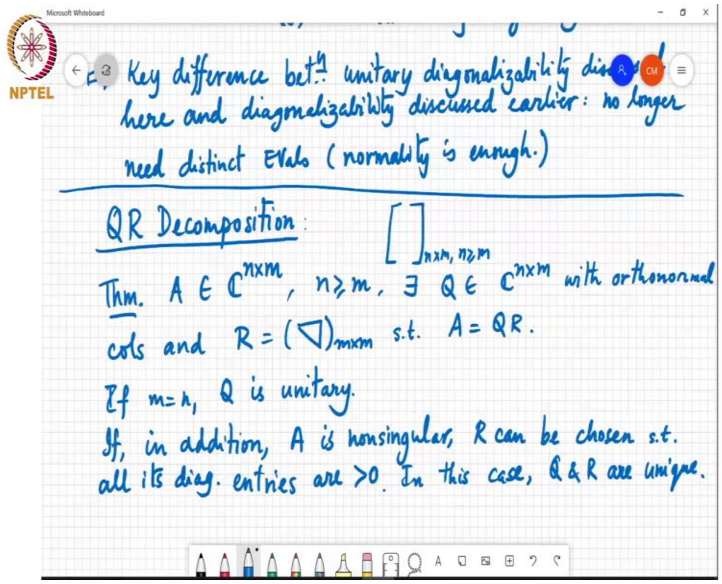
Matrix Theory
Professor Chandra R. Murthy
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore
Lecture 45
QR decomposition and canonical forms

(Refer Slide Time: 00:16)



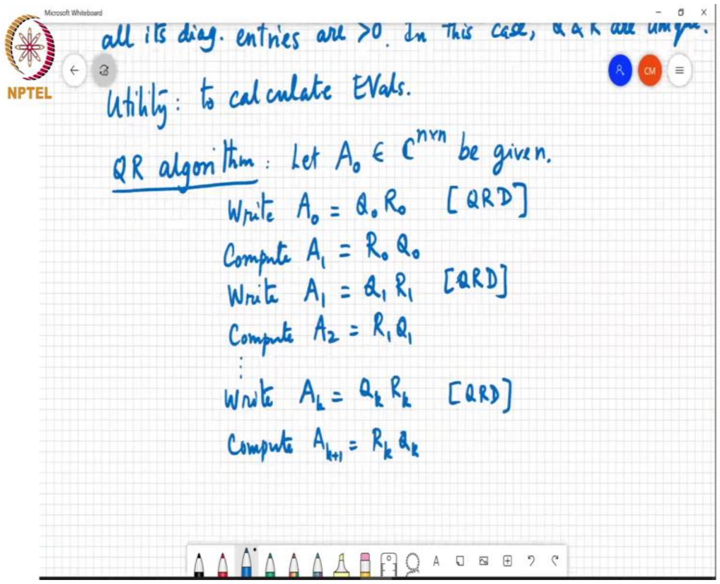
Another way of decomposing a matrix is through QR decomposition and it is very useful in many, many problems and many scenarios. So, basically the QR decomposition theorem says that if you are given a matrix A of size n by m, so need not be square and n greater than or equal to m, so basically it is a matrix that is tall like this.

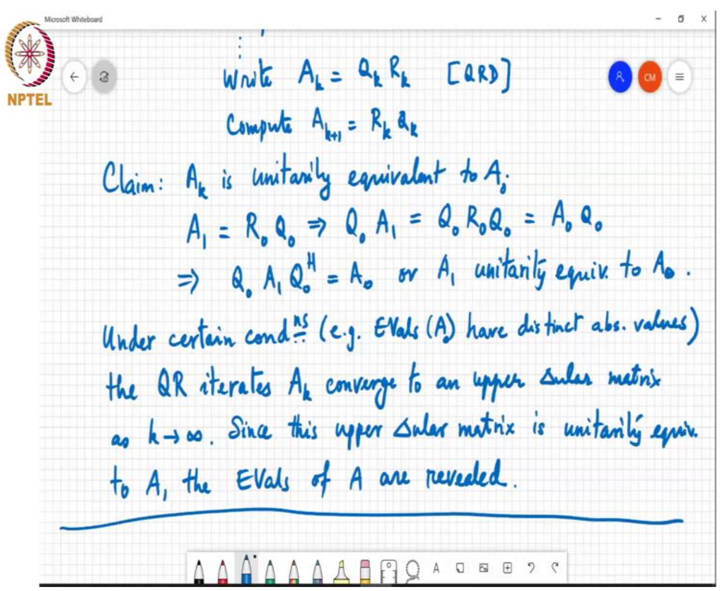
So, this is n by m and n greater than or equal to m. Then there exists a Q which is in C to the n by m and with orthonormal columns and R which is upper triangular of the size m by n such that A equals QR. And if m equals n then Q is unitary. Of course it is orthonormal and it is square, so it must be unitary. And the last part is that if in addition A is nonsingular then R can be chosen such that its diagonal entries are all strictly positive. That means they are real and positive.

So, keep in mind that you cannot compare a complex number to 0 and say that it is greater than 0 or less than 0. You cannot order complex numbers but on the real line you can order things. So, when I say that all its diagonal, R can be chosen such that all its diagonal entries are greater than 0 what I really mean is that it I can choose R such that all the diagonal entries are real and positive. In this case Q and R are unique.

I will not prove this theorem but it is direct consequence of the Gram-Schmidt orthogonalization process. So, essentially all you will be doing to run Gram-Schmidt on the columns of A and you see that the corresponding coefficients that you learn, that you compute can be arranged in a form of an upper triangular matrix R. That is the, I mean that is how it goes. So, I will not write the proof out.

(Refer Slide Time: 04:15)





But one, one important utility of this QR decomposition is to calculate eigenvalues. So, recall that the way to find eigenvalues is to first write out the characteristic polynomial. And then you have to find the roots of this nth order polynomial for a general n cross n matrix. And there is no simple procedure to find the roots for n greater than 2. For n equals 2 it is the quadratic form and we know we can write the roots of the quadratic in closed form. But for n greater than 2 we cannot write the roots in closed form. We will have to use some numerical 0 finding algorithm to find those roots. So, this is called, this algorithm is called the QR algorithm and it basically helps, is useful for finding the eigenvalues of the matrix.

Keep in mind that the QR decomposition by itself does not reveal the eigenvalues of the matrix. In particular the diagonal entries of R are not the eigenvalues of the matrix A. But we can use this decomposition to, in this algorithm to find the eigenvalues of A. So, let the matrix A which I will call A0 in C to the n cross n, it is a square matrix, be given. Then what we do is to first compute the QR decomposition of A. I will state that as write A0 Q0 R0. So, we have computed the QR decomposition.

Student: Sir the dimension of A0 is n cross n

Professor: Yes, so now it is a square matrix because I am trying to show you how this QR decomposition could be used to find eigenvalues. And eigenvalues are things we define only for square matrices. Then what we do is we compute this matrix which I will call A1 which is equal to R0 times Q0. So, all I am doing is I first computed this QR decomposition, and then I am just reversing the order and multiplying it as R0 Q0.

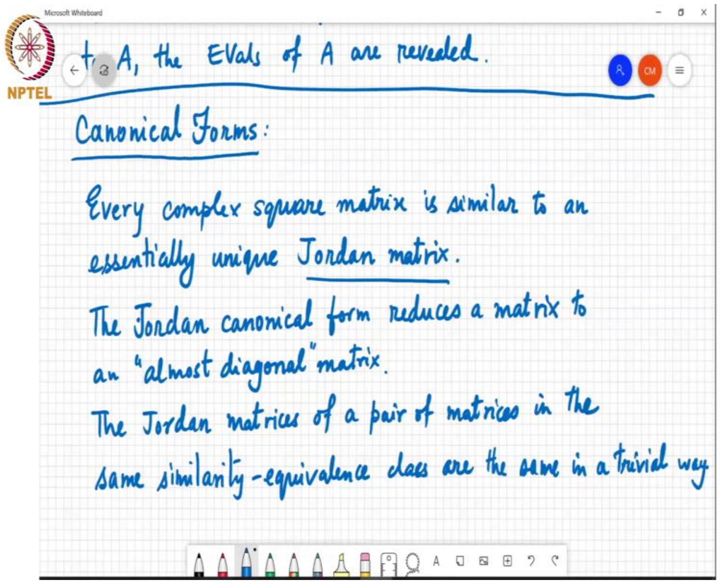
Now, this A1, I will compute its QR decomposition. This is another QR decomposition step and then I will compute A2, do not know how many times I should write this but this is the pattern R1 Q1 and so on. So, the kth step will be we write Ak equals Qk Rk. And k plus 1th step would be to compute Ak plus 1 equals Rk Qk. This is again a QR decomposition step. And now, so this is the algorithm.

And so one, first, before we proceed one claim is that Ak is unitarily equivalent to A. That is easy to show. So, for example if A1 equals R0 Q0 then if I consider Q0 times A1 that is going to be used Q0 R0 Q0, which is equal to Q0 R0 is A0, and this A0 has orthonormal columns so it is unitary. So, which means that Q0 A1 Q0 Hermitian equals A0, or A1 is unitarily equivalent to A0, and so on. And so Ak is unitarily equivalent to A0.

So, it gives you sequence that are all unitarily equivalent. And what one can show, which I am not going to show again here. This is an algorithm. Its one of its properties is that under certain circumstances, so for example if the eigenvalues of A are all distinct, so under certain conditions e. g, eigenvalues of A A0 have distinct absolute values.

The QR iterates Ak converge to an upper triangular matrix as k tends to infinity. So, since this upper triangular matrix is unitarily equivalent to A0 the diagonal entries of this Ak as k goes to infinity are the eigenvalues of A0. So, this is one, another numeric recipe that one can use to find the eigenvalues of matrix A.

(Refer Slide Time: 12:20)



So, now that we started discussing factorizations we will discuss what are known as canonical forms. So, these are basically forms where, I mean processes by which we reduce matrix down to a simpler form. So, the motivation is that, one basic question you can ask is, when are two matrices going to be similar?

We know that similar matrices have the same trace, the same determinant, the same eigenvalue, the same characteristic equation. But it is also possible that matrices can be different without, can be, it is possible to find matrices that are not similar to each other but have the same trace, determinant, eigenvalues and characteristic polynomial.

So, it is still not clear how we will verify that two matrices are actually similar to each other. If you can find the matrix is such that S inverse A S equals B then great, you are lucky. You found this matrix and so you know that A and B are similar. But if you do not, if you are not able to find that matrix how do you prove or otherwise, or disprove that two matrices are similar?

So, one possible approach is to determine similarity is to try and reduce both matrices down to some simple form, for example, a diagonal form and then see if this diagonal forms are similar, are the same upto possibly permutations of the diagonal entries. And so that is one way to determine similarity. If you are able to reduce both matrices down to a diagonal form and check that the two diagonal forms are actually the same then you know that the two matrices are similar.

So, these are what we call canonical forms; reducing a matrix down to its simplest form which will then allow us to test for properties like similarity. So, basically, so if you could reduce things to diagonal matrices or reduce all these matrices to diagonal matrices then that would work. But the problem is that not every matrix is diagonalizable. And so we have an existence problem. If two matrices are both non diagonalizable then it is difficult to know whether those matrices are similar or not.

Now, an alternative could be to try and use Schur's theorem which will allow a matrix to reduce to upper triangular form. And then you can say let me try to compare these upper triangular forms. But in this upper triangular form that you obtained from Schur's theorem the diagonal entries can potentially appear in any order. And two upper triangular matrices with the same, even if the two upper triangular matrices have the same diagonal entries but different off-diagonal entries then those two matrices can still be similar. And so essentially Schur's theorem is insufficient to determine whether or not two matrices are similar.

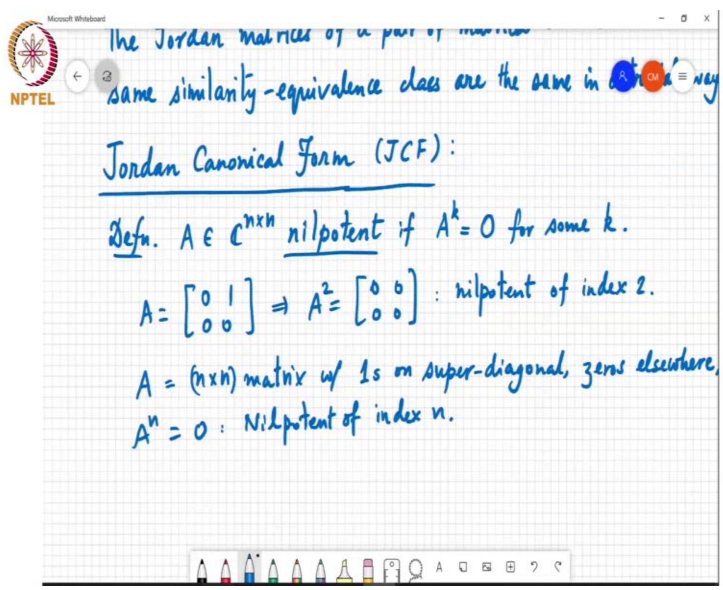
Now, we see if we search for an upper triangular form that is as close to being diagonalizable as possible but it is still attainable for every matrix then that form is called the Jordan canonical form. And this is Jordan canonical form is a set of almost diagonal matrices and in fact if the matrix is indeed diagonal then the Jordan canonical form will return a diagonal matrix. So, in some sense it is a generalization of diagonalizability of matrices.

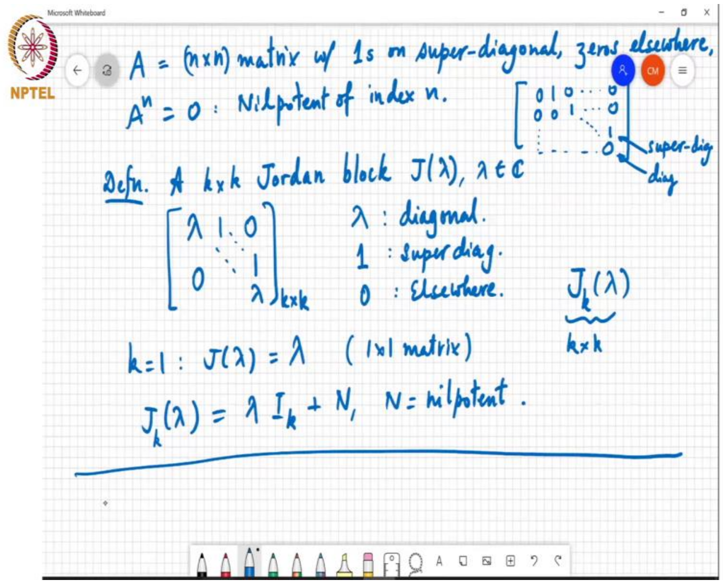
So, the Jordan canonical form is a set of almost diagonal matrices and these matrices are called Jordan matrices, and the Jordan matrices include diagonal matrices. And the punch line is that every equalized class under dissimilarity of square complex matrices includes a Jordan matrix and any two Jordan matrices of same equivalence class are the same in a very trivial way. We will be able to look at the Jordan forms and say, yes these are the same or these are different.

So, the main result we will discuss next is that every complex square matrix is similar to an essentially unique; by essentially unique I mean that these matrices which are called Jordan matrices has a block diagonal structure and those blocks are called Jordan blocks and the only thing that is allowed is the permutation of this blocks, but other than that the matrices, the Jordan blocks will be, the Jordan matrices will be unique. And so this is called Jordan canonical form reduces a matrix to an almost diagonal matrix which is called a Jordan matrix.

And as I said the Jordan matrices of two matrices in the same equivalence class, let me write that is important point. So, the Jordan matrices of a pair of matrices in the same similarity equivalence class are the same in a trivial way; meaning that only the block diagonal, in the block diagonal structure of Jordan matrix some blocks could be exchanged but otherwise they will be the same. So, you can, it is very easy to check whether the Jordan matrices are the same or they are different.

(Refer Slide Time: 20:25)





In order to talk about the Jordan canonical form I need to introduce a couple of definitions. So, the first is a nilpotent matrix. So, A in C to the n cross n is said to be nilpotent if A power k equals what?

Student: 0

Professor: 0

So, the smallest positive k for which this happens is called the index of the nilpotent matrix. Of course if A power k equals 0 then A power k plus 1, A power k plus 2 all that is always equal to 0. So, for example if A is the matrix 0 1 0 0 then A squared is the all 0 matrix. And so we say that it is nilpotent of index 2.

More generally if A is the n cross n matrix with 1s on the super-diagonal and 0s elsewhere then A power n equals 0 meaning that the matrix is nilpotent of index n. So, basically the Jordan canonical form theorem later we will say that every matrix is similar to a matrix of form D plus n where D is a diagonal matrix and n is a nilpotent matrix. So, that is what we are going to go towards.

Student: So, sir

Professor: Yeah

Student: What is super-diagonal?

Professor: Super-diagonal are the entries just about the diagonal.

Student: Ok sir

Professor: The diagonal, not the diagonal, this is the diagonal, and this is the super-diagonal. And similarly this thing would be the sub-diagonal. So, you have 1s on the super-diagonal. If you square this matrix what you would find is that, you can hand-compute it is easy, the 1s will come in the second super-diagonal.

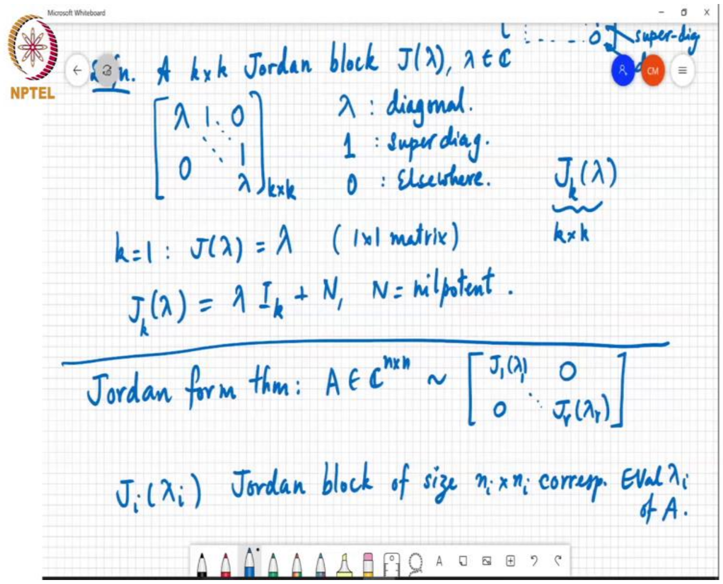
Then if you take this matrix power 3 or if you multiply that by this matrix again it will come in the third super-diagonal. And then fourth, fifth, sixth. Eventually it will come to be a matrix with all 0s except this entry being equal to 1. Then you multiply that one more time by this matrix. You will get rid of everything and you will get 0 matrix.

So, k cross k Jordan block, J of lambda with lambda being a complex number is the following matrix. It has lambdas on the diagonal and 1s on the first super-diagonal and then 0s everywhere else and is of size k cross k. So, lambda is on the diagonal, 1 is on the super-diagonal and 0 everywhere else. So, this matrix is called a k cross k Jordan block with lambda. And of course when k equals 1 J of lambda is just equal to lambda.

So, we will also sometimes use Jk of lambda when we want to indicate the size of the matrix. So, this is for the k cross k. So, we will use both these notations. Hopefully I will not be too confusing to you. But here I am not using the subscript k. So, this is J1 of lambda or J of lambda is just a lambda for a 1 cross 1 matrix or a scalar.

And also J of lambda in the k cross k case is lambda times the k cross k identity matrix. So, I will write it this way. So, it is clear Jk of lambda is lambda times the k cross k identity matrix plus N where N is a nilpotent matrix of index k with 1s on the super-diagonal. So, it is the all 0 matrix with 1s on the super-diagonal.

(Refer Slide Time: 26:49)



So, we have the following theorem which I think since we have only 1 minute left I will just leave it for the next time, so maybe peek your interest a little and say that Jordan form theorem. What it will say is that A in C to the n cross n is similar to a matrix of the form J1 of lambda 1 Jr of, or Jr of lambda r where these are Jordan matrices or Jordan blocks.

So, again I think I am messing a notation a bit but it is a Jordan block of size, not i cross i; ni cross ni. So, this is a bad notation here. I will fix that next time. So, this is corresponding to eigenvalue lambda i of , so this is what we will state and prove in the next class. We will stop here for today.