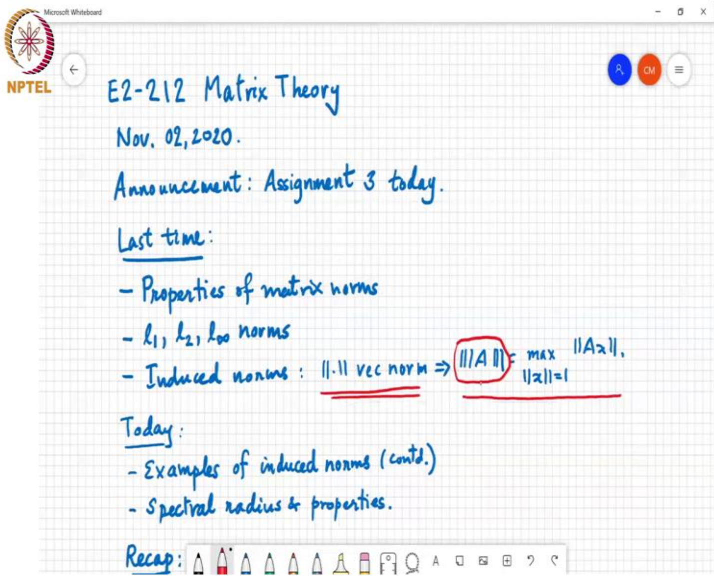
Matrix Theory

Professor. Chandra R. Murthy
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore

Lecture 22

Induced norms and examples

(Refer Slide Time: 00:14)



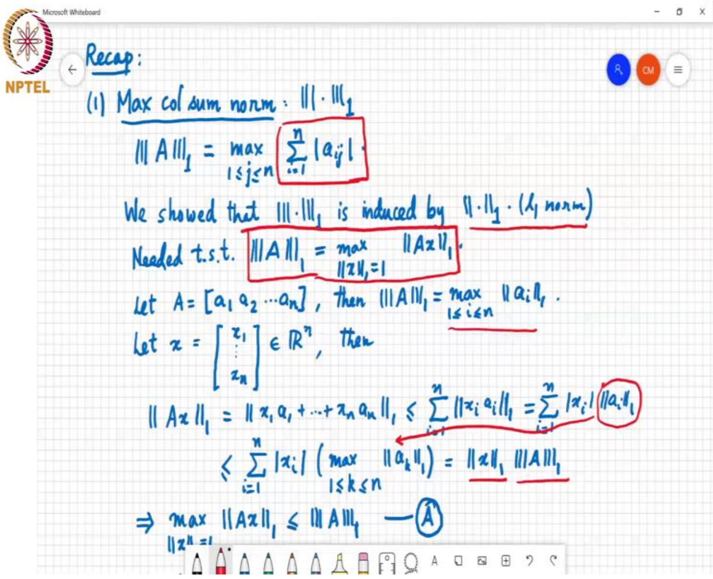
So, the last time we were looking at properties of matrix norms. This is a long chapter in the textbook. So, we will be discussing this for a couple of more classes. So, specifically we discussed l 1, l 2 and l infinity norm and these are three norms that are different from the l 1 l 2 and l infinity norm that we are going to define in this class. These are basically vector, matrix versions of vector norms and in particular, the l 1 norm is the sum of the magnitudes of all the entries of the matrix and that is indeed a matrix norm.

The l 2 norm we defined in the previous class is the sum of the squares of the entries of the matrix to the power half and that is also known as the Frobenius norm and the l infinity norm, if you do n times the maximum magnitude entry of all the entries, the largest magnitude entry among all the entries of the matrix, then that is what we call the l infinity norm and that is also a matrix norm, but these are three different norms compared to what we are going to consider in this class.

So, we discussed about induced norms and the point is that you can start with any vector norm and this could be any vector norm and then if you define the matrix norm to be the maximum value that norm of A x, so A x is a vector here, so you are taking the vector norm of x and you look at the largest value this can take over all x such that, for the same norm, norm of x equals 1.

Now, that quantity, of course, it is not negative and it turns out in fact, we stated and proved the theorem to that effect, it turns out that this is a, this is a matrix norm and so, these are called induced norms, because there is a vector norm underlying vector norm that is inducing a matrix norm and we started discussing about examples of induced norms and today, we will continue and discuss one or two more examples of such induced nodes and then move on to discussing about the spectral radius and its properties.

(Refer Slide Time: 02:40)



So, just to recap what we discussed at the very end of the previous class, we were discussing about induced norms and we particularly discussed this maximum column sum norm and this is what I am writing with three bars and a 1 next to it and this is defined as the maximum column norm among all the columns in A, in the matrix A, written in terms of the entries of the matrix, what you have to do is to take the sum of all the entries of a given column, the magnitude entries of a given column, say the jth column and look which column among all 1 to n columns gives you the largest value and that is what we defined to be the maximum column sum norm of the matrix A.

And there are two ways to show that this is indeed a matrix norm. The first way is to go from first principles, start with the definition of the matrix norm and show that it satisfies the four properties we need for something to be a matrix norm. Namely, the non-negativity positivity, homogeneity, triangle inequality and sub multiplicativity that is five properties, but non-negativity and positivity are lumped as the first property that is the first way.

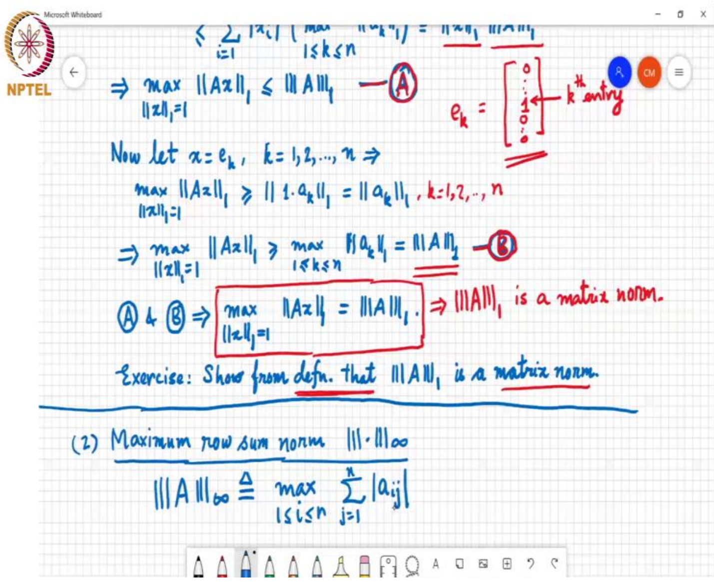
The second way is to show that this norm, whatever we have dysfunction that we are writing here is induced by some other vector norm and so in other words, what we want to show here, so the claim is that this norm is induced by the vector l 1 norm. So, we need to show that the whatever we have defined here is in fact, equal to the maximum of the l 1 norm of A x over all vectors, such that the l 1 norm of x equals 1.

There are multiple ways to show this, but here is one easy way. So, if a1 to a n are the columns of A, then clearly by definition, the maximum column sum of A is the maximum among all 1 to n, i going from 1 to n of the l 1 norm of the eth column of A. Now, if you define a vector x 1 to x n and you look at what norm of A x l 1 is, that is just expanding it out in terms of the columns, it is x 1 a 1 plus etc up to x n a n extend a n l 1 norm.

Then I use triangle inequality of the elbow norm, the l 1 norm is a vector norm and therefore, it satisfies triangle inequality and I take the norm inside the summation. So, I am left with, so I get a less than or equal to summation i equal to 1 to n, the l 1 norm of x i a i, x i here is just a scalar. So, I can bring that out and write it as mod x i times the l 1 norm of a i by the homogeneity property of the vector norm and then I can further upper bound this by replacing all of these guys, norm a i l 1 with its, with the maximum value, which is the max 1 less than or equal to k less than or equal to n, the l 1 norm of A k.

Now, this has no longer depends on i, so it can come right out of the summation and the summation i equal to 1 to n mod of x i is nothing but the l 1 norm of x times, so and then this quantity here is by definition the maximum column sum norm of the matrix A.

(Refer Slide Time: 06:28)



So, the upshot of all this is that, if I am trying to maximize the l 1 norm of A x subject to the l 1 norm of x equals 1, I can upper bound that by the maximum column sum norm of the matrix A. So, this we called as expression A. Then, conversely if you choose specific values of the vector x, specifically if I take e k, e k is a vector with zeros everywhere except 1 in the kth position,

Or you can also think of it as the kth column of the n cross n identity matrix, then if I look at the maximum of this, this e k is a vector which satisfies the constraint the l 1 norm of e k is always equal to 1, as you can see from here and so this maximum, where I am looking across all possible vectors, that has to be at least equal to its value at one particular point, which is e k.

And so, if I substitute e k, then that will only pick off the kth column of A and I will be left with the l 1 norm of A k. So, the maximum over all x such that l 1 norm of x equals 1 of the l 1 norm of A x is at least equal to the l 1 norm of A k for k equal to 1 to n. So, since this is a lower bound for all k, it is of course, if I take the maximum of these numbers over all k from 1 to n, that will still be a lower bound, because n is a finite number here.

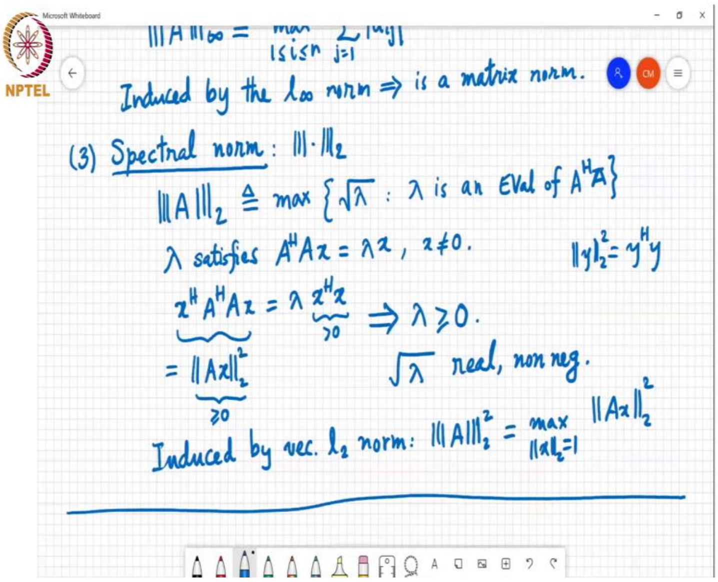
So, we have that the maximum l 1 norm of x equals 1 of the l 1 norm of A x is at least equal to the maximum of the l 1 norm of the columns of A which again, by definition is the maximum column sum norm of the matrix A. So, this is what we call inequality B. So, what have we shown, we have shown that this maximum value of whatever this is, is at least equal to the l 1 norm of A and at most equal to the l 1 norm of A.

So, if that is the case, then it must mean that the maximum of this the l 1 norm of A x subject to l 1 norm of x equals 1 must be equal to the maximum column sum norm of the matrix A and thus the, this norm that we just defined is indeed induced by the vector l 1 norm and therefore A l 1 is a matrix norm.

So, the exercise that I closed the previous lecture with was to show from definition that this is l 1 norm of A, the maximum column sum norm of A is indeed a matrix norm. So now we will continue with a couple of more examples. This is the maximum row sum norm. So, this is very similar to the maximum column sum norm, except that instead of taking the l 1 norms of the columns of A, we will now take the l 1 norms of the rows of A, so if we write A infinity, so I am fixing an i.

So, that is fixing a particular row of the matrix and I am taking the l 1 norm of all the entries in that row and then I am looking for which row from among the rows 1 to n gives me the biggest value and that is what I am defining to be A infinity.

(Refer Slide Time: 11:06)



You can see that this is actually equal to the maximum column sum norm of A transpose. So by that itself, you can see that this is matrix norm. But it is also induced by the vector l 1 norm, by the vector l infinity norm and therefore, it is a matrix norm. Showing this is very similar to the column sum norm. So, I would not do that here.

But we will just move on directly to the next norm, which is the spectral norm. The spectral norm is by far the most important norm that we will look at in this course and we write it like this and it is defined as the A 2 is equal to the maximum square root of lambda such that or the value lambda is an Eigen value of, I will write it for the complex case. But of course, if it is real, this Hermitian would be replaced by transpose. But it is A Hermitian A .

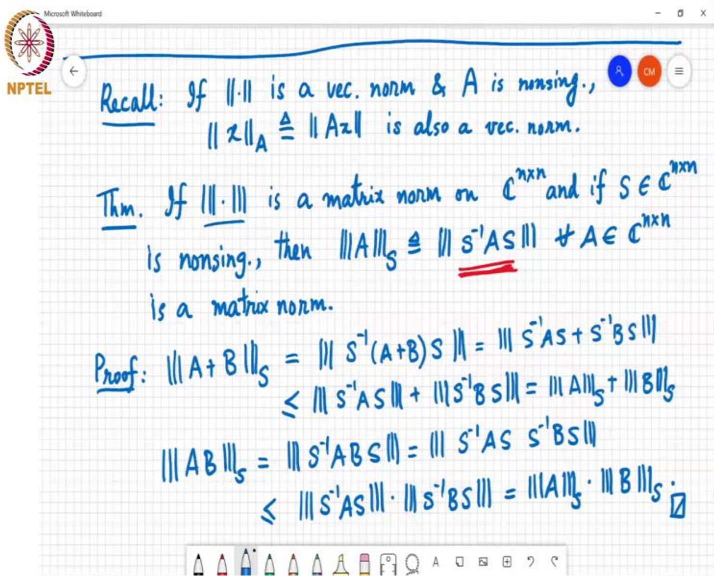
So, we have not formally discussed Eigen values. But unfortunately for these, for many of these results, we will need the notion of Eigen values. But we will connect all of these together later in the course when we discuss Eigen values also. So basically, lambda is a quantity that satisfies A Hermitian A x equals lambda x for sum x naught equal to 0 and so one thing to note here immediately is that if I pre multiply this by x Hermitian, I have x Hermitian A Hermitian A x is equal to lambda is just a scalar. So, I can pull that out and write it as lambda x Hermitian x.

And this quantity x Hermitian A Hermitian A x is actually equal to the l 2 norm of A x square. So, we have seen that already that if I have a vector y, y 2 squared is equal to y Hermitian y. This is another way to write the, it is just the sum of the magnitude squares of all the entries of y, but we can also write that as y Hermitian y.

And so this is, so therefore this quantity is real and positive and this quantity is also real and actually non-negative. So, both these are real and non-negative and so you cannot suddenly have lambda being a complex number and in fact, since both are real and non-negative and in fact, since x is not equal to 0, this is strictly greater than 0 and this is greater than or equal to 0, which implies that lambda is always non-negative.

So, basically square root of lambda is always real and non-negative. Now, this particular norm is in fact induced by the vector l 2 norm. So, that A l 2 squared is actually equal to the max over all x l 2 equals 1 of A x l 2 square. So, this is the Spectral law. So, these are the three examples I wanted to discuss about induced norms. So, the maximum column sum norm, the maximum row sum norm and the spectral norm. We will discuss the spectral norm more later, but first before that. Recall that for vector norms we know that if you if you are given an arm and a non-singular matrix, if you define a new norm to be the norm of.

(Refer Slide Time: 16:43)



So, let me just write that we have seen this property before, if is vector norm and A is non-singular, then if I define this to be the vector norm of A x, then this is a vector now also. So, basically given a, if we know of a particular vector norm, then given any non-singular matrix we can define a new vector norm. Similar there is a similar result for matrix norms.

So, we have this theorem, if is a matrix norm on the space of n cross n matrices and if S is non-singular, then if I define A S to be the matrix norm of S inverse A S for any A in C to the n cross n this quantity A S is a matrix norm. So, this quantity S inverse A S is what is called a similarity transform on the matrix A and it has lots of very nice properties, which we will actually study in quite detail later on.

But for now, we are just observing that if you are given a matrix norm and any non-singular matrix, you can define a new matrix norm using that non-singular matrix. So, how do you show this, the proof is very simple. Of course, you know properties like homogeneity, non-negativity positivity and triangle inequality directly follow from the properties of this matrix, this matrix norm here and so, the only interesting thing we need to show is the sub multiplicativity.

So, for example, I mean just to make my point, just to make my point. So, if you had, if you took take A plus B and then you are looking at this S norm, this is equal to the matrix of S inverse A plus B times S which is equal to the matrix norm of I can take this S and S inverse inside the brackets, so that it is S inverse A S plus S inverse B S, which now this, this is a matrix norm.

So, it will satisfy the triangle inequality and I can write it as S inverse A S plus S inverse B S which is actually equal to the S norm of A plus the S norm of B. So, it satisfies triangle inequality. So, how about the sub multiplicative at that is also very easy. So, if I look at A B as norm, then that is equal to the matrix norm of S inverse A B S and I can just insert an S S inverse in between here because SS inverse is just the identity matrix.

So, this is equal to S inverse A S S inverse B S. Now, I use the fact that this is a matrix norm and so it satisfies this multiplicative property. So, it is less than or equal to S inverse A S times the matrix norm of S inverse B S which is nothing but the S norm of A times the S norm of B.

So, we now know several ways of coming up with different different matrix norms, one way is you start with any vector norm that you know and then look at the induced norm. The other way is start with any matrix norm you know and consider a convenient non-singular matrix and then you do S inverse A S and take the norm of that and that becomes a new norm. So, this great richness in the types of a number of different types of norms that you can construct.