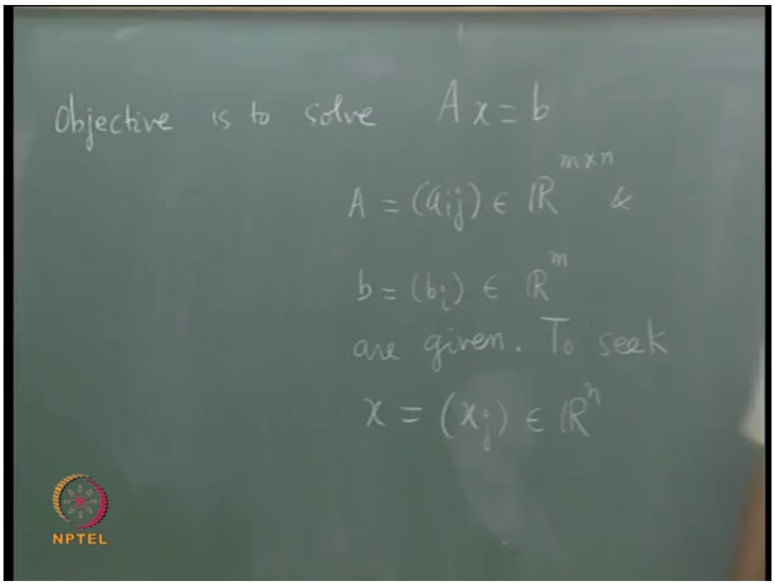
Linear Algebra
By Professor K. C. Sivakumar
Department of Mathematics
Indian Institute of Technology, Madras
Module 1
Lecture 2
Linear Equations

(Refer Slide Time: 0:27)



See we could one could say that the sole objective of linear algebra is to solve a single equation, okay let me write down this equation is to solve primary objective the primary objective of linear algebra is to solve the equation Ax equal to b the linear equation Ax equal to b where what is a what is b let me give you the details, A is an m cross n matrix aij will be the we will denote the ith row jth column entry so this is a matrix of order m by n m rows n columns and entries will be real that is what we mean by this, okay R with a vertical arrow is the set of real field of real numbers.

So A is an m cross n matrix this is given I am also given the right hand side vector b is bi this belongs to Rm, okay these two are given these are given the vector x we are seeking x a column vector, okay let me go back to this location we want x is equal to xj in Rn to seek x that satisfies this equation, okay A x equal to b so one could say that this is the central principle of central objective of linear algebra. So our objective is to any practical problem can be any practical linear problem can be modeled in this manner to solve a system of the type Ax equal to b, okay and towards the end of this course one would hope will be in a position to solve this at least understand the equation Ax equal to b whether it has a solution if there is a solution how to compute the solutions if it does not have a solution we will still have to solve this equation, okay.

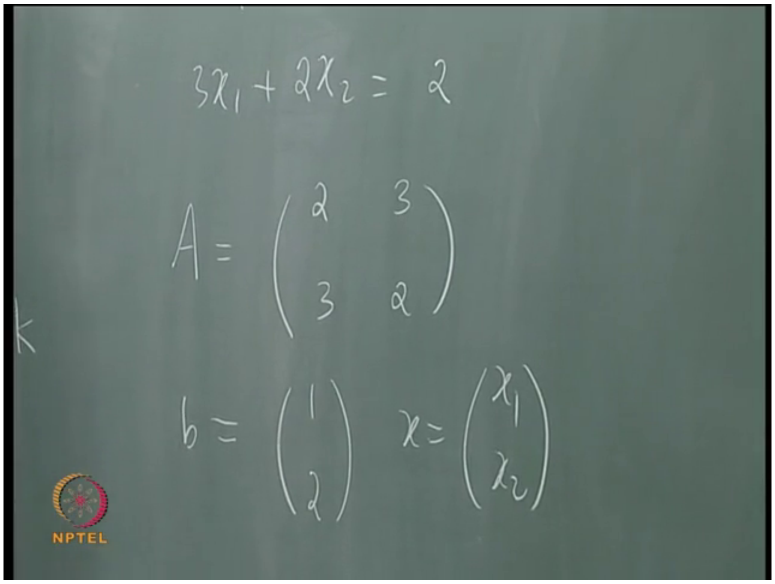
So does this system have a solution that is the first question existence if it has a solution is it unique if the solution is not unique how to at least compute certain solutions and then the extreme case when the system does not have a solution one would still like to solve it so that leads to the notion of best approximate solutions, okay. And you see that this again involves a matrices so for the first few lectures we are going to discuss matrices in particular we are gong to discuss what are called as the elementary row operations on matrices, okay.

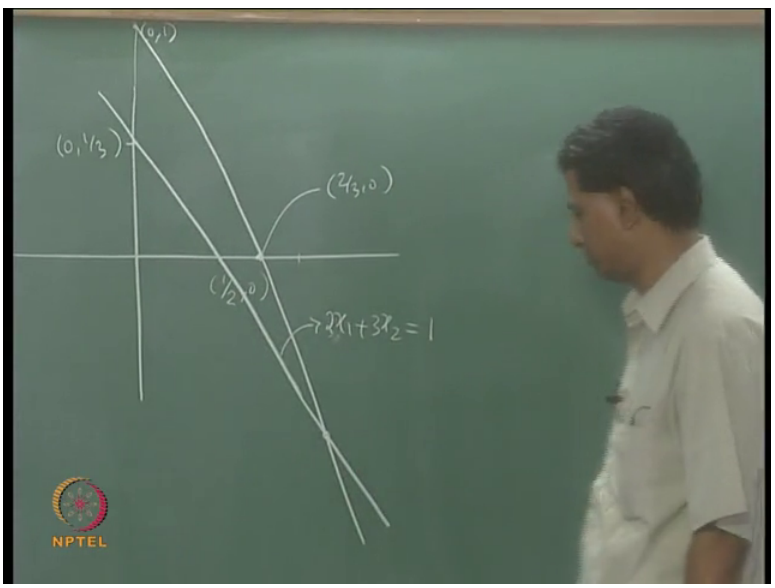
To give a motivation to the notion of elementary row operations and the why one must study those operations let me go back to a 2 by 2 example, okay this example is something that we have seen probably even in our high school higher secondary but I will keep that as a motivating example and also to give a geometric view of solution of linear equations, okay remember that this geometry thing can be applied only if there are two variables for three variables also one could do but it becomes little more difficult. So for two variables we will look at the geometric point of view of what a system represents, what a solution represents etcetera.

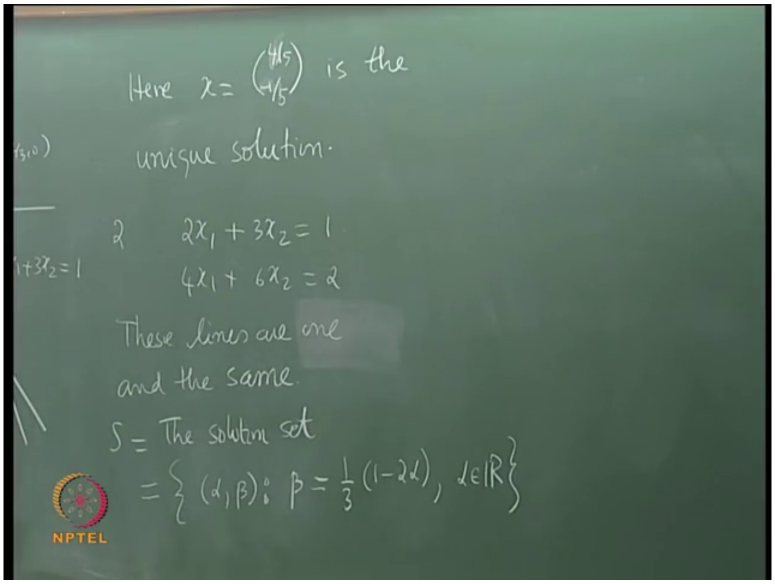
And as you will see when the number of variables increases you will have to take (())(4:58) to a program it is not something that you can solve on the black board, so one needs to write down a program and then enter the inputs in a system and then solve it numerical, okay. Now numerical solutions are again not part of this course but I will at least tell you the theoretical background behind this numerical techniques that is why we need the that is where we need the notion of elementary row operations, okay. So we will discuss elementary row operations some of the properties of these some of the properties of matrices that are so called row equivalent matrices and then see how systems can be solved, okay we will also look at how one could find the inverse of a matrix using the elementary row operations, okay.

So let me go back to the two dimensional example motivating two dimensional example where one could use geometry.

(Refer Slide Time: 5:58)







So we have seen equations of this type I am just taking a hypothetical situation 2 x1 plus 3 x2 equals 1 and let us say I have 3 x1 plus 2 x2 equals 2, okay this is an equation involving this is the system involving two equations in two unknowns this is a particular instance of the system A x equal to b, okay where a A is the coefficient matrix the right hand side vector also called sometime as a requirement vector is 1, 2 now you will see that I am using a column representation, so any vector standing alone for me will a column vector, okay.

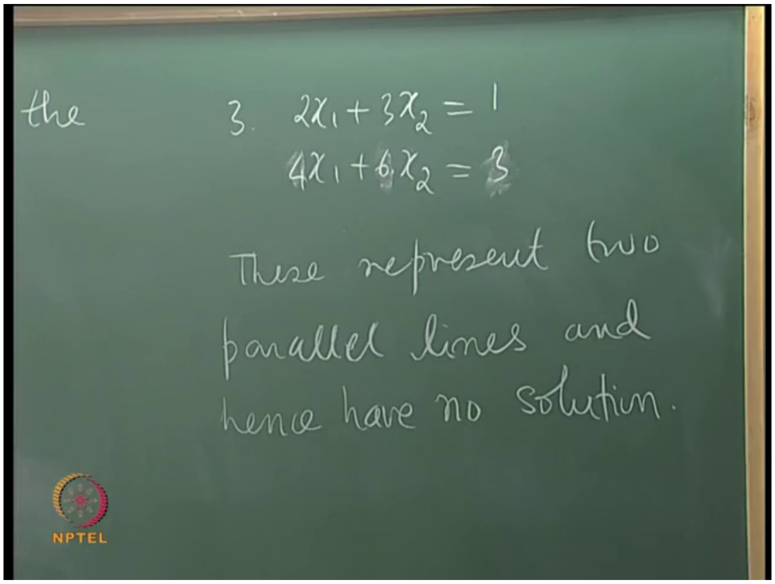
This is the requirement vector the unknown vector has two coordinates the unknown vector has two components x1 and x2 the question is does this system have a solution geometrically what is a meaning of a system having solution. So what one does is to geometrically we solve this problem is to draw these lines, okay 2 x1 plus 3 x2 equals 1 the scaling will be only approximate, okay so what are the points that this line passes through? 1 by 2 comma 0, right? x1 is 1 by 2 let us say this is 1 for me so this is 1 by 2 approximately and the other one is 0 comma 1 by 3, okay so let us this is half so let us say this is my 1 by 3 0 comma 1 by 3, okay and so this is line 1, okay so this represents 2 x1 plus 3x 2 equals 1 and similarly we can draw the other line it passes through 0 comma 1 is that right, 0 comma 1 that is somewhere here, okay and the other one is 2 by 3 comma 0, 2 by 3 comma 0 is somewhere here, okay let us say here.

Point 3 see this is half, isn’t it? And the I, okay so this is 2 by 3 comma 0 this is let us say 0 comma what is the other one, 1, okay then draw the line joining them approximately extend it this is the first line they meet at a point these two lines are not parallel because the slopes are different one has the slope 3 by minus 3 by to the other one has minus 2 by 3 the slopes are different so these two are not parallel so they intersect and so look at this point you can find the coordinates of this point, okay that gives a solution.

So geometric view point is that if the two lines are not parallel then we know in Euclidean geometry the two lines must meet somewhere. The point at which the two lines meet is unique, if two lines are not parallel then they meet at a unique point that unique point is a solution for the system of equations. So one could solve and get the solution for this system, yeah so 4 by 5 minus 1 by 5 that is a unique solution, yeah thanks. Let us look at another system I will take the first equation as it is 2x 1 plus 3x 2 equals 1 the second equation for me will be 4x 1 plus 6x 2 equals 2, okay you can see what I have done the second equation is a multiple of the first one, okay the slopes are the same these two lines are parallel.

So by geometry we know that, that there is no solution, right? Are you sure? These two lines are the same so there are infinitely many solutions, okay these two lines are the same there are infinitely many solutions and one can write down the infinitely many solutions in the form of a solution set these lines the set of solutions let me write S for that this is the set of all alpha comma beta such that beta equals 1 by 3 times 1 minus 2 alpha alpha element of R that is a set of all solutions, is that okay? One could plug in and verify that this satisfies the equation that is 3 beta that is 3 x2 plus 2x 1 equals 1, okay so that is the solution set in this example. So these two lines are one and the same.

(Refer Slide Time: 12:44)

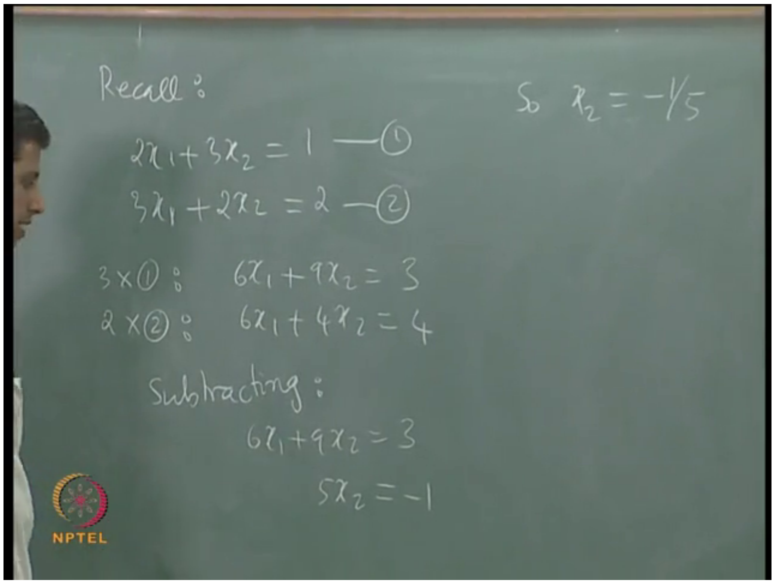


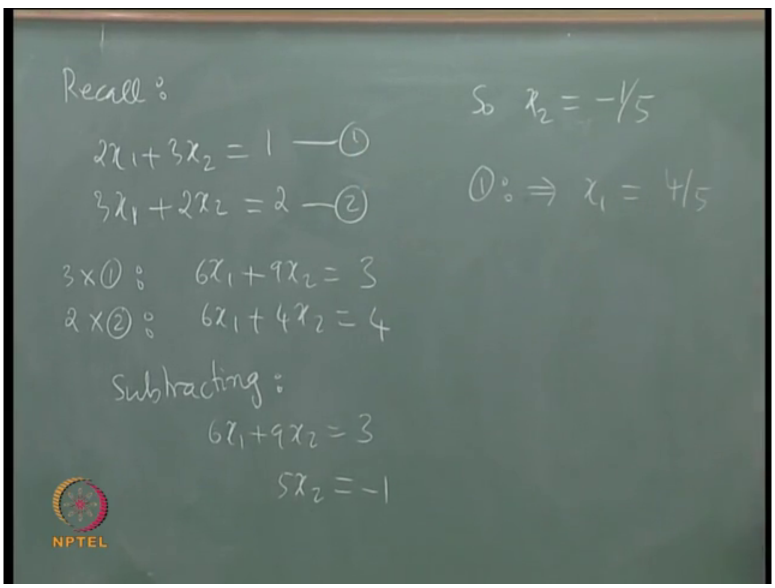
One final example there is a unique solution that is the first example infinitely many solutions second example the last example is where it does not have solution, okay let say I have 3 here on the left it is 2 times the previous equation on the right it is 3 times the previous equation so these do not have a solution, so what is the geometric viewpoint please think it over these are just parallel lines, right? Okay.

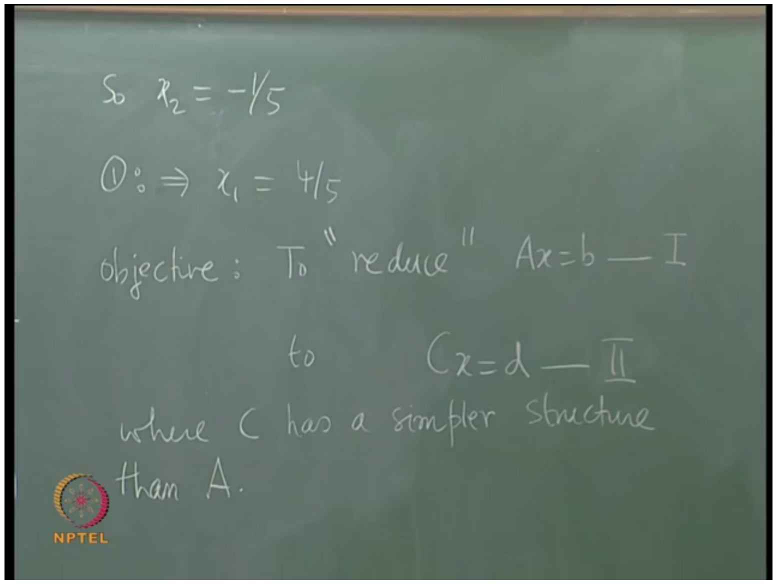
These represent two parallel lines and hence have no solution, okay now these are the three situations that one has when one studies the equation Ax equal to b, okay alright. As I mentioned there are situations when you will have to solve inconsistent systems this is a the typical example of a inconsistence this is an example of an inconsistent system because it does not have a solution so any solution that does not have any system rather which does not have a solution is called an inconsistent system.

In the case of inconsistent system one still has to solve, okay but you need either methods from calculus or methods from inner product spaces to deal with such problem, okay. So that is just to give you a geometric viewpoint of what linear the solutions of a linear equation, okay what is the geometric representation of a solution of linear equations and you see that this can be done only for two dimensions, okay only when there are two variables. Three variables one could still do but it gets little more complicated, four variables is out of question, okay.

(Refer Slide Time: 15:04)







Let us go back to this system that I had written down earlier the first example which is 2 x1 plus 3 x2 equals 1 3 x1 plus 2 x2 equals 2 this is what I wrote down as a first example, how do we solve this in high school? Multiply the first equation by 3, multiply the second equation by 2 and then subtract 1 from the other, okay so let us call this equation 1, equation 2 then 3 times equation 1 gives me 6 x1 plus 9 x2 equals 3 two times equation 2 gives me 6 x1 plus 4 x2 equals 4 then these coefficients are the same so one does a subtraction subtracting we get first equation is 6 x1 plus 9 x2 equals 3 instead of the second equation we have 5 x2 equals minus 1 so I am able to solve for x2 immediately looking at the last equation.

So I have x2 equal to minus 1 by 5 and I go back to this equation and solve for x1, yeah I could do that but I want to specifically lead you into formal Gaussian elimination, see the objective is to formalize elimination Gaussian elimination we have learnt in high school one of the aims of linear algebra is to formalize Gaussian elimination in such a way that one could write a program for instance and use a system for computing the solutions of a system where the number will be huge the number of equations number of unknowns this will be huge, okay so one needs an automated system which can take care of that.

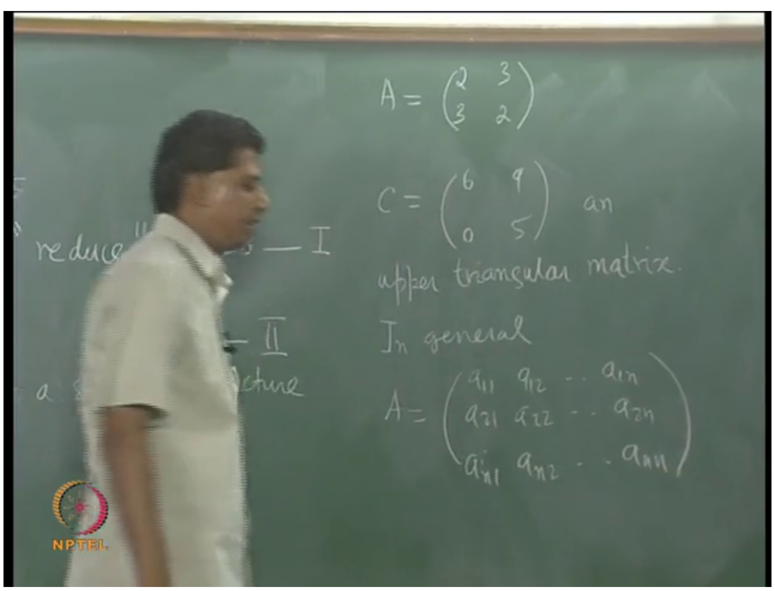
So intentionally I am writing down this equations from this what we normally do is immediately conclude subtracting one from the other that x2 is minus 1 by 5 and then we go back to equation 1 and then solve for x1, okay. So let me say equation 1 gives me x1 is 4 by 5, okay so this is what we have learnt in high school this is Gaussian elimination if you have not heard the name

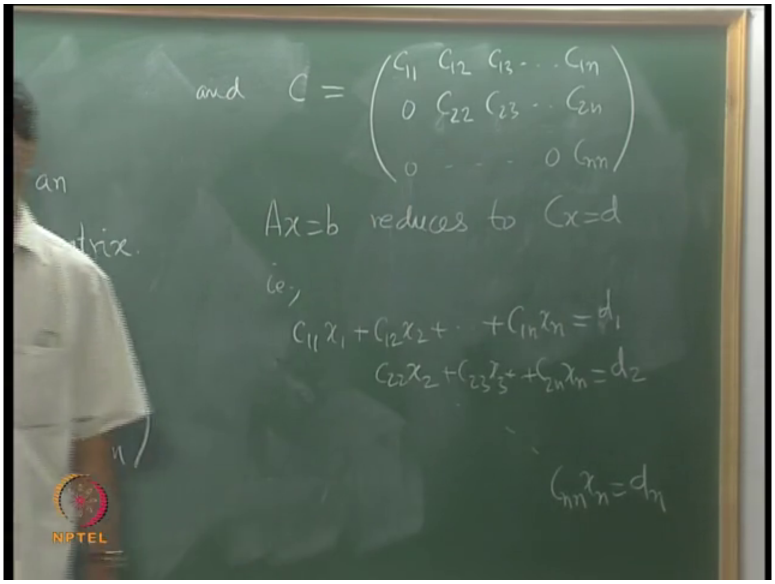
Gaussian elimination before we would like to formalize as I mentioned but before we look at how to formalize this let us look at the structure of the second system that we derived from the system 1, what is the structure of the second system?

The second system has a simpler structure which has allowed me to solve it the simpler structure when you compare with the first one first system is that in the second equation the first variable is not there, okay in general we seek the following in general the objective is to reduce whatever it means to reduce the system Ax equal to b let me call this system 1 to reduce a system Ax equal to b to a system of the type Cx equals d I will call this system 2 I want to reduce system A x equal to b to a system Cx equals d where C has a simpler structure that is what reduction means where C has a simpler structure than A only then we would have reduced, okay the system to a simpler system this is the objective and we have already seen what kind of C we are seeking, what is the type of C that we are seeking? We are seeking C to be an upper triangular matrix, okay.

So let me write down the matrix form of this system 1, this is system 2 for me, okay this is system 1, this is system 2 what is a matrix form of these two systems? To motivate what we mean by a simpler coefficient matrix C.

(Refer Slide Time: 20:27)





So in this example A is 2, 3, 3, 2 C is 6, 9, 0, 5, okay this is a 2 by 2 example where you do not see much of the triangular structure but still it illustrates what we would like to achieve an upper triangular matrix, okay C is an upper triangular matrix a square matrix is called upper triangular if the entry is below the principle diagonal R0, okay.

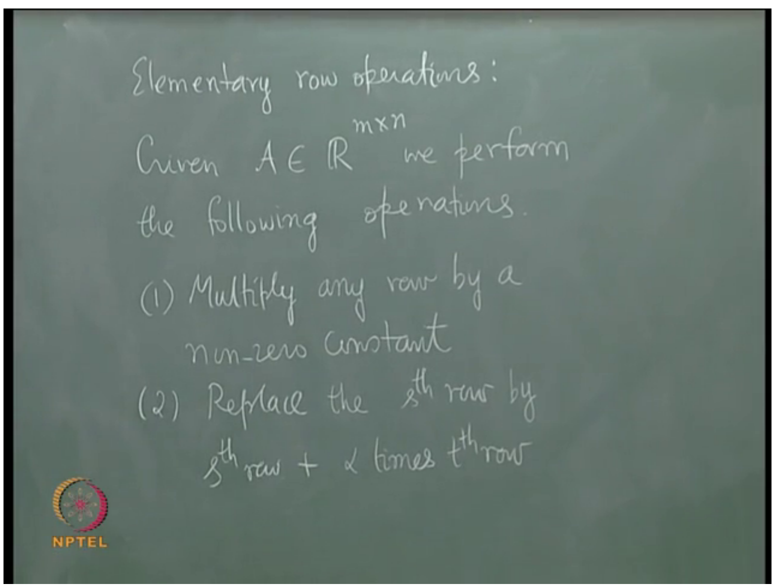
In general what is the structure of C? In general so remember in this case I am only dealing with square systems I have two equations and two unknowns, okay so we are presently dealing only with square systems the number of equations is equal to the number of unknowns. So in general A is A11, A12 etcetera A1n A21, A22, etcetera A2n, An1, An2 etcetera Ann this is A and what is the C that we are seeking upper triangular so must C11, C12, C13 etcetera C1n this entry is 0 C22, C23 etcetera C2n all these entries are 0 the final row is Cnn, okay.

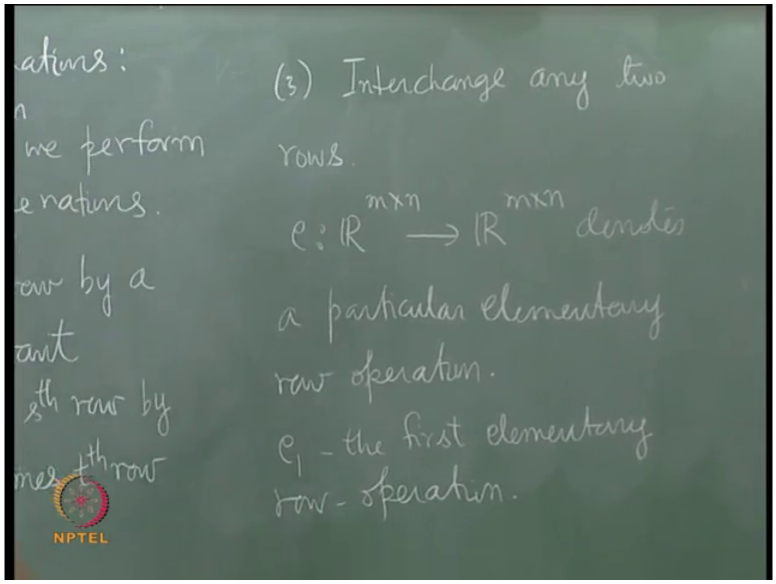
So this of the upper triangular form and so A x equal to b becomes Ax equal to d, right? So let us look at the general problem Ax equal to b reduces to Cx equal to d that is let me now expand matrix multiplication so let me now (())(22:48) what it looks like. C11 x1 plus C12 x2 etcetera C1n xn equals d1, sorry C22 x2 plus C23 x3 etcetera C2n xn equals d2 etcetera the last equation Cnn xn equals dn, okay let me also write down the equation before the last equation C n minus 1 xn minus 1 Cn minus 1 n xn equals d n minus 1 is a last there is a equation before the last equation.

So what we have done is first solve for one of the unknowns from the last equation the last unknown the unknown xn so we get xn as dn by Cnn provided Cnn is not 0, okay see we are dealing with those technicalities right now what is a guarantee that Cnn is not 0 etcetera we do not know right now there is no such there is no guarantee that we have till now but it is working in this example, okay. So if Cnn is not 0 then Xn is dn by Cnn go to the previous equations substitute for this Xn you determine Xn minus 1 then the previous equation will determine Xn minus 2 etcetera.

This is called backward substitution, okay this method is called backward substitution this is what we have done in the case of two variables in this particular problem. So we need to derive a C, okay from the matrix A we must do certain operations in order to get to the matrix C which has this particular upper triangular structure that is where elementary row operations come into the picture so let me give you the details of what elementary row operations are, okay and then actually do a problem where the systems reduces to a system where you have an upper triangular structure derive the coefficient matrix has upper triangular structure, okay.

(Refer Slide Time: 25:51)



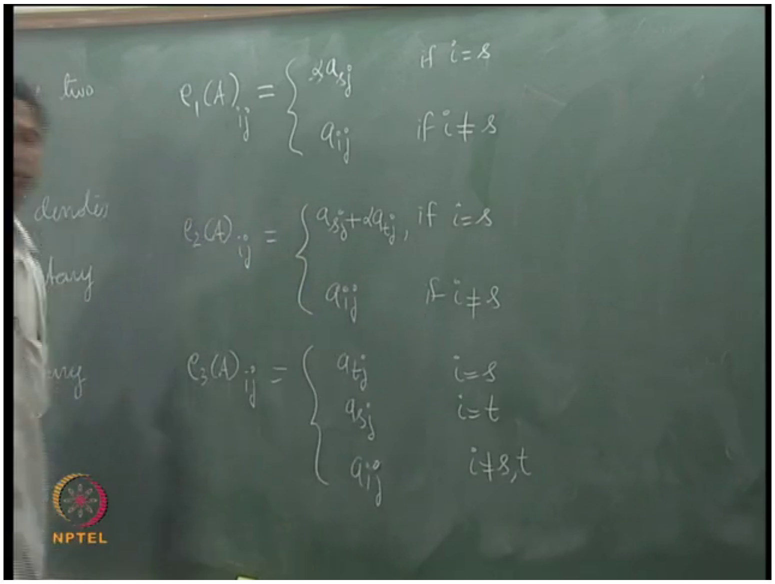


So let me first tell you what this elementary row operations are, okay so we are specifically interested in only three row operations let me list these operations, okay so given a matrix A we perform the following operations, the first one is multiply a row by a non-zero constant, okay multiply any row by a non-zero constant that is a first operation, second operation is replace any row let us say replace the sth row by sth row plus a constant alpha times tth row replace the sth row by sth plus a constant times another row I am calling that as the tth row that is a second operation.

Third operation is interchange any two rows, these are the three elementary row operations, let us formalize this let us introduce a function and then write down these operations in terms of this function let me call e as a function from R m cross n to R m cross n will be a particular elementary row operations e denotes a particular elementary row operation, okay e denotes a particular row operation I will use e1, e2, e3 to denote the three row operations, so what is e1 of A, okay I need to write down e1 A completely in terms of the entries of A that is what I will do next.

What is e1 of A? So let us say e1 corresponds to the first elementary row operation let us say e1 is the function corresponding to the first elementary row operation then what is formula for e1, see e1 is a function a function is known if its action on each element is known, okay.

(Refer Slide Time: 30:15)



So what is e1 of A? Multiply any row by a non-zero constant let us fix that row let us fix that row and say that it a sth row that we are multiplying by a non-zero constant alpha, okay so this is if see I want the entries of e1 of A so let me say e1 A ij, what is ith row jth element of e1 of A? e1 A ij is this is equal to tell me if this is alright this is A ij if i is not equal to S if i is equal to S I am multiplying the sth row by non-zero constant alpha so if y is equal to S then it is S times I am sorry alpha times A Sj this describes the action of e1 completely, is that okay the sth row has been replaced by a constant times sth row.

So this is e1 of A for me this gives a complete description can we write down e 2 of A similarly e2 ij replace the sth row by sth plus alpha times tth row, okay if i is not equal to S then there is no change so it is a ij as before the entries of a corresponding to these rows except the sth row for the sth row what do we have if i is equal to S the sth row has been replaced by sth row plus alpha times tth row it is a sj plus alpha a tj if i is equal to s, okay so this describes the second operation completely, third operation interchange any two rows let us say we are interchanging s and t row s and row t, okay.

So if it now row or row t if i is not equal to s or t then there is no change if i is equal to s it is a tj and if i is equal to t to a sj so this is the complete formula for e3 of a, okay. So what I have done is to rewrite the elementary row operations in every i which will be useful for us to apply to verify certain properties, okay specifically I want to demonstrate that each of these elementary row operations has an inverse operation and that the inverse operations are also elementary row operations of the same type, okay this is an important observation, okay let me repeat each of these elementary row operations is invertibles you see we have we are viewing this elementary row operations as functions, okay.

So we would like to know whether these functions are invertible do they have inverses, okay so what we will show by using this definition by using this definition we will be able to show that these are the inverses of these elementary row operations are also elementary row operations not only that they are of the same type, okay that will be useful in reducing a system to a system of the type Cx equal to d with a specific structure for C, okay. So I will stop with this for today.