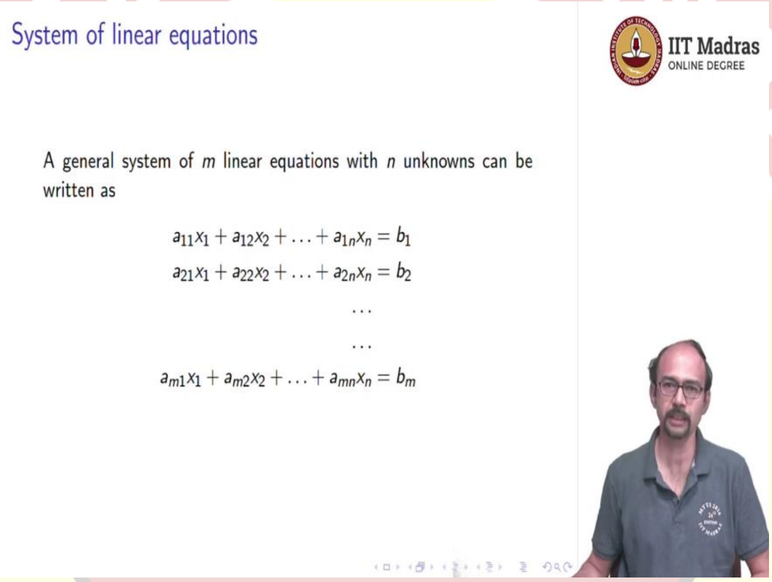
IIT Madras

ONLINE DEGREE

Mathematics for Data Science 2
Professor. Sarang Sane
Department of Mathematics
Indian Institute of Technology, Madras
Lecture No. 09
The echelon form

Hello and welcome to the maths 2 component of the online BSc course on Data Science. In today's video we are going to talk about the echelon form. So, the Echelon form is a particular form for a matrix and when the matrix is in this form, we can read off solutions quite easily.

(Refer Slide Time: 00:34)

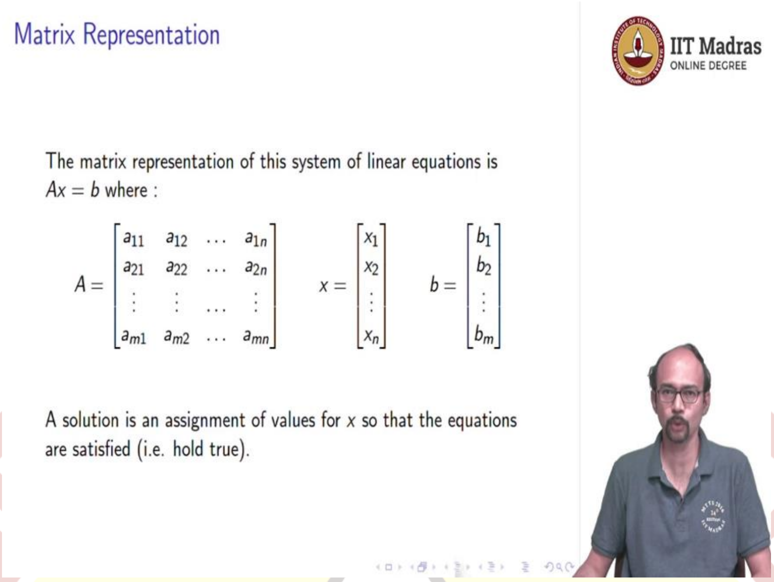


So, let us look at, let us recall first what is the system of linear equations. So, a general system of m linear equations within n unknowns is like this

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$$
$$\dots$$
$$\dots$$
$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m$$

So, there are n unknowns, which is the x_i's and there are m equations. So, for each equation we have a constant on the right-hand side. So, the i'th equation that constant is b_i.

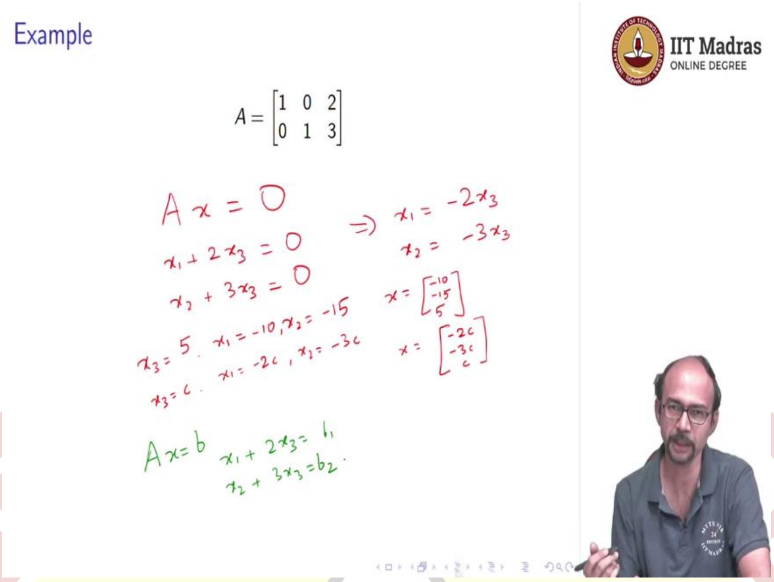
(Refer Slide Time: 01:16)



We have seen earlier that, we can write this in a vector form in a matrix form that is, so we consider the m by n matrix A, which is a matrix of coefficients and x is the vector of unknowns x_1, x_2, x_n and b is the vector of constants, the column matrix b_1, b_2, b_m. And then what is the solution?

So, solution is an assignment of values for x so that the equations are satisfied. So, when you put x_i is equal to some particular number and for each x_i I put in a number and when the equation holds true, that means, indeed when you compute a_{11} times whatever value you have substituted for x_1 plus a_{12} times whatever value of substitute for x_2 + a_1 n times whatever value you have substituted for x_n is b_1 and the same thing happens for all the other equations, then we say that those particular values that we have substituted constitute a solution for x_1, x_2, x_n.

(Refer Slide Time: 02:23)



So, let us look at the example of this matrix. So, we have A = $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{bmatrix}$ So, let us see how we can find out solutions for this. So, here, I am going to write down what the corresponding equations are, of course, here I have not mentioned what the constants are. So, let us take the equation $Ax = 0$. So, this is called a homogeneous system, which we will study later. So, the corresponding equations are going to be $x_1 + 2x_3 = 0$ and $x_2 + 3x_3 = 0$.

So, now, typically, what do we do, we try to manipulate around and find values for x_1, x_2, x_3. If you look at this, this is not really much we can do to manipulate this and the reason is that we can have many, many solutions. So, how do we find solutions? So, you can see from these equations that $x_1 = -2x_3$ and $x_2 = -3x_3$ and that is all that we need, in order to get a solution. So, whenever we have, we substitute a value for x_1 and x_2.

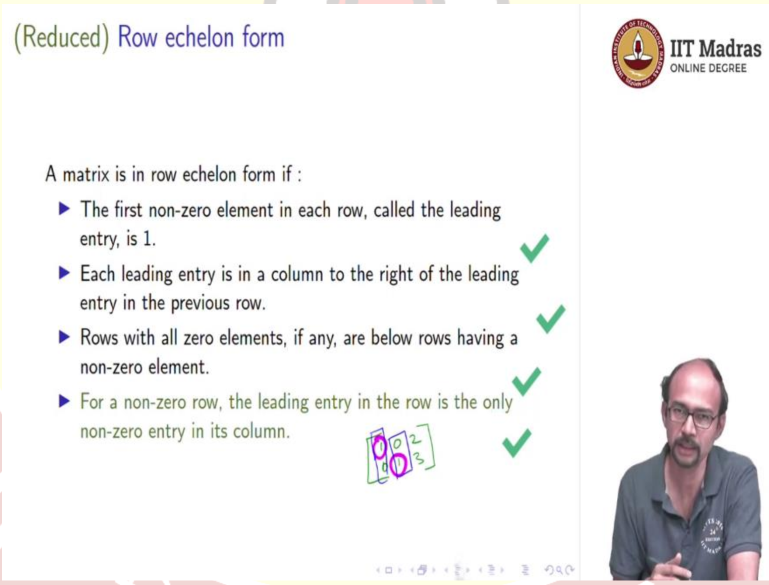
So, you can put $x_3 = c$ and then let us say I take $x_3 = 5$. So, if $x_3 = 5$, then to get my solution, what I can do is I can put $x_1 = -10$ and $x_2 = -15$ and indeed, the one possible solution then is $x = \begin{bmatrix} -10 \\ -15 \\ +5 \end{bmatrix}$, so 5 was an arbitrary choice for x_3, I could have used some other number. So, for example, if I use $x_3 = c$, then I get $x_1 = -2c$ and $x_2 = -2c$.

So really, the set of solutions for this, we can write this as $x = \begin{bmatrix} -2c \\ -3c \\ +c \end{bmatrix}$, where c can be any real number. So, we can have c is 5, c is 20, c is 1 million, whatever you want and all of these, for each value of c we will get a solution. So, we have infinitely many solutions, that is the first thing we observed and not only that, we in fact see that we can get these all these solutions in a very particular way, we can really do this for any b.

So, if you carefully observe what we did, instead of $Ax = 0$, if we had $Ax = b$, we can obtain solutions for Ax is b as well, maybe I will describe this very fast. So, for $Ax = b$, what do we do is we write down again, the same set of equations, the equations are $x_1 + 2x_3 = b_1$ and $x_2 + 3x_3 = b_2$. And from here, we can move x_3 to the other side and then whatever value of x_3 we have, we can read off the values of x_1 and x_2 and each time we will get a solution. So again, we have an infinite number of solutions.

So, there was something particularly easy about this matrix, we could read off all the solutions not only to $Ax = 0$, but $Ax = b$ any b. So, matrices of this form, have a particular, there is a name for this form, it is called the Row echelon form and that is what we are going to study in this video.

(Refer Slide Time: 06:28)



So, when it is said to be in Row echelon form. So, it is said to be in Row echelon form, if the first non-zero element in each row called the leading entry is 1. So, whenever you have a row and it is a non-zero row, so you could have 0 rows of course, if it is a non-zero row, then there will be a first element which is non-zero as you move from left to right. And the first such 1 has to be 1.

The next axiom for the Row echelon form is that each leading entry is in a column, which is to the right of the leading entry in the previous row. So, like you saw in the previous example, in the first row, the 1 was in the 1 1'th place, in the second row, the 1 was in the 2 2'th place. So, if you consider the second column and you ask, or rather, you consider the second row and you ask where is the leading entry, which is the first 1 that you have? It is in the second column and that is to the right of the leading entry in the previous column, which was in the first row.

So, broadly speaking, you have to go downwards and to the right, like this, the 1s have to be like this. Let us look at the next requirement. So, rows with all zero elements, so if there are rows, which are entirely 0, then they must lie at the bottom of the matrix, they must come at the end, so they must be below those rows, which have a non-zero element.

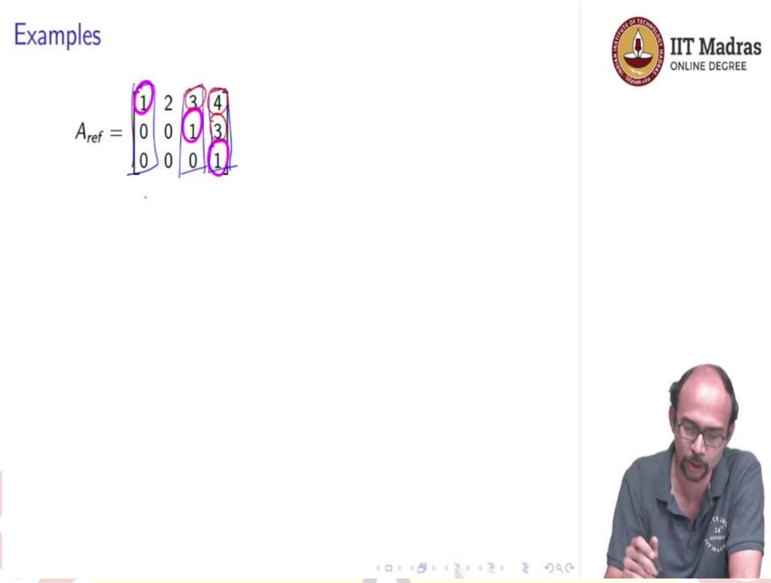
So, these 3 are the requirements for a matrix to be in Row echelon form and then, if you want the matrix to have one further property, which is the following, that for a non-zero row, if you look at the leading entry, you look at that first one that you have and then you look at the column in which the entire column. So, already the requirement that the second requirement that each leading entry in the column to the right, tells you that below this 1 everything is 0, you cannot have anything non-zero below this 1. But you could have non-zero entries above this 1. So, we call it, we say it is in the reduced Row echelon form, if the entries above that leading 1 are 0.

So, the previous example that we had, which was $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{bmatrix}$, this is indeed in reduced Row echelon form. So, first of all, it is in Row echelon form. Let us check that. So, the first non-zero element of each row is a 1 indeed, that is the case because what is the first non-zero element in each row, well this is the first non-zero element, this is the first non-zero element and both of them are 1s.

And then the second requirement is that if you have a leading entry, you will look at the column and all the leading entries in rows before that must be before it, so this one over here this is in the second column, so if you look at the previous row, the 1 is in the first column. So, that is okay, that is exactly what we want. So, this is satisfied and this is satisfied and rows with all zero elements, well, there is no such row, so this is satisfied.

And now let us talk about reduced Row echelon form. So, it says, if you take these ones, then the entire column other than that 1 is 0. Indeed, below this, the first 1 there is a 0 and above the second 1, there is a 0, so that is also satisfied. So, we have seen that in our previous example, indeed, satisfied this. So, it was not the reduced Row echelon form and we could read off the solutions to $Ax = b$ very easily.

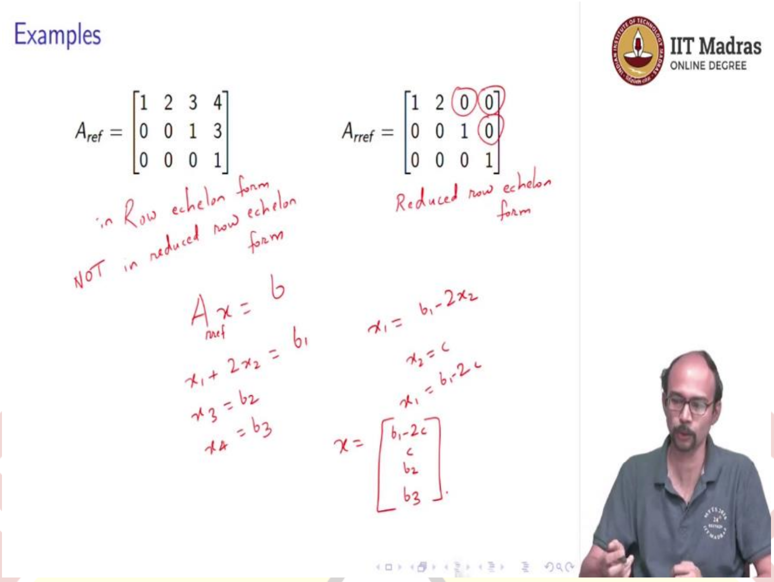
(Refer Slide Time: 10:54)



So, let us do another example. So, here is a matrix, which is in Row echelon form. So, the first in each non-zero n row. So, all 3 rows are non-zero, the first non-zero term is indeed 1 is that the case indeed, that is the case. 1, 1 and 1 and then whenever you have a leading term, so you have this 1 over here, this is in the fourth column, so the 1s which are in the previous columns, so the 1 in the previous column, it is before this it is in the third column. So, that matches up and then the 1 above this is in the first column, so that matches up.

So, your 1s have to sort of go diagonally, but like this not, there cannot be 1 below 1, 1 to the below and left, it has to be below and right. So, this is in reduced echelon form. So, let us ask, this is in Row echelon form? Is it in reduced Row echelon form? So, for it to be in the reduced Row echelon form? For the leading terms, we have to look at the columns and ask, are all the other entries 0? Well, that is certainly not the case over here because here is a non-zero entry, here is a non-zero entry, here is a non-zero entry, so this is not in reduced Row echelon form, this is in Row echelon form. So, let us look at another example.

(Refer Slide Time: 12:33)



And we can do the same analysis for the second example. In fact, these matrices look very similar and there is a way of going from one to the other, which we are going to study in the next video. So, let us look at this matrix over here. Is this in Row echelon form? Well, certainly all the leading terms are 1, all the 1s which are leading terms come to the right of the ones which are in the previous rows and there is no non-zero row, so the third axiom is requirement is satisfied easily.

But this has a further property that not only is it in Row echelon form, it is in the reduced Row echelon form, because now, these entries over here are indeed zeros and that is a requirement for reduced row echelon form. So, this matrix is in row echelon form, this matrix, the first one is in row echelon form. But it is not in reduced row echelon form and this is in reduced row echelon form. And of course, if it is in reduced row echelon form, it is already in row echelon form because reduced is a stronger requirement.

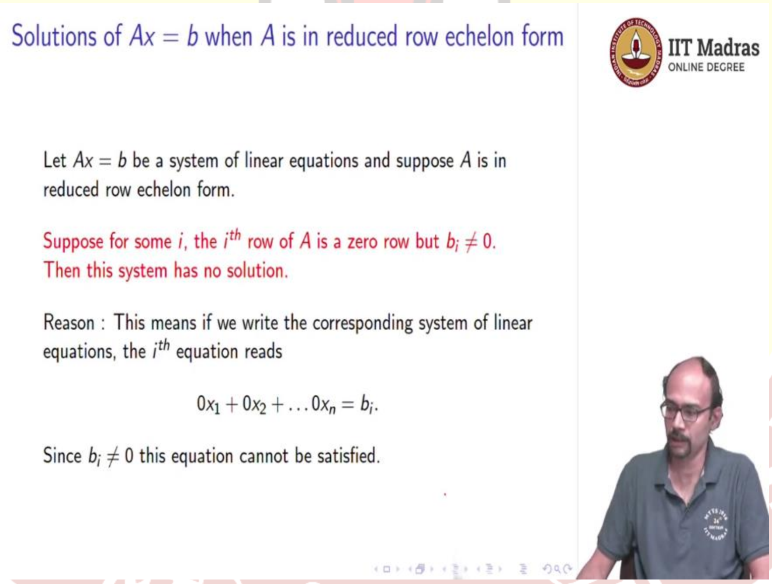
So, let us do the analysis that we did in the example that we saw earlier in this video, so let us ask what is the set of solutions for this? So, if we want, let us say $A_{rref}x = b$, and then $x_1 + 2x_2 = b_1$, $x_3 = b_2$ and then $x_4 = b_3$. So, now I want the set of solutions for this. So, already we know what is x_3 and x_4 are fixed they are b3 and b4 respectively and for x_1, what do I do? Well, I can do what we did before.

So, I can say $x_1 = b_1 - 2x_2$ and then if I choose a value for x_2, automatically, I can get a solution by substituting that value in this equation and getting and substituting x_1 to be that particular number. So, if I take $x_2 = c$, then I can get $x_1 = b_1 - 2c$. So, this gives me a solution.

So, what is the solution? The solution is $x = \begin{bmatrix} b_1 - 2c \\ c \\ b_2 \\ b_3 \end{bmatrix}$, this is my solution for $Ax = b$ and now whatever value of c you substitute, will give you a solution.

So, if you take $c = 0$, for example, then your solution will be $\begin{bmatrix} b_1 \\ 0 \\ b_2 \\ b_3 \end{bmatrix}$. If you take $c = 5$, let us say then your solution will be $\begin{bmatrix} b_1 - 10 \\ 5 \\ b_2 \\ b_3 \end{bmatrix}$. So, you can see that I can read off solutions for these matrices in this way and not just some solutions, we can in fact get read off all solutions. This is the only way we can get solutions for these matrices and this is how, this is the use of the row echelon form or the reduced row echelon form.

(Refer Slide Time: 16:25)



So, solutions of $Ax = b$ when A is in reduced Row echelon form. We did not do one small case in the examples, but I will state it here before I talk about the general situation. So, let $Ax = b$ be a system of linear equations and suppose A is in reduced row echelon form. Suppose for some i, the i'th row of A is a 0 row, but $b_i \neq 0$. So, the i'th row of A is 0, but the corresponding constant $b_i \neq 0$, then this system has no solution.

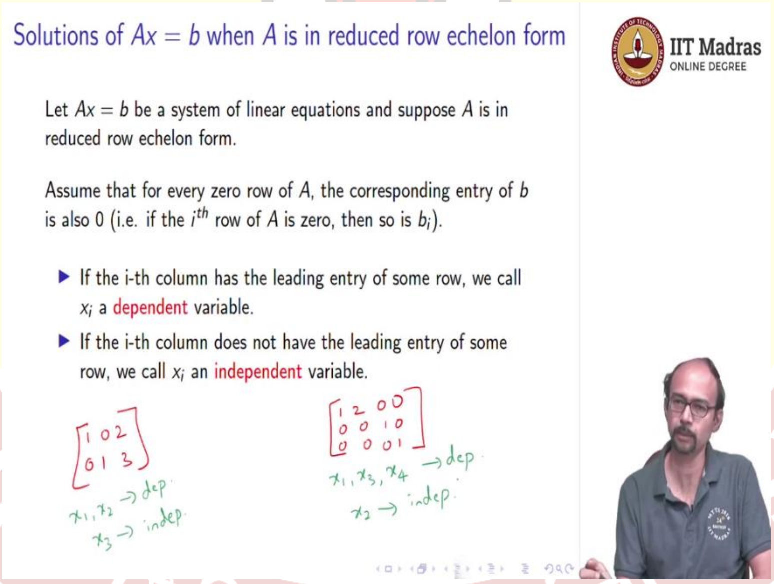
You can think of for a second and if you still cannot see why then here is the reason. So, we can write down the corresponding system of equations and if you write down the system of equations, what will the i'th equation look like, the i'th equation will look like

$$0x_1 + 0x_2 + \cdots + 0x_n = b_i$$

Now the left-hand side, it does not matter what the x_i's are, what values you put inside x_i's, the left-hand side will always evaluate to be 0. So, if $b_i \neq 0$, then there is no hope of getting a solution. Whatever values of x_1, x_2, x_n and you take this equation cannot be satisfied.

So, there is no possible solution for this system of equations. So, the key point is that there if row of $A = 0$, then the corresponding constant $b_i = 0$, in order for there to be a solution, otherwise there cannot be a solution, there will not be a solution, we have seen why? So now, let us look at the case where such a thing does not happen. So, whenever we have 0 rows in A then the corresponding constants $b_i = 0$.

(Refer Slide Time: 18:41)



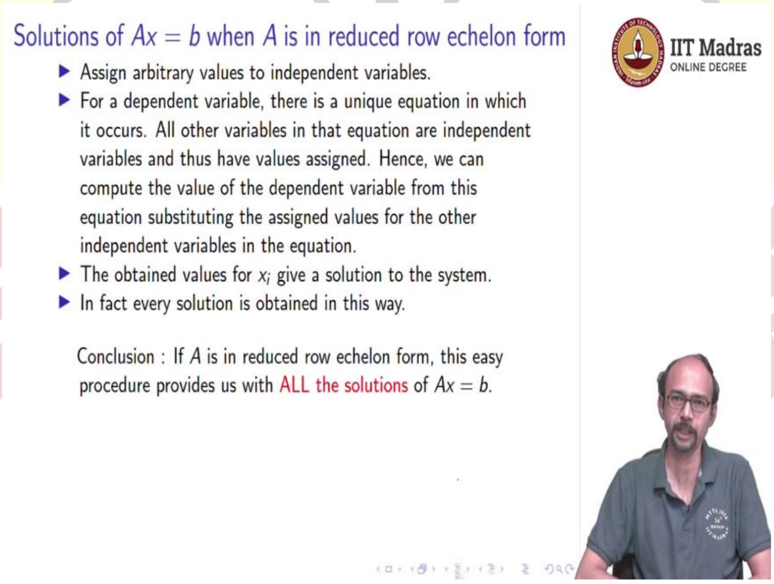
So, solutions again. So, suppose we have $Ax = b$ and A is in the reduced Row echelon form. So, assume that for every zero row of A, the corresponding entry of b is also 0. So, that means if the i'th row of A is 0, then b is also 0, then we can apply the procedure that we saw in the examples that we did before. So, we are going to give some names here, so that we make this procedure very algorithmic.

So, if the i'th column has the leading entry of some row there are going to be some columns, which have leading entries. In our first example of where we had $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{bmatrix}$, the leading entries were in the first column and the second column respectively. So, then we will call x_i as a dependent variable. So, in that first example, since the leading entries were in columns 1 and 2, x_1 and x_2 are going to be dependent variables.

And if the i'th column does not have the leading entry of some row, we call x_i an independent variable. So, all the other variables are independent variables. So, in that first example, the independent variable is x_3. In the second example, the independent variables where x_2. So, let us maybe write that down. So, we had $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{bmatrix}$. So, here x_1 and x_2 are dependent and x_3 is independent. And in the second example, well, we had I think $\begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. So, here are the leading columns the leading entries are in columns 1, 3 and 4 respectively.

So, x_1, x_3, x_4 are dependent variables and x_2 is an independent variable and you saw this reflected in the solutions, why we are calling it independent. So, if you remember how we solve this in the first one, we took x_3 to be, fixed x_3 to be some arbitrary c and then from there we got what is x_1 and x_2. And similarly, in the 2nd example, we fixed x_2 to be some arbitrary c and then x_1, x_3, and x_4 for fixed, so that is what we are going to do.

(Refer Slide Time: 21:28)



Assign arbitrary values to the independent variables and then for a dependent variable, what happens? There is a unique equation in which it occurs. So, if you look at the first example, $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{bmatrix}$, x_1 occurs exactly in 1 equation, it occurs in no other equation. Similarly, for the second example, if you look at x_1 it occurs in only 1 equation, x_3 occurs and only 1 equation, x_4 occurs in only 1 equation.

So, from those equations, we can find out their values. So, all other variables in that equation are independent variables. That is why that is a power of the reduced row echelon form. Because all the other column with a leading 1, the entries, all the other entries are 0, so no dependent variables can occur together. In a single equation, you will have only 1 dependent variable, if at all. So, all the other equations, variable in that equation are independent variables and we have assigned values to them. So, from there, I can move them to the right-hand side and get the value of the dependent variable and then if we put all these together the obtained values for x_i give a solution to the system.

And, in fact, every solution can be obtained in this way, that is what we saw in the 2 examples that we did. So, the conclusion here is, if A is in the reduced row echelon form, this very easy procedure, it is an algorithm. Let us recall what it is, you find out what are the leading ones, the corresponding columns, the column numbers will tell you what are the dependent variables, all the other ones are the independent variables.

For the independent variables, you assign arbitrary values and then calculate the values of the dependent variables from the corresponding equations. So, this procedure provides us with all the solutions of $Ax = b$. Thank you.