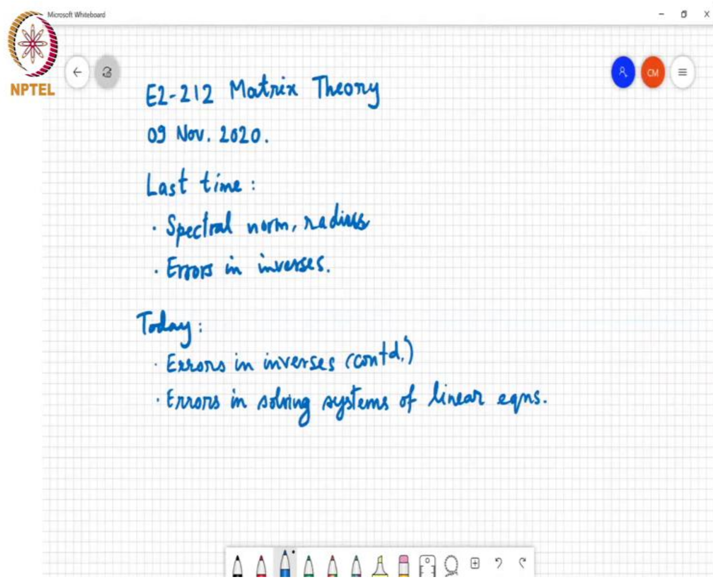
Matrix Theory

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

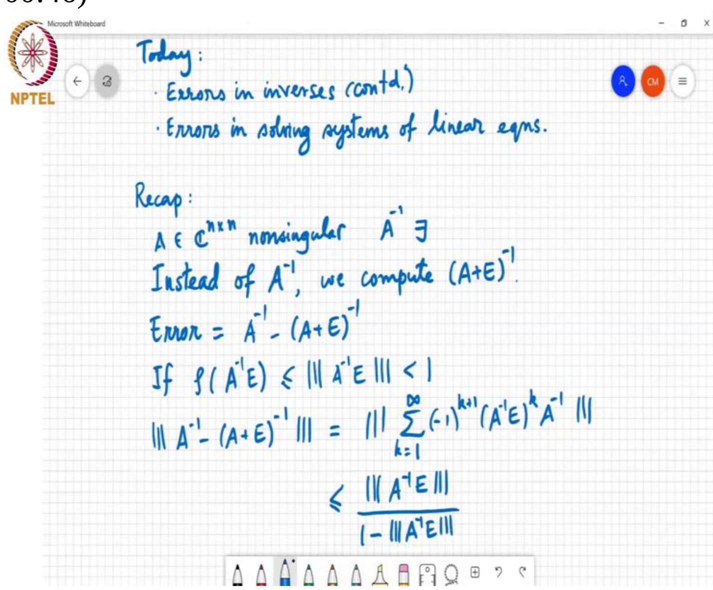
Error in Inverse of Matrices

(Refer Slide Time: 00:14)



The last time we looked at spectral norm, and some of its, and spectral radius and some, r a d i u s, and some of their properties. And we started discussing about the error in inverses. So, today we will continue the discussion about the error in computing matrix inverses, inversions. And we will also talk about the errors in solving systems of linear equations.

(Refer Slide Time: 00:46)



So, just to recap, so we were looking at, in a matrix A and C to be n cross n, which is non singular. And we want to compute A inverse. So, A inverse exists. But because of various reasons, we do not end up computing A inverse. Instead we compute, instead of A inverse, we compute.

Student: Sir your voice is breaking.

Professor: A plus E inverse. So, I understand if my voice is breaking, but I think I am connected to the ISCW WLAN. So, this is the best connection I can get.

Student: No sir your voice is okay actually, maybe you might be.

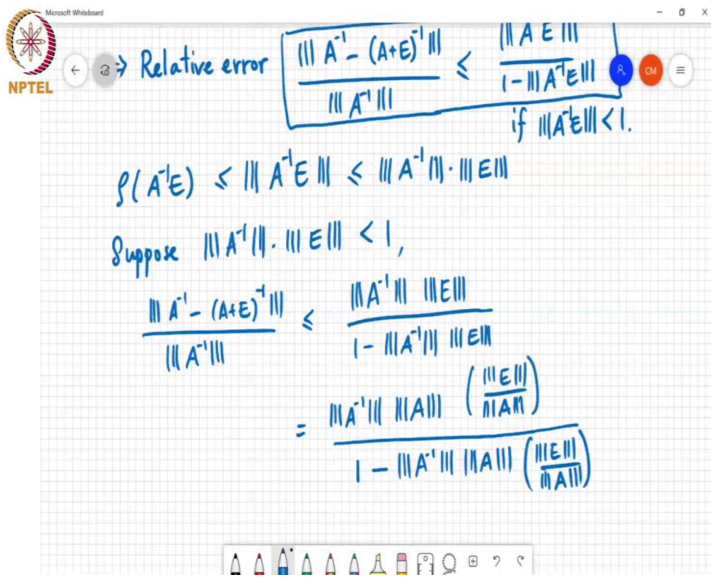
Student: Sir your voice is okay.

Professor: Okay thanks, yeah this is the best network I can get. So, I cannot really do much more than this. So, so the last so, we were looking at the error, which is equal to A inverse minus A plus E inverse. And we wanted to sort of get a feel for how this error, how big this error can be. And we did some algebraic manipulations, which I will not repeat again today, you can go back and look at the notes from the previous class.

But what we showed is that if the norm of. So, let me put it this way, if I look at the spectral radius row of A inverse E, this is going to be less than or equal to the norm of, some norm, it does not matter which one. So, some matrix norm, A inverse E is less than 1, then these infinite summations can be evaluated in closed form. And so, we showed that the norm of A inverse minus A plus E inverse can be written as the norm of the summation k equal to 1 to infinity.

Because the first term drops off, because it cancels with this A inverse term, minus 1 power k plus 1 times A inverse E power k times A inverse. And this by using the sub multiplicative property, we split that as the norm of this thing power k times the norm of A inverse. And so, we can show that this is equal to, actually less than or equal to norm of A inverse E divided by 1 minus norm of A inverse E times the norm of A inverse.

(Refer Slide Time: 04:47)



So, we said that the relative error, which we defined to be norm of A inverse minus A plus E inverse divided by norm of A inverse is less than or equal to the norm of A inverse E divided by 1 minus norm of A inverse E, if norm of A inverse E is less than 1. So, this gives us a way to bound the relative error in computing the inverse of a matrix, in terms of the norm of matrix A inverse times E.

Student: Sir?

Professor: Yes?

Student: Sir, in order to calculate the norm of A inverse E, would not we need to know A inverse?

Professor: You would need to know A inverse and E. So, as written, it is not very useful. So, but, so it will be useful if there is some way you can, in some independent manner, obtain a bound on the norm of A inverse E. So, for example, we have a row of A inverse E is less than or equal to the norm of A inverse E which can be further bounded as the norm of A inverse times the norm of E. And suppose this was also, norm of A inverse times norm of E was also less than 1.

Now, in general, you actually do not know A inverse, and you do not know the, you do not know E. And so, you would not, so but then if you have some other way of obtaining a bound on the norm of A inverse and the norm of E, then you can find their product and utilize something like this. Yeah, but this is the nature of the results in a lot of the error analysis, where it gives you a way to bound the error at least when you know what is the error matrix. So, for example, if you are doing quantization of the entries of A, then you do know what the quantization error matrix is.

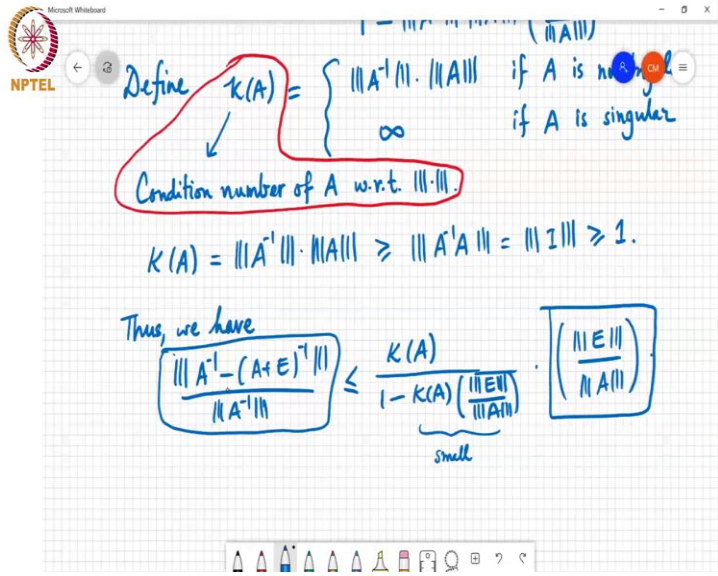
And so, if you knew the correct A matrix and then you know the quantized A matrix, you know the quantized error matrix. The other way these things can be useful is if you know that the error matrix comes from a certain distribution, then you can look at what values this right hand side can take as you take different values, different E matrices from that distribution. And use that to obtain some kind of bounds on this.

So, you will have to bring in probability on top of this and say what is this, what is this bound, can this bound be shown to be less than something with high probability over all possible E matrices. So, you will have to use those kinds of techniques to further make these useful. But, if this is also less than 1, then we can simplify this further. See, this is the, this A, norm of A inverse times norm E is something that is bigger than this.

So, if I substitute norm of A inverse times norm E over here. It only makes the numerator (())(08:42) it only makes the denominator smaller. So, it makes this overall ratio actually bigger. And so, we have that norm of A inverse minus A plus E inverse divided by norm of A inverse is less than or equal to norm of A inverse times norm of E divided by 1 minus norm of A inverse times norm of E.

And I can multiply and divide by norm of A to write this as norm of A inverse times norm of A times norm of E divided by norm of A divided by 1 minus norm of A inverse norm of A times the same thing, norm of E divided by norm of A.

(Refer Slide Time: 10:09)



So, we will define this quantity here to be kappa of A norm of A inverse times norm of A if A is nonsingular, and infinity if A is singular. So, this thing is called the condition number of A. So, if we define it like this. So, so how many of you have heard of this concept of conditional number? Have any of you heard of it before?

Student: Sir, it was discussed in the last tutorial.

Professor: Very good, okay. But other than that in your undergraduate program, has any of you heard of the condition number of a matrix?

Student: Something like the ratio of maximum minimum Eigen values.

Professor: Correct, correct. So, if you take the spectral norm of the matrix, it is the largest magnitude Eigen value of the matrix. And therefore, one other side result is that if you know the Eigen values of a matrix, the Eigen values of the inverse of the matrix are the inverses of the Eigen values of the original matrix. And so, this condition number reduces to the ratio of the maximum to the minimum magnitude Eigen values of that matrix.

So, yes, so, yeah so this is called the condition number. And note that if we consider this condition number K of A, this is for an invertible matrix, it is the norm of A inverse. So, but the more general definition is valid for any norm, it is specific to the norm. So, depending on which norm you choose here, you will get different values for this condition number. And this norm that you are choosing here is the norm under which (())(12:51) relative error over here.

So, this is norm of A times the norm of A inverse, which is greater than or equal to the norm of A inverse A, this is just sub multiplicativity, which is equal to the norm of the identity matrix. And we know that for any matrix norm this is greater than or equal to 1. So, the condition number of, for any matrix A is going to be a number which is greater than or equal to 1, for any matrix norm.

And further because it is lower bounded by 1, K of A even though it is mapping a matrix to the real line, can never be a matrix norm. Because K of A equal to 0 will never happen. And so, it will never satisfy this positivity constraint. So, what is the, so for, in particular, if I take the all 0 matrix, what is its condition number?

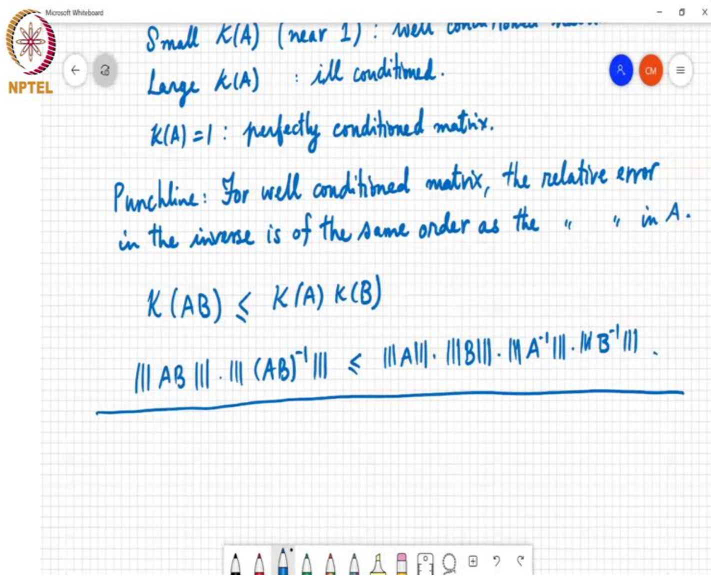
Student: Infinite.

Professor: Infinity, correct. So, so thus we have this norm of A inverse minus A plus e inverse over norm of A inverse is less than or equal to K of A times, I will write it like this, 1 minus K of A times norm E over norm A times norm E over norm A. So, from this we can see a lot of interesting properties. So, for example, if this number is small then we can neglect this term here, it need not be a lower bound anymore, because then you are making this thing smaller.

But assuming that this is small enough that neglecting it does not break this upper bound, you have kappa of A times norm E over norm A, which means that this K of A or the condition number of A, it represents the, it allows you to bound the relative error in computing the inverse in terms of the relative error in the matrix itself. And if K of A is close to 1, that is it is small, then we say that the matrix A is a well conditioned matrix.

Because it only amplifies the norm E over norm A by a small amount, when you compute the inverse. Whereas if K of A is a large number, the matrix A is ill conditioned and, or poorly conditioned, and it will lead, it can lead to potentially a large increase in the relative error or a large value of the relative error in computing the inverse even for small perturbation E.

(Refer Slide Time: 16:40)



So, I will just mention (())(16:39) words that I dropped, small k of A, it is called well conditioned matrix, large K of A say that it is an ill conditioned matrix. And for completeness, if k of A equals 1, we say that it is a perfectly conditioned matrix. So, what is an example of a perfectly conditioned matrix?

Student: Identity.

Professor: Yes. How about unitary matrices? Are they perfectly conditioned?

Student: Yes, sir.

Professor: It depends on.

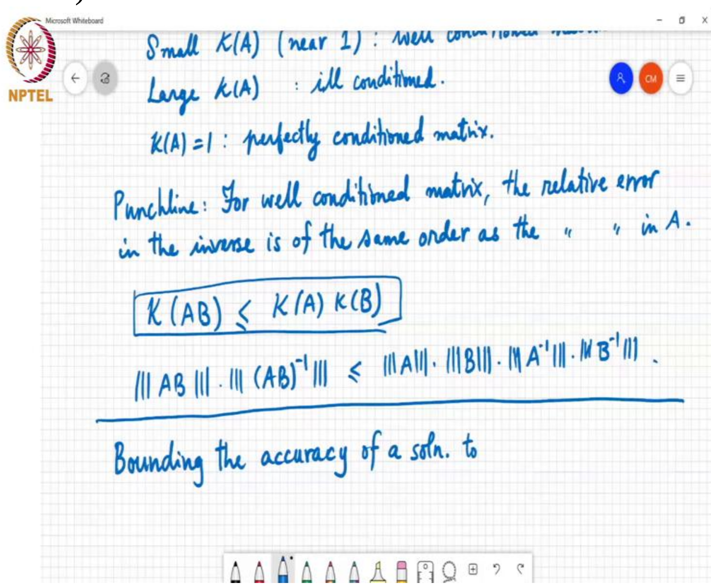
Student: Yes, sir they are perfectly.

Professor: It depends on which norm you choose to use. For unitary matrices, U transpose or u hermitian is the same as U inverse. And so, it becomes K of A becomes norm of U, U hermitian times norm of U. And so, depending on which norm you are using, they could be perfectly conditioned. So, in particular, if you are using the spectral norm, then they will be perfectly conditioned.

So, the punchline is that, for well conditioned matrices, the relative error in the inverse is the same as the relative error in the data. Now, one other property of the condition number is that, if you take the product of two matrices, this condition number is less than or equal to the product of the two condition numbers. That is simply because the left hand side is norm of AB times the norm of AB inverse, which is less than or equal to.

Now, I just use sub multiplicativity, norm of A times norm of B times norm of A inverse times norm of B inverse. So, actually, there is a lot more one can say about this condition number. I will leave some of them as homework exercises. And I will come back to this later in the course, if there is enough time. But right now, I do not want to distract ourselves from discussing about computational errors.

(Refer Slide Time: 21:03)



And so let us now discuss about bounding the accuracy.

Student: Sir?

Professor: Yeah?

Student: Sir in the condition number, when will the equality exist in the above inequality?

Professor: In this one, you mean?

Student: Sir in the K AB is less than KA into KB.

Professor: Yeah that is hard to say, when it will hold with equality. So, obviously, if A or B is the identity matrix, it will hold with equality. But for, I mean, it is not easy to come up with very general conditions under which equality will hold. It depends on the norm also. So, not easy. I mean, there is no straightforward answer to that question.

Student: Sir, even if A and B are I, the equality will hold only when the norm considered is the greater norm, right?

Professor: Correct.

Student: Okay, thanks.