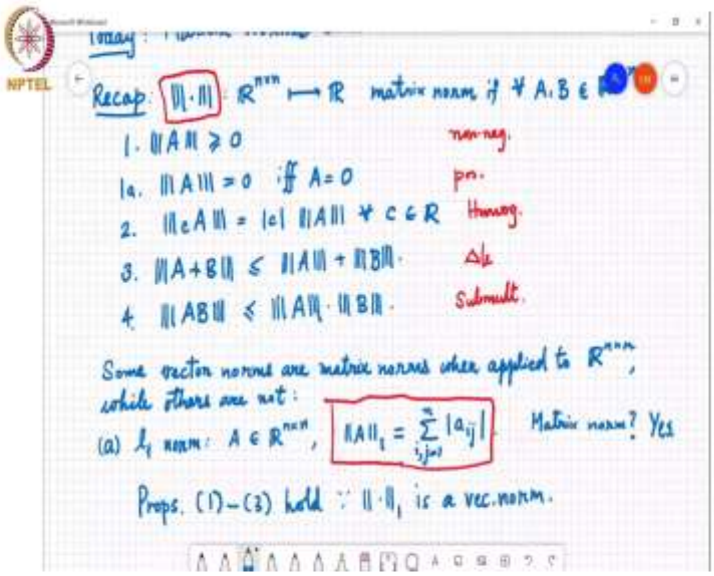
Matrix Theory
Professor Chandra R. Murthy
Department of Electrical Communication Engineering
Indian Institute of Science Bangalore
Matrix norms: Properties

So, last time we were discussing norms and matrix norms in particular. Today we will continue our discussion on Matrix Norms.

(Refer Slide Time: 00:24)



Just to recall, matrix norm is a mapping from the space of square matrices R to the n cross n to the real number line R and we say that this matrix norm, which is denoted by this symbol with three lines around it, this is a matrix norm if for any A and B in R to the n cross n, it satisfies these four properties. The first is that it is non-negative.

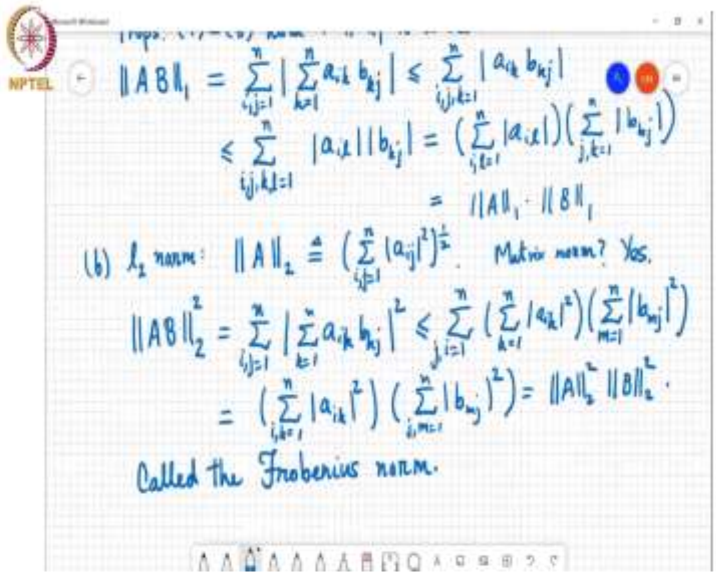
The second is that it is positive, meaning that it can only be equal to 0 if the matrix itself is equal to 0. The third is that it is homogeneous and the fourth is that it satisfies triangle inequality. So basically, any vector norm I start with will satisfy all four of these properties. So, if I think of an n cross n matrix as a big vector containing n squared entries in it and I define a vector norm on that, that will certainly satisfy these four properties.

The only question mark is whether it satisfies this last property, which is called the sub-multiplicativity property or not. So, this is what we will examine today. So, some vector norms are in fact matrix norms when you apply them on R to the n cross n, whereas others are not. So we will start with the l 1 norm.

So, remember that the l 1 norm of a vector is the sum of the absolute value of all of its entries. If I simply extend that to an n cross n square matrix, then I would define the l 1 norm of a matrix A to be the summation of all of its entries in magnitude. So the question is this a matrix norm or not? So, the answer is, yes.

Of course, this norm as defined here, does satisfy the non-negativity, positivity, homogeneity and triangle inequality by virtue of the fact that it is a, it is a l 1 norm in a vector space. So, it directly satisfies those four properties and so we only need to check whether it satisfies this sub-multiplicativity property or not. So, I will just write that here. 1 to 3 hold because it is already a vector norm. So, now about sub-multiplicativity?

(Refer Slide Time: 03:22)



So, what I need to show is that if I take any two matrices and I find A B, l 1, this is going to be less than or equal to the l 1 norm of A times the l 1 norm of B, that is what I need to show. Now this quantity a b l 1 is the sum of the magnitude of the entries of A B which is equal to sigma i j equal to 1 to n. So, this is the sum over all the entries of this product matrix.

But each entry in the product is the inner product between a row and a column of A and B respectively. So, I will write that like this. So if you remember we discussed this different ways of writing out a matrix product. This is the way by which we write out every entry of the product matrix a i k times b k j. This is the i, kth entry of this product matrix A B.

Now this quantity itself, I can upper bound by taking the magnitude inside the summation, this is less than or equal to summation. Now because the summation will go over i j and k, I will just write them together i j k equal to 1 to n, mod of a i k , k b k j. This in turn is less than or equal to. What I will do is, say here it is index k and then there is k repeating here, I will just replace this k with an l and take the summation l going from 1 to n as well.

What I am doing there is I am introducing a whole bunch of non-negative terms into the summation and so that can only increase its value, it cannot decrease it. So, I will write this says sigma i j k l equal to 1 to n mod of a i l, mod b k j. So, this is a double summation over two indices, just make this a little neater i j k l equal to 1 to n.

Now, this is just a, so notice that in the first term, it depends on i and l, it has no j and k in it, the next term has j k in it, but no i or l. So, this is actually equal to the product of these two terms, sigma i l equal to 1 to n mod of a i l times sigma j k equal to 1 to n mod of b k j and this is nothing but by our definition here, this is nothing but the l 1 norm of a times the l 1 norm of below.

And thus A B l 1 is less than or equal to the l 1 norm of A times the l 1 norm of B. So, it satisfies sub-multiplicativity at and hence, this A l 1 as defined here is indeed a matrix norm. So, we can go to the next possibility. So, we did l 1 now we can look at l 2.

Student: Sir?

Professor: Yeah.

Student: The notation for matrix norm will have three vertical bars. Why have you ignored the one, one of the bars?

Professor: Yeah, good question. So, it turns out that I am going to define the l 1 matrix norm a little differently in a few minutes. Once we discuss something called induced norms. So, it turns out that this is not quite the definition of the matrix l 1 norm that I am interested in and so I have used a different notation.

So, for these norms that I am discussing here, I will use only two bars to distinguish this norm from another notion of l 1 norm on matrices that I am going to define momentarily.

Student: Okay, sir.

Professor: So, the l 2 norm also, I am going to denote it with two bars, because again, this is not quite the l 2 norm on matrices that I am going to later be interested in. But I will define this to be.

TA: Sir, there is another question.

Professor: Yeah.

Student: Sir, in the previous one, you have written like a i k, b i k and you have introduced a i. So, what does it mean? I mean, for every a i k, you are replacing it with bunch of a i l? I mean, you are adding a is.

Professor: Yeah.

Student: So, a i k and you are adding some other like a is, a i 1 is it?

Professor: Yeah, there are many a i. I am adding n into n minus 1 terms into the summation, which will only increase its value, but they are all in magnitude. So, none of them is negative. So, it can only increase the value.

Student: I got it, thank you.

Professor: So, the l 2 norm I defined like this. So, it is so just like the vector l y l 2 norm, which is the square root of the sum of the squares of all the entries, I defined the l 2 norm to be the sum of the squares of all the entries in the matrix and then you take the square root. So, is this a matrix norm.

So once again, by virtue of the fact that it is, this is the l 2 norm of the vectorised version of the matrix. This is this will satisfy the non-negativity positivity, homogeneity and triangle inequality. Again, the only thing we need to look at is whether it satisfies the sub-multiplicativity property or not and again for this norm, it turns out the answer is yes. So, we will see that very quickly.

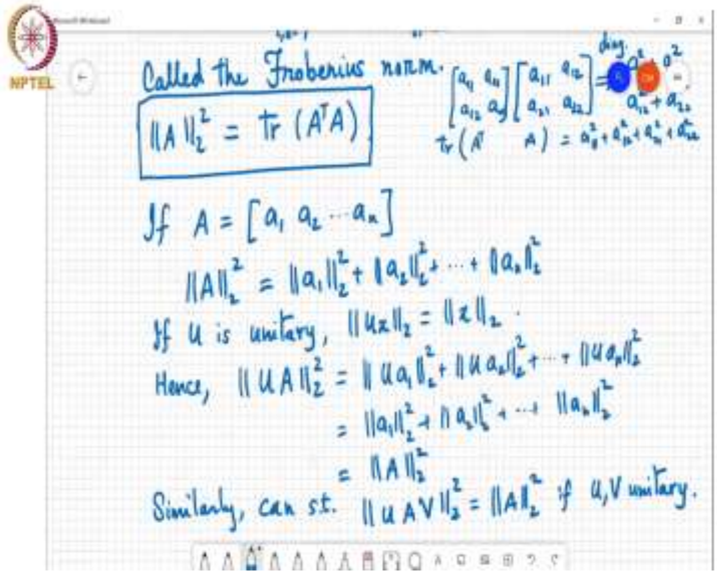
So, once again if I take the l 2 norm of a product, A B, this is equal to sigma i j equal to 1 to n, so I will consider the square so that I do not have to keep writing square roots. So, this A B squared, if I show that this is less than or equal to l 2 norm of A squared times l 2 norm of B squared, that is also good enough for me. So, this is i j equal to 1 to n, I have to take a mod square of this term here.

So, I will just write the same term here, sigma k equal to 1 to n, a i k b k j square and now this is actually the square of the inner product between the ith throw of a and the jth column of B and I can use Cauchy Schwarz inequality here and write this as less than or equal to sigma i equal to i j equal to 1 to n sigma k equal to 1 to n mod a i k square times sigma, just to just for convenience, I will replace k with m and write m equal to 1 to n, mod of b m j square. This is just from Cauchy Schwarz inequality.

Then notice that when I take this, if I take this the, if I look at these two terms, this term has a summation over k, but it has no dependence on j and this term has a summation over m but it has no dependence on i. So, I can write this as, this is equal to sigma i k equal to 1, equal to 1 to n mod a i k square times sigma j m equal to 1 to n mod b m j square.

These two are exactly equal, just another way of writing this double summation with the summation inside of it and so this is equal to A 2 squared times B squared. So, it does satisfy this, sub-multiplicativity property. This particular matrix norm is actually called the Frobenius norm. So, I will maybe mentioned just one property of this norm and the property is that this particular norm is invariant to left or right multiplication by a unitary matrix.

(Refer Slide Time: 14:03)



So, one useful trick in this in this direction is that this norm A l two squared is actually equal to the trace of A transpose A. So, if you think about the entries of A transpose A along the diagonals, you will get the sum of the squares of each column, the entries of each column of A. So, I will because we are doing this online.

I just maybe show you that very easily, very quickly. If I have a11, a12, a21 a22 and if I take its transpose and left multiply it, that would be a11 a21 a12 a22 and then if I look at the diagonal entries, they will be the first diagonal entries will be a11 square plus a21 square. The second diagonal entry will be a12 squared plus a22 square and so if I take the trace of this matrix, this is A transpose A.

So, if I take the trace of this, this is equal to a11 square plus a12 square plus a21 square plus a22 square and so it is exactly this Frobenius norm as we defined it here. So, this is one useful formula. Now, what I wanted to say though, is that if you have a matrix A and I write out its columns, a2 up to a n, then if I look at what this a2 squared is, this is actually equal to the sum of the squares of all the entries in this matrix.

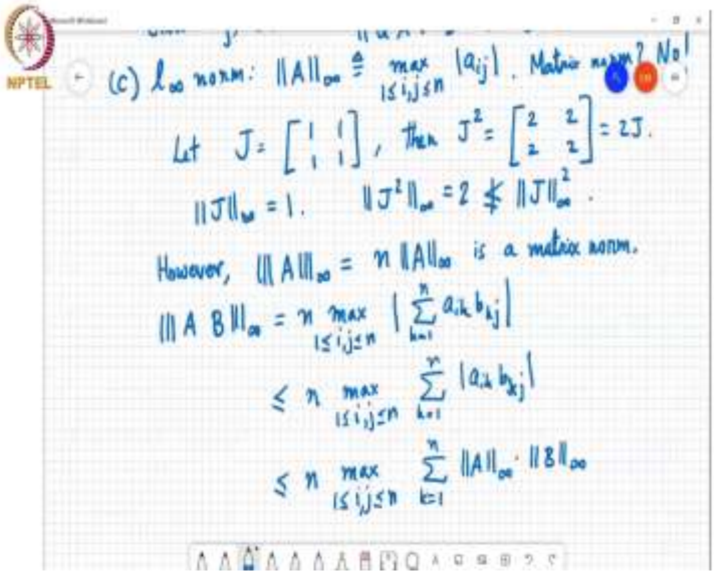
But I can also write that as a1 l 2 norm squared, which is a sum of the squares of the entries in the first column of A plus A 2 l 2 norm square plus plus a n l 2 norm squared. This is another way of writing it. Now, the l 2 norm by itself is unitary invariant meaning that if I have, if U is unitary then U x l 2 norm is equal to the l 2 norm of x itself.

The way to see this immediately is that recall again that the l 2 norm square in the vectors for vectors the l 2 norm square is equal to x transpose x. So, if I take the l 2 norm of U x that is equal to x transpose U transpose U x and U transpose U is the identity matrix. So, that is the same as x transpose x. So, this is true for any unitary matrix and for the l 2 norm.

So, we have that if I take the l 2 norm of U times A square then that is equal to by using this formula here it is equal to U a1 l 2 norm square plus U a2 l 2 norm square plus etc plus U a n l 2 norm square. So, now, I am using the column view of matrix multiplication, when I multiply U with a matrix A the columns of the product are equal to the multiplication of U with each of the individual columns of A.

And so this is because this is equal to a1 square plus, this is equal to l 2 norm of a1 square, this is equal to the l 2 norm of a2 square plus etc plus the l 2 norm of an square, which is equal to the l 2 norm of A square. So, this Frobenius norm is invariant to left multiplication by a unitary matrix. You can similarly show that if I have, if I take you U A V l 2 norm square this is equal to A l 2 norm square when or i will say if you U V are unitary. So, basically this Frobenius norm is invariant to left or right multiplication by a unitary matrix.

(Refer Slide Time: 19: 47)



So, the next thing I want to consider is the l infinity norm looked at l 1 l 2 and l infinity. So, we are basically covering most of these l p norms for p not equal to 1, 2 or infinity, it is a little more difficult to figure out what is going on. So, we do not and usually in applications, you do not encounter those norms.

So, we do not worry about too much about those norms. So, this l infinity norm, if I simply extend the definition, it would be the max 1 less than or equal to i j less than or equal to n, mod a i j the l infinity norm of a vector is the max magnitude entry and so I am extending that and saying the l infinity norm of a matrix is the max magnitude entry across the matrix. Is this a matrix norm?

So, for this particular example, it turns out that the answer is no. So, there is a, so once again, when you want to show that it is not a matrix, now, all you need to do is to provide one counter example, where it does not work and that is enough. So, if I let the matrix, so if I consider J equal to the all ones 2 cross 2 matrix, then J square, what is J squared equal to it is the Alto matrix, which is equal to two times J.

So, if I look at the l infinity norm of J, as per this definition, it is the max magnitude entry, it is 1, but if I look at the l infinity norm of J square, it is 2 which is not less than or equal to the l infinity norm of J square, this was one of the requirements, I mean this see if sub multiplicativity were to hold, then the l infinity norm of J squared must be less than or equal to the l infinity norm of j times the l infinity norm of J, which is l infinity norm of J squared, but that is equal to 1 here, not bigger than 2 and so, this is as written here.

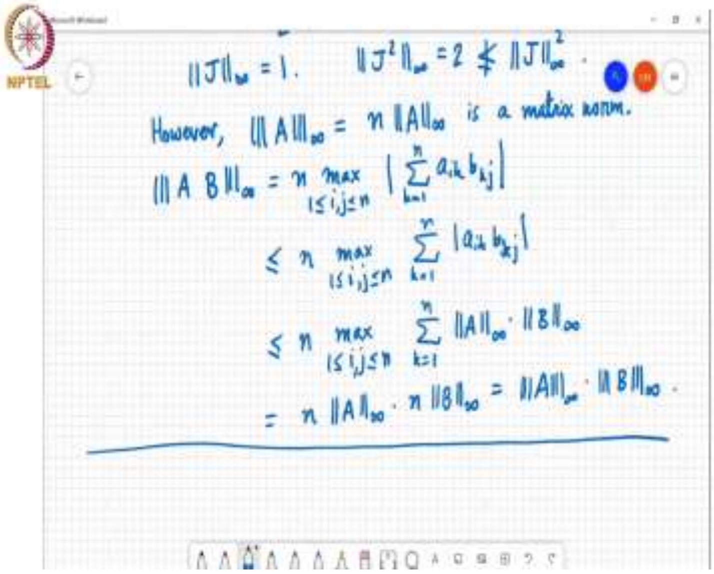
This is not a matrix now. However, a slight modification to this norm, if I define, now I am using three bars, to distinguish it from what I wrote about, if I write it to be n times A infinity, n is the dimension of A, if A is an n cross n matrix. This is a matrix norm. So, if I consider, so once again, because it is just a scaled version of a vector norm, the first four properties namely non-negativity positivity, homogeneity and triangle inequality are naturally satisfied by this.

So, the only property we need to check is the sub multiple creativity property. So, if I consider A B infinity, then this is equal to n times the maximum entry in magnitude of the i jth entry of A B and as we wrote above, this is equal to sigma k equal to 1 to n a i k b k j. Now, first of all, I can, if I read, if I take the modules inside the summation, I cannot decrease the value of the summation.

So, it is less than or equal to n max 1 less than or equal to i j less than or equal to n of sigma k equal to 1 to n mod a i k b k j. And that in turn is less than or equal to. So, what I can do is, I can replace all these a i ks with the largest a i ks value the largest magnitude entry in the entire matrix that will only increase the value of this solution, I can replace this b k js with the largest magnitude entry of the matrix B that will also further only increase the value of this summation.

And so that is less than or equal to n max 1 less than or equal to i j less than or equal to n of sigma k equal to 1 to n A infinity, where this is the maximum magnitude entry, which is the norm A infinity with two bars as I defined it here times B infinity and now there are n such terms here, so I get an extra factor of n by removing this summation here.

(Refer Slide Time: 25:39)



And so this is equal to n times A infinity times n times B infinity which in turn is equal to the norm as I defined here times the norm as I defined here. So, with a small modification to the definition I can get, I can get a definition which is indeed a matrix now. Any other, Any questions so far.