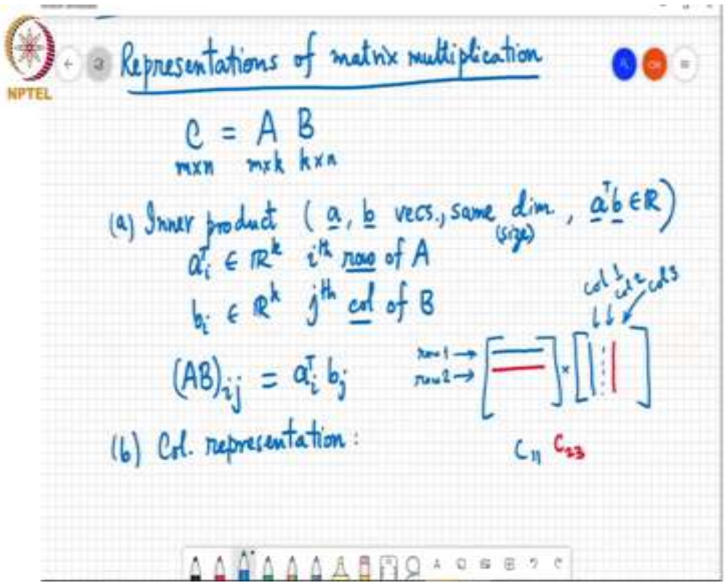
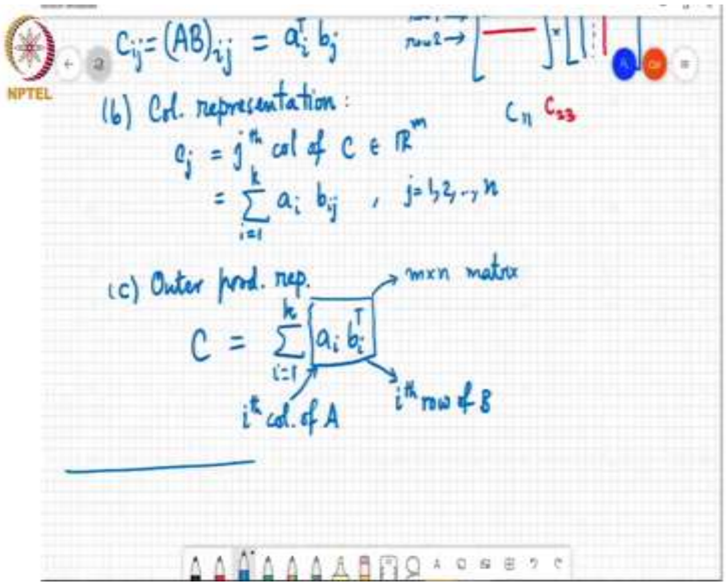
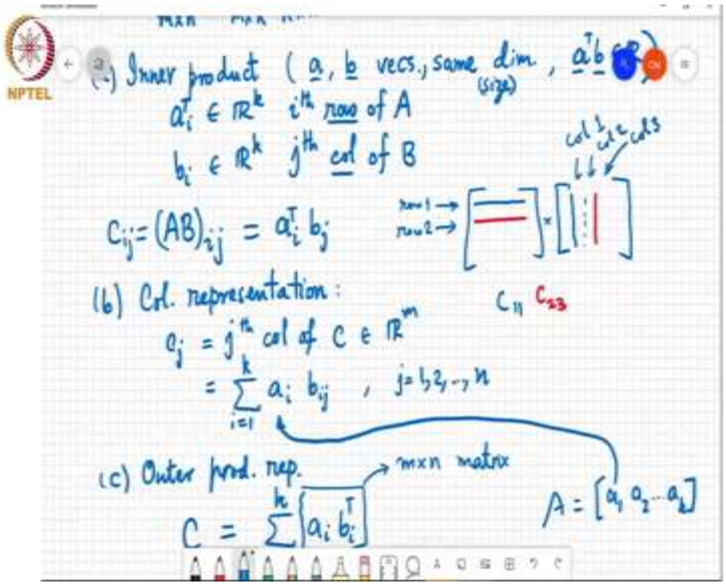
Matric Theory
Professor. Chandra R. Murthy
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore
Inner Product

(Refer Slide Time: 0:15)







So, suppose you write C as A times B where A is a matrix of size m by k and B is a matrix of size k by n. Then one way to view this is in terms of an inner product, which we have seen in the previous, in one of the previous classes already, but we will also formally define it in a few minutes, but inner product, so suppose I have A and B as vectors of the same length, then the inner product is defined as a transpose B, which is just a real number. Then if I consider ai transpose.

Student: Sir?

Professor: Yeah!

Student: Sir, what is the length of a vector?

Professor: Length of a vector is the number of elements in it.

Student: Sir, would not it just be equal to the dimension?

Professor: Like dimension like in vectors.

Student: It is okay.

Professor: No, but you are right, in that we will momentarily define length of a vector also and so same dimension or size is what I really meant here. So, suppose ai is the ith row of A and bi also in R power k is the jth column of A, jth column of B.

Student: Hello sir?

Professor: Yeah, please go ahead.

Student: Sir, what is v-e-c-s in bracket written?

Professor: Vectors.

Students: Vectors, okay sir.

Professor: So, the i jth entry of the product a b is nothing but the inner product between the ith row of A and the jth column of B, so this is one way to view matrix multiplication. Every element of the product matrix is an inner product between a row of A and a column of B. The other representation is the column representation.

Student: Sigma ai transpose should be there…

Professor: So there is no summation here, this is the inner product representation. So, if I have two matrices and I am taking the product, recall that if you want the 1 comma 1th element of their product you have to take the first row of this matrix, first column of this matrix, take their inner product and that gives you the 1 comma 1th element of this product.

If I want the, say 2 comma 3th entry of this of this product matrix, then I must take the second row and then the third column, so there is one more here and then I take the inner product of the two red vectors and that gives me the two comma 3th entry. So, this is called the inner product representation. The other thing is the column representation.

So, Cj if I define to be the jth column of C which is going to be in R to the m, then this is equal to summation i equal to 1 to k ai b ij, this is true for j equal to 1, 2, up to n. So, each column of this matrix C is a linear combination of the columns of A and what are the coefficients, they are given by bij, the jth column of b gives you the coefficients of the linear combination that will form the jth column of C.

So, it is a linear combination, I repeat it is a linear combination of the columns of A with the coefficients given by the jth column of B. And the third is the outer product representation. So, we can also write the entire matrix C, so we are going from small to big, here we looked at how to write an ijth entry of C, a single entry of C, this is for getting you an entire column of C and this is a representation which will give you the entire matrix C itself.

This is equal to summation k, I will write it in terms of i, i equal to 1th row k ai bi transpose, so this is the ith column of A and this is the ith row of B. Now, since A is of size m by k, the ith column of A is of size m by 1 and the ith row of B is of size 1 by, so when I take this outer product I will get an m by n matrix. And so the matrix C is the sum of k rank 1 matrices of size m by n. So, now we move on to the inner product.

Student: Sir, what is ai and b, like column?

Professor: Sorry, I did not understand your question.

Student: Sir in column representation session what is ai in that summation, inside summation?

Professor: These are the columns of a and a i is the ith column of A.

Student: Then the summation (())(8:12)…

Professor: Just hold on one second, A is of size m by k, so A has k columns, yeah so the summation goes from 1 to k. So, what is the question? There are k columns in the matrix A, A is of size m by k, a1 is the first column, a2 is the second column, ak is the kth column.

Student: So, now I got it.

Student: jth column of C belongs to R power of m.

Professor: jth column of C is an R to the m, yes. But every column of C is in Rm. C is of size the m by n.

Student: But it is written R to the power of m there.

Professor: R to the power?

Student: M, in column representation.

Professor: No, Cj is the jth column of C, it is in R to the m, n.

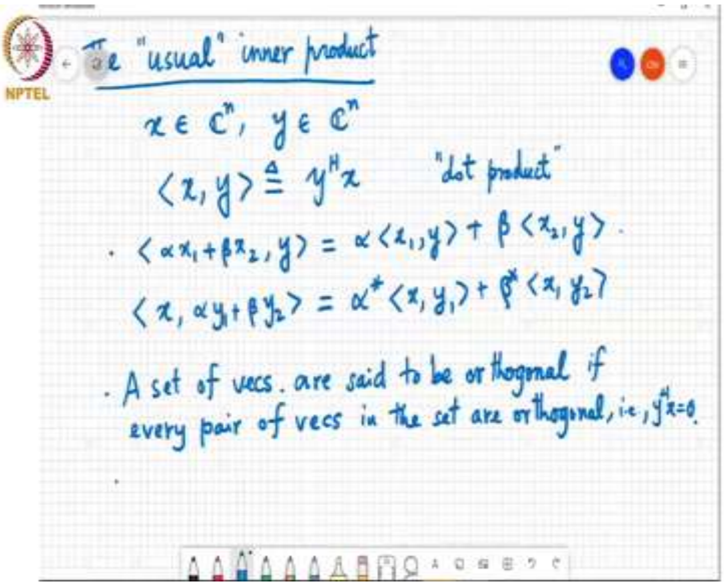
Student: Okay, sir thank you.

Professor: You are right, yeah each column is in R to the m and there are n such columns, is it clear?

Student: Yes sir.

Professor: These things are a little tricky, it is good to think about it independently., but sometimes it can be a little bit confusing.

(Refer Slide Time: 9:59)



So, the usual inner product, so given two vectors x in C to the n and y in C to the n. So, I am using complex vectors here because it is just the more general definition. So, the inner product between x and y is written like this and it is equal to y Hermitian x or it is defined as y Hermitian x and this is also called the dot product.

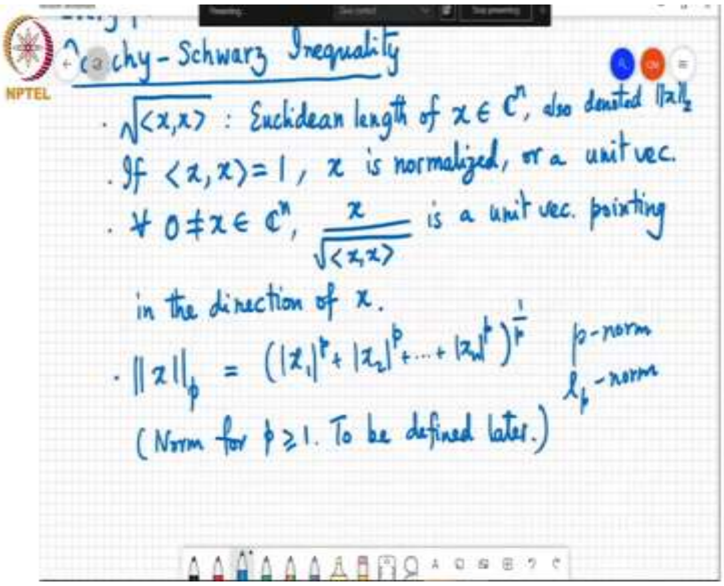
So, Hermitian is nothing but the conjugate transpose, so you take the transpose of y and then you take the complex conjugate of every entry, that is y Hermitian, y Hermitian inner product with or dot product with x is defined to be the inner product between x and y. There are other definitions possible but we will revisit that later in the course. But this definition here, it satisfies two properties that I can immediately tell you.

The first is that if I take alpha x1 plus beta x2 inner product with y this is going to be equal to alpha inner product of x1 and y plus beta inner product of x2 with y. But if I take the inner product between x and alpha y1 plus beta y2, what will I get?

Student: A conjugate.

Professor: x, y1 plus beta star x, y2, so it is linear in the first argument and it is conjugate linear in the second argument. Now, we can immediately define orthogonality. A set of vectors are said to be orthogonal if every pair of vectors in the set are orthogonal. So, you take any two vectors, you take their inner product, you get 0, then we say that, if that happens for every pair of vectors in the set, we say that the set of vectors are orthogonal.

(Refer Slide Time: 13:10)



Now, one inequality, very famous one related to the inner product is what, there is one very famous inequality associated with the inner product.

Student: Cauchy-Schwarz.

Professor: Exactly! So, before I put down the Cauchy-Schwarz inequality, let me just say one small thing, if I take square root of the inner product of x with itself this is called the Euclidean length of x in C to the n and also, it is also denoted by norm of x2. Now if the inner product of x with itself is 1, then the vector x is said to be normalized or that x is a unit vector. So, basically if I take any vector x in C to the n, then x over square root of the inner product of x with itself is a unit vector pointing in the direction of x.

Student: Sir?

Professor: Yes please.

Student: Sir, what is the subscript to the above statement? What does it mean?

Professor: Can you take a guess?

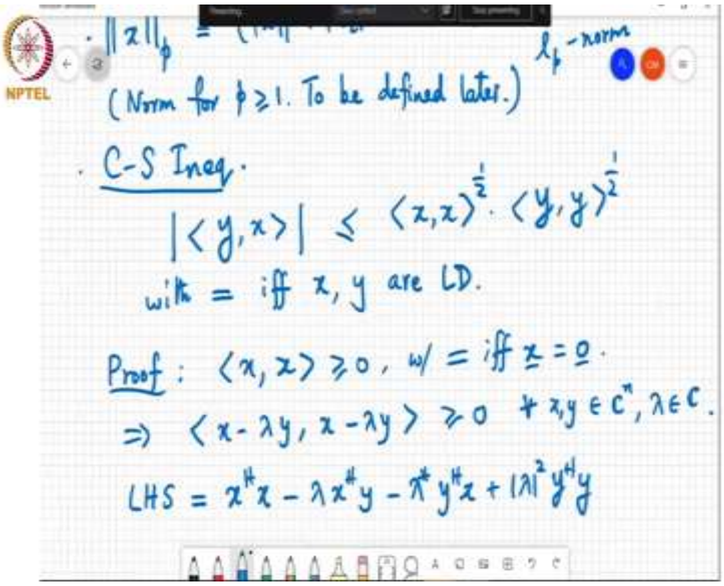
Student: Inner product of two quantities.

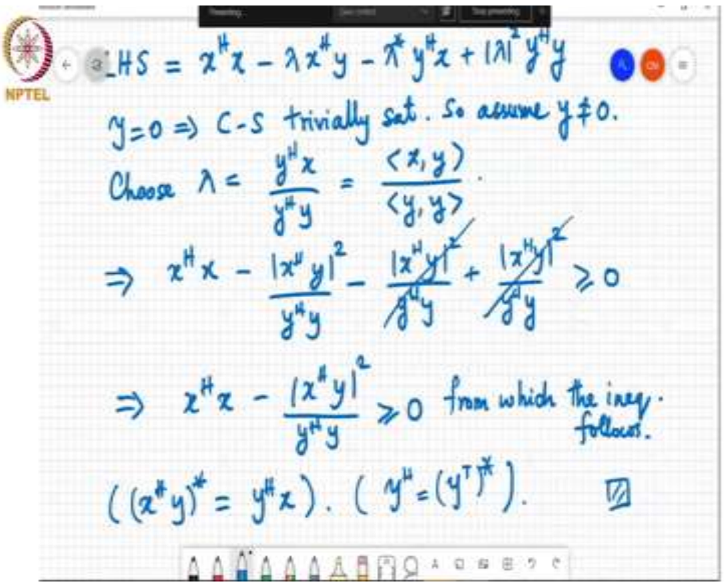
Professor: No, the square root, it is square root, right?

Student: No, not that. So, the point is that if you look at what this thing is doing, it is taking every entry of x, taking the mod square of it, adding that up across the entries and then you are finally taking a square root, because you are taking the square of each of the entries and then eventually taking the square root, it is denoted by norm x2. So, if I, mean since you ask the question I will say that norm x p is going to be equal to… So, this is a definition of what is called the p-norm of x. It is also known as the lp-norm and it turns out it is a norm for p greater than or equal to 1.

But I will, I have not defined, I have not defined what a norm is, so for now you take it on faith that if for any p greater than or equal to 1 I can define norm xp to be this quantity here and it actually is a norm. So, we will study that later, so there are other kinds of inner products and norms that you can define, which we will study later.

(Refer Slide Time: 17:37)





But now I am ready to state the Cauchy-Schwarz inequality.

Student: Sir, I have a question, we cannot define 0 norm, because…

Professor: So, you can you can define norm xp for p less than 1 and you can get, you can take p equals 0, you can even take negative p, but for p less than 1 this will, this norm xp will not be a norm and so I do not want to get into the details right now, because I have not defined norm yet, we will define it in one or two classes and then we can discuss these properties, but for now I just want to relate it to this.

Somebody asked me why it is norm x2 here, and so I am just answering that, it is actually coming from a more general definition of the length of a vector which can be defined like this and this is a valid definition of the length of a vector for p greater than or equal to 1. So, let us not worry about what happens when p is less than 1 or equal to 0 or even negative, just for the moment, we will come back to that point later.

Student: Okay, sir, thank you.

Professor: So, the Cauchy-Schwarz inequality simply says that the inner product between y sand x in magnitude is less than or equal to the inner product between x and itself power half times the inner product between y and itself power half. And with equality, if and only if x and y are linearly dependent, basically they point in the same direction.

So, the proof is very quick. So, I am just going to, in the dying minutes of this class run through the proof, all you do is you start with noting that x, inner product with x is always greater than or equal to 0 with equality if and only if x equals 0. As I wrote it, here you can see that the inner product of x with itself is the sum of the squares of the entries of x power 1 by 2 and so if this equals 0, the sum of the squares of the entries are is 0, which is only possible if every one of these entries is 0 because they are all non-negative quantities.

So, now we will take this x minus lambda y inner product with itself and this is because I am taking the inner product of a vector with itself and lambda is some scalar here and since it is a inner product of a vector with itself, this is always greater than or equal to 0 with equality if and only if x minus lambda y equals 0. So, from this you can already see that the equality condition will only be satisfied if x equals lambda y.

So, that this difference is 0 which means that x and y are linearly dependent. And this is true for every x, y belonging to C to the n and lambda belonging to C. So, the we just expand the left hand side and if you expand it out you will see, you just have to take this Hermitian times this and expand it out, you will see that this is equal to x Hermitian x minus lambda x Hermitian y minus lambda star y Hermitian x plus mod lambda square y Hermitian y.

Now, if y equals 0, then this inequality is trivially satisfied, y, x inner product is less than or equal to this times this, if y equals 0, the left hand side is 0, the right hand side is also equal to 0 and so this inequality is clear, is trivially satisfied. So, I will just say y equal to 0, the Cauchy-Schwarz is trivially satisfied.

So, assume y is not equal to 0 and in that case I can choose lambda equal to y Hermitian x over y Hermitian y which is equal to the inner product between x and y divided by the inner product between y and y. If I substitute this value of lambda here you see that I get y Hermitian x times x Hermitian y, which is the same as mod of x emission y squared divided by y Hermitian y and this term is also y Hermitian.

The conjugate of y Hermitian x is nothing but x summation y times y Hermitian x divided by y Hermitian 1. So, I will just and in this last case I get mod of y Hermitian x square and one of the y Hermitian y's will cancel and I will be left with y Hermitian y in the denominator. So, just writing it out the above becomes x Hermitian x which stays as it is minus x Hermitian y squared over y Hermitian y minus x emission y squared over y Hermitian y plus x summation y squared divided by y Hermitian y.

So, things I have used here, so this is greater than or equal to 0 and these two terms obviously cancel and so which implies that x Hermitian x minus x Hermitian y squared over y Hermitian y is greater than or equal to 0, from which the inequality follows. All I have to do is to take this to the other side, bring y Hermitian y to the top and then take the square root throughout.

So, I used a couple of things here one is that x summation y complex conjugate equals y Hermitian x, this is one, I guess this is the only side result that I use and you can see that, you can, this is not difficult to see, so you can see that this is true. Just write out what it is and, so just keep in mind that y Hermitian is equal to y transpose complex conjugate. So, that is the proof of this inequality. We will stop here.