

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

Lecture – 05
Inner Product Space and its Properties: Cauchy Schwarz Inequality

Welcome to another module in this massive open online course. So, we are looking at the inner product of the matrices and we already have defined the inner product of two matrices in the previous module.

(Refer Slide Time: 00:24)

The slide shows handwritten mathematical derivations on a whiteboard background. At the top left, there is a green asterisk. The main equation is $\langle A, B \rangle = \text{Tr}(B^T A)$. Below this, matrix B is defined as $B = \begin{bmatrix} b_1 & b_2 \end{bmatrix}$ with a 3×2 dimension label above it. Matrix A is defined as $A = \begin{bmatrix} a_1 & a_2 \end{bmatrix}$. Arrows point from the text '1st column' and '2nd column' to the columns of A. Below these, the product $B^T A$ is shown as $B^T A = \begin{bmatrix} b_1^T \\ b_2^T \end{bmatrix} \begin{bmatrix} a_1 & a_2 \end{bmatrix}$. The slide is from a presentation, with a toolbar at the top and a footer showing '61 / 125'.

In fact, it is shown that the inner product of two matrices A and B of the same size is equal to the trace of $(B^T A)$. That is

$$\langle A, B \rangle = \text{Tr}(B^T A)$$

For instance, for

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}$$

Let us say that first and second columns of A as a_1 and a_2 ; and similarly first and second columns of B as b_1 and b_2 . Then;

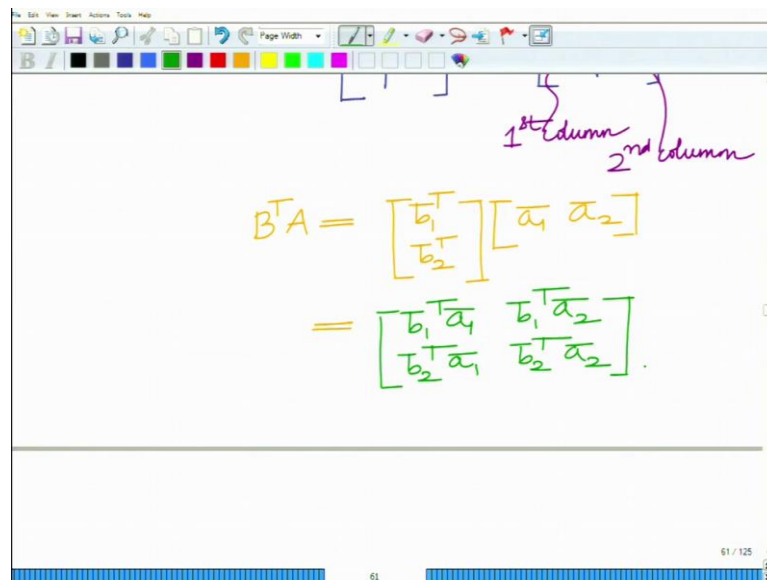
$$A = [\bar{a}_1 \mid \bar{a}_2]$$

$$B = [\bar{b}_1 \mid \bar{b}_2]$$

And hence;

$$B^T A = \begin{bmatrix} \bar{b}_1^T \\ \bar{b}_2^T \end{bmatrix} [\bar{a}_1 \mid \bar{a}_2]$$

(Refer Slide Time: 01:54)



The image shows a digital whiteboard with a toolbar at the top. The handwritten text in yellow and green ink shows the matrix multiplication:

$$B^T A = \begin{bmatrix} \bar{b}_1^T \\ \bar{b}_2^T \end{bmatrix} [\bar{a}_1 \mid \bar{a}_2]$$

Handwritten purple notes with arrows point to the columns of the second matrix: "1st column" for \bar{a}_1 and "2nd column" for \bar{a}_2 .

$$= \begin{bmatrix} \bar{b}_1^T \bar{a}_1 & \bar{b}_1^T \bar{a}_2 \\ \bar{b}_2^T \bar{a}_1 & \bar{b}_2^T \bar{a}_2 \end{bmatrix}$$

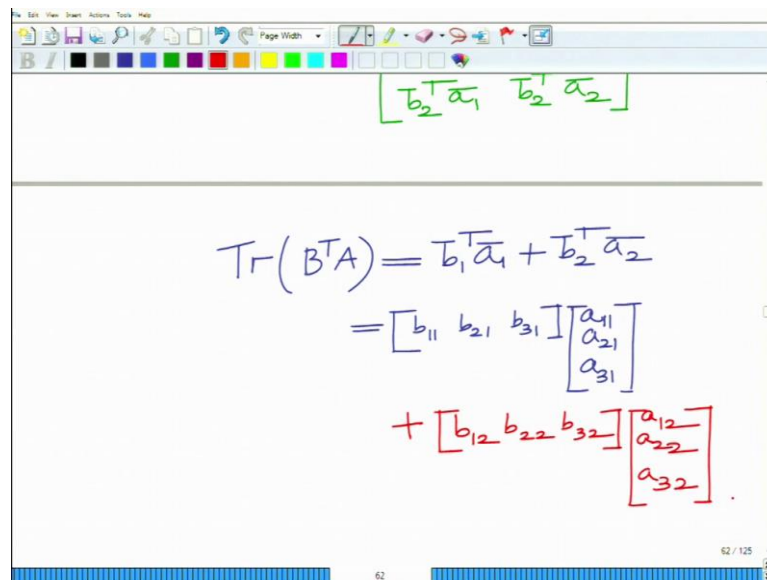
The bottom status bar of the whiteboard shows "61 / 125".

And it will be a 2×2 matrix and is as follows.

$$B^T A = \begin{bmatrix} \bar{b}_1^T \\ \bar{b}_2^T \end{bmatrix} [\bar{a}_1 \mid \bar{a}_2]$$

$$= \begin{bmatrix} \bar{b}_1^T \bar{a}_1 & \bar{b}_1^T \bar{a}_2 \\ \bar{b}_2^T \bar{a}_1 & \bar{b}_2^T \bar{a}_2 \end{bmatrix}$$

(Refer Slide Time: 02:21)

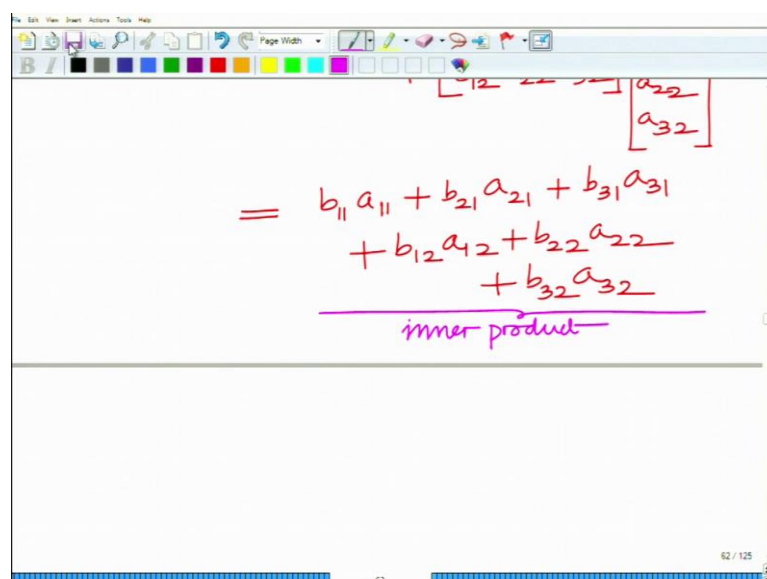


A screenshot of a presentation slide showing handwritten mathematical derivations. At the top, the vectors $\begin{bmatrix} b_1^T a_1 & b_2^T a_2 \end{bmatrix}$ are written in green. Below, the trace of $B^T A$ is calculated in blue ink: $\text{Tr}(B^T A) = b_1^T a_1 + b_2^T a_2$. This is then expanded into matrix multiplication: $\begin{bmatrix} b_{11} & b_{21} & b_{31} \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} + \begin{bmatrix} b_{12} & b_{22} & b_{32} \end{bmatrix} \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix}$, with the second part written in red ink. The slide number 62/125 is visible in the bottom right corner.

And therefore, the trace of $(B^T A)$ is

$$\begin{aligned}
 \text{Tr}(B^T A) &= \bar{b}_1^T \bar{a}_1 + \bar{b}_2^T \bar{a}_2 \\
 &= \begin{bmatrix} b_{11} & b_{21} & b_{31} \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} + \begin{bmatrix} b_{12} & b_{22} & b_{32} \end{bmatrix} \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} \\
 &= b_{11}a_{11} + b_{21}a_{21} + b_{31}a_{31} + b_{12}a_{12} + b_{22}a_{22} + b_{32}a_{32} \\
 &= \langle A, B \rangle
 \end{aligned}$$

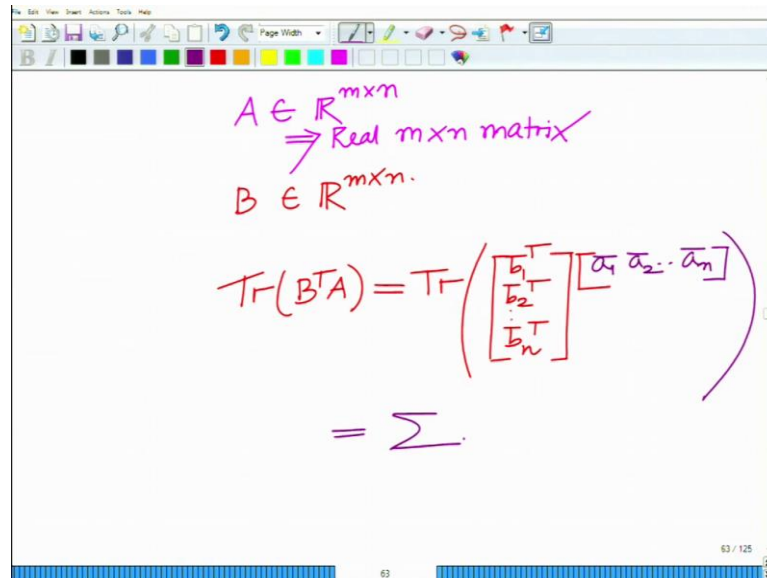
(Refer Slide Time: 03:45)



A screenshot of a presentation slide showing the expansion of the inner product. The expression $b_{11}a_{11} + b_{21}a_{21} + b_{31}a_{31} + b_{12}a_{12} + b_{22}a_{22} + b_{32}a_{32}$ is written in red ink. A horizontal line is drawn under the last three terms, and the words "inner product" are written in purple ink below the line. The slide number 62/125 is visible in the bottom right corner.

So this is the inner product of A and B. And in general this can be generalized for $m \times n$ matrices.

(Refer Slide Time: 04:47)



Handwritten notes on a whiteboard:

$$A \in \mathbb{R}^{m \times n} \Rightarrow \text{Real } m \times n \text{ matrix}$$

$$B \in \mathbb{R}^{m \times n}$$

$$\text{Tr}(B^T A) = \text{Tr} \left(\begin{bmatrix} \bar{b}_1^T \\ \bar{b}_2^T \\ \vdots \\ \bar{b}_n^T \end{bmatrix} \begin{bmatrix} \bar{a}_1 & \bar{a}_2 & \cdots & \bar{a}_n \end{bmatrix} \right)$$

$$= \sum$$

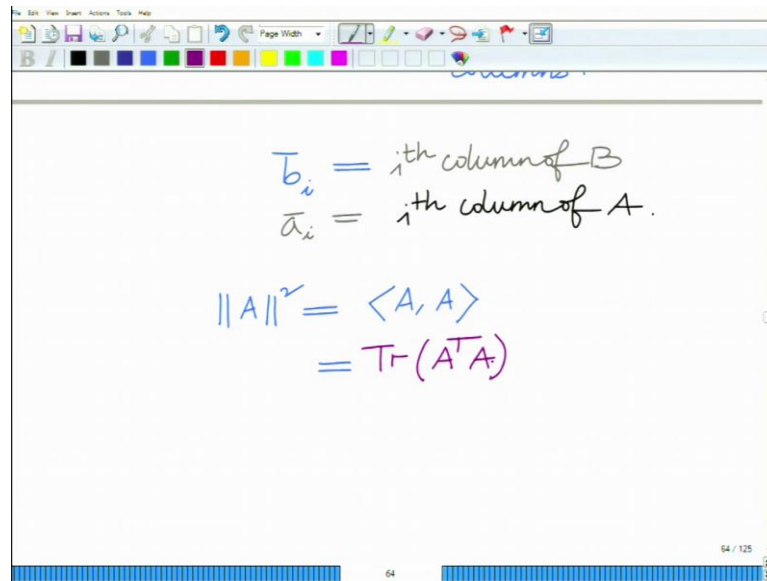
So, In general for $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{m \times n}$; we can say that

$$\text{Tr}(B^T A) = \text{Tr} \left(\begin{bmatrix} \bar{b}_1^T \\ \bar{b}_2^T \\ \vdots \\ \bar{b}_n^T \end{bmatrix} \begin{bmatrix} \bar{a}_1 & \bar{a}_2 & \cdots & \bar{a}_n \end{bmatrix} \right)$$

$$= \sum_{i=1}^n \bar{b}_i^T \bar{a}_i$$

Here, n equals the number of columns.

(Refer Slide Time: 06:29)



The image shows a digital whiteboard interface with a toolbar at the top. The whiteboard contains the following handwritten text in blue and purple ink:

$$\bar{b}_i = i^{\text{th}} \text{ column of } B$$
$$\bar{a}_i = i^{\text{th}} \text{ column of } A.$$
$$\|A\|^2 = \langle A, A \rangle$$
$$= \text{Tr}(A^T A)$$

At the bottom right of the whiteboard, the text "64 / 125" is visible.

The various quantities are

$$\bar{b}_i = i^{\text{th}} \text{ column of } B$$

$$\bar{a}_i = i^{\text{th}} \text{ column of } A$$

And as we know that every inner product induces a norm. Therefore,

$$\begin{aligned} \|A\|^2 &= \langle A, A \rangle \\ &= \text{Tr}(A^T A) \end{aligned}$$

(Refer Slide Time: 07:37)

The image shows a digital whiteboard with a toolbar at the top. The handwritten text in purple and red ink is as follows:

$$\begin{aligned} &= \text{Tr}(A^T A) \\ &= \sum_{i=1}^n \bar{a}_i^T \bar{a}_i \\ &= \sum_{i=1}^n \|\bar{a}_i\|^2 \\ &= \|\bar{a}_1\|^2 + \|\bar{a}_2\|^2 + \dots + \|\bar{a}_m\|^2 \\ &= \text{sum of magnitude squared of all elements of } A \end{aligned}$$

At the bottom right of the whiteboard, the text "64 / 125" is visible.

This further can be written as

$$\begin{aligned} \|A\|^2 &= \text{Tr}(A^T A) \\ &= \sum_{i=1}^n \bar{a}_i^T \bar{a}_i \\ &= \sum_{i=1}^n \|\bar{a}_i\|^2 \\ &= \|\bar{a}_1\|^2 + \|\bar{a}_2\|^2 + \dots + \|\bar{a}_n\|^2 \end{aligned}$$

which is sum of the squares of the norms of all the columns of matrix A. This is also equal to the sum of magnitude square of all elements of A. This is termed as the Frobenius norm and is denoted as $\|A\|_F$. So we can it as;

$$\|A\|_F^2 = \|\bar{a}_1\|^2 + \|\bar{a}_2\|^2 + \dots + \|\bar{a}_n\|^2$$

(Refer Slide Time: 09:00)

The whiteboard shows the following content:

$$= \sum_{i=1}^n \|\bar{a}_i\|^2$$
$$\|A\|_F^2 = \|\bar{a}_1\|^2 + \|\bar{a}_2\|^2 + \dots + \|\bar{a}_m\|^2$$

sum of magnitude squared of all elements of A

Frobenius Norm

This is termed as Frobenius Norm

65 / 125

This is the matrix Frobenius norm which is the sum of the norm square of all the columns of the matrix or sum of the magnitude square of all the elements of the matrix. So, this is norm that is also induced by the definition of the inner product corresponding to matrices. The next important aspect is the Cauchy Schwarz inequality.

(Refer Slide Time: 10:13)

The whiteboard shows the following content:

CAUCHY-SCHWARZ

INEQUALITY :

Fundamental Property for inner product

$$\langle \bar{u}, \bar{v} \rangle^2 \leq \langle \bar{u}, \bar{u} \rangle \cdot \langle \bar{v}, \bar{v} \rangle$$
$$= \|\bar{u}\|^2 \cdot \|\bar{v}\|^2$$
$$\Rightarrow |\langle \bar{u}, \bar{v} \rangle| \leq \|\bar{u}\| \cdot \|\bar{v}\|$$

66 / 125

Cauchy Schwarz inequality is very fundamental property of a inner product. It states that

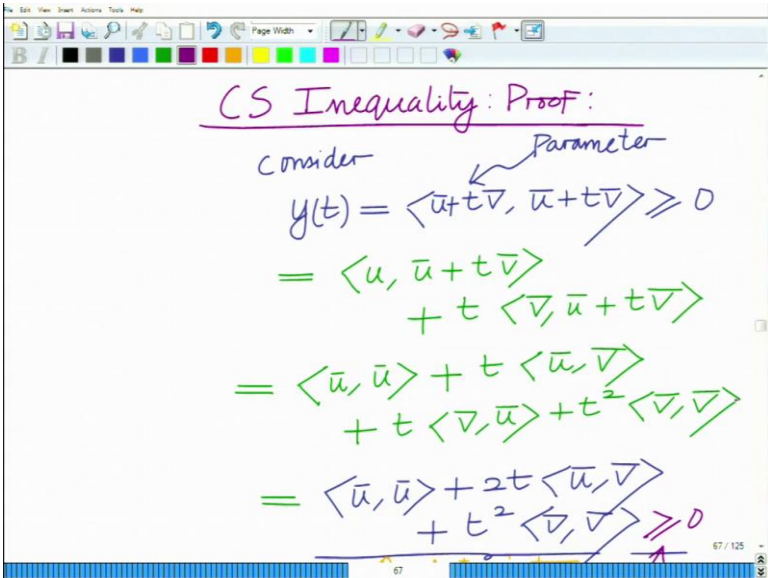
$$\begin{aligned}\langle \bar{u}, \bar{v} \rangle^2 &\leq \langle \bar{u}, \bar{u} \rangle \cdot \langle \bar{v}, \bar{v} \rangle \\ &\leq \|\bar{u}\|^2 \cdot \|\bar{v}\|^2\end{aligned}$$

This basically implies that the magnitude of the inner product of two matrices is less than or equal to the product of square of individual norms of two matrices.

$$|\langle \bar{u}, \bar{v} \rangle| \leq \|\bar{u}\| \cdot \|\bar{v}\|$$

This is the Cauchy Schwarz inequality. This is valid for either the inner product for the vectors or inner product of the functions. So, this is valid for any general definition of the inner product.

(Refer Slide Time: 12:41)



The image shows a whiteboard with a handwritten proof of the Cauchy-Schwarz inequality. The title is "CS Inequality: Proof:". Below it, the word "consider" is written, followed by the definition of a function $y(t) = \langle \bar{u} + t\bar{v}, \bar{u} + t\bar{v} \rangle \geq 0$. An arrow points from the word "Parameter" to the variable t in the expression. The proof then expands the inner product:
$$\begin{aligned}y(t) &= \langle \bar{u}, \bar{u} + t\bar{v} \rangle + t \langle \bar{v}, \bar{u} + t\bar{v} \rangle \\ &= \langle \bar{u}, \bar{u} \rangle + t \langle \bar{u}, \bar{v} \rangle + t \langle \bar{v}, \bar{u} \rangle + t^2 \langle \bar{v}, \bar{v} \rangle \\ &= \langle \bar{u}, \bar{u} \rangle + 2t \langle \bar{u}, \bar{v} \rangle + t^2 \langle \bar{v}, \bar{v} \rangle \geq 0\end{aligned}$$
The final expression is underlined and marked with a red arrow pointing to the inequality sign.

The Cauchy Schwarz inequality can be proved as follows. Consider a function $y(t)$ such as;

$$y(t) = \langle \bar{u} + t\bar{v}, \bar{u} + t\bar{v} \rangle \geq 0$$

This t is a parameter. And now on expanding the inner product, we get

$$\begin{aligned}y(t) &= \langle \bar{u}, \bar{u} + t\bar{v} \rangle + t \langle \bar{v}, \bar{u} + t\bar{v} \rangle \\ &= \langle \bar{u}, \bar{u} \rangle + t \langle \bar{u}, \bar{v} \rangle + t \langle \bar{v}, \bar{u} \rangle + t^2 \langle \bar{v}, \bar{v} \rangle \\ &= \langle \bar{u}, \bar{u} \rangle + 2t \langle \bar{u}, \bar{v} \rangle + t^2 \langle \bar{v}, \bar{v} \rangle\end{aligned}$$

(Refer Slide Time: 15:07)

Handwritten derivation on a whiteboard:

$$\begin{aligned}
 &= \langle \vec{u}, \vec{u} \rangle + t \langle \vec{v}, \vec{u} \rangle + t^2 \langle \vec{v}, \vec{v} \rangle \\
 &= \langle \vec{u}, \vec{u} \rangle + t 2 \langle \vec{u}, \vec{v} \rangle + t^2 \langle \vec{v}, \vec{v} \rangle \geq 0
 \end{aligned}$$

Quadratic in t

For all values of t

Holds true only when discriminant of quadratic ≤ 0

$$\begin{aligned}
 &\Rightarrow b^2 - 4ac \leq 0 \\
 &\Rightarrow 4 \langle \vec{u}, \vec{v} \rangle^2
 \end{aligned}$$

67 / 125

Now observe that this is a quadratic equation in t and according to the property of inner product space, this is always greater than or equal to 0 for all values of t . So, this holds true only when discriminant of the quadratic is less than or equal to 0.

(Refer Slide Time: 17:44)

Handwritten derivation on a whiteboard:

$$\begin{aligned}
 &\Rightarrow b^2 - 4ac \leq 0 \\
 &\Rightarrow 4 \langle \vec{u}, \vec{v} \rangle^2 - 4 \langle \vec{u}, \vec{u} \rangle \langle \vec{v}, \vec{v} \rangle \leq 0.
 \end{aligned}$$

$$\Rightarrow \langle \vec{u}, \vec{v} \rangle^2 \leq \langle \vec{u}, \vec{u} \rangle \langle \vec{v}, \vec{v} \rangle = \|\vec{u}\|^2 \|\vec{v}\|^2$$

$$\Rightarrow |\langle \vec{u}, \vec{v} \rangle| \leq \|\vec{u}\| \cdot \|\vec{v}\|$$

CS Inequality

68 / 125

This implies that

$$b^2 - 4ac \leq 0$$

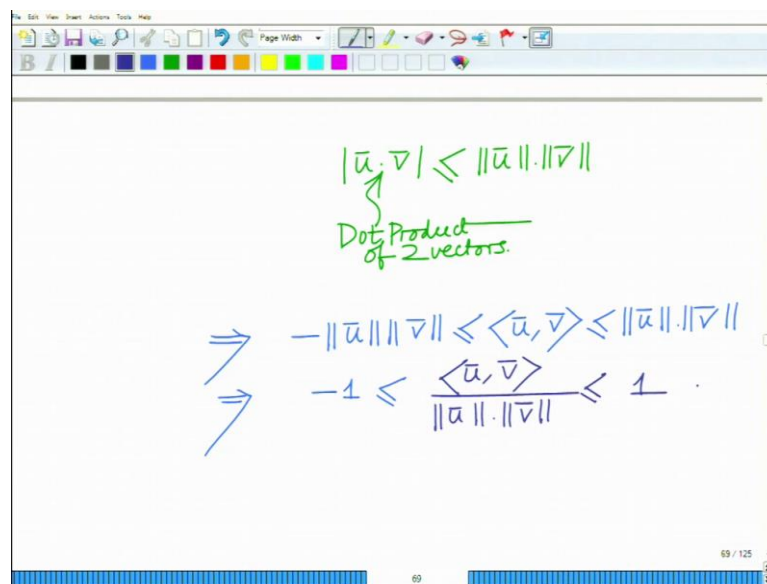
$$4\langle \bar{u}, \bar{v} \rangle^2 - 4\langle \bar{v}, \bar{v} \rangle \langle \bar{u}, \bar{u} \rangle \leq 0$$

$$\langle \bar{u}, \bar{v} \rangle^2 \leq \langle \bar{v}, \bar{v} \rangle \langle \bar{u}, \bar{u} \rangle$$

$$|\langle \bar{u}, \bar{v} \rangle| \leq \|\bar{u}\| \cdot \|\bar{v}\|$$

This is basically Cauchy Schwarz inequality. So, it verifies the Cauchy Schwarz inequality that the magnitude of the inner product between two vectors is less than or equal to the product of their individual norms.

(Refer Slide Time: 19:32)



The image shows a whiteboard with handwritten mathematical derivations. At the top, the equation $|\bar{u} \cdot \bar{v}| \leq \|\bar{u}\| \cdot \|\bar{v}\|$ is written in green. Below it, a bracket points to the dot product term with the text "Dot Product of 2 vectors." in green. Further down, two blue arrows point to the inequalities $-\|\bar{u}\| \cdot \|\bar{v}\| \leq \langle \bar{u}, \bar{v} \rangle \leq \|\bar{u}\| \cdot \|\bar{v}\|$ and $-1 \leq \frac{\langle \bar{u}, \bar{v} \rangle}{\|\bar{u}\| \cdot \|\bar{v}\|} \leq 1$.

Also we can write it as the magnitude of the dot product. Therefore;

$$|\langle \bar{u}, \bar{v} \rangle| \leq \|\bar{u}\| \cdot \|\bar{v}\|$$

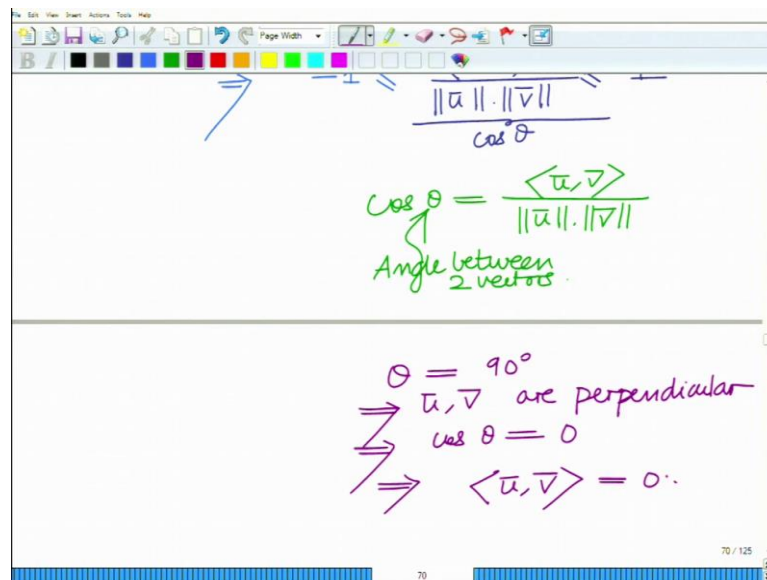
$$|\bar{u} \cdot \bar{v}| \leq \|\bar{u}\| \cdot \|\bar{v}\|$$

And hence we can also conclude that

$$-\|\bar{u}\| \cdot \|\bar{v}\| \leq \langle \bar{u}, \bar{v} \rangle \leq \|\bar{u}\| \cdot \|\bar{v}\|$$

$$-1 \leq \frac{\langle \bar{u}, \bar{v} \rangle}{\|\bar{u}\| \cdot \|\bar{v}\|} \leq 1$$

(Refer Slide Time: 20:45)



So, this quantity can be defined as the cosine of an angle θ because cosine θ lies between minus 1 and 1.

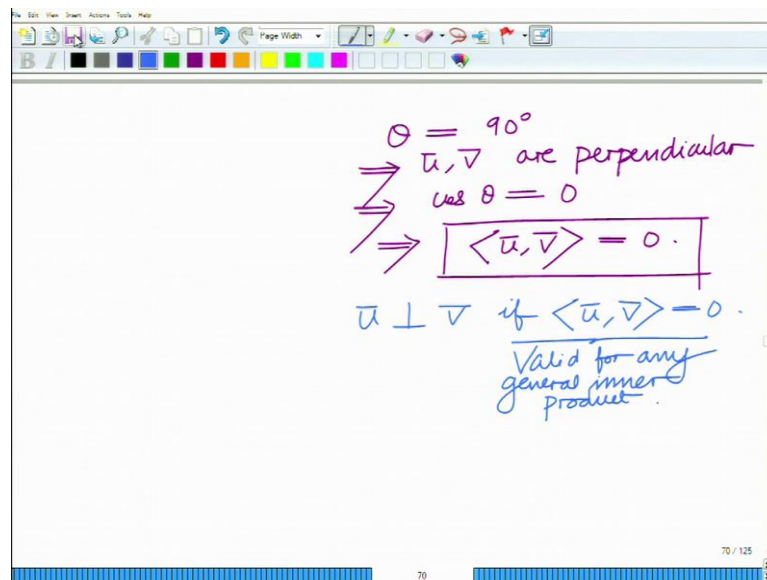
$$\cos \theta = \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{u}\| \cdot \|\vec{v}\|}$$

And here θ is the angle between the two vectors \vec{u} and \vec{v} .

So, if $\theta = 90^\circ$ that implies the vectors are perpendicular. This implies \vec{u} and \vec{v} are perpendicular to each other. This implies that

$$\langle \vec{u}, \vec{v} \rangle = 0$$

(Refer Slide Time: 22:23)



This shows that if the inner product of two vectors is equal to 0 then this means both the vectors are perpendicular to each other. This is valid for any general inner product. This concept of inner product is a very interesting and powerful concept which has a large number of applications and yields several interesting insights. So, we will stop here and continue with other aspects in the subsequent modules.

Thank you very much.