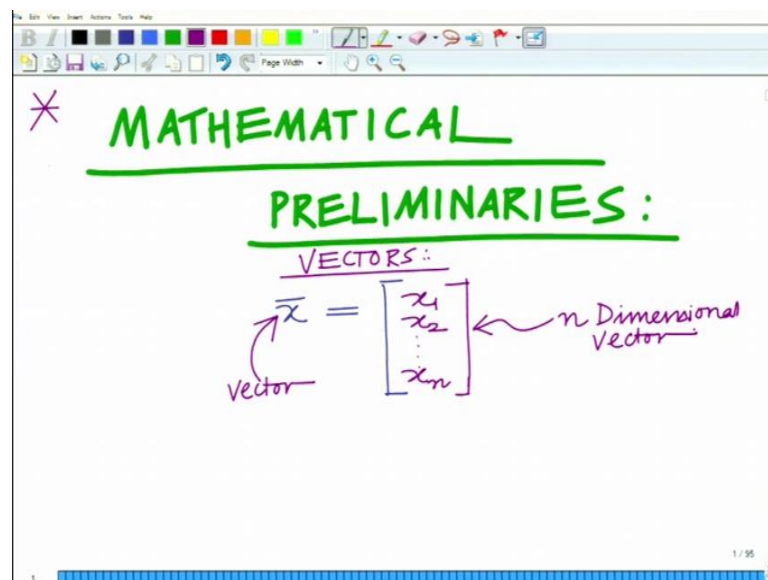


**Applied Optimization for Wireless, Machine Learning, Big Data**  
**Prof. Aditya K. Jagannatham**  
**Department of Electrical Engineering**  
**Indian Institute of Technology, Kanpur**

**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  $\vec{x}$  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  $\vec{x}$  is an n-dimensional object which contains n components as follows.

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

(Refer Slide Time: 02:28)

**VECTORS:**

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

Annotations:

- column vector  $n \times 1$
- $n$  Dimensional Vector
- Real numbers.
- $x_1, x_2, \dots, x_n \in \mathbb{R}$
- $\Rightarrow \bar{x} \in \mathbb{R}^n$
- $n$  Dimensional Real vectors.

$\bar{x}^T = [x_1, x_2, \dots, x_n]$

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

(Refer Slide Time: 03:16)

$\bar{x}^T = [x_1, x_2, \dots, x_n]$

Annotations:

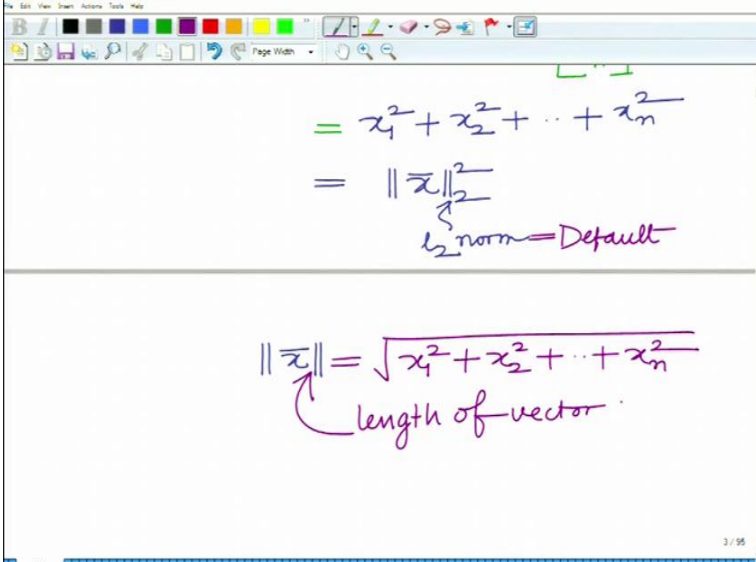
- Row vector  $1 \times n$

$\bar{x}^T \bar{x} = [x_1 \ x_2 \ \dots \ x_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

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

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

(Refer Slide Time: 04:29)



The image shows a whiteboard with handwritten mathematical expressions. The top part shows the dot product of a vector  $\bar{x}$  with itself, resulting in the sum of squares of its components, which is then equated to the square of the L2 norm of the vector. A note indicates that the L2 norm is the default norm. The bottom part shows the L2 norm of the vector as the square root of the sum of squares of its components, with a note indicating this is the length of the vector.

$$\begin{aligned}&= x_1^2 + x_2^2 + \dots + x_n^2 \\ &= \|\bar{x}\|_2^2 \\ &\quad \downarrow \\ &\quad l_2 \text{ norm} = \text{Default}\end{aligned}$$


---


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

length of vector

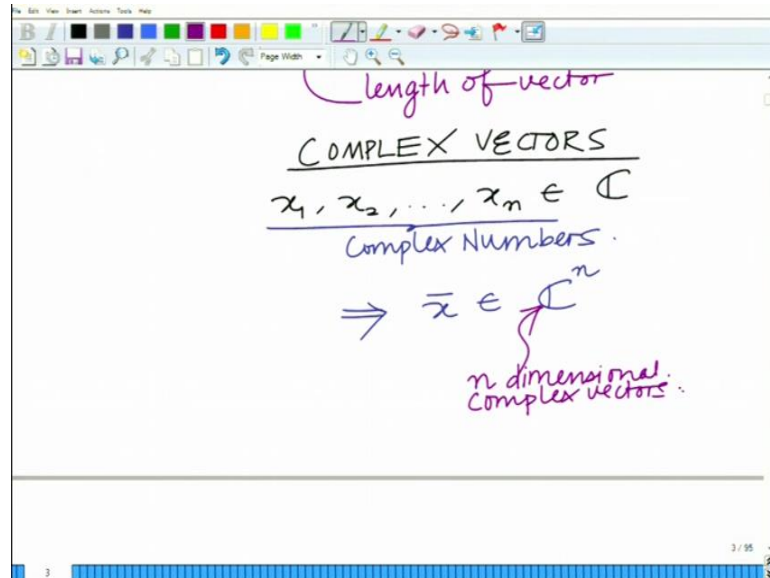
Here we can see,  $\bar{x}^T \bar{x}$  is basically equal to the summation of  $x_1^2, x_2^2, \dots, x_n^2$  which is also denoted by the norm square  $\|\bar{x}\|_2^2$ . In fact, this is a specific case of a norm, the  $l_2$  norm  $\|\bar{x}\|_2$ .  $\|\bar{x}\|_2$  is the norm of  $\bar{x}$ . Here 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

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

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,

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

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  $x_1, x_2, \dots, x_n$  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)

The image shows a whiteboard with handwritten mathematical expressions. At the top, the Hermitian of a vector  $\bar{x}$  is defined as a row vector of complex conjugates:  $\bar{x}^H = [\bar{x}_1^* \ \bar{x}_2^* \ \dots \ \bar{x}_n^*]$ . A blue arrow points from this expression to the text "Row vector + complex conjugate of elements". Below this, the product of the Hermitian vector and the original column vector is shown:  $\bar{x}^H \bar{x} = [\bar{x}_1^* \ \bar{x}_2^* \ \dots \ \bar{x}_n^*] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ . This is then simplified to the sum of squared magnitudes:  $= |x_1|^2 + |x_2|^2 + \dots + |x_n|^2$ . The whiteboard interface includes a toolbar at the top and a status bar at the bottom indicating "4 / 95".

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

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

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  $x_1^*, x_2^*, \dots, x_n^*$ . Now, the product of vectors  $\bar{x}^H$  and  $\bar{x}$  is

$$\begin{aligned} \bar{x}^H \bar{x} &= [x_1^*, x_2^*, \dots, x_n^*] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= |x_1|^2 + |x_2|^2 + \dots + |x_n|^2 \\ &= \|\bar{x}\|_2^2 \end{aligned}$$

As the summation of squares of the magnitudes is

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

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

(Refer Slide Time: 09:09)

The image shows a handwritten derivation on a whiteboard. The top part shows the calculation of the squared L2 norm of a vector  $\vec{x}$  with components  $x_1, x_2, \dots, x_n$ . It starts with  $= |x_1|^2 + |x_2|^2 + \dots + |x_n|^2$  and then states  $= \|\vec{x}\|_2^2$ . The bottom part shows the definition of the L2 norm:  $\|\vec{x}\| = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2}$ . The whiteboard has a toolbar at the top and a status bar at the bottom indicating slide 5 of 95.

$$= |x_1|^2 + |x_2|^2 + \dots + |x_n|^2$$
$$= \|\vec{x}\|_2^2$$
$$\|\vec{x}\| = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2}$$

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

$$\|\vec{x}\| = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2}$$

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)

The image shows a handwritten definition on a whiteboard. The top part shows the L2 norm formula:  $\|\vec{x}\| = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2}$ . An arrow points from this formula to the text "General Definition For real & complex vectors." The bottom part shows the formula for the unit norm vector:  $\tilde{\vec{x}} = \frac{\vec{x}}{\|\vec{x}\|}$ . An arrow points from this formula to the text "unit - Norm vector." The whiteboard has a toolbar at the top and a status bar at the bottom indicating slide 5 of 95.

$$\|\vec{x}\| = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2}$$

General Definition  
For real & complex  
vectors.

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

unit - Norm  
vector.

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)

The image shows a handwritten derivation on a presentation slide. At the top, there is a green arrow pointing to the text "unit - Norm vector." Below this, the derivation starts with the inner product of the unit vector with itself:

$$\tilde{x}^H \cdot \tilde{x} = \frac{\bar{x}^H}{\|\bar{x}\|} \cdot \frac{\bar{x}}{\|\bar{x}\|}$$

$$= \frac{\|\bar{x}\|^2}{\|\bar{x}\|^2} = 1$$


---

Below the horizontal line, the derivation continues with two steps:

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

$$\Rightarrow \boxed{\|\tilde{x}\| = 1}$$

The slide has a toolbar at the top and a status bar at the bottom indicating "6 / 95".

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

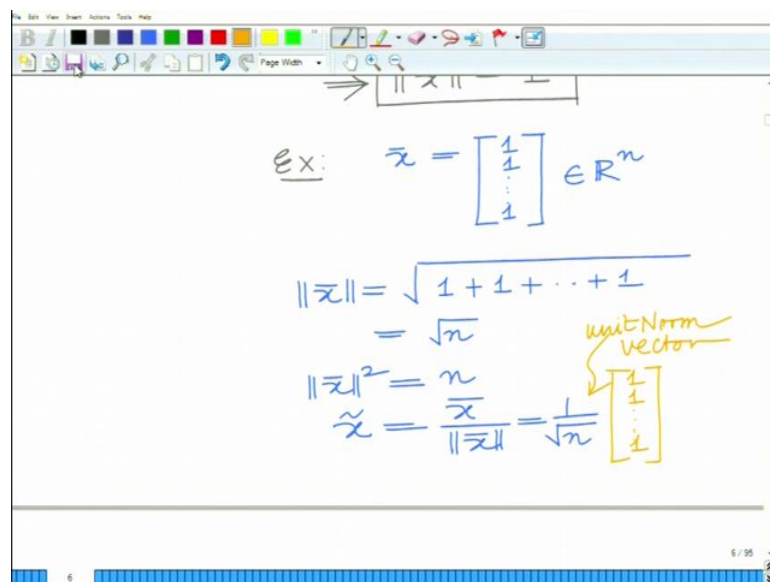
$$\|\tilde{x}\|^2 = 1$$

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

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)



The image shows a digital whiteboard with a toolbar at the top. The handwritten text on the board is as follows:

ex:  $\bar{x} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^n$

$\|\bar{x}\| = \sqrt{1+1+\dots+1}$

$= \sqrt{n}$

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

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

A yellow arrow points from the text "unit Norm vector" to the vector  $\tilde{x}$ .

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

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

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

$$\|\bar{x}\| = \sqrt{1+1+\dots+1}$$

$$= \sqrt{n}$$



Therefore,

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

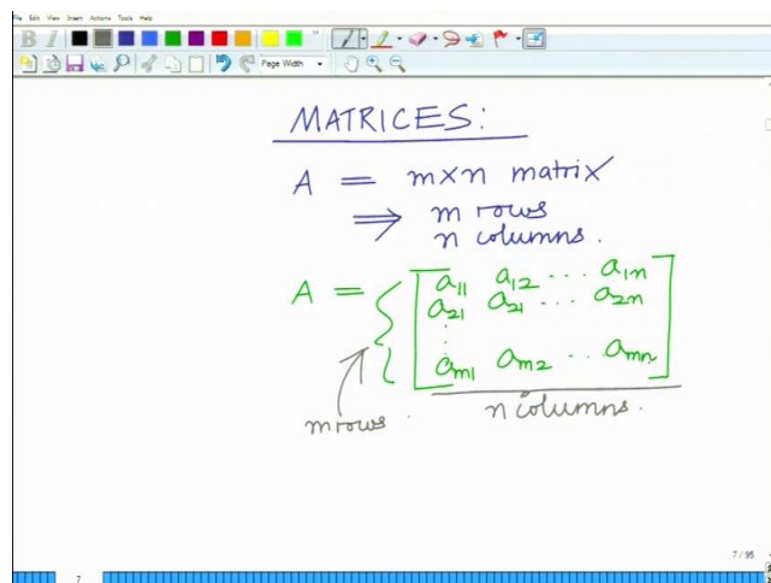
$$\text{And in fact, } \tilde{x} = \frac{\bar{x}}{\|\bar{x}\|} = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

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

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

(Refer Slide Time: 16:05)

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

m rows      n columns

$a_{ij}$  = element in  $i^{\text{th}}$  row &  $j^{\text{th}}$  column

If  $m = n$   
Then,  $A = \text{Square matrix}$

Here,  $ij^{\text{th}}$  element  $a_{ij}$  equals to the element in  $i^{\text{th}}$  row and the  $j^{\text{th}}$  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)

$\vec{w}_1, \vec{w}_2, \dots, \vec{w}_m$   
m vectors.

Linear Independence (LI)

Linearly independent if  
there do NOT exist  $c_1, c_2, \dots, c_m$  (NOT all zero) such that

$$c_1 \vec{w}_1 + c_2 \vec{w}_2 + \dots + c_m \vec{w}_m = 0$$

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

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

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  $C_1, C_2, \dots, C_m$ , such that all of them cannot be 0. So, in other words, for linear independent  $\bar{W}$ , there cannot be set of constant  $C_1, C_2, \dots, C_m$  such that

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

(Refer Slide Time: 20:01)

Linear Independence (LI)

Linearly independent if  
there do NOT exist  $C_1, C_2, \dots, C_m$  (NOT all zero) such that  
 $C_1 \bar{W}_1 + C_2 \bar{W}_2 + \dots + C_m \bar{W}_m = 0$   
Linear Combination

8 / 95

(Refer Slide Time: 20:40)

Linear Dependence:

Linearly dependent if ~~there~~ exist  $c_1, c_2, \dots, c_m$  ~~NOT~~ all zero, such that

$$c_1 \bar{w}_1 + c_2 \bar{w}_2 + \dots + c_m \bar{w}_m = 0$$

Linear Combination

So, if there exists these weights  $c_1, c_2, \dots, c_m$  such that not all of them are 0 and the linear combination of the vectors  $\bar{w}_1, \bar{w}_2, \dots, \bar{w}_m$  is 0, then these vectors  $\bar{w}_1, \bar{w}_2, \dots, \bar{w}_m$  are linearly dependent. For instance, let us take a very simple example to understand this.

(Refer Slide Time: 22:49)

exist  $c_1, c_2, \dots, c_m$  ~~NOT~~ all zero, such that

$$c_1 \bar{w}_1 + c_2 \bar{w}_2 + \dots + c_m \bar{w}_m = 0$$

Linear Combination

Ex:  $\bar{w}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$   $\bar{w}_2 = \begin{bmatrix} -2 \\ -2 \\ -2 \end{bmatrix}$

$$2 \cdot \bar{w}_1 + 1 \cdot \bar{w}_2 = 0$$

$\Rightarrow \bar{w}_1, \bar{w}_2$  are Linearly Dependent

Consider the vectors  $\bar{w}_1$  and  $\bar{w}_2$  equals to

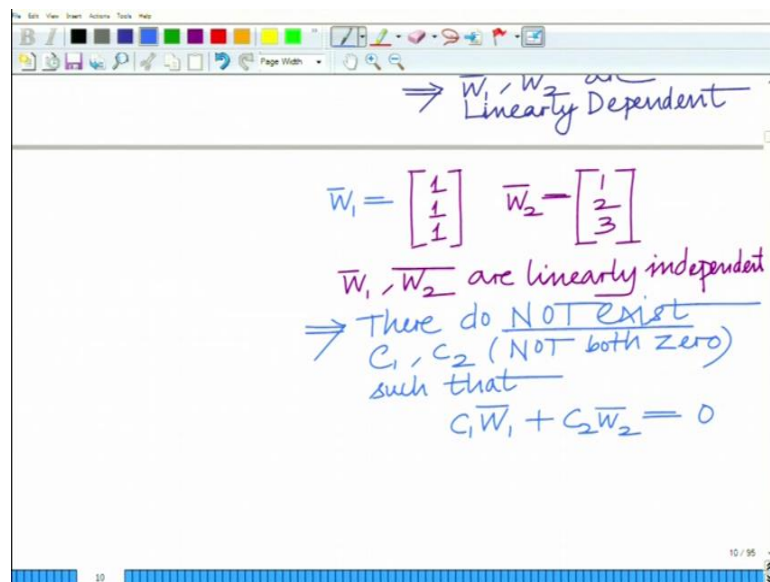
$$\bar{W}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ and } \bar{W}_2 = \begin{bmatrix} -2 \\ -2 \\ -2 \end{bmatrix}$$

then clearly we can write

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

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

(Refer Slide Time: 23:44)



On the other hand, we consider another example.

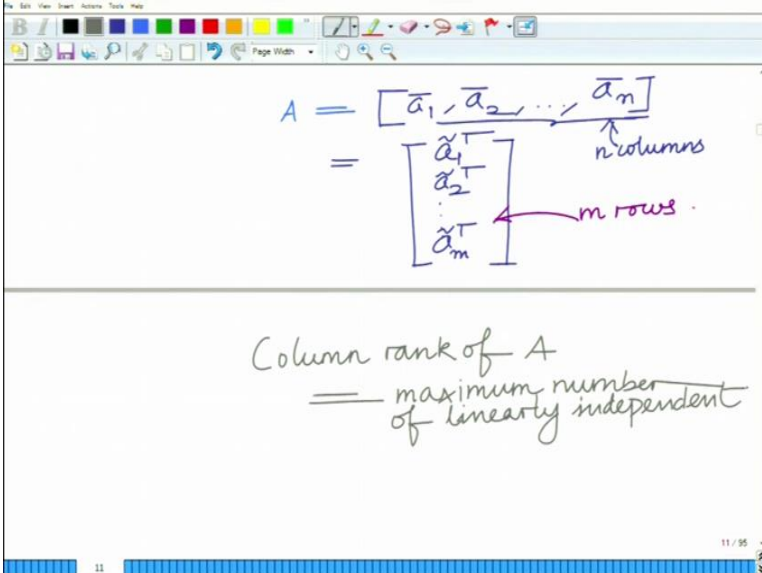
$$\bar{W}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ and } \bar{W}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

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  $C_1\bar{W}_1 + C_2\bar{W}_2 = 0$ . So we can quickly verify that  $\bar{W}_1$  and  $\bar{W}_2$  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)



The image shows a handwritten slide with two parts. The top part defines a matrix A as a row matrix  $A = [\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n]$  and as a column matrix  $A = \begin{bmatrix} \tilde{a}_1^T \\ \tilde{a}_2^T \\ \vdots \\ \tilde{a}_m^T \end{bmatrix}$ . An arrow points to the  $\bar{a}_n$  term with the label 'n columns', and another arrow points to the  $\tilde{a}_m^T$  term with the label 'm rows'. The bottom part defines the column rank of A as the 'maximum number of linearly independent' columns.

$$A = [\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n]$$

$$= \begin{bmatrix} \tilde{a}_1^T \\ \tilde{a}_2^T \\ \vdots \\ \tilde{a}_m^T \end{bmatrix}$$

Column rank of A  
= maximum number of linearly independent

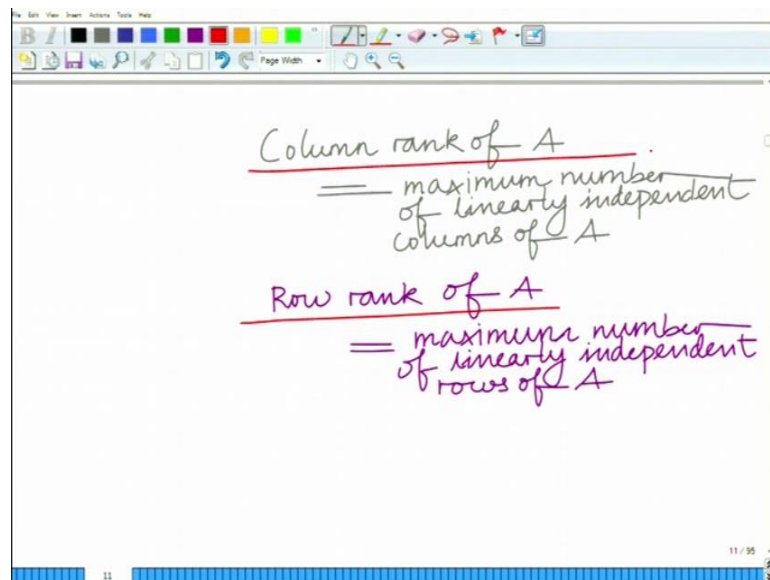
To define the rank of a matrix, let us consider matrix A as

$$A = [\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n]$$

$$= \begin{bmatrix} \tilde{a}_1^T \\ \tilde{a}_2^T \\ \vdots \\ \tilde{a}_n^T \end{bmatrix}$$

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  $\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n$ . 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)

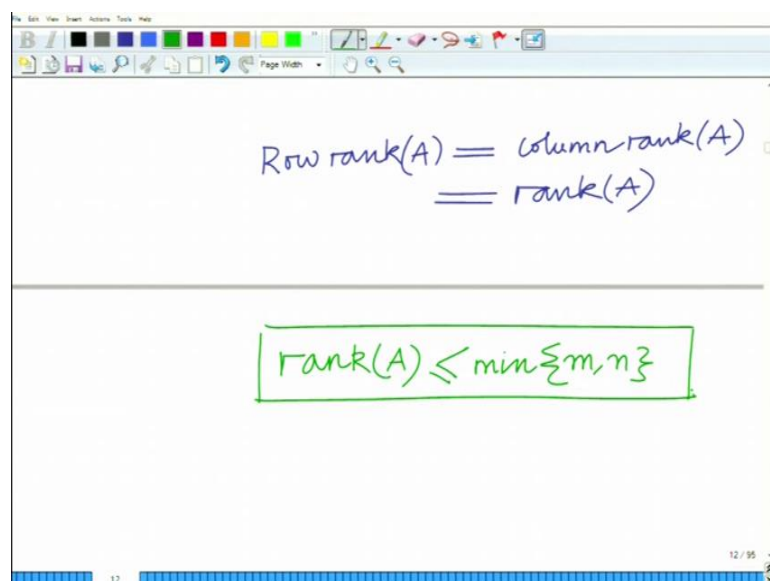


Similarly the Row Rank of a matrix  $A$  is the maximum number of linearly independent rows of matrix  $A$ .

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;

$$\text{Row rank}(A) = \text{Column rank}(A) = \mathbf{\text{Rank}(A)}$$

(Refer Slide Time: 28:45)

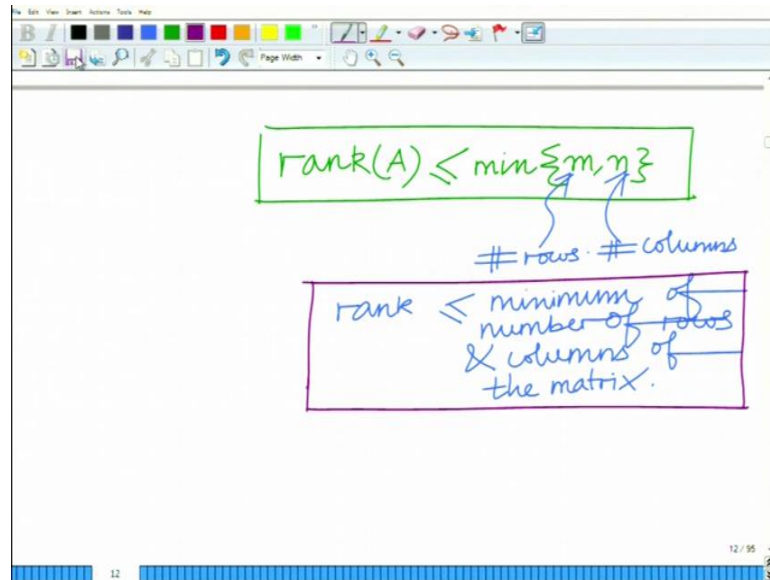


And in addition to this, rank of a matrix satisfies the property that

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

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.