution earlier, that still stays the solution.

And thirdly, if you multiply an equation by a non-0 scalar, you are changing the, scaling both sides equally, so the solution does not change. So but an important thing is here it should be non-0 scalar, multiply by non-0 scalar. So these 3 type of operations are operated upon the equations and they are called elementary row operations. So we observe that the elementary row operations when performed on a system of linear equations, do not change the solution set.

(Refer Slide Time: 04:28)



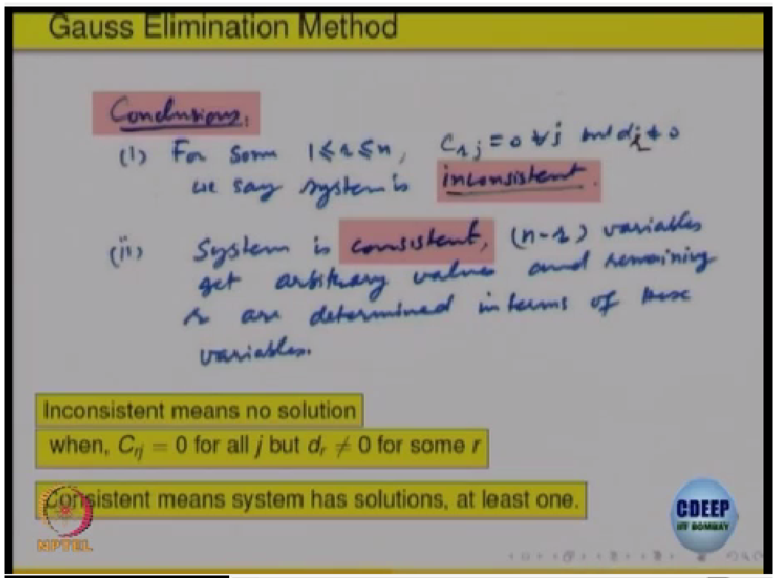
And what is the aim? The aim is by applying this operations, elementary row operations, m*n system that is m in linear equations in n variables can be transformed to a system of equations of the following type where the first equation has coefficients c11 c12 and so on. In the next one, so in the next equation, at least one of the variables is eliminated. If here c11 was not 0, then in the next one, that coefficient is eliminated and you get an equation which starts with x2.

Possibly x2 also can be 0 but x1 is positively missing in this. So 1 variable has been eliminated. So using those elementary row operations, the system, given system can be changed, can be transformed to a system of the following type. So we write this as there is a number r. So r is less than or equal to the number of the variables, right. n is the number of variables, so that is number r between 1 and n such that the first r equations possibly have non-0 coefficients c1 ci's and on the right hand side, possibly these are non-0 but after the r+1 equation, everything is 0=0.

So what we have done is? We have reduced the number of equations which are required to find the solution, okay, to r equations possible. Here remaining equations n-r equations are all 0 equations 0. So this is the method suggested by Gauss that do those row operations, elementary row operations on the system and transform into this form. But what is the advantage of this? So let us, now supposing this last equation, the remaining equations are always true but look at this equation, okay.

In this supposing it so happens, okay, that one of the coefficients crr is not equal to 0, is 0, all the coefficients here are 0 but dr is not equal to 0, suppose this happens, right. It is possible, we did not say that these coefficients need not be 0. They all can be 0 but dr is not 0. That will be in what? You will get 0=a non-0 number and that will be in, there is some inconsistency in the system. So the system will not have any solution. So that is called an inconsistent.

(Refer Slide Time: 07:21)



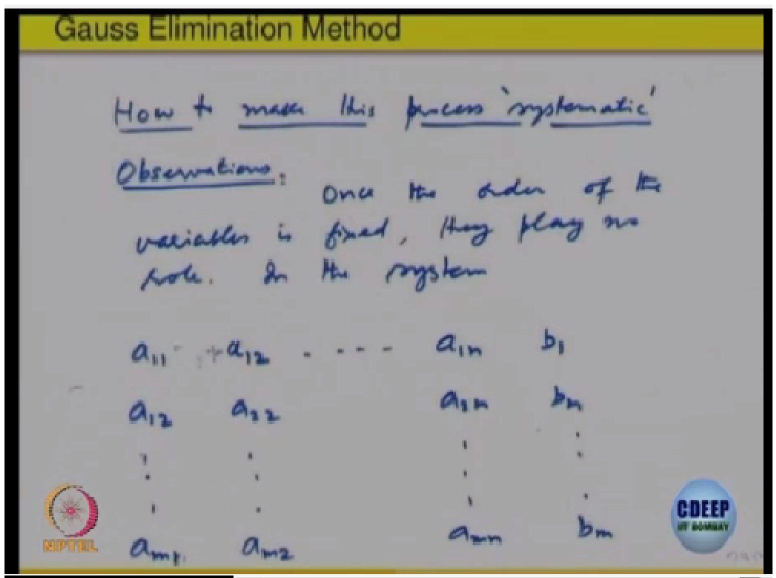
So if conclusion is, if for some r between 1 and n, n is the number of variables, all the coefficients crj=0, so that means for all j, that equation on the left hand side is 0 but the right hand side is not equal to 0. Then the system is inconsistent. So the second possibility is, system is consistent, that means this does not happen. So then there are solutions. So then n-r variables get arbitrary values.

So let us just look at the previous equation. Supposing this is a consistent system. So this involves the variable xr xr+1 and xn. So in this equation, this is an equation involving these variables, n-r variables are involved in this. We can solve one of the variables in terms of the other variables using this equation.

So in this, we can give the remaining variables, are arbitrary variables, find values for one of the variables. So that what it says that n-r variables get arbitrary values, okay. And the remaining can be determined in terms of these variables from the backward substitution. So this, we are summarizing this.

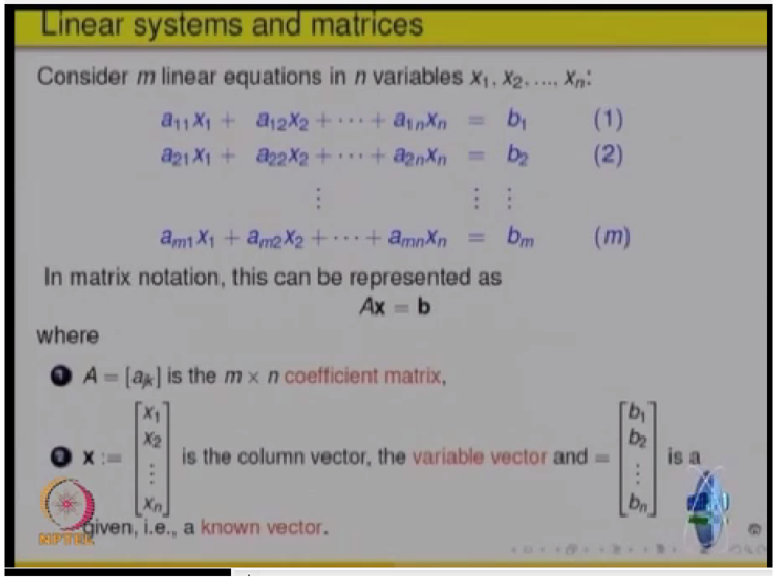
The system is inconsistent means no solution and that happens when there is something like 0=non-0 in one of those equations which have been reduced to a specific form using the row operations. Consistent means there are solutions at least 1 solution and you can get between infinite number of solutions when putting various values for some variables and calculating the other in terms of that. So this is the method suggested by Gauss to solve system of linear equations.

(Refer Slide Time: 09:27)



To make the system more systematic, we realize that the coefficients are important and not the variables. So a11x1+ so on, so what we do is, we will just look at the coefficients of these equations. Here it was equality equal to b1. So we look at this array of numbers. So it becomes important to study array of numbers. That gave us a concept of what is a matrix.

(Refer Slide Time: 10:00)

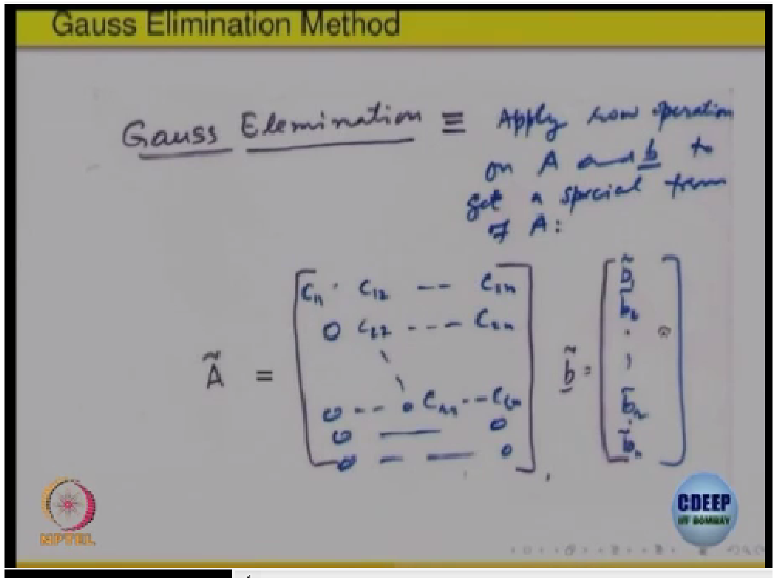


So such a thing is called a matrix. So this system we saw it yesterday in the previous lectures that this system of linear equations m equations in n variables can be written in terms of matrices. So a11 a12 a1n, if you take the other as the first row, second row a21 a2n and so on, and call that matrix as A. So A=ajk is the m*n coefficient matrix.

So the matrix consisting of the coefficients of this m equations for getting the variables and forgetting the arithmetic symbols that is A, x is the variable which is unknown vector, x1 x2 xn, that is the variable vector you can call it X. So this is n rows, 1 column, this is m rows n columns, so when you multiply, you will get m*, right.

So b1 b2 bm, so there is a typo here. It should have been bm, okay. So in the matrix form, you get this AX=b in the matrix form of the equation. The idea writing in the matrix form is we are not really concerned with this X, we are only concerned with A and b when we transform, apply elementary transformations.

(Refer Slide Time: 11:18)

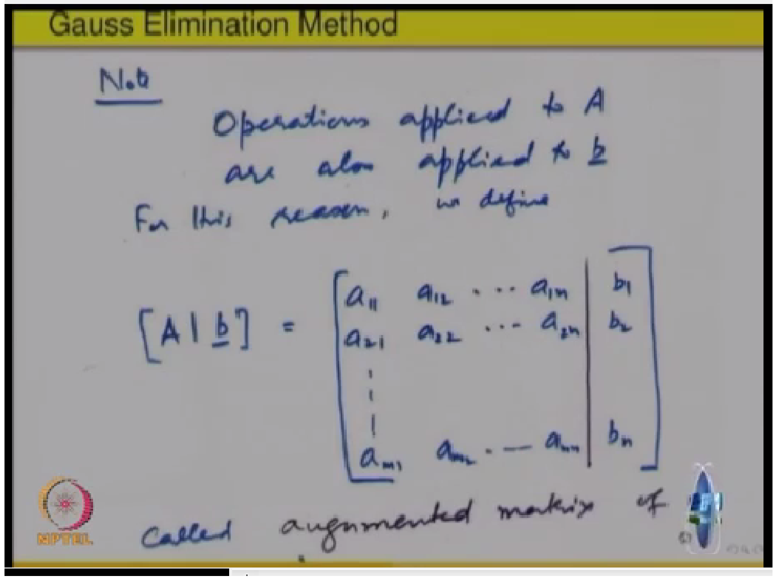


So and by Gauss elimination method says that the coefficient matrix and that vector b, after those elementary row operations, get transformed into matrices of this type. A gets transformed into something like c11 c12 and so on and there are some bottom rows which are equal to 0, right. That is a form after applying elementary row operations. And of course the vector b will get changed to something like this.

So that is a vector b~. So the transformed matrix A which was original is transformed to A~ which is of this form and the vector b gets transformed to b~ because whatever operations you are doing on the left hand side, you have to do it on the right hand side also, right. So when A changes B automatically is going to change according to those operations, right. So this is the new coefficient matrix for the new system which is equivalent to the earlier one.

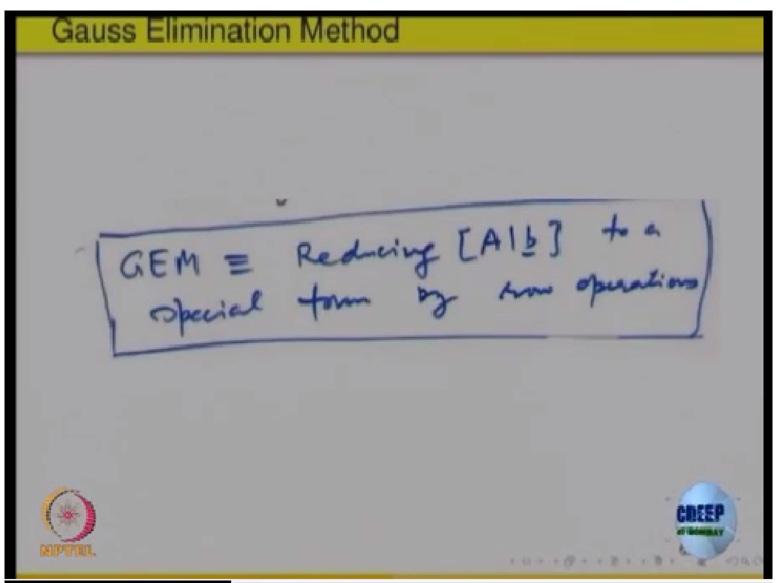
And this has a special form. Look at the bottom, there are some rows which are all equal to 0 and the number of possibly non-0 coefficients, right, the place where they occur is increasing. This is special form of a matrix. We will spend some time on this form of the matrix because this is going to be important for all future calculations.

(Refer Slide Time: 12:43)



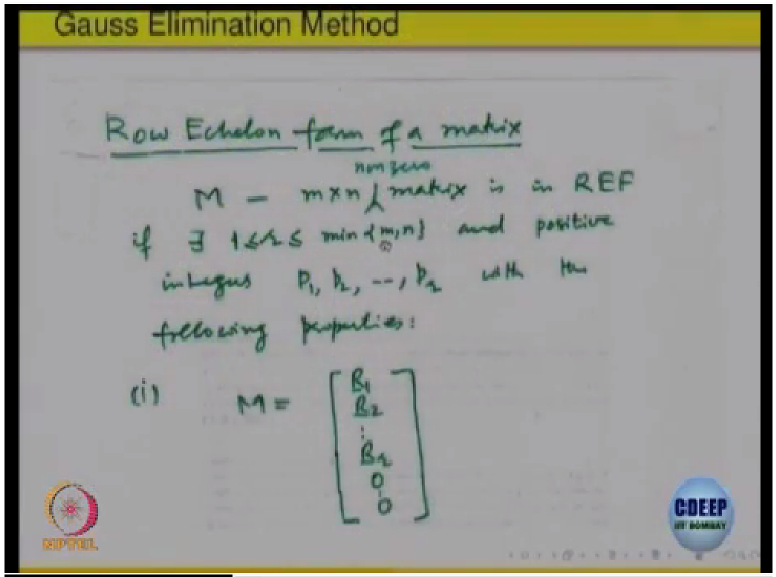
So to keep track of A and the vector b together, we form a new matrix, A is m*n, b is n*1, right. b is the coefficient, b is on the right hand side, b1 b2 bm. So again there is a typo, here it should have been bm, right. Because there are m equations, right. So the row operations whatever we do on A are also going to be done on b. So this is the matrix on which we are going to operate. And the idea is do the elementary row operations and change it to that form. This part of the matrix should be changed to the special form.

(Refer Slide Time: 13:31)



So this matrix is called the augmented matrix because we have added 1 more, we have added 1 more column to A. So it is called augmented matrix of the system of linear equations. And Gauss elimination method is summarized as reducing this matrix to a special form by elementary row operations. So what is that special form? We will discuss that.

(Refer Slide Time: 14:00)

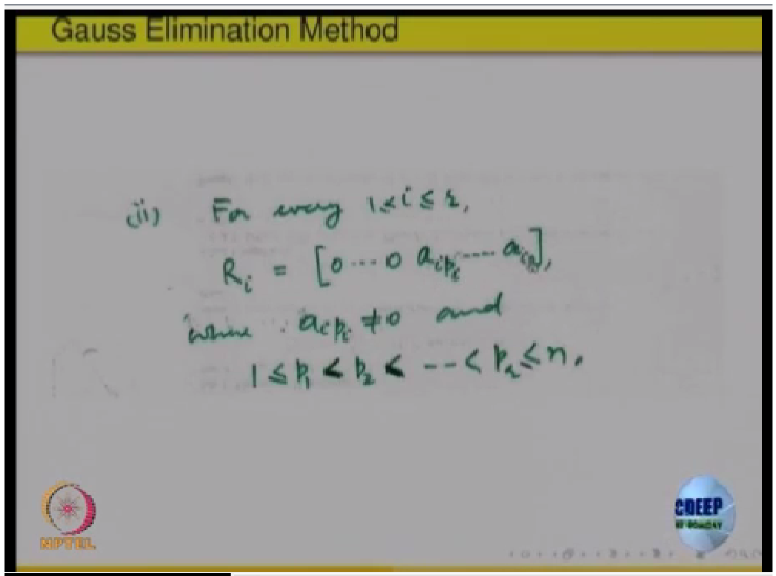


Here is the special form. It is called the row echelon form of a matrix. For any matrix m which is m*n, non-0, of course, the 0 matrix has nothing to do with it. It will not given anything. So we assume that it is at least 1 non-0 entry in the matrix. So m is m*n, non-0 matrix. We say this matrix is in row echelon form. REF is short for row echelon form if there is a number r between 1 and the minimum of m and n.

What is m? M is the number of rows, n is the number of columns. So there is a number r which is between 1 and the minimum of the 2, okay. And there are integers which are p1 p2 and pr, right, natural numbers p1 p2 pr. Here is the number r, right. So this is between minimum and there are natural numbers 1, p1 p2 pr with the following properties, right.

So we are going to describe the matrix in terms of this r and p1 p2 and pr. The first property says the matrix looks like the first r rows possibly are non-0, r1 r2 rr. The first r rows are possibly non-0. So this is a number r which is appearing here. Remaining are all 0 rows. Bottom is just 0, okay. So that is the property of this, characteristic property of this number r, okay.

(Refer Slide Time: 15:44)

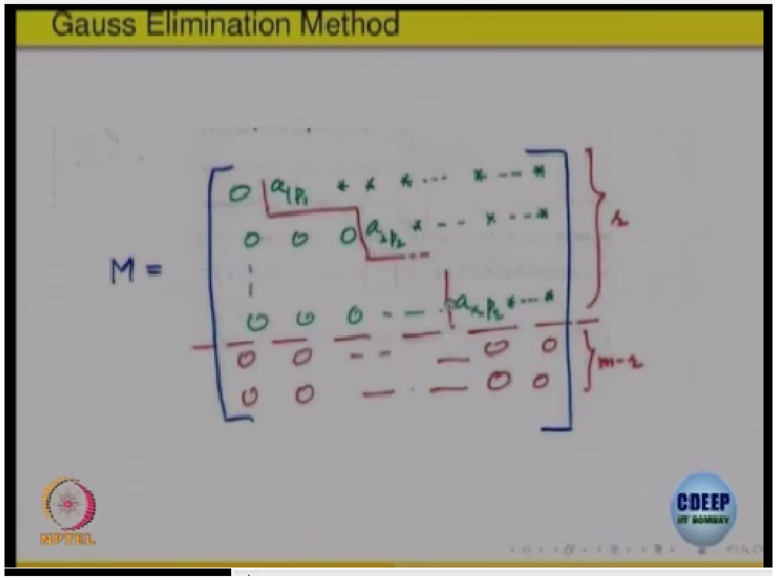


So what is the next one? Now we add those, now we want to describe those, these are non-0 rows, right. So I am going to describe these non-0 rows now. So pickup anyone of them. So call it Ri, I is between 1 and r. This is a non-0 row. How does it look like? It look likes 0 0 0 0. The first non-0 entry should be aipi. So that pi which was there, right that was minimum, between 1 and; so is less than the number of columns.

So that means in the ith, pith column, the first non-0 entry appears. This is the ith column. This is the, right, pith column, ith row, this is the ith row going. At the pith column, the first non-0 entry appears. We do not bother about what are the remaining entries. Only we bother that the first non-0 entry in the non-0 row, it should come at the pith place, pith column. And these numbers, they should be in this order. p1 should be strictly less than p2, right; p2 should be strictly less than p3.

That means what? These are (()) (16:58) in the first r rows, the non-0 entries, they cannot go back kind of a thing. If it is, in some row it is coming at a place, then the next rows, it has to be on the right side of it, column, right. The previous ones are all 0. So the first non-0 entry comes at pith column. So in p1 somewhere comes, so p2 non-0 entry should not come on the left side of that column. It should be only on the column number on the right hand side of it. So that is the property of this p1 p2 pr, okay.

(Refer Slide Time: 17:32)



So this is what the matrix will look like. Maybe the first row, first entry is 0, it does not matter. So there is a number r. Here is the number r, so that the bottom rows are all 0. So what are the non-0 rows left? These top r rows are non-0. So it has a non-0 entry. Each of these rows has a non-0 entry somewhere. We are only describing which place it has non-0 entry. For the first one, it comes at the p1th column.

This is p1 column. For the second one, it should not come on the left side column. It should go somewhere on the right side. So it comes at a place called number as p2. So the column for the second row is p2 where the first non-0 entry in that row occurs. That should be on the right side of p1 and so on. So in the rth one, it comes at a place we do not know what are these entries remaining.

We are not bothered about them also. The only condition we want is, there should be a number r, okay. M*n matrix, so that the bottom rows are all 0 and the top r rows have got at least 1 non-0 entry somewhere, right. Now we want to describe where it has it. So we will start looking at the first one. First one comes in the column p1. For the second row, the non-0 entry should come at the column p2 which is bigger than p1.

We are not saying it should be the next one. No, we are not saying even that. If it comes in p1, p2 should be on the right side. So p2 should be the column. p2 should be bigger than p1. It could be here somewhere also, possibly, right. It can go beyond. But positively, it should be on the right side of the column for the first one. Each next row the non-0 entry should come on the right side, right.

The column in which it appears, should be, number should be bigger than the previous one. So it should look like some kind of a staircase kind of a thing and then bottom are all 0, right. So this is called the row echelon form of a matrix and the Gauss said if you apply elementary row operations, every matrix, he did not say exactly in the matrix form, he was doing it for linear equations 2 and 3 variables and so on.

He said you can bring it to that form, right. So later on it was made more rigorous. For a matrix, of course non-0, m*n, there is a number r. Why this has to be written minimum of m*n? Because r rows are non-0, right. So r has to be less than the number of rows, r has to be less than m but the non-0 entries going to come in a column. So r has to be less than or equal to number of columns, right.

Is it clear. That is why this condition is put r should be between 1. It should be between the first column and the minimum of m and n. If it is a 3*2 matrix, right, this number r cannot be more than 2. That means it cannot be 3. It has to be 1 or 2 only, right. So the first, so r tells us how many rows are non-0. Bottom rows are 0, right. And each row which is non-0 is described, sorry it is described by numbers pi.

For the ith row which is non-0, look at the column where the non-0 entry is coming. So look at the place where the non-0 entry, first non-0 entry. This is the first non-0 entry, is coming. So that place is pi, then we should have p1<p2<pr. Of course, it has to be less than or equal to n, the number of columns. So that is row echelon form and that is what it says. The matrix will look like this. Now note, we are not saying that is the only one row echelon form of a matrix, right?

For example, if I multiply this whole thing by some scalar, still the non-0 will remain non-0. So that also will be in row echelon form, right. So when we change a matrix to a row echelon form, we are not saying there is only 1 possible answer, right. The numbers could be different but that pattern has to be same, right. In 2 row echelon forms, r has to be unique and p1 p2 pr have to be unique, right. Entries are not important in the row echelon form, right. Is it clear? We will do more examples to understand that.

(Refer Slide Time: 22:33)