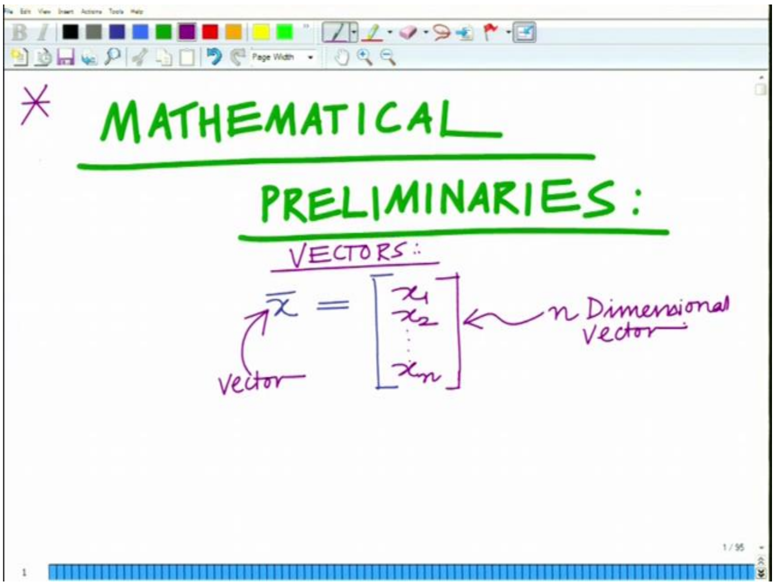
Applied Optimization for Wireless, Machine Learning, Big Data
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture – 01
Vectors and Matrices – Linear Independence and Rank

Hello. Welcome to this module in this massive open online course. So, let us start with the mathematical preliminaries that are required to understand the framework of optimization that is which form the basis of building the framework for optimization, the various tools and techniques for optimization.

(Refer Slide Time: 00:37)

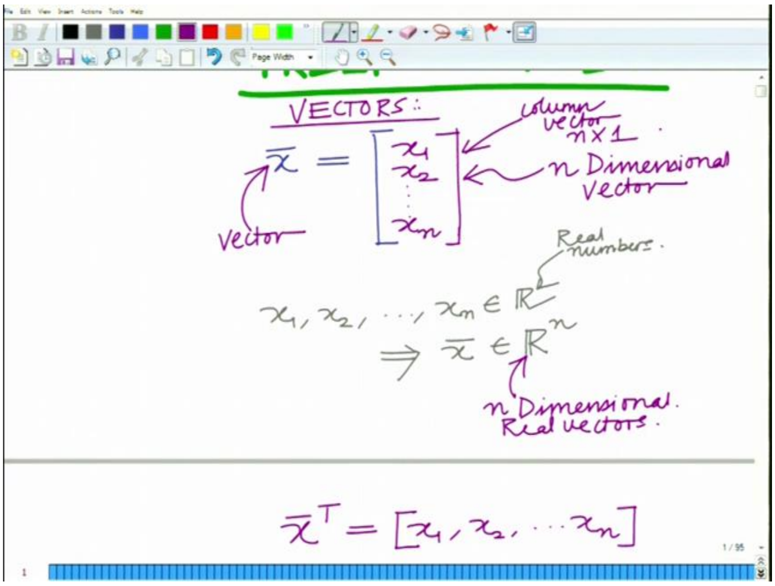


So, we want to start with the mathematical preliminaries, the notation and so on that we are going to use frequently in our treatment of optimization in order to illustrate or in order to basically describe the various concepts of optimization.

Now, the first thing that we are going to use is that as you must all be familiar of vector $\bar{x}$ which is denoted by a bar on the top of the quantity. Let us start with the concept of vectors and a vector is denoted by the quantities like the bar on the top. So, vector $\bar{x}$ is an n-dimensional object which contains n components as follows.

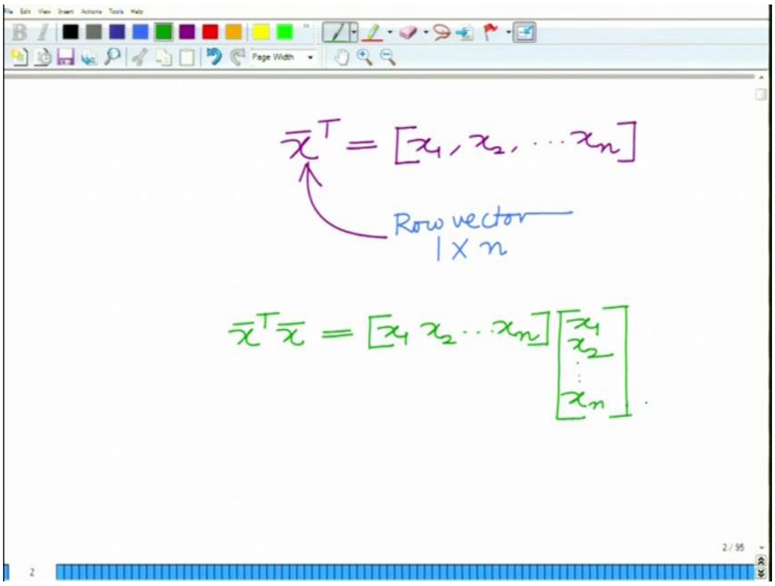
$$\overline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

(Refer Slide Time: 02:28)



Where elements $x_1, x_2, \ldots x_n \in \Re$ where $\Re$ is the set of real numbers, which means $\overline{x}$ belongs to the set of n-dimensional real vectors. So, this is a space of n-dimensional real vectors.

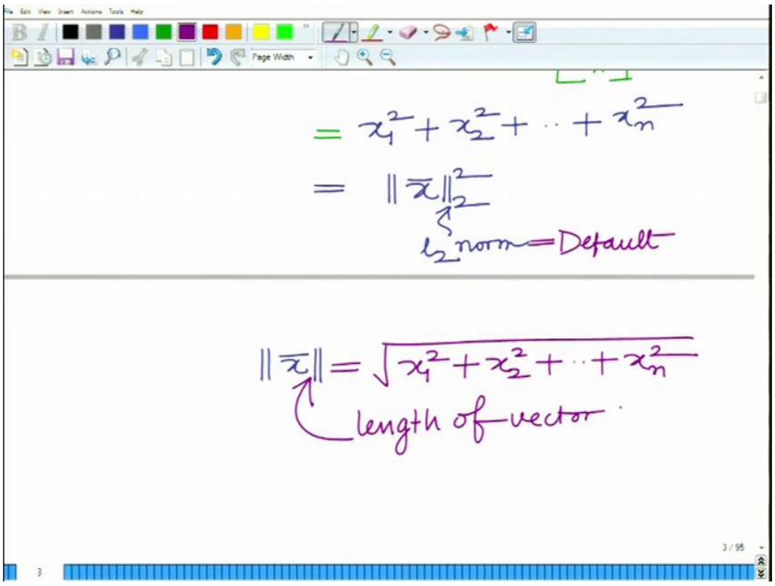
(Refer Slide Time: 03:16)



And, if we consider $\bar{x}$ is the column vector and therefore, transpose of this vector i.e. $\bar{x}^T$ will similarly be a row vector. So, $\bar{x}$ is a column vector which has dimension $n \times 1$. Similarly, $\bar{x}^T$ is a row vector which is of dimension $1 \times n$ i.e. 1 row and n columns. Further, the product of vectors $\bar{x}^T$ and $\bar{x}$ is

$$\bar{x}^T \bar{x} = [x_1, x_2, \ldots x_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$
$$= x_1^2 + x_2^2 + \ldots + x_n^2$$

(Refer Slide Time: 04:29)



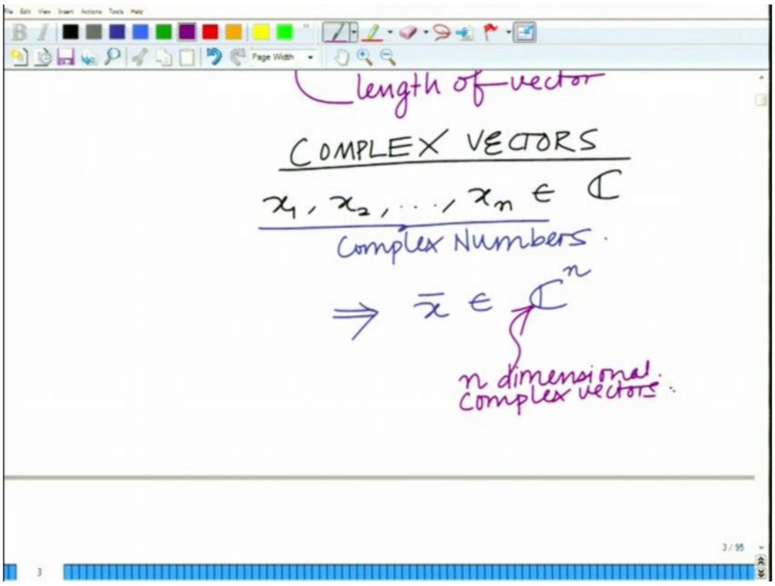
Here we can see, $\bar{x}^T \bar{x}$ is basically equal to the summation of x_1^2, x_2^2, …, x_n^2 which is also denoted by the norm square $\|\bar{x}\|_2^2$. In fact, this is a specific case of a norm, the l_2 norm $\|\bar{x}\|_2^2$. $\|\bar{x}\|_2$ is the norm of $\bar{x}$. Here subscript 2 indicates the l_2 norm of the vector. l_2 norm is the default so $\|\bar{x}\|_2^2$ can be written as $\|\bar{x}\|^2$. Therefore, if the norm is not explicitly specified, it indicates the l_2 norm.

l_2 norm of a vector is basically the length of a vector in n-dimensional space. So, $\|\bar{x}\|$ is simply something that most of you might be very familiar with that

$$\|\bar{x}\| = \sqrt{x_1^{\ 2} + x_2^{\ 2} + \cdots + x_n^{\ 2}}$$

which is basically the length of the vector $\bar{x}$.

(Refer Slide Time: 06:10)

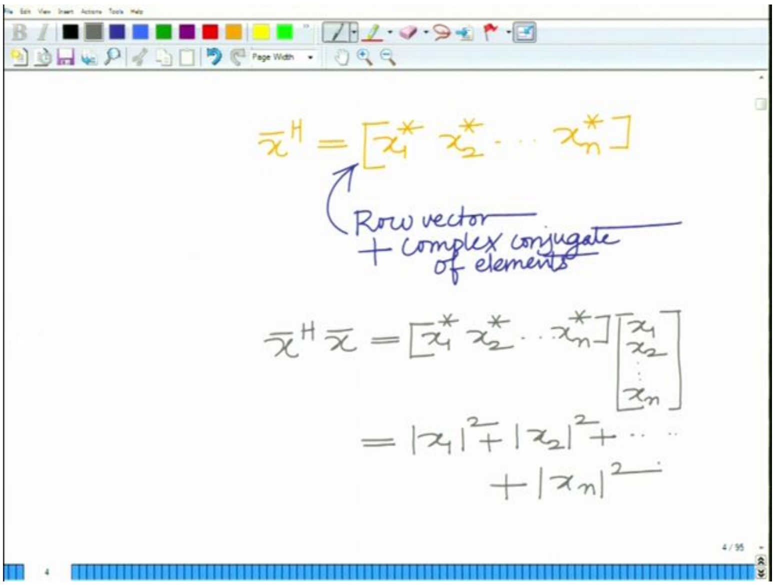


So till now we have Real vector. Similarly, if, on the other hand,

$$x_1, x_2, \ldots x_n \in \mathbb{C}$$

where $\mathbb{C}$ is the set of complex numbers. So, now, we want to see the notion of a Complex vector. So, a complex vector with elements $x_1, x_2, \ldots x_n$ belong to $\mathbb{C}$ implies that the vector $\bar{x} \in \mathbb{C}^n$ that is, $\bar{x}$ has the space of n dimensional complex vectors.

(Refer Slide Time: 07:18)



Now, Hermitian of vector $\bar{x}$ that is $\bar{x}^H$ is basically equal to

$$\bar{x}^H = \left[x_1^*, x_2^*, \ldots, x_n^* \right]$$

Now, in this case when taking the Hermitian of a vector, the column vector becomes a row vector and its elements are the complex conjugate of each corresponding elements of the vector $\bar{x}$ that is $x_1^*, x_2^*, \ldots x_n^*$. Now, the product of vectors $\bar{x}^H$ and $\bar{x}$ is

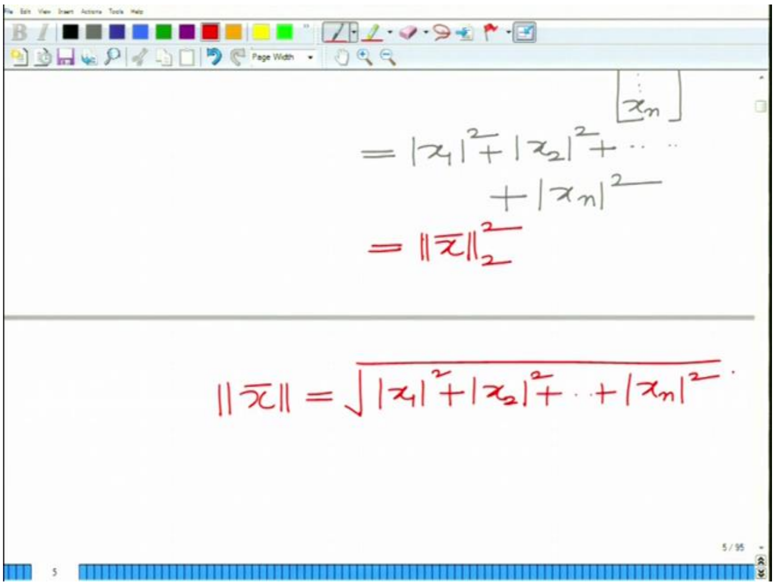
$$\bar{x}^H \bar{x} = \left[x_1^*, x_2^*, \ldots x_n^* \right] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$
$$= |x_1|^2 + |x_2|^2 + \cdots + |x_n|^2$$
$$= \|\bar{x}\|_2^2$$

As the summation of squares of the magnitudes is

$$|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2 = \|\bar{x}\|^2$$

which is equal to the norm square or in fact, the l_2 norm of $\|\bar{x}\|^2$.

(Refer Slide Time: 09:09)

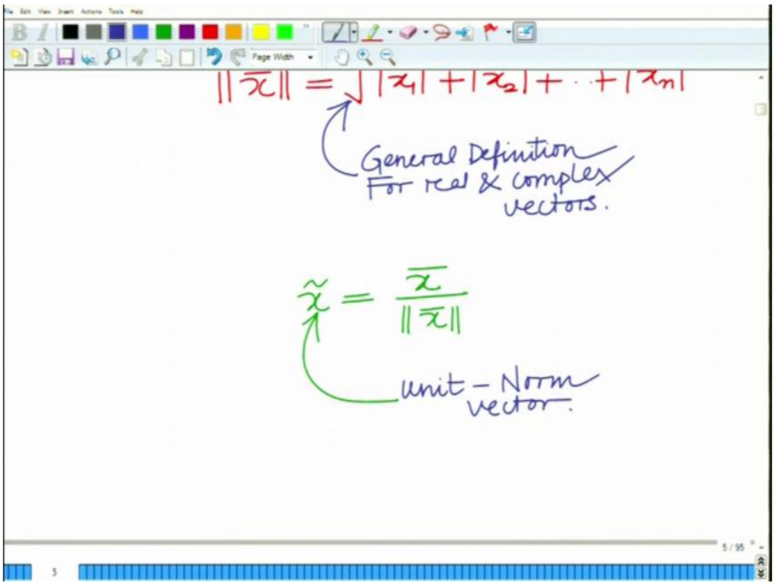


And therefore, once again now you see that, in this case, the norm of a complex vector $\bar{x}$ is

$$\|\bar{x}\| = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2}$$

Here we have replaced x_i^2 with $|x_i|^2$ which is the general definition of a norm. Therefore, It works for both the real as well as complex vectors.

(Refer Slide Time: 10:07)



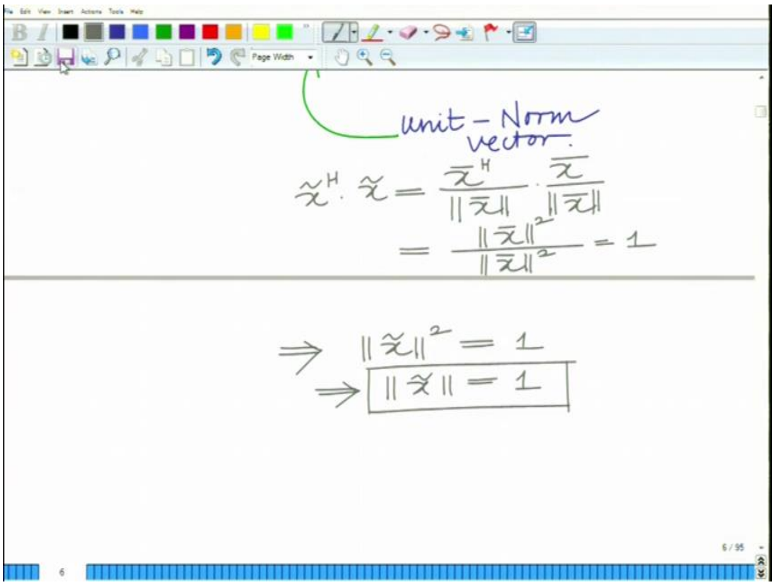
Hence in case of the real vector, one can simply replace the magnitude squares by the squares of the elements.

Now, a special kind of a vector $\tilde{x}$ is obtained as following

$$\tilde{x} = \frac{\bar{x}}{\|\bar{x}\|}$$

Here we are taking the vector $\bar{x}$ and dividing it by its norm and that gives a unit norm vector. So, vector $\tilde{x}$ is basically unit norm vector because the norm of $\tilde{x}$ is unity which is an interesting property.

(Refer Slide Time: 11:27)



In fact, if you look at $\tilde{x}^H \tilde{x}$; that is

$$\begin{aligned} \tilde{x}^H \tilde{x} &= \frac{\bar{x}^H}{\|\bar{x}\|} \cdot \frac{\bar{x}}{\|\bar{x}\|} \\ &= \frac{\|\bar{x}\|^2}{\|\bar{x}\|^2} = 1 \end{aligned}$$

And we know that $\tilde{x}^H \tilde{x} = \|\tilde{x}\|^2$

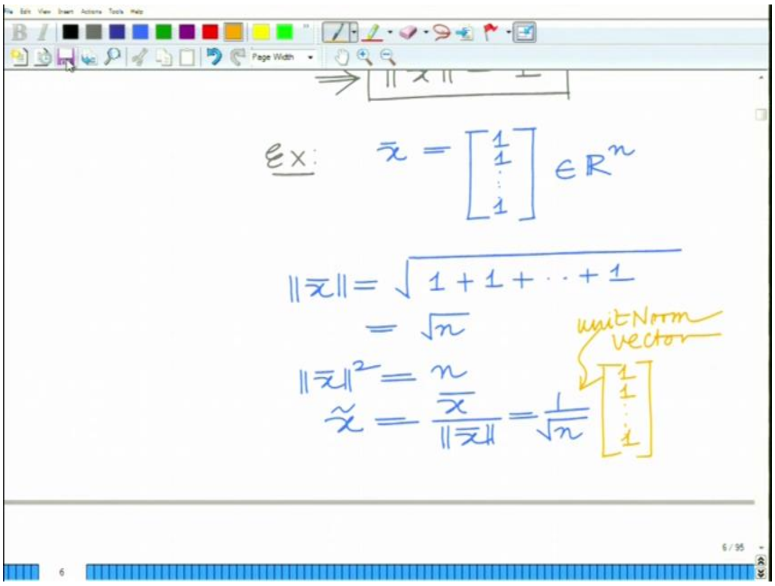
So, this implies now that

$$\|\tilde{x}\|^2 = 1$$
$$\|\tilde{x}\| = 1$$

So, $\tilde{x}$ is basically a unit norm vector in the direction of $\bar{x}$. Therefore, if a n-dimensional vector $\bar{x}$ is representing a particular direction in n-dimensional space, the unit norm vector can be thought of as a unit vector basically pointing in the same direction in n-dimensional space. So, $\bar{x}$ and $\tilde{x}$, both are a line except that $\tilde{x}$ is a unit norm vector that is it has norm equal to unity.

Let us take a simple example to understand this.

(Refer Slide Time: 12:49)



For instance, let's consider the vector $\bar{x}$, equals to the n-dimensional all 1 vector that is

$$\bar{x} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \in \Re^n$$

Then, we have norm of $\bar{x}$ equals to

$$\begin{aligned} \|\bar{x}\| &= \sqrt{1+1+\cdots+1} \\ &= \sqrt{n} \end{aligned}$$

Therefore,

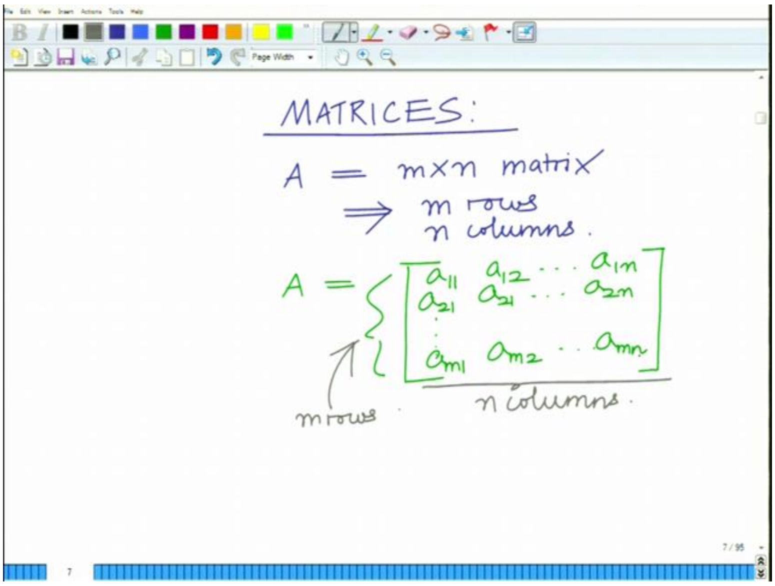
$$\|\bar{x}\|^2 = n$$

And in fact, $\tilde{x} = \frac{\bar{x}}{\|\bar{x}\|} = \frac{1}{\sqrt{n}}\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$

So, this is basically the corresponding unit norm vector.

So, that completes a brief summary of the properties of the various properties of vectors. Now we will quickly refresh some aspects of the matrices.

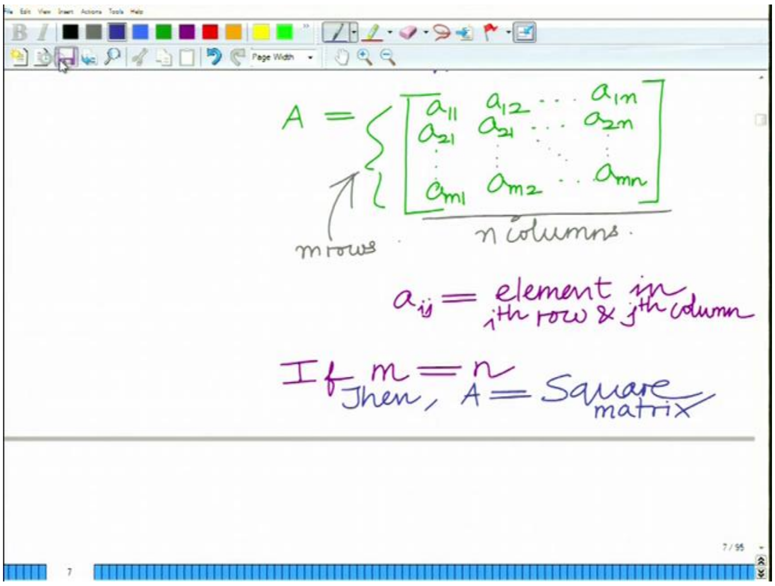
(Refer Slide Time: 14:39)



Once again, a brief review of various concepts in linear algebra and matrices. So, let us consider $m \times n$ matrix A. This implies A has m rows and n columns and you can represent A as

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

(Refer Slide Time: 16:05)

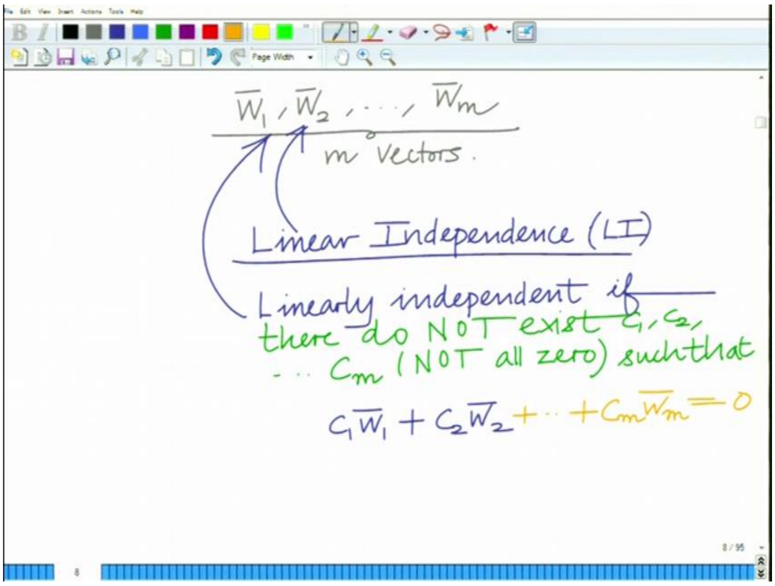


Here, ij^{th} element a_{ij} equals to the element in i^{th} row and the j^{th} column of matrix A.

And, when the number of rows is equal to number of columns that is $m = n$ then A= square matrix.

Let us now look at an important concept of the row space and column space. To first understand this concept of a row space and column space of a matrix, we have to understand what we mean by the space and what we mean by the rank of a set of vectors.

(Refer Slide Time: 17:36)



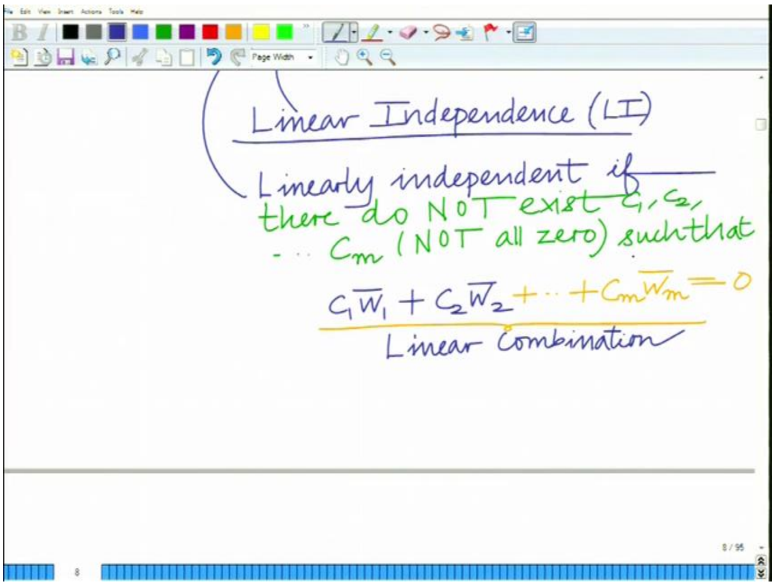
So, let's start with this notion of rank. Consider vectors $\bar{W}$ such that,

$$\bar{W} = \bar{W}_1, \bar{W}_2, \cdots, \bar{W}_m$$

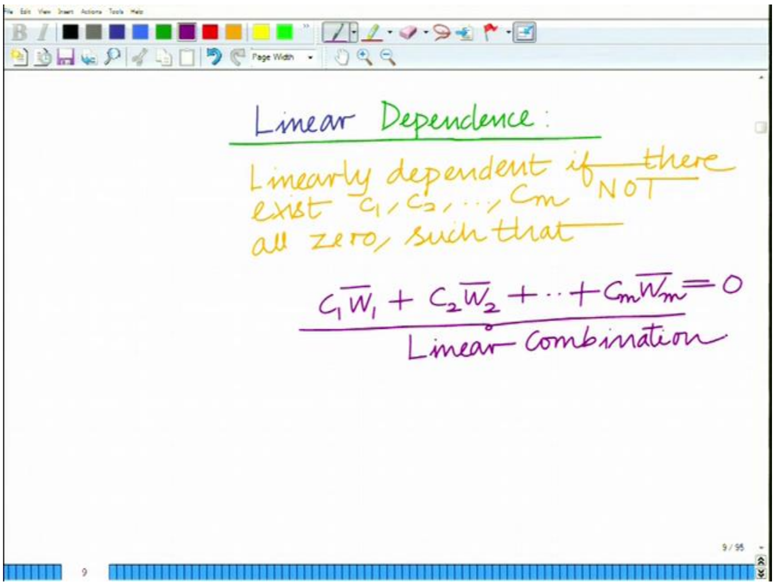
This is a set of m vectors considering all vectors are linearly independent. So, these vectors are linearly independent, only if there do not exist $C_1, C_2, \cdots C_m$, such that all of them cannot be 0. So, in other words, for linear independent $\bar{W}$, there cannot be set of constant $C_1, C_2, \cdots C_m$ such that

$$C_1\bar{W}_1 + C_2\bar{W}_2 + \cdots + C_m\bar{W}_m = 0$$

(Refer Slide Time: 20:01)

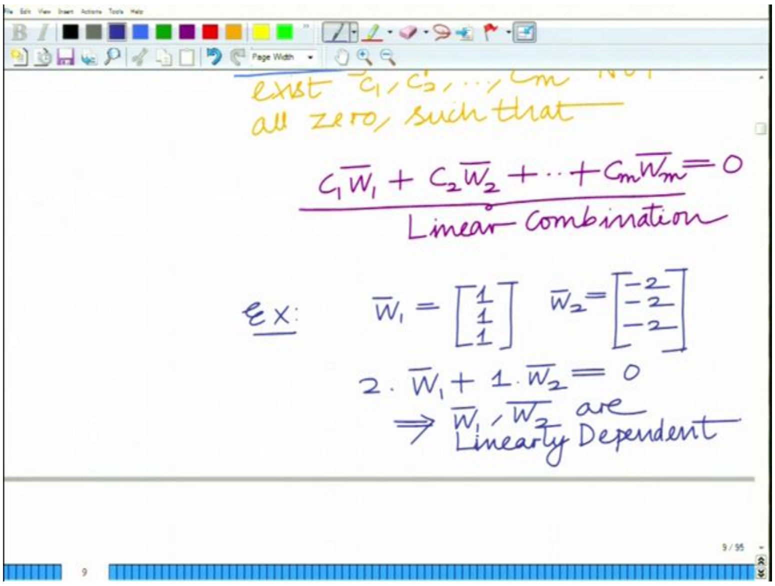


(Refer Slide Time: 20:40)



So, if there exists these weights $C_1, C_2, \cdots C_m$ such that not all of them are 0 and the linear combination of the vectors $\overline{W}_1, \overline{W}_2, \cdots, \overline{W}_m$ is 0, then these vectors $\overline{W}_1, \overline{W}_2, \cdots, \overline{W}_m$ are linearly dependent. For instance, let us take a very simple example to understand this.

(Refer Slide Time: 22:49)



Consider the vectors $\overline{W}_1$ and $\overline{W}_2$ equals to

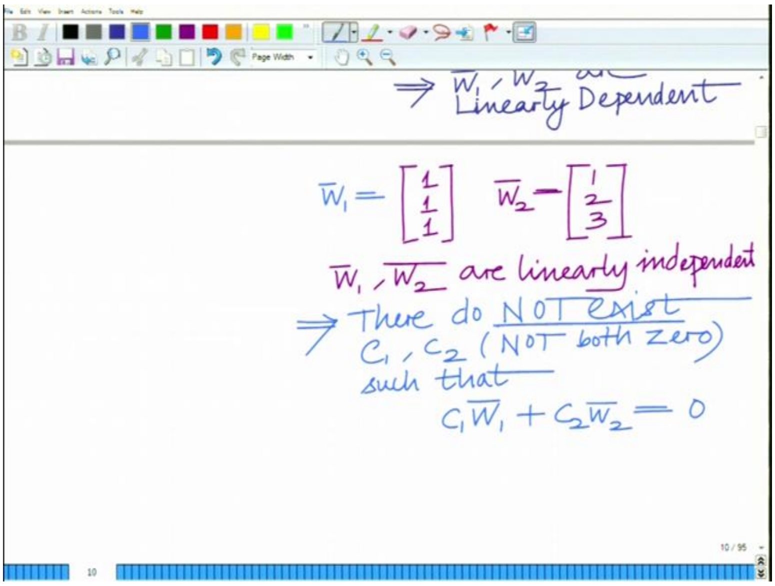
$$\overline{W}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ and } \overline{W}_2 = \begin{bmatrix} -2 \\ -2 \\ -2 \end{bmatrix}$$

then clearly we can write

$$2 \cdot \overline{W}_1 + 1 \cdot \overline{W}_2 = 0$$

So, clearly we can infer that $\overline{W}_1$ and $\overline{W}_2$ are linearly dependent.

(Refer Slide Time: 23:44)



On the other hand, we consider another example.

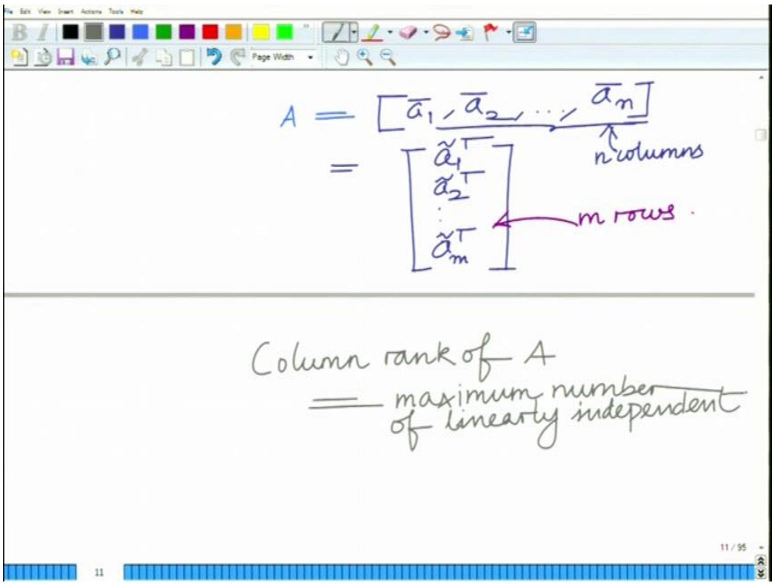
$$\overline{W}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ and } \overline{W}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Now, we can see that there do not exist any constant C_1 and C_2 where both of them are not 0 (but one of them can be 0) such that $C_1\overline{W}_1 + C_2\overline{W}_2 = 0$. So we can quickly verify that $\overline{W}_1$ and $\overline{W}_2$ are linearly independent.

So, basically this is the concept of linear dependence and linear independence of a set of vectors.

Now, we can reduce the concept of linear independence to define the rank of the matrix A.

(Refer Slide Time: 25:34)

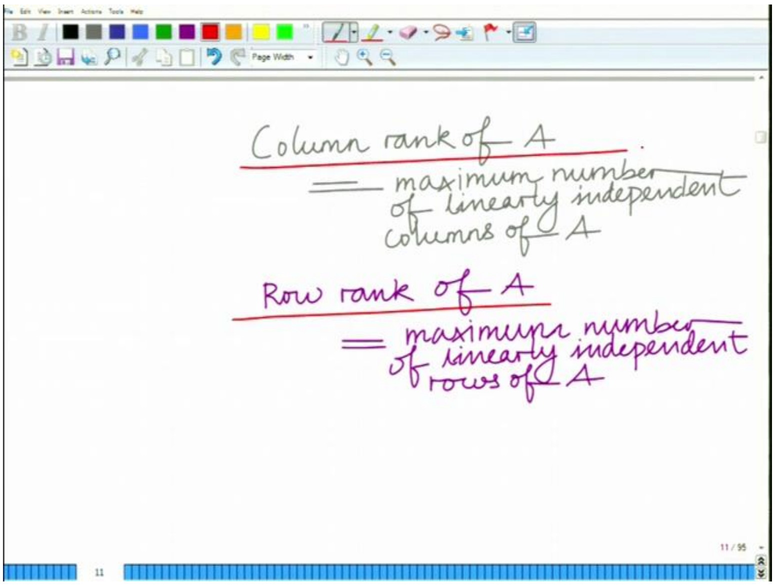


To define the rank of a matrix, let us consider matrix A as

$$A = [\bar{a}_1, \bar{a}_2, \cdots, \bar{a}_n]$$
$$= \begin{bmatrix} \tilde{a}_1{}^T \\ \tilde{a}_2{}^T \\ \vdots \\ \tilde{a}_n{}^T \end{bmatrix}$$

Here matrix A is defined both as a row matrix and also as a column matrix. Now column rank of A equals the maximum number of linearly independent columns that is $\bar{a}_1, \bar{a}_2, \cdots, \bar{a}_n$. This means that the maximum number of columns that we can choose from A such that there does not exist any linear combination which is 0 is the column rank of matrix A.

(Refer Slide Time: 27:24)

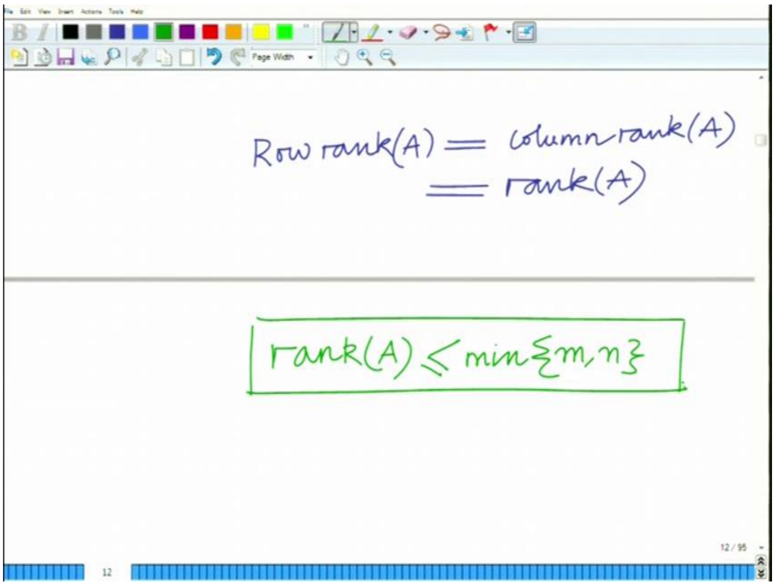


Similarly the Row Rank of a matrix A is the maximum number of linearly independent rows of matrix A.

Now, one of the fundamental results in linear algebra or matrix theory is that the row rank of any matrix equals the column rank and this quantity is simply denoted by rank of the matrix. That is;

Row rank(A) = Column rank(A) = **Rank(A)**

(Refer Slide Time: 28:45)

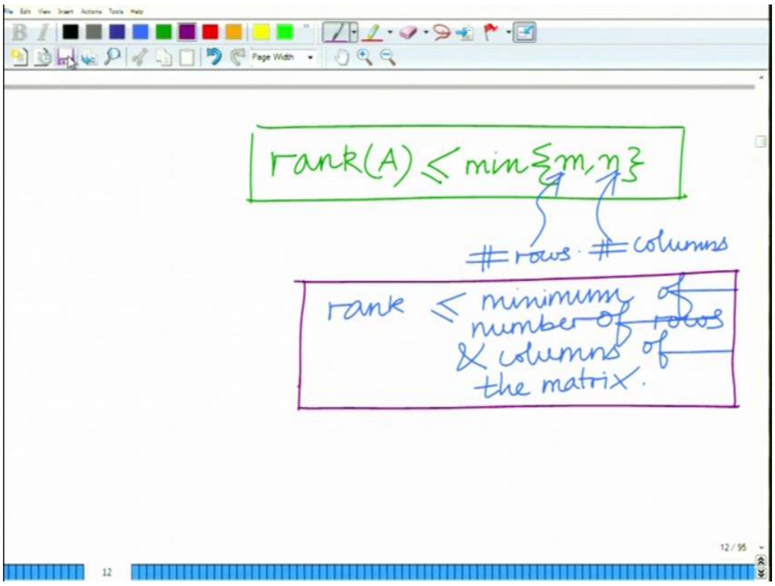


And in addition to this, rank of a matrix satisfies the property that

$$rank(A) \leq \min\{m, n\}$$

Where m and n are the number of rows and columns of matrix A.

(Refer Slide Time: 29:48)



So, on summarizing these properties, we see that the rank of any matrix is less than or equal to minimum of the number of rows and columns of the matrix and this is the fundamental property of the matrix. Also this rank has to be less than or equal to the minimum of the number of rows and columns of the matrix.

So, we have covered some of the mathematical preliminaries required to develop the various tools and techniques for optimization. We will continue this discussion in the subsequent modules.

Thank you very much.