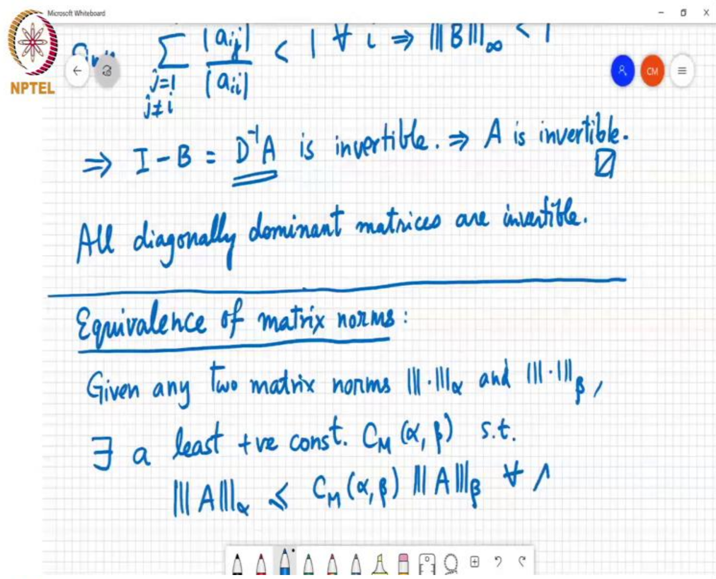
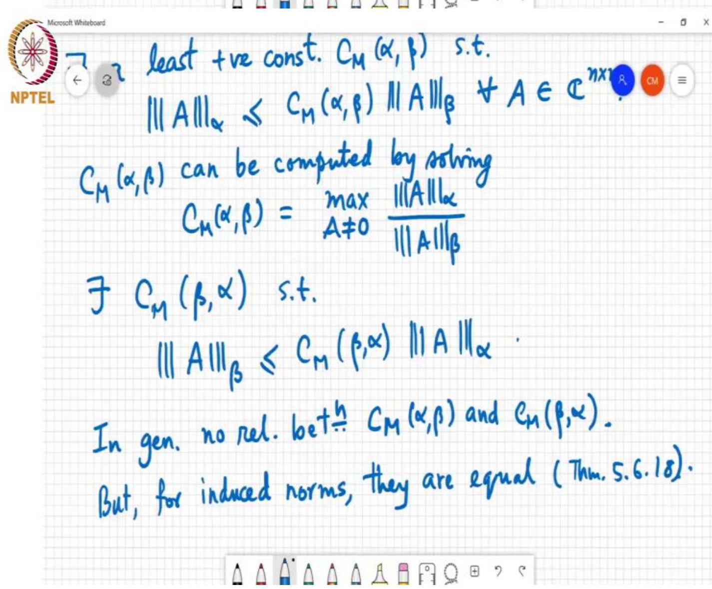
Matrix Theory
Professor Chandra R. Murthy
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore
Equivalence of Matrix Norms and Error in
Inverse of Linear System

(Refer Slide Time: 00:14)





So, equivalence of matrix norms basically says that, if you have two different norms you can bound the norm of any matrix with respect to the first norm in terms of the norm of the matrix with respect to the second norm. And this is useful again because you may be interested in showing convergence of certain algorithms, and in these algorithms you will get a sequence of matrices.

And once again you, we have seen that, if you are able to show that the sequence of norms of these matrices, if that converges, then the matrix itself will converge. And so, we are often interested in showing that the sequence of norms of a matrix converges, but then for different problems it may be more easy to take different types of norms and show that, that particular norm converges.

But that may not be the norm in which you are actually executing this optimization problem that you are interested in. Fortunately, for us, these norms are all equivalent. So, if one particular norm is converging, then every norm will converge, maybe not to the same value, but to some other value. And so, it is useful to have this kind of result on equivalence of matrix norms, because it means we can use any norm that is convenient for us in trying to show convergence properties.

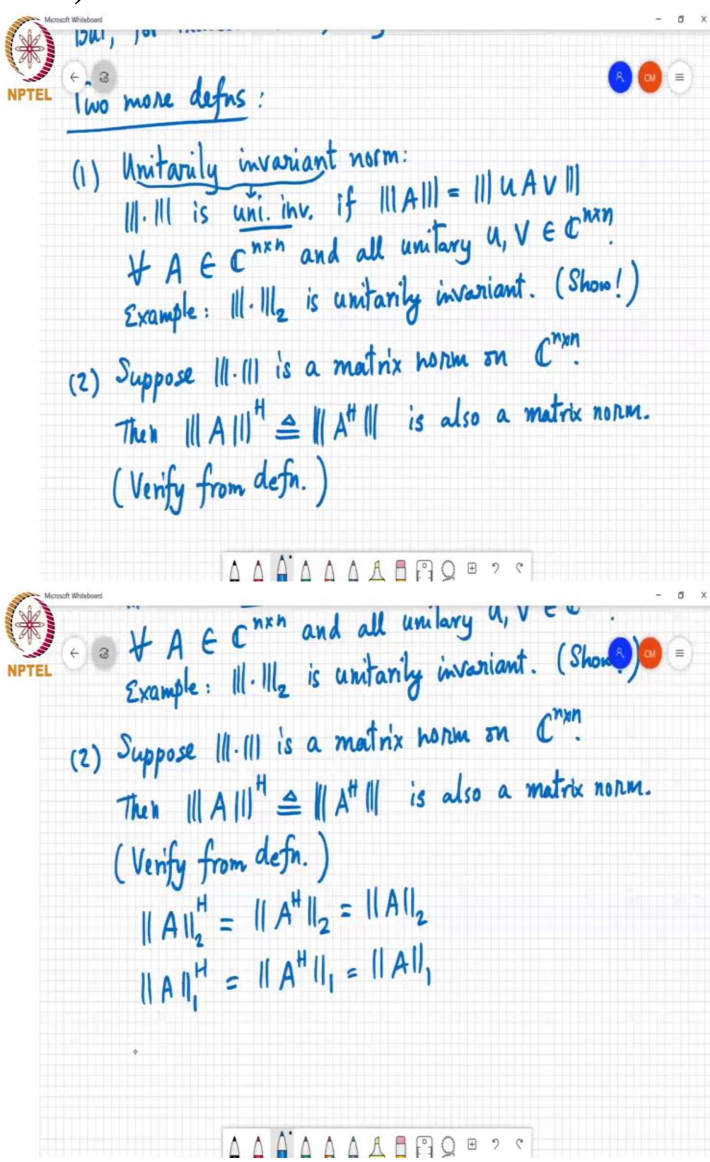
So, so, basically given any two norms, two matrix norms alpha and, then there exists a least positive constant, CM of alpha beta such that, the alpha norm of A is less than or equal to CM of alpha beta times the beta norm of A for every A. In fact, CM of alpha beta can be computed by solving CM of alpha beta is equal to the maximum over all A not equal to 0 of norm A alpha divided by norm A beta.

So, you take this ratio and you find the biggest number it can take, then of course, it must be true that norm A alpha is less than or equal to CM of alpha beta times norm A beta for any other A. Because that is like a sub optimal A, that you are choosing which must satisfy this inequality. So, by, if I, in this, I just said alpha and beta and the, there is no, I can always just exchange alpha and beta.

And the way to say that is that there exists this, this CM of alpha beta there exists similarly a CM of beta, alpha, such that A beta is less than or equal to CM of beta alpha, times A alpha. And this CM of beta alpha is computed as the maximum of norm A beta divided by norm A alpha. So, that is a different, completely unrelated sort of different optimization problem. So, it is not the I mean, so basically in general there is no relationship between CM alpha beta and CM of beta alpha.

But for induced norms, that is both alpha norm and beta norm must be an induced norm, they are equal. So, this is actually theorem 5.6.18. Now, the textbook has, a lot, a lot, a lot of theorems I obviously cannot cover all of them. But where, where appropriate, I will indicate some theorems. Mainly, I will focus on stating and proving theorems that we will actually try and use later in the course. But this is just an interesting result that is there and in the text that you can go take a look if you are interested.

(Refer Slide Time: 06:05)



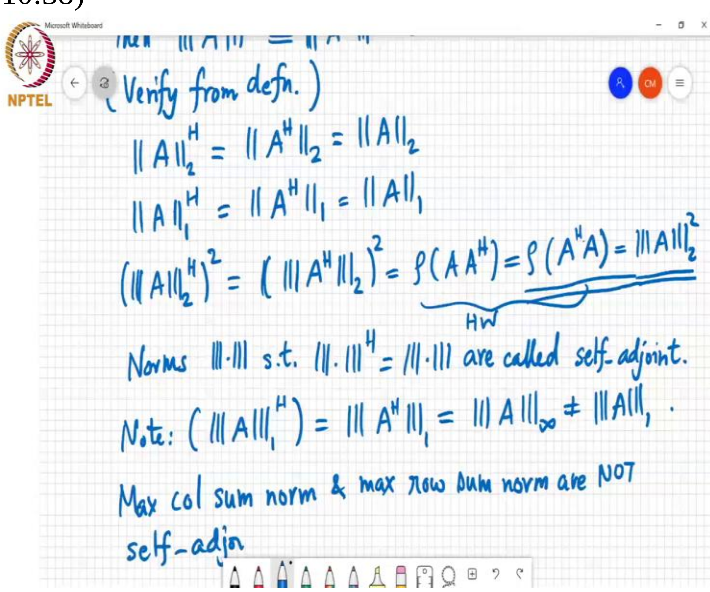
Now, there are just two more definitions. The first is the notion of a unitarily invariant norm. So, we say that this norm is unitarily invariant. I will write this in short here, this is the same as this statement here. If norm of A equals norm of u A v for every A belonging to C to the n cross n and, C to the n cross n. So, one example is the spectral norm, is unitarily. This is a small exercise, show this, start from the definition and you will be able to show that the norm of u A v spectral norm is equal to the spectral norm of A, for any A and all possible unitary matrices.

The second definition is, is that. So, suppose, is a matrix norm on C to the n cross n. Then so, this is a new function I am defining with an H on top, which is defined to be the norm of A hermitian is also a matrix norm. Now, this norm that I am defining here, may not be an induced norm, it is true for any norm, any matrix norm that I can define on C to the n.

If I instead of computing the norm of A, if I say the norm of A is this, whatever norm I have defined, operated on A hermitian, that is also a matrix norm. It is actually straightforward to show this from the definition. So, in particular, if you take the frobenius norm, this is the square root of the sum of the squares of all entries in the matrix A. This is equal to by definition, the A hermitian frobenius norm.

So, this is the hermitian or the H norm of the frobenius norm, which is the frobenius norm of A hermitian. Of course, a conjugate transpose does not change the sum of the squares of all, of the magnitudes of all the entries in A. So, this is exactly equal to A 2. And similarly, we defined this with two bars, A1 to be the sum of the magnitudes of all the entries of A. So similarly, if I define this H norm to be the, one norm of A hermitian, then again, taking the conjugate transpose does not change the magnitudes of the entries in A. So, this is also equal to A1. So, this is what we call the l1 norm and the spectral norm.

(Refer Slide Time: 10:38)



So, if I take the H of the spectral norm, I will work with the square because that is easier. This is equal to, I have to compute. So, when I take the H norm, I have to compute the spectral norm of A Hermitian instead of the spectral norm of A. And then I have to square it. Now, the spectral norm of A, spectral norm of a matrix is the square root of the largest Eigen value of A hermitian A.

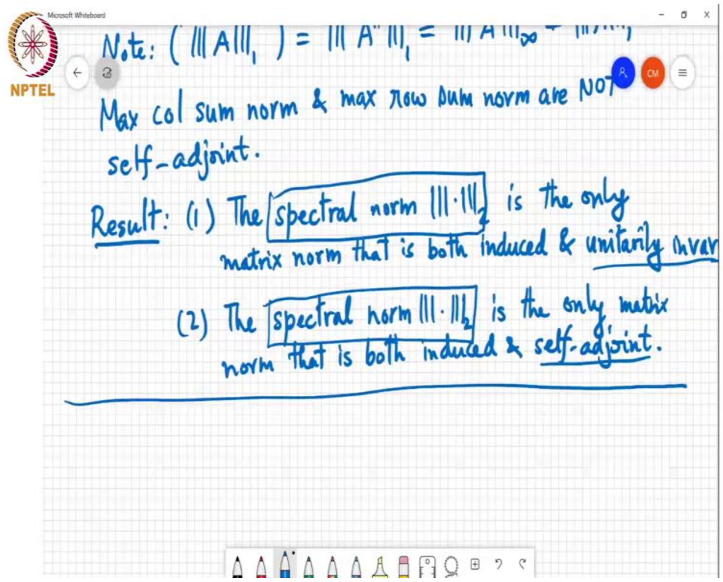
And so if I apply that to this, to A hermitian, I get that this is equal to the square of this is going to be row of, the spectral radius of AA hermitian. And this is another result that you have to show in your homework. That row of AA hermitian is the same as row of A hermitian A, which is equal to the spectral norm of A square. So, I am slipping in two different things here.

One is that I am showing that, the spectral norm of A is invariant to, while doing this H operation, the H norm of the spectral norm is the spectral norm itself. And the second thing is I am saying that the spectral radius is, has a nice relationship with the spectral norm. In that, the spectral norm squared norm A2 squared, spectral norm squared is exactly equal to the spectral radius of A hermitian A.

So, that is the relationship between the spectral radius and the spectral norm. That is A relation between the spectral radius and the spectral norm. So, this is homework. So, you see that here, when I take this H norm, it gives you the same norm, here also when I took H norm it gives, gives me the same frobenius norm. And the H norm of the l1 norm is also equal to the l1 norm. Such norms, for which H is the same as the norm itself are called self adjoint norms.

And by the way note that, if I compute the max column sum norm A H, that is equal to the max column sum norm of H A hermitian, which by definition is the max row sum norm of A. And that is different from A l1. The max column sum norm of A. So, this is not a self adjoint norm. So, max column sum norm and max row sum norm are not self adjoint.

(Refer Slide Time: 14:28)

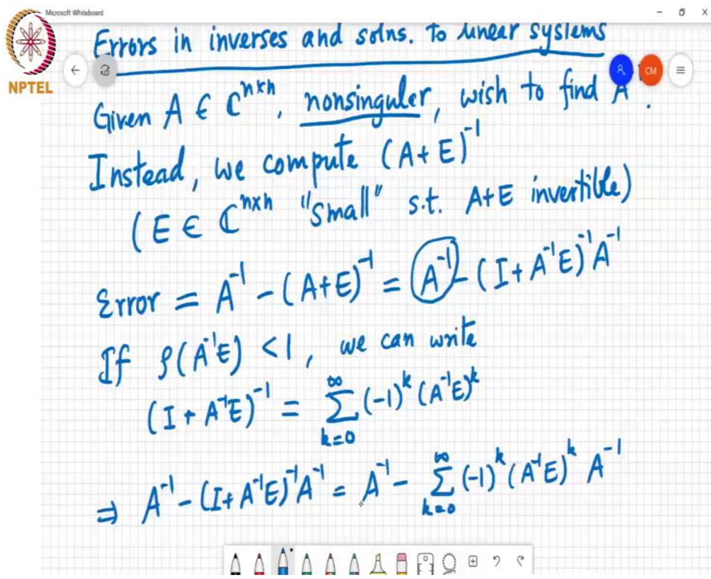


So, there is one very nice result that says that one, the spectral norm is the only matrix norm that is both induced and unitarily invariant. Second property is that, the spectral norm is the only norm that is both induced and self adjoint. The proof of this is in the text. But these are again two special properties of the spectral norm, which is one of the reasons why, and multiplying by unitary matrices or taking the conjugate transpose operations are fundamental operations that arise in many, many signal processing and engineering applications.

And that is why for many of these applications, we are interested in working with the spectral norm. Because it is invariant to these two operations. It is a very special norm in that way. So, it is we still have about twelve minutes in the class. But the next thing I want to talk about is, about some uses of these norms, I discussed several theorems. But now we can maybe talk about solving linear systems.

And how these norms help us for example, in bounding the error in computing inverses or solving linear systems. Now, if I start on the next thing I need about, let me see. So, let us, let us just maybe make a few remarks and see how far we can go.

(Refer Slide Time: 17:31)



Inverses and solutions to linear systems. So basically, if we are given a matrix A, which is non singular, we know that A inverse exists. But when we want to try to compute A inverse, then we may have to compute it on a finite precision arithmetic machine, or maybe we do not get to observe A, we only get to observe a noisy version of A. And then we compute the inverse on a noisy version of A.

It turns out that under appropriate modeling, you can actually generally consider both of these as a system where, or a general simple mathematical model, under which you can consider both these types of errors, is to consider that what you have inverted is some other matrix A plus E. So, given A and C to the n cross n, non singular. So, pay attention to this, this is important, I am starting out by assuming that, the original matrix I wish to invert is actually non singular.

And that is why I am brave enough to try to compute its inverse. And we wish to compute A inverse. So instead, we compute A plus E inverse, where E is an error matrix. So, basically, here E is small, such that A plus E is also invertible. So, the error in, the error I have incurred will be, be this like an error matrix, which is A inverse minus A plus E inverse, which I can write as, I will pull out an A inverse out of this. So, I will write it as A inverse minus I plus A inverse E inverse times A inverse. I am just using the fact that A B inverse is B inverse A inverse, to write this.

Student: Sir?

Professor: Now, Yeah, go ahead.

Student: So, here we are trying to come to the inverse of a non singular matrix. So, this matrix E, is it deliberately added or is it observed? I mean, is it deliberately added to make it invertible?

Sir: No, no, no, if A is non singular to begin with.

Student: Okay.

Professor: So, I can see right quantities like A inverse. Otherwise, this would be a meaningless thing to write, if A could be singular. So, A is invertible, A plus E is also invertible.

Student: Okay.

Professor: It is just that I could not compute, or I did not have A exactly in my hands, what I got to observe was A plus E, where E was a small perturbation on A.

Student: Okay, okay.

Professor: That is that A plus E was also invertible. So, I wanted A inverse, but I have A plus E inverse in my hand, the difference between these two is the error.

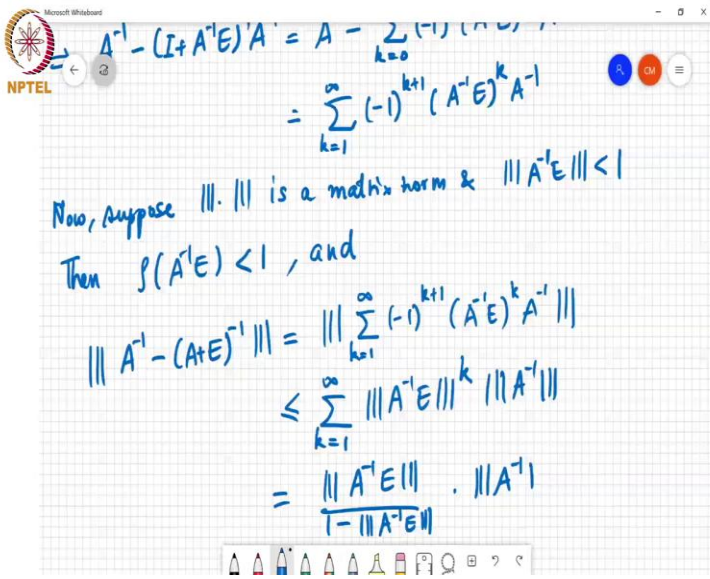
Student: Yes Sir.

Professor: Now, what we have seen is that if the spectral radius of A inverse E is less than 1, then we can write I plus A inverse E inverse is equal to, we can write this as a series, sigma k equal to 0 to infinity minus 1 power k A inverse E power k. So, this is what we have seen, we just recently, we saw this, this result. So, so we can write it like this.

So, now I will substitute this in here, which means that A inverse minus I plus A inverse E inverse times A inverse is equal to A inverse minus sigma k equal to 0 to infinity minus 1 power k A inverse E power k times A inverse. Now, if I take the k equal to 0 term, I will get minus 1 power 0, this thing, power 0 is the identity matrix times A inverse. So, the first term here exactly cancels this A inverse.

And so I will be left with, this is equal to all the other terms k equal to 1 to infinity. And I will observe this minus 1 into this and write it as minus 1 power k plus 1 times A inverse E power k times A inverse. So, this is true if. So, keep in mind my starting point is if row of A inverse E is less than 1.

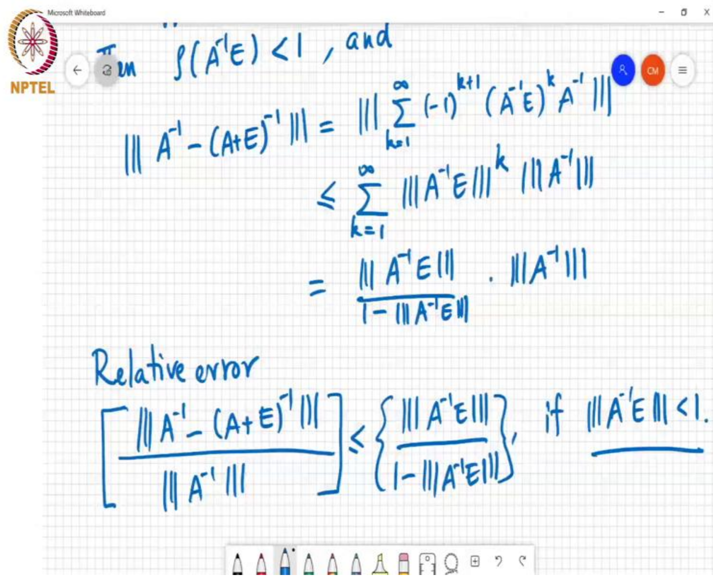
(Refer Slide Time: 23:47)



Now, suppose this is a matrix norm. And A inverse E measured according to this row of A inverse E is also less than 1. And if I compute the norm of A inverse minus A plus E inverse, so that will be equal to the norm of the summation, k equal to 1 to infinity minus 1 power k plus 1 A inverse E power k A inverse, which is less than or equal to, I will take the norm inside and I will apply the sub multiplicativity, sigma k equal to 1 to infinity norm of A inverse E power k times norm of A inverse.

Now, this is not dependent on k. So, it can come out of the summation, and norm of A inverse is less than 1. So, this is summable. And so, I can write this to be equal to norm of A inverse E divided by 1 minus norm of A inverse E times norm of A inverse.

(Refer Slide Time: 25:26)



So, we now know, thus we know that the relative error, if I define it to be norm of A inverse minus A plus E inverse divided by the norm of the guy I wanted to compute. So, this is the relative error in computing A inverse, this can be upper bounded by norm of A inverse E divided by 1 minus norm of A inverse E. If norm of A inverse E is less than 1. That was the assumption we made.

So, we see that norms are useful to help us bound the relative error in computing things like A inverse. And many more uses, which we will see in the next class. Also with respect to (())(26:27) of linear equations, but we will stop here for today. And we will continue on Monday. Thank you.