Applied Linear Algebra for Signal Processing, Data Analytics and Machine Learning

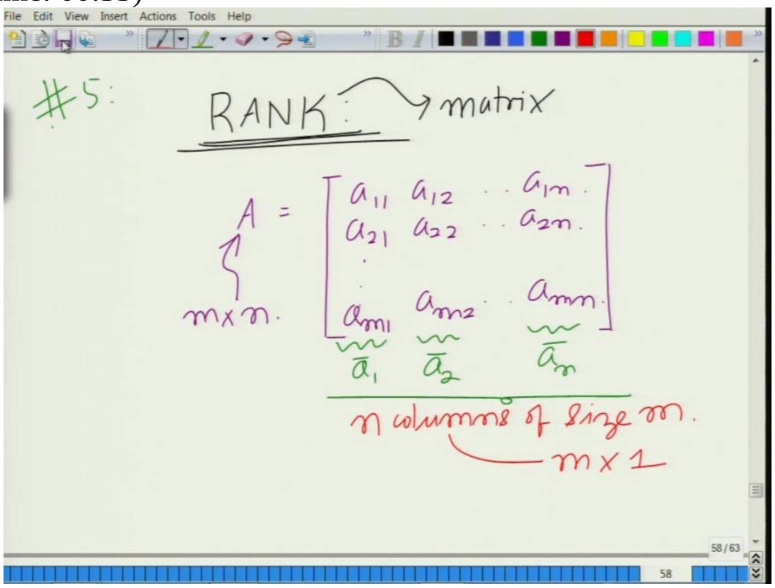
Professor Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture- 5

Matrix: Column Space, Linear Independence, Rank of Matrix, Gaussian elimination

Hello, welcome to another module in this massive open online course. In today's module, let us start our discussion or let us start looking at another very important concept or property of a matrix that is known as the rank of a matrix, alright?

(Refer Slide Time: 00:35)

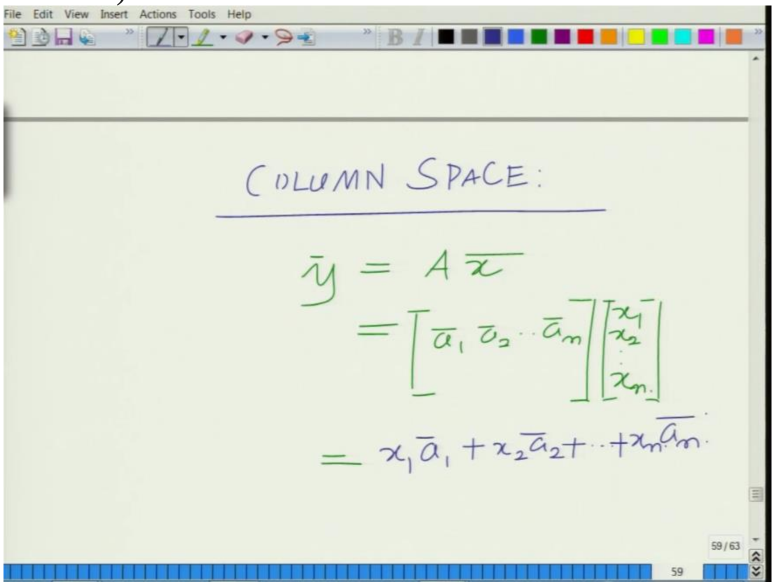


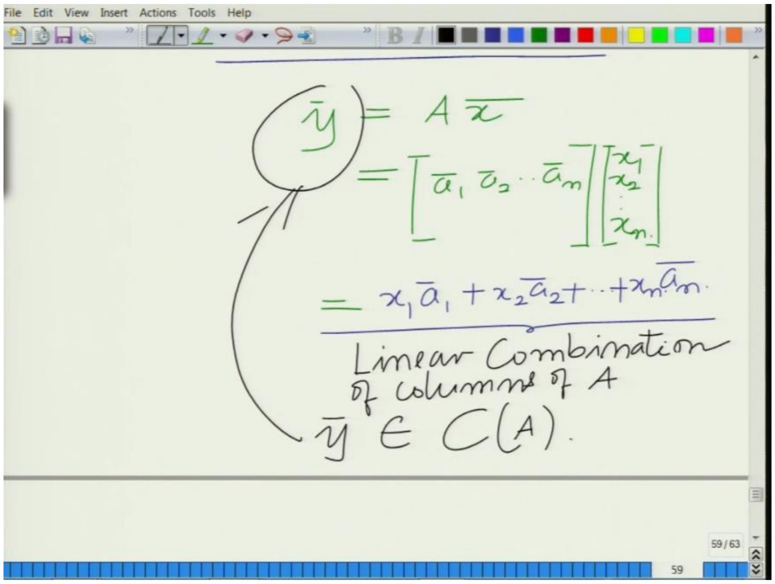
So, let us start our discussion on another very important aspect which is the rank of a matrix. So, in order to answer and this is a very important concept or this notion of rank is a very important concept and is one of the fundamental property. So, let us start this discussion, let us again begin with an $m \times n$ matrix. I hope all of you remember we have this $m \times n$ matrix $\mathbf{A}$ as

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

And this is your $\mathbf{A} = [\bar{\boldsymbol{a}}_1 \quad \bar{\boldsymbol{a}}_2 \quad \ldots \quad \bar{\boldsymbol{a}}_n]$. Let us recall that we can represent matrix $\mathbf{A}$ as a collection of n columns where each column is basically represents a vector of size m where m denotes the number of rows of the matrix $\mathbf{A}$. So, these are n columns of size m or basically you can think of these are column vectors of size $m \times 1$.

(Refer Slide Time: 2:22)





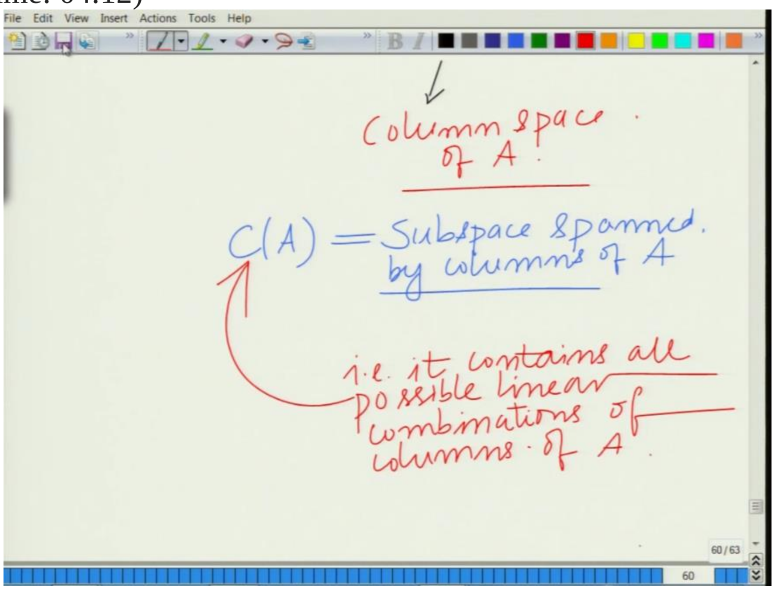
Now, we define this notion of what is known as the column space which is essentially

$$\bar{\mathbf{y}} = \mathbf{A}\bar{\mathbf{x}} = [\bar{\boldsymbol{a}}_1 \quad \bar{\boldsymbol{a}}_2 \quad \ldots \quad \bar{\boldsymbol{a}}_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

This can be now essentially represented as a linear combination of the columns of the matrix $\mathbf{A}$ with the elements of the vector $\bar{\mathbf{x}}$. You can view it as nothing but a linear combination of the columns of the matrix $\mathbf{A}$ as $x_1\bar{\boldsymbol{a}}_1 + x_2\bar{\boldsymbol{a}}_2 + \cdots + x_n\bar{\boldsymbol{a}}_n$.

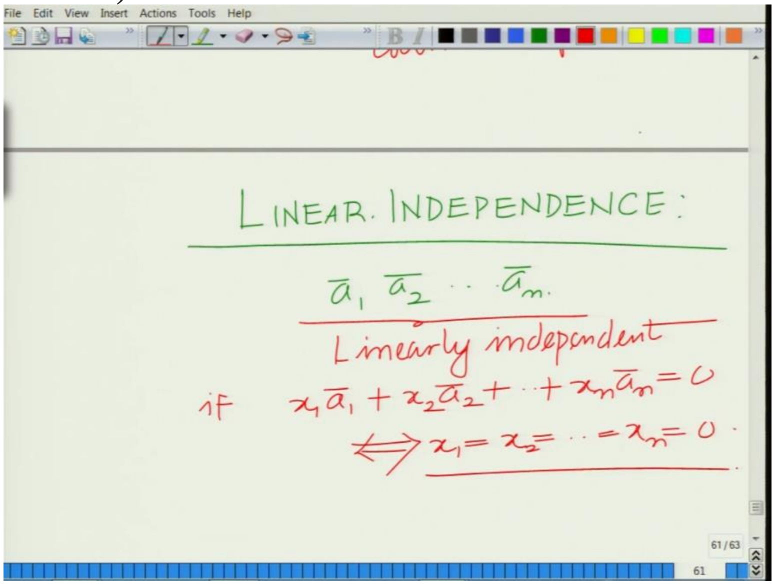
So, this is basically nothing but you can see, this is essentially a linear combination. You have already seen this concept of linear combination. This is a linear combination of columns of $\mathbf{A}$ and we call this vector $\bar{\mathbf{y}} = \mathbf{A}\bar{\mathbf{x}} = x_1\bar{\boldsymbol{a}}_1 + x_2\bar{\boldsymbol{a}}_2 + \cdots + x_n\bar{\boldsymbol{a}}_n$, which is basically a linear combination of columns of **A,** and is belonging to the column space, that is, we say this belongs to the column space of $\mathbf{A}$.

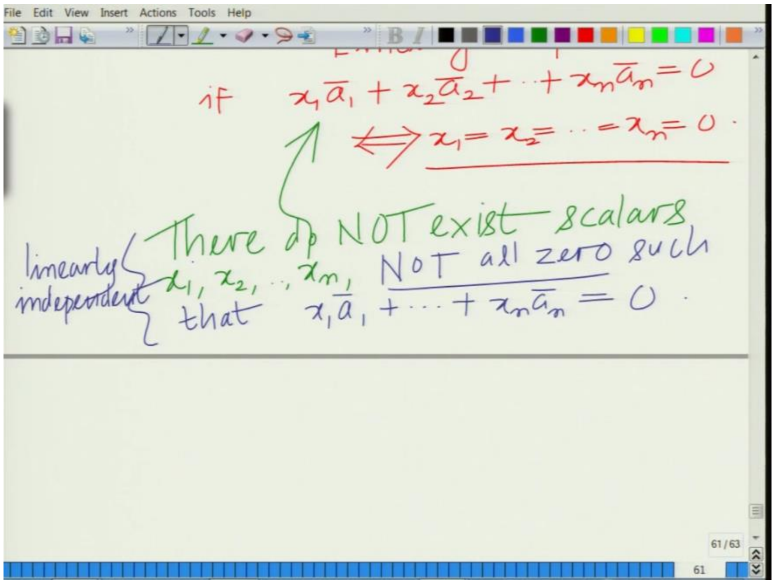
(Refer Slide Time: 04:12)



So, this $C(\mathbf{A})$, this is essentially the column space of the matrix $\mathbf{A}$. This is essentially is a subspace. So, $C(\mathbf{A})$ this is the subspace spanned by columns of $\mathbf{A}$ that is essentially it contains all possible linear combinations of the columns of $\mathbf{A}$.

(Refer Slide Time: 5:55)



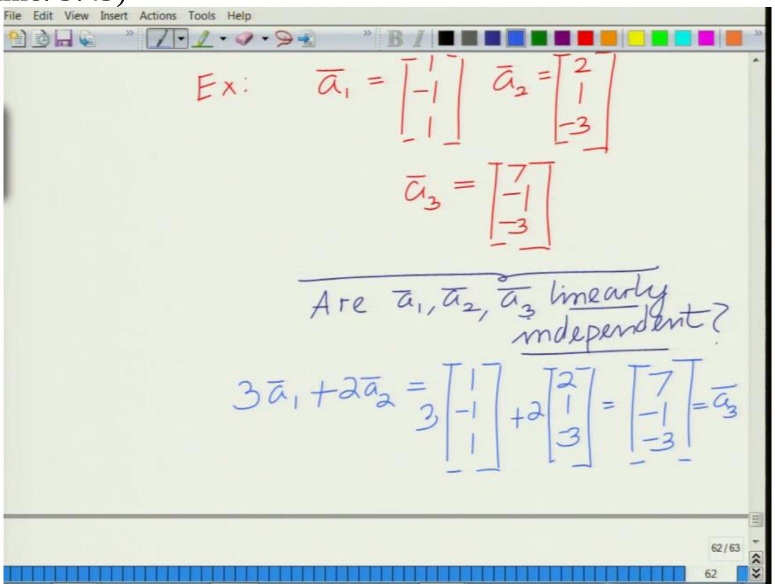


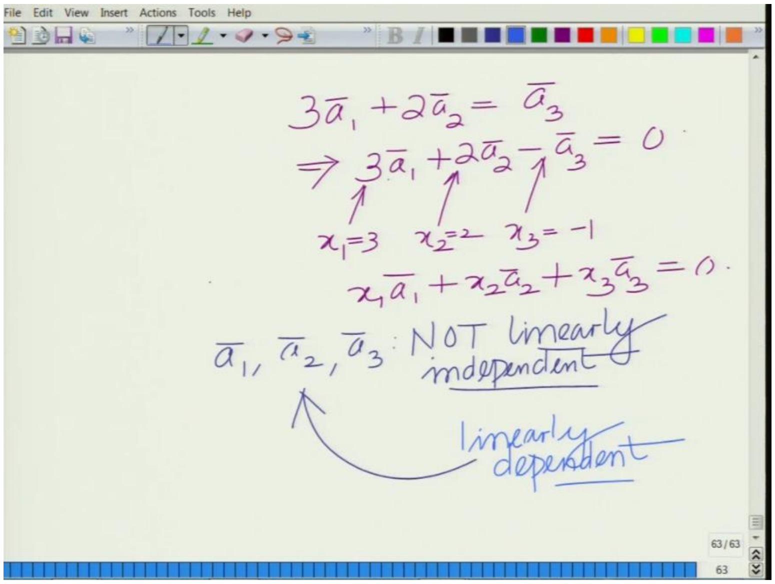
Now, we come to another important concept that is linear independence. What is linear independence? Now linear independence, that is, if we have these vectors $\bar{\boldsymbol{a}}_1, \bar{\boldsymbol{a}}_2, \ldots, \bar{\boldsymbol{a}}_n$, and we say that this set of vectors is linearly independent. This is linearly independent if there is a linear combination $x_1\bar{\boldsymbol{a}}_1 + x_2\bar{\boldsymbol{a}}_2 + \cdots + x_n\bar{\boldsymbol{a}}_n$=0. This is possible if and only if all the x_i's are identical and equal to 0, that is, what it means essentially saying is, that is, any linear combination $x_1\bar{\boldsymbol{a}}_1 + x_2\bar{\boldsymbol{a}}_2 + \cdots + x_n\bar{\boldsymbol{a}}_n$, this is equal to 0, only if each x_i that is $x_1 = x_2 = \cdots = x_n = 0$, and if any of these coefficients is nonzero then the linear combination is not 0. So, essentially it means that the only linear combination of these vectors $\bar{\boldsymbol{a}}_1, \bar{\boldsymbol{a}}_2, \ldots, \bar{\boldsymbol{a}}_n$ that is 0, when all the coefficients are identical and equal to 0.

So, to put the same thing in other words it says, that is, there do not exist scalars $x_1, x_2, \ldots, x_n$ not all 0, this is an important step, not all zeros such that $x_1\bar{\boldsymbol{a}}_1 + x_2\bar{\boldsymbol{a}}_2 + \cdots + x_n\bar{\boldsymbol{a}}_n = 0$.

That is there does not exist any set of any scalars $x_1, x_2, \ldots, x_n$ not all 0 such that $x_1\bar{\boldsymbol{a}}_1 + x_2\bar{\boldsymbol{a}}_2 + \cdots + x_n\bar{\boldsymbol{a}}_n = 0$, that is, the only time this is 0 is when all the scalars are 0. Which means even if one or a few of them are nonzero, then the linear combination is not 0. Such a set of vectors is known as a linearly independent set of vectors. We call this as basically linearly independent. So, this is essentially known as your linearly independent set of vectors.

(Refer Slide Time: 9:49)





And for example, let us take a look at a simple example. Let us take

$$\bar{\boldsymbol{a}}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \bar{\boldsymbol{a}}_2 = \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}, \text{and } \bar{\boldsymbol{a}}_3 = \begin{bmatrix} 7 \\ -1 \\ -3 \end{bmatrix}.$$

Now, let us ask the question are these set of vectors $\bar{\boldsymbol{a}}_1, \bar{\boldsymbol{a}}_2,$ and $\bar{\boldsymbol{a}}_3$, we asked the question that $\bar{\boldsymbol{a}}_1, \bar{\boldsymbol{a}}_2,$ and $\bar{\boldsymbol{a}}_3$ are these linearly independent?

And we will show that they are not linearly independent. In fact, there exists a nonzero linear combination that exists, a linear combination with nonzero combining coefficients such that the linear combination equals 0. That is essentially how we are going to demonstrate that these are not linearly independent.

For instance, if you look at $3\bar{\boldsymbol{a}}_1 + 2\,\bar{\boldsymbol{a}}_2$, which is basically $3\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}$, then you can see, this is equal to essentially

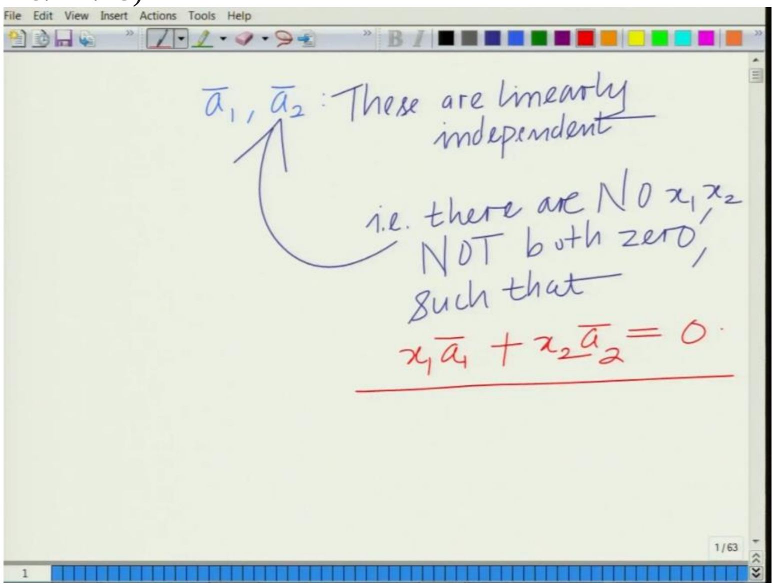
$$3\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix} = \begin{bmatrix} 7 \\ -1 \\ -3 \end{bmatrix} = \bar{\boldsymbol{a}}_3,$$

which is essentially nothing but which is equal to $\bar{\boldsymbol{a}}_3$. So,

$$3\bar{\boldsymbol{a}}_1 + 2\,\bar{\boldsymbol{a}}_2 = \bar{\boldsymbol{a}}_3 \Rightarrow 3\bar{\boldsymbol{a}}_1 + 2\,\bar{\boldsymbol{a}}_2 - \bar{\boldsymbol{a}}_3 = 0.$$

So, if your set $x_1 = 3$, $x_2 = 2$, and $x_3 = -1$. So, we have $x_1\bar{\boldsymbol{a}}_1 + x_2\bar{\boldsymbol{a}}_2 + \cdots + x_n\bar{\boldsymbol{a}}_n = 0$, where not all are 0. In fact, you can see neither is 0, in fact you can say all are not 0 in this case. So, therefore, these are not linearly independent.

(Refer Slide Time: 14:18)

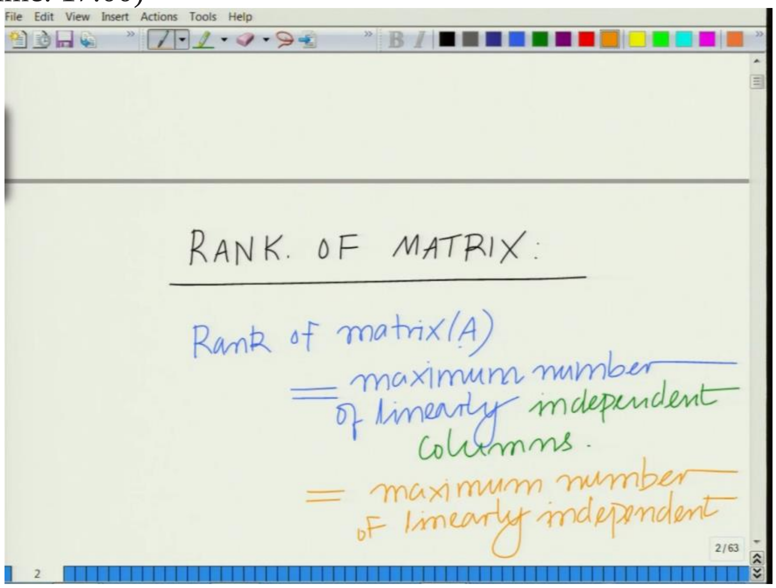


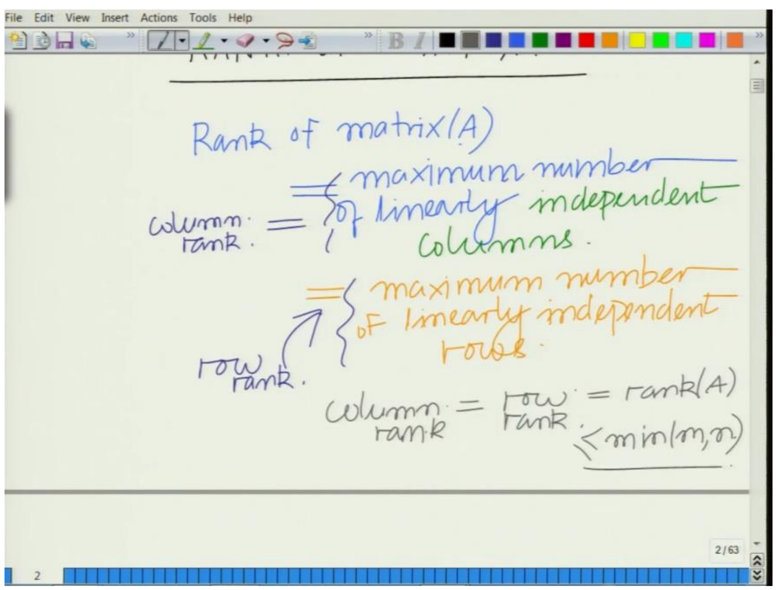
On the other hand, if you take $\bar{\boldsymbol{a}}_1$ and $\bar{\boldsymbol{a}}_2$ now, you can show that these are linearly independent, these are linearly independent. Not very difficult to show that these are linearly independent, that is, there do not exist, there are no scalars x_1, x_2 not both 0 such that $x_1\bar{\boldsymbol{a}}_1 + x_2\bar{\boldsymbol{a}}_2 = 0$. So, that is, there do not exist scalars x_1, x_2 not both 0 so that is at least one of them is nonzero, says that $x_1\bar{\boldsymbol{a}}_1 + x_2\bar{\boldsymbol{a}}_2 = 0$.

Now, in fact how do we systematically determine the rank? And this is why our row and column elimination, this is what we are going to come to. So, for a system, so, in just a little bit of course, we have not illustrated at this point a clear procedure to determine if the columns are linearly dependent or the columns are linearly independent. Previously, we simply showed a linear combination of the column vectors which reduces to 0. In this case, we are saying that there is no linear combination which is equal to 0 with at least one of the coefficients nonzero.

So, how to systematically determine linear dependence or linear independence that is why a row and column elimination is useful, which we are going to see shortly. So, this brings us to the next important concept which is essentially that of the rank which is what we set out to discuss.

(Refer Slide Time: 17:00)





So, now, we come to the notion of the rank of a matrix **A**. This is equal to the maximum number of linearly independent columns which is equal to the and this can be shown, we are not going to explicitly show this, that is, the maximum number of linearly independent columns is also equal to the maximum number of linearly independent rows.

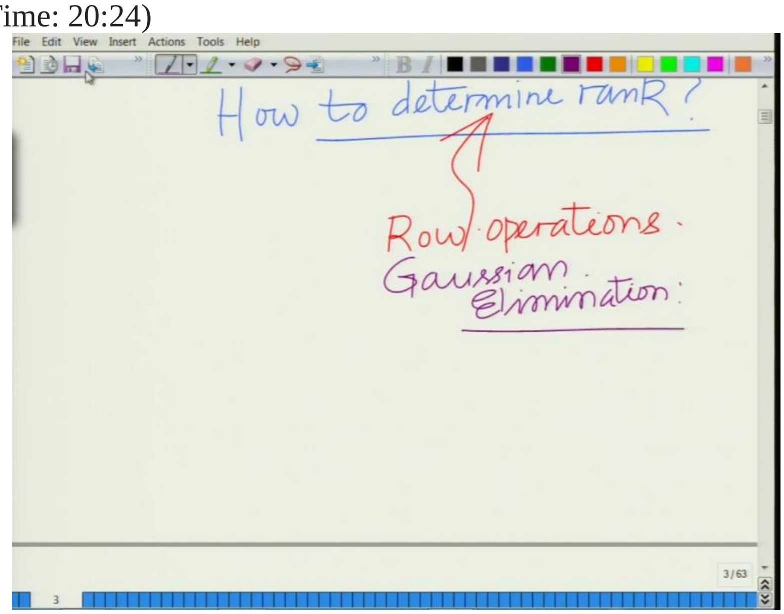
The maximum number of linearly independent columns, this is what we call as the column rank. And the maximum number of linearly independent rows, this is what we call as more specifically the row rank. And we can show that for any matrix **A,** the column rank equals row rank which is what we call as simply the rank of matrix **A**, which is less than or equal to the minimum of the number of linearly independent rows and columns.

So, the rank has to be less than or equal to minimum of the number of rows or columns. So, therefore, it has to be less than both the number of rows as well as the number of columns, the maximum it can achieve is the minimum of the number of rows and columns it cannot be greater than that. The rank of a matrix any matrix **A**, the maximum value of that can be the minimum of the number of rows and columns.

So, it is always less than or equal to the minimum of both the number of rows and columns. And what is the rank? The rank is the minimum number. It is basically the maximum number of linearly independent columns or the maximum number of linearly independent rows and both these are actually the same.

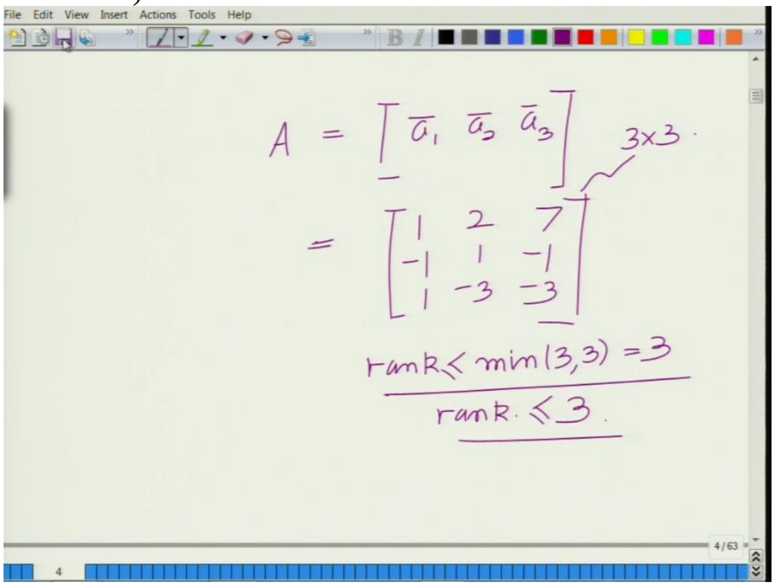
This basically, maximum number of linearly independent columns is the column rank maximum number of linearly independent rows is the row rank. And actually, it can be shown that both these quantities are the same. And this is what we are terming as the rank of the matrix **A**.

(Refer Slide Time: 20:24)



How to find the rank, we now come to this concept of row operations. How to determine rank? This is a very important concept. This is why are the row operations you can simply do column operations to get what is known as the row, so this is basically known as the Gaussian elimination procedure.

(Refer Slide Time: 21:19)



For instance, let us again go back to our matrix **A** . Let us take the three columns that we looked at before, let us put them in, does look at our $\bar{\boldsymbol{a}}_1, \bar{\boldsymbol{a}}_2,$ and $\bar{\boldsymbol{a}}_3$ that we previously looked at, let us put them together as this is the first column as

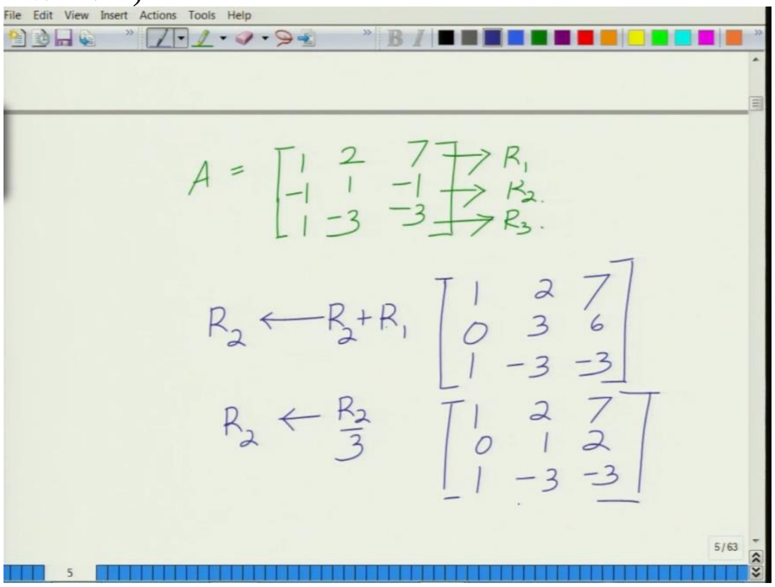
$$\bar{\boldsymbol{a}}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \bar{\boldsymbol{a}}_2 = \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}, \text{and } \bar{\boldsymbol{a}}_3 = \begin{bmatrix} 7 \\ -1 \\ -3 \end{bmatrix}.$$

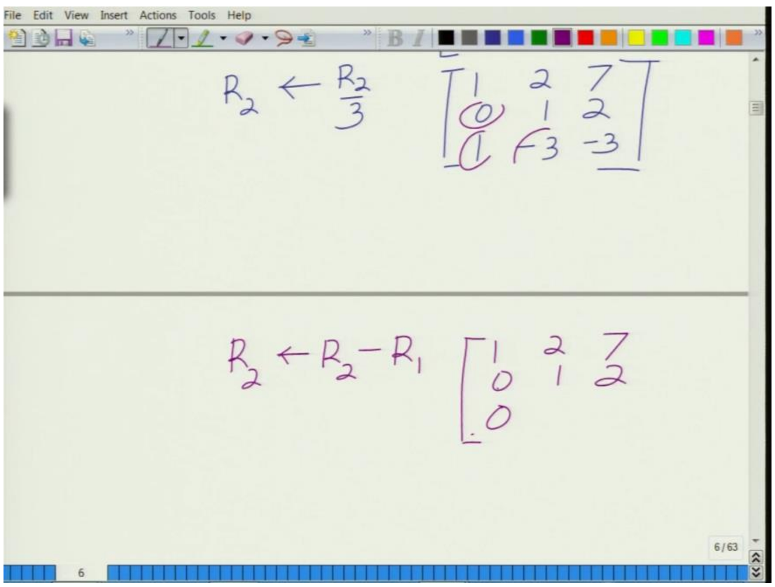
So, these are your three columns. So, this essentially becomes your 3 cross 3 matrix. So the matrix **A** is

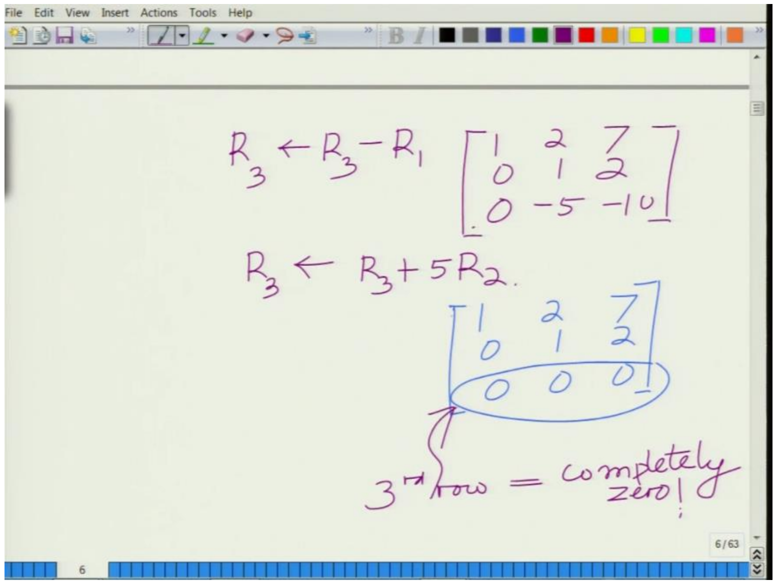
$$\mathbf{A} = [\bar{\boldsymbol{a}}_1\ \bar{\boldsymbol{a}}_2\ \bar{\boldsymbol{a}}_3] = \begin{bmatrix} 1 & 2 & 7 \\ -1 & 1 & -1 \\ 1 & -3 & -3 \end{bmatrix}.$$

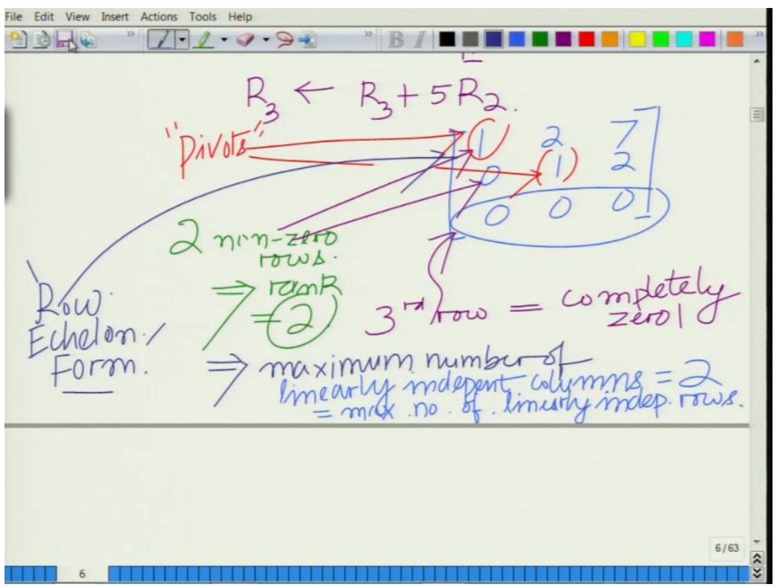
So, you know that the rank has to be less than or equal to minimum of number of rows and number of columns, $rank(A) \leq min(3,3)$, which in this case, essentially 3, so the maximum rank equals 3, so rank can equal to three, but cannot be more than that.

(Refer Slide Time: 22:22)









Now, let us start with the row operations. Now, what we are going to do is we have the matrix

$$\mathbf{A} = [\overline{\boldsymbol{a}}_1 \; \overline{\boldsymbol{a}}_2 \; \overline{\boldsymbol{a}}_3] = \begin{bmatrix} 1 & 2 & 7 \\ -1 & 1 & -1 \\ 1 & -3 & -3 \end{bmatrix} \begin{matrix} \rightarrow R_1 \\ \rightarrow R_2. \\ \rightarrow R_3 \end{matrix}$$

The first operation that we are going to do, is to add R_1 and R_2 and replace R_2 with this, that is, $R_2 \leftarrow R_1 + R_2$. Essentially, what we are going to see is, we have to essentially make the leading entries of each row as 0, so essentially that this gives me, what does this give me. So, you have

$$\begin{bmatrix} 1 & 2 & 7 \\ 0 & 3 & 6 \\ 1 & -3 & -3 \end{bmatrix}.$$

Third row R_3 remains as it is. Now, you replace R_2 by R_2 divide by 3, that is, $R_2 \leftarrow \frac{R_2}{3}$. So this becomes

$$\begin{bmatrix} 1 & 2 & 7 \\ 0 & 1 & 2 \\ 1 & -3 & -3 \end{bmatrix}$$

Now, we are essentially going to replace, now this leading entry in R_2 is reduced to 0. Now, the next row we can target to reduce the leading entry of R_3 to 0. So we can do $R_3 \leftarrow R_3 - R_1$.

And that essentially means that you are going to have

$$\begin{bmatrix} 1 & 2 & 7 \\ 0 & 1 & 2 \\ 0 & -5 & -10 \end{bmatrix},$$

and now we do $R_3 \leftarrow R_3 + 5R_2$ which is essentially becomes

$$\begin{bmatrix} 1 & 2 & 7 \\ 0 & 1 & 2 \\ 0 & -5 & -10 \end{bmatrix},$$

and you can clearly see the entire third row R_3 becomes 0. So, the third row is completely 0 and therefore, you have two nonzero rows essentially, plus a rank equal to 2. Or essentially and we had seen this before that in fact, there are 3 columns only two of them are linearly independent. So, maximum number of linearly independent columns, which will also be equal to maximum number of linearly independent rows, this is equal to 2. We had seen that earlier.

So, let me write it clearly, maximum number of linearly independent columns which is equal to maximum number of linearly independent rows and in fact, these leading nonzero elements of each nonzero row these have an interesting name these are what are known as the pivots.

So, you essentially have 2 pivots you can also see it in terms of the pivots. So, you have 2 nonzero pivots essentially the rank of this matrix is 2, maximum number of linearly independent columns is 2, maximum number of linearly independent rows is 2 and when you do the row elimination, you can see that the number of nonzero rows is also 2.

Similarly, you can go to column elimination also, although row elimination is what is usually preferred and this is known as the Row Echelon form, the reduced row echelon form. This is known as the Row Echelon form of the matrix which you get via Gaussian elimination with the pivots.

So, from the Row Echelon form, the number of nonzero pivots essentially gives you the rank of the matrix which is a very important property as we go into subsequent modules. So, let us stop our discussion here. We will continue our discussion on matrices. In fact, as I told you matrices form a very important component.

In fact, linear algebra is sometimes also known as matrix algebra. So, our discussion of matrices and various types of matrices, the properties and the various manipulations, the use for the various uses of these is going to form essentially the bulk of this course. So, let us stop here and we will continue our discussion in the next month. Thank you very much.