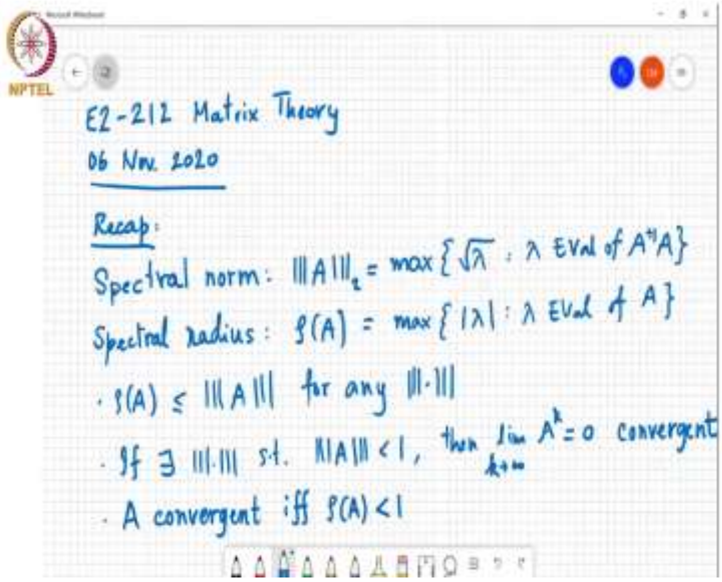
Matrix Theory
Professor Chandra R. Murthy
Department of Electrical Communication Engineering
Indian Institute of Science Bangalore
Recap of Matrix Norms and Levy-Desplanques Theorem

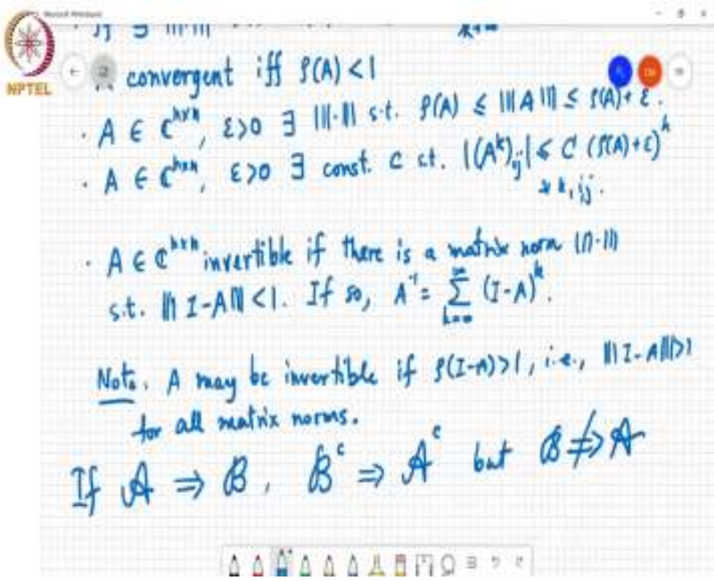
(Refer Slide Time: 00:17)



So, just to recap where we were at the last time we discussed, so just let me maybe you know we discussed the spectral norm which we defined to be maximum square root of lambda, where lambda is an Eigen value of A Hermitian A and the spectral radius which we defined to be rho of A is the maximum magnitude eigenvalue, where lambda is an Eigen value of A.

We also saw that rho of A is less than or equal to the norm of A for any matrix norm and if there is a norm such that norm of A is less than 1, then limit k tending to infinity A power k is equal to 0 and such matrices were called convergent and we saw that A is convergent, if and only if rho of A is less than 1.

(Refer Slide Time: 02:44)



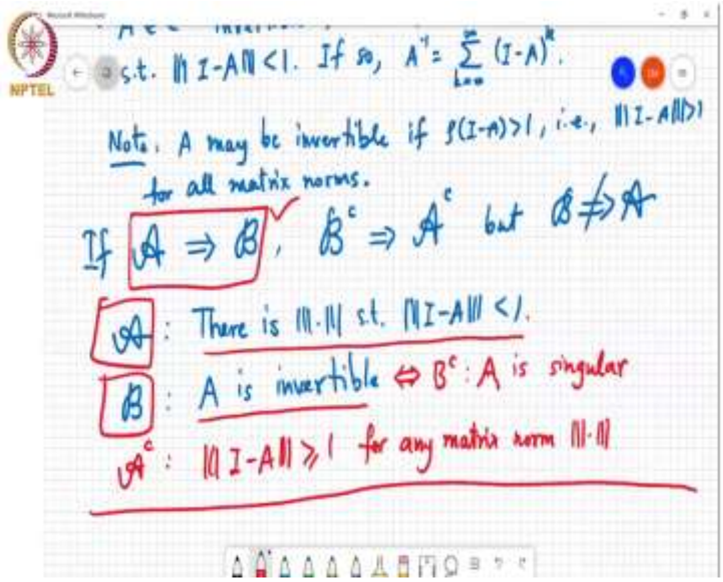
And further this row of a it is a lower bound on any norm of A and it is also the greatest lower bound, in the sense that given A in C to the n cross n and epsilon greater than 0, there exists a norm such that rho of A or the norm of A is between rho of A and rho A plus epsilon. So, in other words, they can find a norm such that the norm of A will be as close to rho A as I wish.

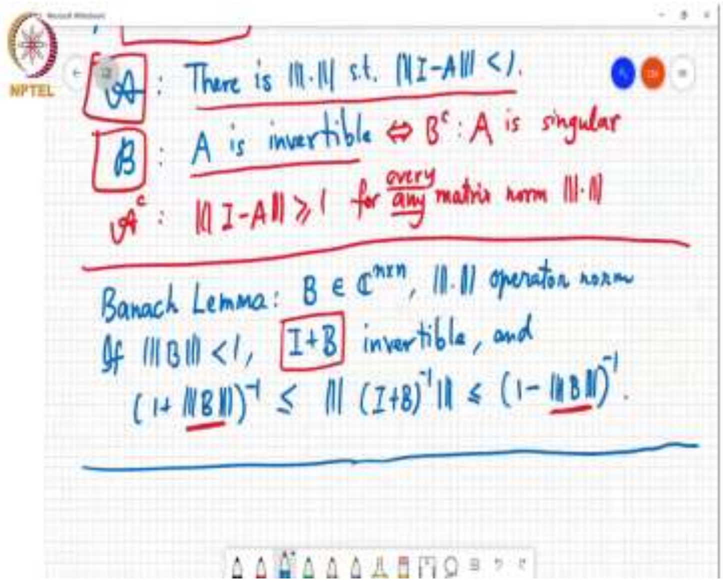
And further the spectral radius can be used to bound the entries of A power k. So, if A is in C to the n cross n and epsilon is greater than 0 there exists constant C such that A power k i j is less than or equal to C times rho of A plus epsilon power k and this is for true for every k i and j. We also saw that A is invertible, if there is, or there exists a matrix norm such thing norm of I minus A is less than 1.

And if so A inverse is equal to sigma k equal to 0 to infinity I minus A power k. So, about this I made a note the last class that A maybe invertible, if say rho of I minus A is greater than 1, i.e. norm I minus A is greater than 1 for all matrix norms. So, I just want to maybe make this point a little more clear, because I think the last time it was not entirely clear the way I said it.

So, basic logic says that if you have a statement A and if we say A implies B, then it means. So, if A implies B, then not B, which I will write as B compliment implies A compliment but B does not imply A, this is the basic rules of logic.

(Refer Slide Time: 06:14)





Now, in this case, what is A, A is the statement that there is norm such that norm of I minus A is less than 1, the statement B is that A is invertible and so, now let us look at the statements and see what they mean. So, what the result says is that A is invertible, if there is a matrix norm such that norm of I minus A is less than 1.

So, that means, if it is true that there is a matrix norm such that I minus A is less than 1, then A is invertible that implies that A will be invertible. So, this A norm of I minus A less than 1, this A implies B. So, this is true for these two statements, as I have written here. Now, what is B complement here, B complement is that A is singular.

So, B complements, so A is singular it means the compliment of A, what is A compliment? A is the statement that there is a norm such the norm of I minus A is less than 1, the opposite of this statement is that norm of I minus A, there is norm such that I minus A norm is less than 1, which means I minus A is greater than or equal to 1, for all matrix norms or for any let us say.

So, it means. So, B compliment implies A compliment is the statement that if A is a singular matrix, then if you compute norm of I minus A, you will always get a number that is greater than 1, no matter which norm you take and what is B does not imply A here. B is the statement that A is invertible, it does not imply that there must be a norm such that norm of I minus A is less than 1.

So, B does not imply A. In other words, it is possible that there is a norm such that I minus A is greater than 1, but A is invertible. So, invertibility of A does not imply the existence of a norm such that I minus A norm is less than 1, but if this condition is satisfied, there is a norm such that I minus A is less than 1, then A will be invertible. I hope that is a little more clear than how we discussed it the last time.

And at the end of the previous class, we saw the Banach lemma. We said that if B is an n cross n matrix and this is an operator norm, then if norm of B is less than 1 then I plus B is invertible and 1 plus norm of B inverse less than or equal to norm of I plus P inverse is less than or equal to 1 minus norm of B inverse. So, one thing I will say is that a form like I plus B arises very frequently in many problems.

And I will show you an example today of where you will end up with a form that looks like I plus B and then what happens is you know something about this matrix B, in particular, if you know some norm of that matrix, this allows you to bound the norm of I plus B inverse in terms of how big the norm of, norm of B is. So, that is why these kinds of results are useful. Now, so this is basically about the recap of what we did last time, so.

Student: Sir?

Professor: Yeah.

Student: Sir, would you explain the difference between any and every?

Professor: In this context?

Student: Yes, yeah.

Professor: There is no difference. So, the thing is that now. So, in this kind of contexts, there is no difference between any and every. Now, where it will make a difference is, for example if A were a random matrix, then you would be associating a probability with which this event will happen, we will say I minus A is greater than or equal to 1 with probability greater than 1 minus epsilon or something like that.

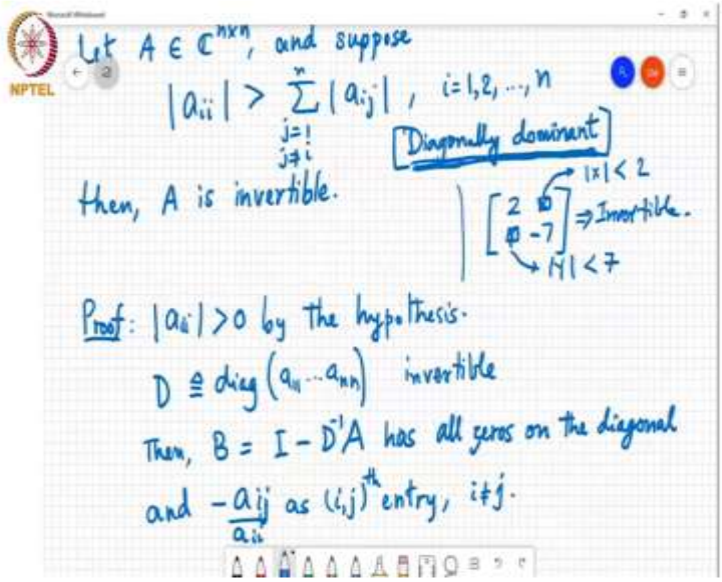
In those kinds of contexts, there is a difference between saying any and every. What you mean by any is that you will fix a norm first and then you will take different different instantiations of this matrix A and you will look, you will try to see what is the probability that or what is the number or what is the percentage number of cases where I minus A norm is greater than or equal to 1 and that is the probability that you are bounded, when you say for any matrix norm.

But when you say for every matrix norm, what you will do is, you will pick an instantiation of A and you will ask what is the probability that all matrix norms of I minus A are greater than 1 greater than or equal to 1 and a matrix fails this, if there is a norm for which norm of I minus A is less than 1 and then you ask if I take different different instantiations of this matrix A, what is the probability of success of this event?

This is well beyond what we want to cover in this course, but I mentioned this only to tell you that in this context any and are actually this one and the same because we have fixed the matrix A, we are only looking at whether this norm is greater than or equal to 1 for this given single matrix A, but when we start talking about random matrices, there is a difference between any and every.

So, there is just one or two more results I want to mention to. So, one use of this result, which says that if there is a norm such that the norm of I minus A is less than 1 then the matrix A is invertible is in showing the following corollary which is called the Levy-Desplanque theorem.

(Refer Slide Time: 14:25)



So, which says A n cross n and suppose mod a i i is greater than sigma j equal to 1 to n j not equal to i, mod a i j for i equal to 1, 2 up to n. What does this mean? It means that if I take the diagonal entry and I compare that with the sum of the magnitudes of all other entries in the same row, but all across all other columns, then the diagonal entries magnitude is strictly bigger than the sum of all other entries in magnitude in the same row.

And such matrices are called diagonally dominant matrices and they are very important from a matrix theory point of view, we are going to see a lot more about such, because in many signal processing manipulations, you end up with matrices which are diagonally dominant.

So, for example, if you think about computing the covariance matrix of a random vector, the diagonal elements of the covariance matrix represent the individual variances of these random variables, while the off diagonal entries represent the cross covariances between these random variables and typically the cross covariances are small compared to the variances of the random variable.

And so, under some assumptions, of course, about the underlying six set of random variables that you are looking at these cross covariances could be small enough that the matrix satisfies a property like this and so then everything we say about diagonally dominant matrices apply to covariance matrices such kind of, kind of a dominant, then A is invertible.

So, this is interesting, because what it says is that, if so if I take a 2 cross 2 example, let us say, and let us say just for fun, I write 2 and then 7 here, now I am allowed to write, for variety, I say minus 7, I am allowed to write any number here or and any number here with the only requirement that the magnitude of this number, so I call it x, the magnitude of x must be smaller than 2 and I can I call this y, then magnitude of y must be smaller than seven, you can try your luck with putting down any number here even be a complex number.

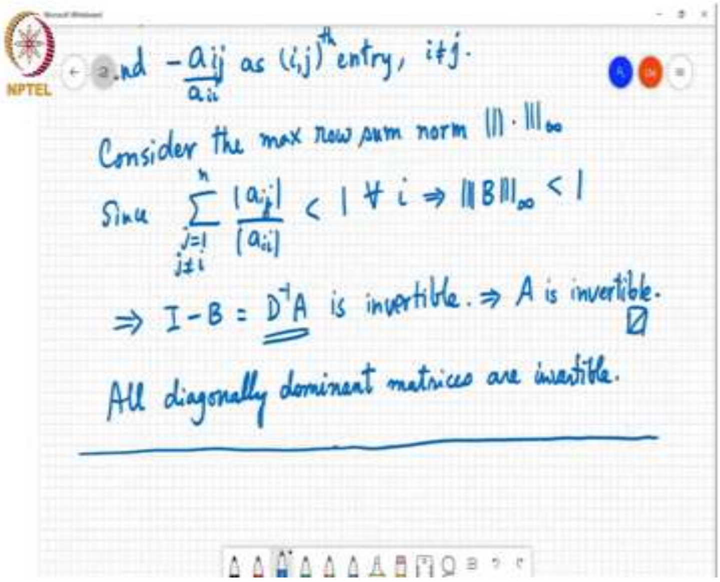
Here, this will always be invertible. That is just to give you an idea of what this theorem is saying. So, let us show this. So, a i i is strictly bigger than the sum of all the off diagonal terms in the same row and these are all non-negative because the magnitude of some number. So, what this means is that this itself is greater than or equal to 0, which means a i i is strictly bigger than 0.

By hypothesis, that means that all the diagonal entries of this matrix are strictly positive. So, if I take the matrix D, defined to be diag of a 11 through a n n, this will be invertible. It is a diagonal matrix with strictly positive entries, strictly positive magnitude entries. In fact, the inverse of this is just a 11 inverse a 22 inverse up to a n n inverse.

So then, if I consider the matrix B equal to I minus D inverse A, what this will do is, this is a diagonal matrix which is operating on A. So, what happens to the diagonal entries, the diagonal entries get scaled, the first diagonal entry will get scaled by a 11 inverse, the second diagonal entry of this matrix will get scaled by a 22 inverse and so on.

And so, when I consider D inverse A, all its diagonal entries will be equal to 1 and so this matrix B will have this identity matrix of course has ones on the diagonal, so when I subtract these out, the one and one will cancel and this matrix will have has zeros on the diagonal. And for the off diagonal entries, it will have minus a i j over a i i as i jth entry i not equal to j. This is just the fact that when I am pre multiply by a diagonal matrix, every row of A gets scaled by the corresponding entry of the D inverted.

(Refer Slide Time: 20:59)



Now consider the of course the off diagonal terms in the identity is 0, so it becomes 0 minus a i j divided by a i i, which is minus a i j over a i i. Now let us consider specifically the max row sum norm. So, you can already probably start seeing how this proof works out. I am going to add up the entries across each row in magnitude and then I will take the maximum value.

Now, since sigma j equal to 1 to n, j not equal to i, mod a i j over mod a i i is less than 1 for every i. Now I am doing j not equal to i, but the corresponding i, ith entry of B is just 0. So when I compute the row sum norm, each of them will be some number like this and it is less than 1 and why is this less than 1, it is because of this condition here, I am just taking mod a i i to the other side.

So, this summation, the summation of j equal to 1 to n, a i j over a i i, for A is going to be less than 1 and this is true for every i, which implies that the norm of B infinity which is just the max of all these guys is also going to be less than 1. So, B infinity is less than 1 that is this L infinity norm of this is less than 1.

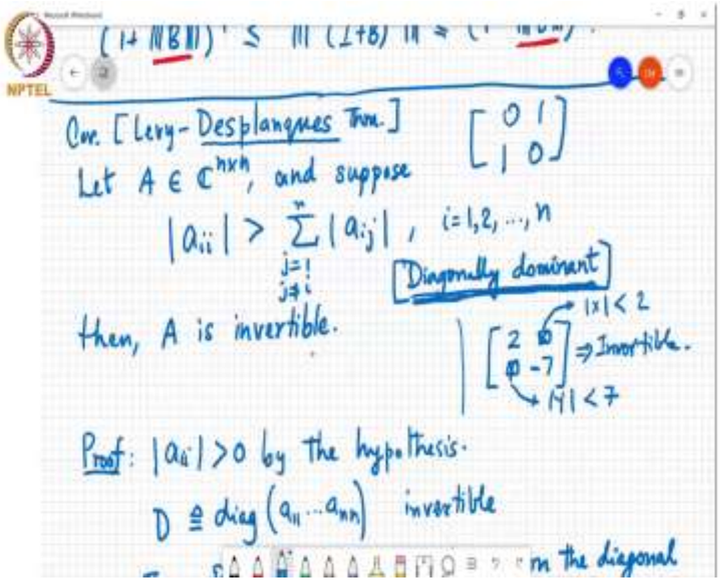
So, there is a norm such that the norm of I minus D inverse A is less than 1 which then means that I minus B which is equal to D inverse A is invertible. But D is a non-singular matrix, so which implies that A is invertible. So, basically all diagonal dominant matrices are invertible.

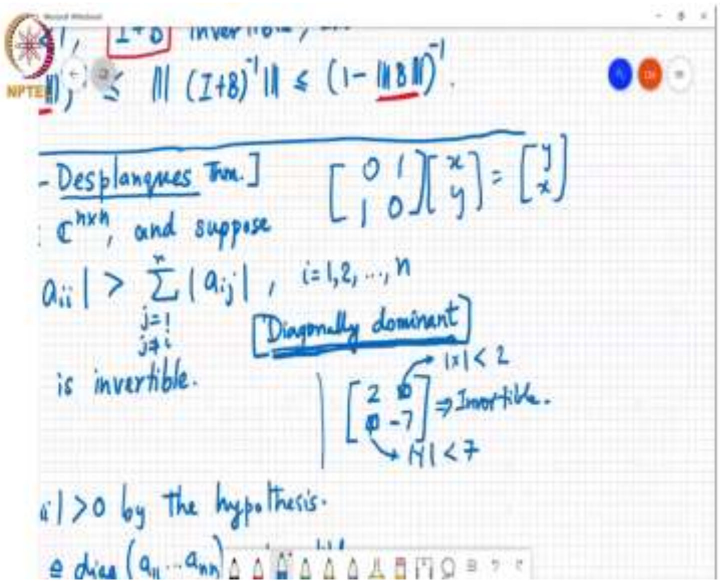
Student: Sir can you please repeat the point why is a (())(24:47)?

Professor: See, this product is the product of two matrices, a product of two matrices this will be invertible only if both the matrices are invertible. You have seen that already in previous classes. Also you have done homework problems on it, that if you should write a matrix as the product of two matrices and if that is invertible then each of the constituent matrices must be invertible.

Student: Okay, sir.

(Refer Slide Time: 25:19)





Professor: So, now that there are just couple of more points before I close out this particular line of discussion, which has to do with remember we talked about equivalence of vector norms, there is a similar thing we can see about equivalence of matrix norms.

Student: Sir, the converse is valid or not in above theorem?

Professor: Which theorem?

Student: Do all invertible matrices are need to be diagonally dominant?

Professor: What do you think? Is this matrix invertible?

Student: Yes.

Professor: Of course, its determinant is minus 1.

Student: Yes.

Professor: In fact, its inverse itself, you multiply this matrix by itself you will get the identity matrix.

Student: Yes, sir got it.

Professor: It is not diagonally dominant. So, again, it is a good point because all these theorems are usually very carefully stated. If there is any error in the, if there is any, anything weird or inconsistent in the statement, that is purely my mistake, in writing the theorem down, the textbook actually is very, very good.

The theorems are all extremely carefully stated and so, when it says if. So, what it is saying is if a matrix is diagonally dominant, it will be invertible, it does not mean that every invertible matrix will be diagonally dominant. The converse is not true.

Student: Hello?

Professor: Yes, please.

Student: That matrix which you have represented above that is 0 1 1 0. So is there any name for it? Anti-diagonal? No, sorry?

Professor: It is a permutation matrix. See if I multiply x y. What I will get is y x. It permutes the entries of the vector. It is a permutation matrix and permutation matrix have lots of very nice properties and this is a, an example of a 2 cross 2 permutation matrix. Of course, the only other permutation matrix and 2 cross 2 is the identity matrix, which does not do any permutation actually, it is the identity permutation. Because you have only two entries, you can either keep them where they are or you can exchange them. Nothing else you can do.