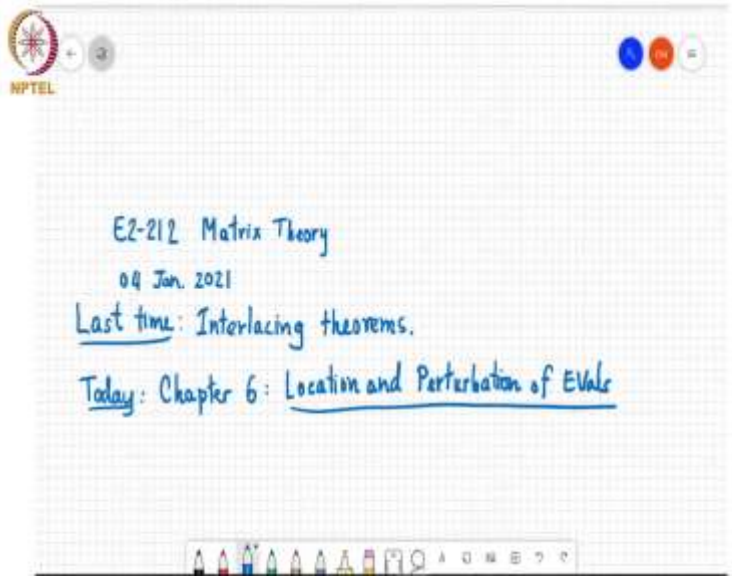
Matrix Theory
Professor Chandra R. Murthy
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore
Location and Perturbation of Eigenvalues
Part 1: Dominant Diagonal theorem

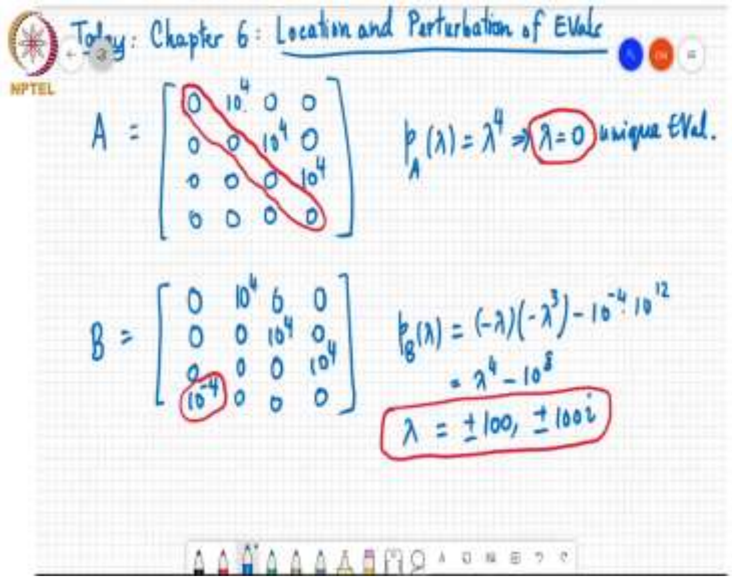
(Refer Slide Time: 00:14)



So, the last time we looked at some interlacing theorems, so that concludes this chapter 4 of the Horn and Johnson textbook today we will start with chapter 6, which has to do with location and perturbation of eigenvalues. So, if you recall we looked at, we have already studied a little bit about the condition number and its relationship with the sensitivity of solutions to linear systems.

And the sensitivity of inverses of matrices to perturbations in these matrices, so and vectors B that is the right hand side of a linear system of equations. So, also of interest is the question of how sensitive are the eigenvalues and eigenvectors of a matrix to the perturbation of its values. So, to take a simple…

(Refer Slide Time: 01:17)



So, consider a matrix A, 0 0 0 0 and say this is 10 power 4 0 10 power 4, 0 0 0 0 10 power 4, 0. So, what are the eigenvalues of this matrix?

Student: 0's.

Professor: Yes. So, basically pA of lambda equals lambda power 4 all the diagonal entries are 0's and this is upper triangular matrix, so lambda equal to 0 is the only distinct eigenvalue of A. Now, consider the matrix which is a slightly perturbed version of this matrix, so I will call it B, this is 0 0 0 and say 10 to the minus 4 and the other entries are the same, so this is 10 power 4 0 0 0, 0 10 power 4 0 0, 0 0, 10 power 4 and 0.

So, now what are the eigenvalues of B? So, this you cannot say directly by looking at the matrix because it is not upper triangular, but it is not difficult to work out pB of lambda, so if you do determinant of lambda i minus B and say expand along first column, you will get minus lambda times minus lambda power 3 minus 10 to the minus 4 times 10 power 12, three determinant of this matrix which is equal to lambda power 4 minus 10 power 8.

So, that means lambda if you solve is this equal to 0, you will get lambda equal to plus or minus 100 and plus or minus 100i. So, we see that a small perturbation by adding a 10 to the minus 4 term to the bottom left here has changed the eigenvalues from lambda equal to 0 to lambda equal to plus or minus 100 plus or minus 100i, so it caused a large deviation in the eigenvalues.

So it would be, what we would like to do is to understand when a matrix would exhibit such a property and when its matrix, when a matrix is such that a small perturbation in the matrix will lead to a small perturbation in the eigenvalues. In fact, the latter point is more important because then we can say things about stability of systems, linear systems or otherwise, so that we can be assured that even if we have the matrix slightly wrong and say slight the true matrix is a slightly perturbed version, the Eigen space of that matrix is not severely perturbed by the perturbation.

And what are such matrices which would be less sensitive to perturbations? So, for example, if you take diagonal matrices or upper triangular matrices like the one we have considered here, the eigenvalues are the diagonal entries, and eigenvalues are in fact continuous functions of these diagonal entries. And so, small changes in the diagonal entries will result in small changes in the eigenvalues.

So, the natural question you can ask is so just to illustrate this from this matrix this point of view, if I had perturbed any of these entries slightly by 10 to the minus 4, then the corresponding eigenvalue would have perturbed by 10 to the minus 4, that is it. It would not have led to such a large perturbation in the eigenvalues. So, a natural question when is if in what if a matrix is nearly diagonal or nearly upper triangular?

Then that means that the off diagonal entries are small, which is not the case here the off diagonal entries are quite big, but if the off diagonal entries were small compared to the diagonal entries, then can we say something about perturbations in the matrix and how that affects the perturbations in the eigenvalues.

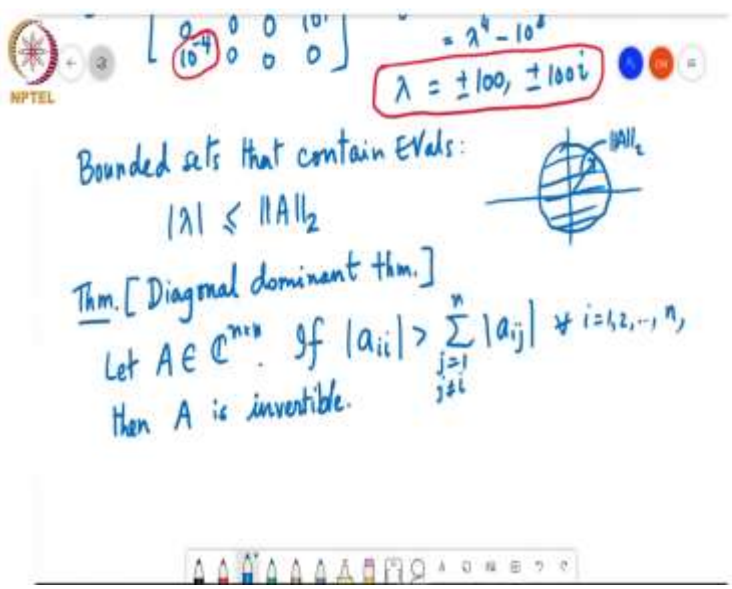
Such matrices arise for example, in when you are computing the covariance of a random process that is nearly wide, so the diagonal entries will be dominant and the other entries will be smaller. A related question to this is can we approximately figure out where the eigenvalues are. So, this is useful for example in linear system theory, where there is a notion of stability that is the stability of system of differential equations for example.

And you have a matrix that describes the state of the system and how the state evolves and you look at the eigenvalues of this matrix and if the real part of those eigenvalues are all negative, then we say that the system is stable. So, in these kinds of scenarios where we are interested in understanding the stability of linear systems, we are interested in knowing whether the real part of the eigenvalues are negative, we do not really need to know exactly what the eigenvalues are.

And similarly, if you want to show that a matrix is positive definite, we need to show that the say its matrix which is Hermitian symmetric, and you want to see whether it is positive definite or not, we just need to show that the eigenvalues are strictly positive, we do not necessarily need to show that the eigenvalues are I mean we need not, we may not want to know exactly what the eigenvalues are, but just show that the eigenvalues are positive.

So, how do we approximately locate eigenvalues? One way to do this is to find a bounded set which is guaranteed to contain these eigenvalues and if this bounded set is in the strictly positive real line then we know that the eigenvalues are all positive. So, one trivial bounded set.

(Refer Slide Time: 08:10)



So, let me say this year bounded sets that contain, so for example, one trivial bounded set is that we know that all the eigenvalues are less than or equal to the induced two norm or in fact for this purpose I can even just use the Frobenius norm. So, the way to look at this is that if I take the complex plane, and if I draw a circle whose radius is equal to the l2 norm of A, then all the eigenvalues will be contained inside the circle. And so, this is one way to bound eigenvalues.

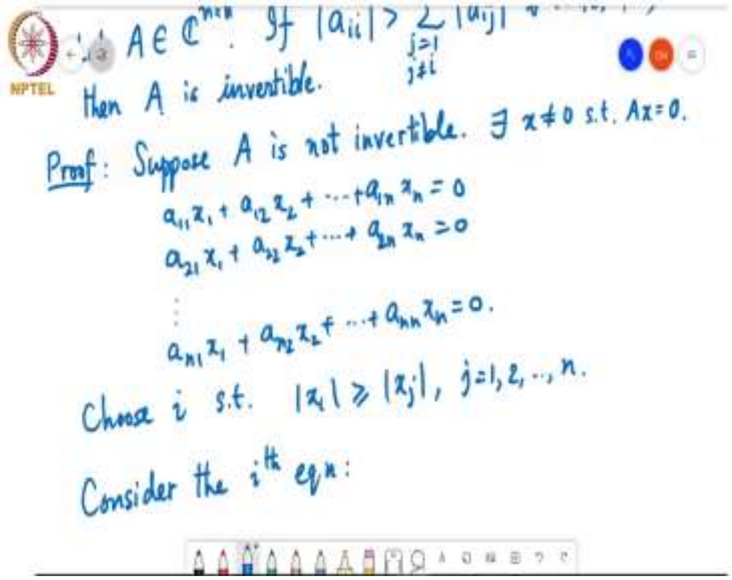
But this is not the kind of bounds that we are looking for, we are looking for something a little more precise than this. The key point is that eigenvalues are ultimately continuous functions of the entries of a matrix and this is something that we will take on faith and we will not prove that here, it requires a different kind of mathematics that we have not really encountered or we have not used very much in this course.

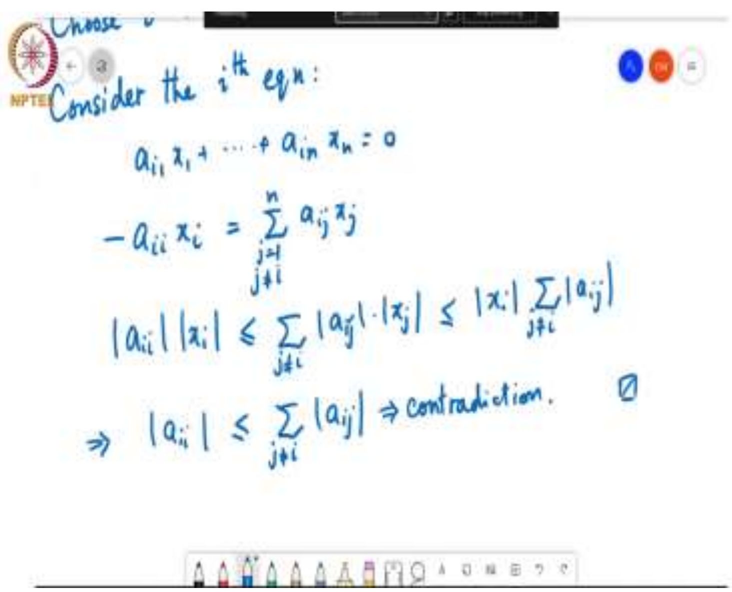
But eigenvalues are actually continuous functions of the entries of a matrix and so, if we perturb a matrix by a small enough quantity, then the eigenvalues will not change too dramatically, that is the essence of what we are going to discuss. Now, in order to state or prove a core theorem on the location approximate location of eigenvalues, we need one interesting theorem, which is called the diagonal dominant theorem.

So, what does theorem says is that, so let A in C to the n cross n be a matrix, so it is not so we have moved away from this thing of Hermitian symmetric matrices, this matrix need not be Hermitian symmetric, is just squared. If mod of aii, the magnitude of the ith diagonal entry of a is strictly bigger than sigma j equal to 1 to n, j naught equal to i mod of aij.

So, I am adding up all the entries in the same row, but in all other columns of the matrix except the diagonal entry and the sum of all the entries in the same row is strictly smaller than the magnitude of the ith diagonal entry. And this is true for i equal to 1, 2 up to n, then A is invertible. So, the condition is basically telling us that the diagonal entry of each row is dominant dominates over all the other entries. So, let us just quickly see how why this is true.

(Refer Slide Time: 12:05)





The proof is by contradiction, in other words will show that not this implies not this, so if A is not invertible, then there must be some i for which this condition is getting violated. So, suppose A is not invertible, that means A is a singular matrix, so in other words its rank is less than n, so there exists an x not equal to 0, non-zero x, such that Ax equals 0.

So, I will just write that out in full form. So, what that is saying is the first row is saying a11 x1 plus a12 x2 plus et cetera plus a1n xn equals 0. The second row is saying a21 x1 plus a22 x2 plus a2n xn equals 0 and we proceed like this. And the last equation we will read an1 x1 plus an2 x2 plus ann xn equals 0. Now, what we will do is this x is some vector which is in the null space of A, one of its centuries must be bigger than or equal to all the other entries in magnitude.

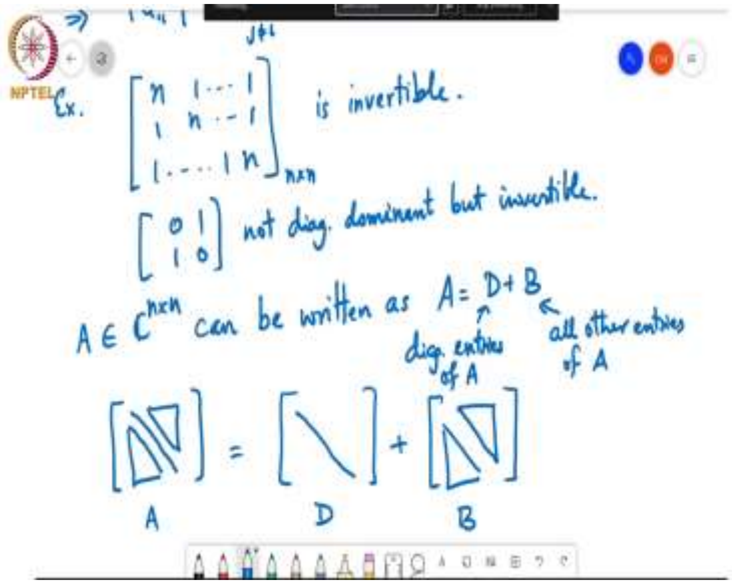
So, we will just choose suppose that is the ith entry, so choose i such that mod of xi is greater than or equal to x mod of xj, j equal to 1, 2 up to n. So, i is the largest magnitude entry in x and note that this mod of xi must be strictly greater than 0, because x is a non-zero vector, so some entry will have magnitude which is not equal to 0.

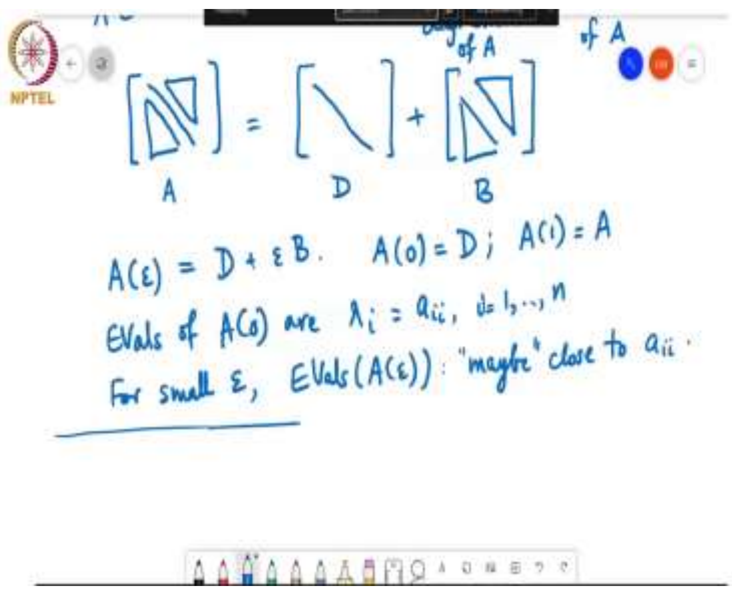
So, now we will consider the ith equation. What does it saying? Its ai1 x1 plus et cetera plus ain xn equals 0. Now, I will just take the ith term to the other side, so minus aii xi equals sigma over j equal to 1 to n, j naught equal to i, aij xj. Now, we will take the magnitude and so if you take the magnitude on both sides mod aii times mod xi.

And if I take the magnitude inside the summation I will get the less than or equal to inequality sigma over j naught equal to i mod aij times mod xj. But mod xj is less than or equal to mod xi, so if I replace mod xj with mod xi for all for if I replace mod xj with mod xi for all j then I am only increasing the value, so this is in turn less than or equal to mod xi times sigma j naught equal to i times mod ai of mod aij.

And as I already remarked mod xi is strictly positive and so this means that mod aii is less than or equal to sigma over j naught equal to i mod aij, which contradicts what we said in the beginning that mod aii must be greater than the summation over j naught equal to I, of aij for all i. But for this i we see that the inequality goes the other way. And so, this is a contradiction.

(Refer Slide Time: 16:58)





So, for example, if I take the matrix n 1 1, 1 n 1, 1 1n and this is an n cross n matrix, then the sum of all these guys is n minus 1 and that is strictly less than n, the sum of all the other terms here is n minus 1 strictly less than n and so on. So, this matrix is invertible by the diagonal dominance theorem.

Now, obviously the converse of the diagonal dominant theorem is not true, meaning that a matrix may be invertible without being diagonally dominant. So, a trivial example is 0 1 1 0 this matrix is not diagonally dominant, but it is invertible. Now, let us proceed, so the the thing is that any matrix A in C to the n cross n can be written as A equal to D plus B where D is a matrix containing the diagonal entries of A and this has all other entries.

So, pictorially I can write this as if A is a matrix which has some lower diagonal part, a diagonal part and an upper diagonal part, and I can write this as this diagonal part which I will call D plus everything else, which is this and I call this B. Now, let us consider this matrix A of epsilon which I will define to be D plus epsilon times B.

Now, A of 0 is equal to this diagonal matrix D, and A of 1 is the matrix A, it is just D plus B. And so as epsilon goes from 0 to 1, the matrix A of epsilon transitions from D to A. Now, if I take A of 0 it is a diagonal matrix and its eigenvalues are easy to find, they are just equal to the diagonal entries of A. So, eigenvalues of A of 0 are lambda i equals aii i equal to 1 to n.

And for small epsilon the eigenvalues of A of epsilon there would be maybe I will say maybe, but in fact it is true, close to aii, so the diagonal entries of A. And as epsilon increases they just go further and further away. So, the theorem that I am going to state now is going to make this motion much more precise, this is one of the very central theorems of related to approximate location of eigenvalues.