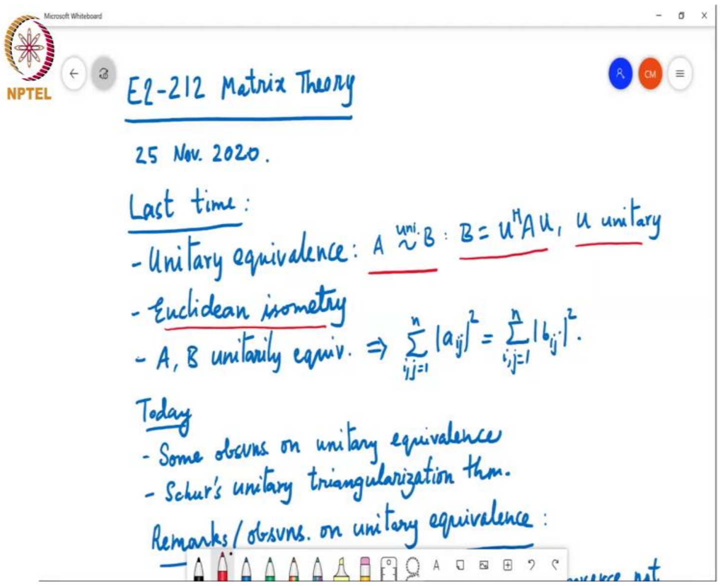
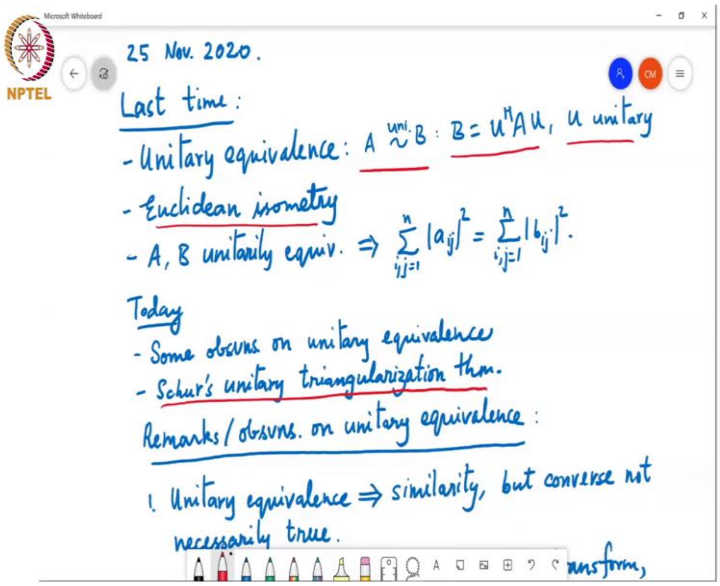
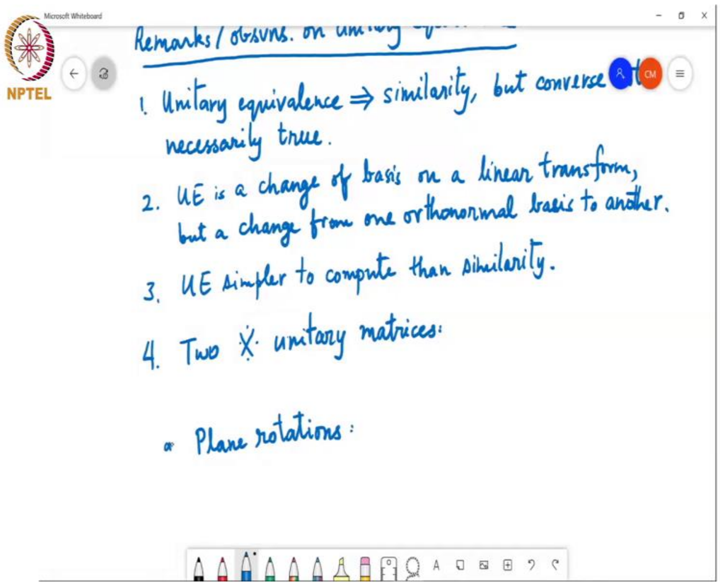
Matrix Theory
Professor Chandra R Murthy
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore
Lecture 40
Schur's Triangularization Theorem

(Refer Slide Time: 0:15)







So, the last time we looked at unitary equivalence. So, A and B are unitarily equivalent if B is equal to u Hermitian A u where u is a unitary matrix. And we also looked at the notion of Euclidean isometry. We finally showed one small result which said that A and B if A and B are unitary equivalent then the Frobenius norm square of the two matrices which is the sum of the squares of all the elements of the matrix will be equal.

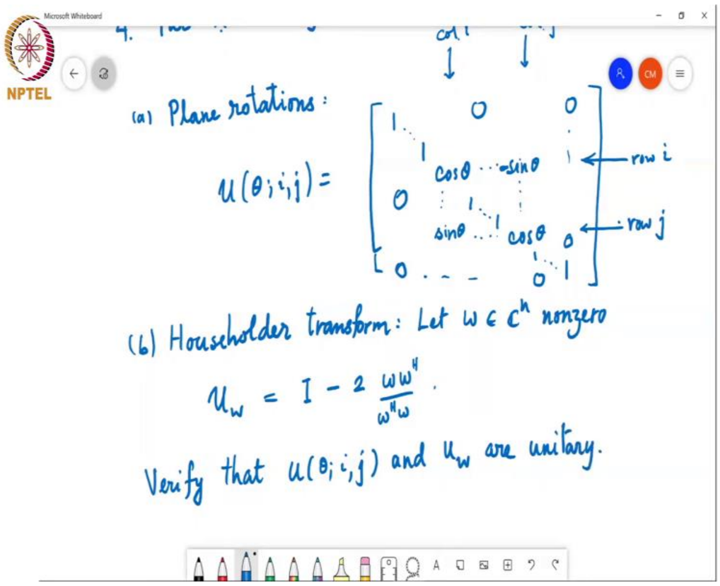
So, today we will discuss some concluding observations about unitary equivalence. And then we will cover this shows unitary triangle realization theorem. So, couple of remarks one is that unitary equivalence two matrices are unitarily equivalent it means that they are similar. Because for a unitary matrix u Hermitian equals u inverse. So, they satisfy the definition of similarity but the converse is not necessarily true.

So, two similar matrices need not be universally equivalent. That also means that unitary equivalence partitions the complex n cross n matrix space into a final equivalence class compared to similarity-based equivalence. So, within each similarity-based equivalence there could be many matrices that are unitarily equivalent to each other. But many sub classes which are not unitarily equivalent to each other.

But any pair of matrices that are unitary equivalent are also similar. So, they belong to the same similarity class. And we observed that the similarity transform is essentially it corresponds to a change of basis that is if you have a linear transform and you change the basis then you ask what is the linear transforming according to the new basis that is given by the similarity transform. However, the unitary equivalence is also a change of basis but especially one is the change of basis from one orthonormal basis to another.

So, to continue oftentimes we like unitary equivalents because they are simpler to compute. So, I mean simpler to compute than similarity. So, for example there is no matrix inversion involved it says u Hermitian A u. So, and it is also numerically more stable because these unitary matrices are well conditioned. And so numerically computing the similarity transform is more stable than computing a similarity transform. Now, I will give you two examples of important unitary matrices that show up in many applications.

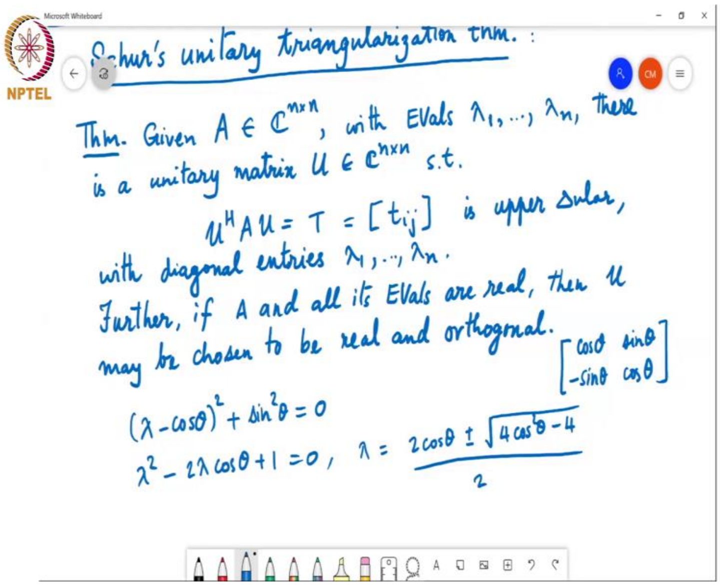
(Refer Slide Time: 4:08)

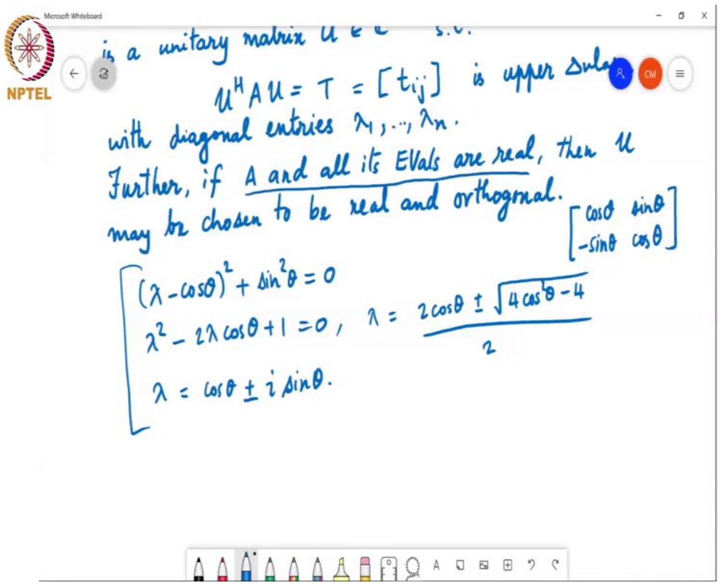


So, the first is called plane rotations. So, we write u of theta i j is this matrix which has once on the diagonal and somewhere in between it has a cos theta. And then some more 1's and then another cos theta then maybe some more 1's. And here it has sin theta or maybe minus sin theta and sin theta here. And then 0's everywhere else. So, basically this is the ith row and this is the jth row. And this is the ith column and this is the jth column.

So, we ith row so it is basically the identity matrix except in the i and jth column where the column i column j sub matrix form set 2 cross 2 matrix with cos theta minus sin theta minus sin theta and cos theta as its four elements. So, this is called a plane rotation. The second one is called the householder transform is named after the mathematician who came up with it not because it is it has anything to do with householders. So, let w be a vector in C to the n which is nonzero. Then u w is defined as I minus 2 w w Hermitian divided by w Hermitian w. So, an exercise for you is to verify that these two are indeed unitary matrices.

(Refer Slide Time: 7:43)





So, this is what I wanted to say about unitary matrices and unitary equivalence. But we are going to use that immediately in this next result which we are going to discuss which is shows the unitary triangularization theorem.

Student: Sir.

Professor: Yes.

Student: Sir we understand the use of a plane rotation matrix but what is the use of this householder transform? I mean can you tell some application.

Professor: So, it is useful for example in computing the Q R decomposition of a matrix it is also useful in so one way to decompose a matrix is to write the matrix as the product of a series of householder transforms and a diagonal or an upper triangular matrix followed by another series of householder transforms. So, it is mainly used in at least one of the uses is in matrix factorizations into simpler forms.

Student: Okay sir thank you sir.

Professor: So, this shows unitary triangularization theorem is a very important result in linear algebra. In fact, what Hon and Johnson says about this theorem is that it is perhaps the most fundamentally useful fact of elementary linear algebra. So, if there is one theorem that you want to take away from this course this could be one of them. So, essentially the theorem says that any complex n cross n matrix is unitarily equivalent to an upper triangular matrix.

And so, it so while this kind of decomposition of a and showing it to be unilaterally equivalent to an upper triangular matrix is far from unique it still represents the simplest form that one can achieve using unitary equivalence. And of course because the unitary equivalence preserves all the eigenvalues obviously if you can find the equivalent unitary equivalent upper triangular matrix.

Then the diagonal entries of that upper triangular matrix are the eigenvalues of this matrix. So, the decomposition also reveals the eigenvalues of the matrix. So, let us formally write down what the theorem is. So, given A in C to the n cross n with Eigen values lambda 1 to lambda n there is a unitary matrix u such that u Hermitian A u is equal to T which we will write as its elements as t ij this matrix T is u Hermitian A u equal to T is upper triangular with diagonal elements lambda 1 so lambda n.

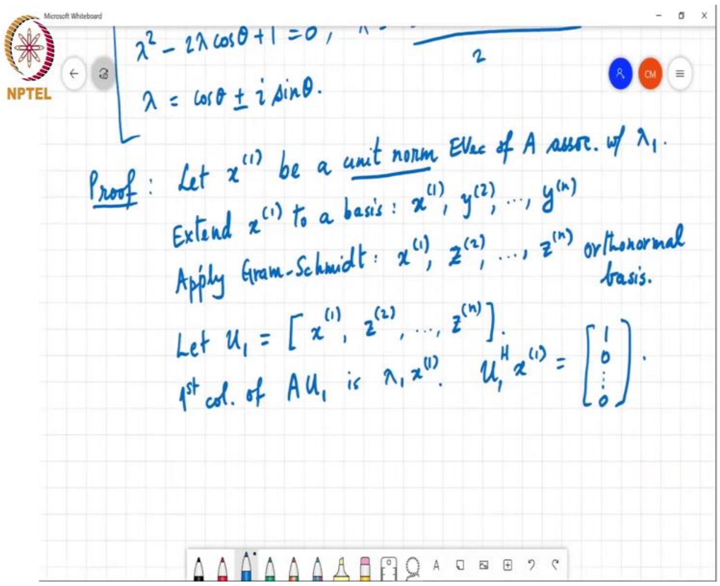
Further if A and all its eigenvalues real then u may be chosen to be real and orthogonal. So, this is the theorem I will just sort of wait for a few seconds for you to read it again. Because it is such an important theorem. So, you can look at it for a few seconds. So, quick question if a matrix is the earth all the entries of a matrix are real value is not it necessary that all its eigenvalues are real? What is an example of a very simple the simplest 2 cross 2 matrix you can think of which has real valued entries but complex valued Eigen values?

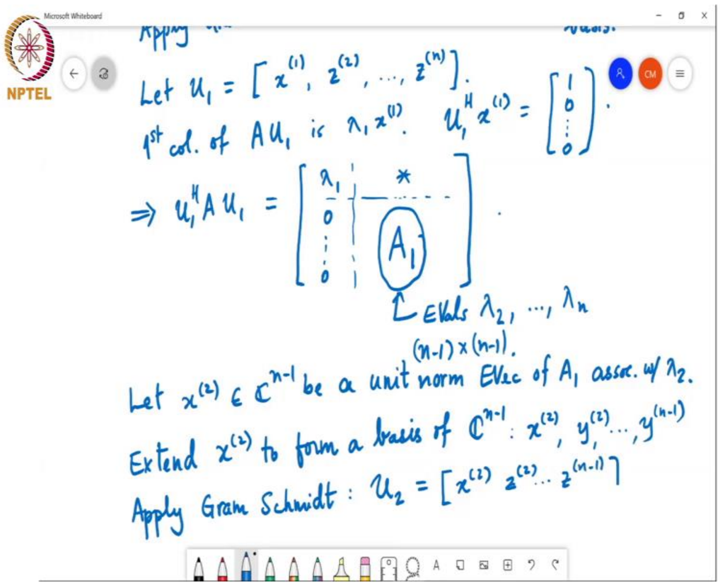
Student: Sir, I think cos theta sin theta minus sin theta cos theta.

Professor: So, all the entries are real valued and what are its Eigen values? You have to actually find it. So, let us do it just for the fun of it. So, if I do determinant of lambda I minus this matrix equals 0. I will get lambda minus cos theta the whole square minus and then it becomes plus sin squared theta equals 0 check me check and let me know if I make a mistake. So, if I simplify this, this gives me lambda squared minus 2 lambda cos theta plus 1 equals 0. And its roots are lambda equals 2 cos theta plus or minus square root of 4 cos square theta minus 4 divided by 2. Is that correct?

So, let me know if I make a mistake. So, that is lambda equals cos theta plus or minus i times square root of 1 minus cos squared theta sin squared theta. So, correct? So, basically it is Eigen values are complex even though the matrix is real valued. So, you cannot run away from complex numbers if you want to study matrices and its Eigen value properties. So, this is an aside. So, now let us prove this theorem. So, that is why we need to say if a and all its eigenvalues are real then it is true that you can choose you to be real orthogonal matrix.

(Refer Slide Time: 16:40)





So, let x1 be a unit norm eigenvector of A associated with lambda 1. Now, so this is a unit norm vector that what we will do is we will extend x1 to a basis. That is, you find other vectors that are linearly independent of x1 such that x1, y2 et cetera up to y n form a basis of C to the n.

Then we will apply Gram Schmidt. This will give me a set of vectors which are all unit norm. Which I am going to call x1. Say of course the first vector is already unit norm. So, when you apply gram Schmidt that does not change the first factor. But it will change the second third all that. So, we will call it z2 up to zn which are orthonormal vectors forming a basis of C to the n so this is an orthonormal basis. Then what we will do is we will let u1 be the matrix x1 z2 zn. So, now if I so we will start with this matrix.

Now, if I consider what happens so the first column of A u1 that is the same as u x1 times A the first column of u1 is x1. So, the first column of this product A u1 will be x1 times A which is equal to lambda 1 times x1. And also, if I did u1 Hermitian times x1, what I will get is u1 is this matrix, it is an orthonormal matrix. So, if I do u1 Hermitian x1, this is going to be a vector whose first element will be x1 Hermitian x1 which is equal to 1.

The second element will be z2 Hermitian x1 which is 0 the last element will be zn Hermitian x1 which is also 0. So, this is actually just the vector 1 0 0. So, this means that if I do u1 Hermitian A u1, then I get the first column of this product is actually this lambda 1 times this column here, so that will just give me lambda 1 0 0.

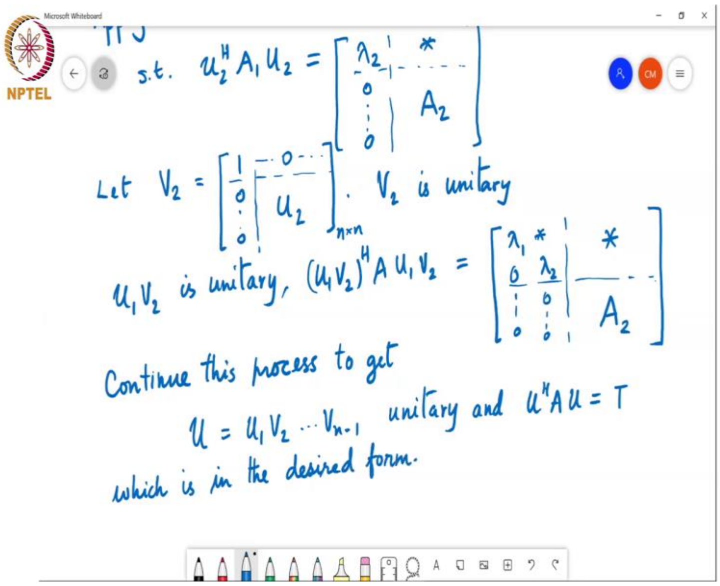
And over here in the first row, I will get some entries, I actually do not care about them. And I will call whatever I get down here as the matrix A 1. Now, lambda 1 is here, so what can I say about the eigenvalues of A 1, the eigenvalues of A lambda 1 to lambda n. And I found A unitary matrix u, I computed u1 Hermitian A u1. And this matrix has lambda 1 followed by zeros here. So, it is a block upper triangular matrix with lambda 1 up here and an A 1 matrix down here. And So, what will be the eigenvalues of A 1?

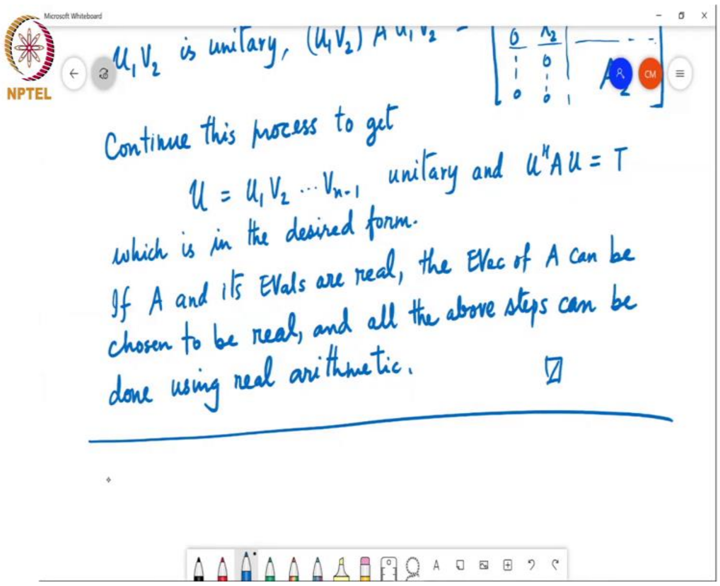
Student: Lambda 2 to lambda n.

Professor: Exactly the other Eigen values will be (())(22:08) A 1. Now, what we will do is the same idea exactly but repeated with respect to A 1. So, let x2 in, now x2 will be in C to the n minus 1, this A 1 is of size n minus 1 cross n minus 1. So, let this be, let x2 n minus 1 be unit norm eigenvector of A1 associated with lambda 2. Then, again as before we extend x2 to form a basis of C to the n minus 1.

So, the proof is constructive, so it shows that you can construct a matrix satisfying u Hermitian Au is equal to t, where t is an upper triangular matrix with diagonal entries lambda 1 to lambda m, we are just constructing this matrix. So, this is x2 and say y2 up to yn minus 1, so I need to find another n minus 2 vectors, there is one vector already. So, these n minus 2 vectors together form a basis for C to the n minus 1. And then we will apply Gram Schmidt and we will get an orthonormal matrix which we will call u2 which is equal to x2, z2, zn minus 1.

(Refer Slide Time: 25:05)





And this u 2 is such that, u 2 Hermitian A 1 u 2 will be equal to what, it will be exactly in this form here. So, it will have lambda 2 here and 0s somewhere else in the first column and a matrix which we will call A 2 down here and here it has something which we do not care about.

And again, this matrix will have Eigen values lambda 3 to lambda n, and it is of size n minus 2 cross n minus 2. But ultimately, we want to multiply with the matrix A. So, it is not enough if we can find a u 2 such that this happens, we need to find a matrix that I can multiply with A. So, let V2 be the big matrix 1 0 0 and then 0s in the first row and u 2 down here. Then V2 is unitary, you can verify this by direct multiplication V2 Hermitian V2 will give you the identity matrix.

So, if v2 is unitary, clearly rather remember maybe that u1 v2 the product of unitary matrices is also unitary and u1 v2 Hermitian A u1 v2 will give us the matrix lambda 1 0s and lambda 2 in the second row 0s and this can be something and here also you may have something out here, but you will have the matrix A 2 here. So, now we got lambda 1 lambda 2 on the title. Yes?

Student: Sir, could you explain v 2 once more please?

Professor: v 2 is just A matrix where I have padded 1 0 0 0 1 on the left and then this the all 0s vector on this part on the right on the top. So, it is now this matrix is now of size n cross n.

Student: Okay, okay. So, we have just extended u 2 to n cross n.

Professor: Yes. So, basically on the side on the left-hand top you insert 0s, but in the top left you insert A 1. So, basically, now we see the pattern. So, we continue this process, to get this matrix u which is equal to u1 v2 up to v n 1 which is unitary and u Hermitian A u is equal to T which is in the desired form. Basically, upper triangular with all the eigenvalues going on the desired form.

Now, for the last part of the theorem, if A and its eigenvalues are real then the Eigen vectors can be chosen to be real. And all the above arguments can be applied with real arithmetic so that clarifies the last assertion. So, if A and its Eigen values all the above, let me put it this way, the eigenvector of A can be chosen to be real. And all the above steps can be done with real arithmetic. So, that is the proof.

There is another version of this theorem for considering strictly real matrices. And basically, as we have seen, the strictly real matrices do not necessarily have real eigenvalues. But it turns out that if the (eigen), so here is the statement, if a matrix is real valued, then its Eigen values always occur in complex conjugate pairs. Why is that true?

Student: Sir, can you please repeat?

Professor: If a matrix is real valued, then the Eigen values, even if they are complex valued, always occur in complex conjugate pairs.

Student: Sir, (())(32:14) is sum of all Eigen values.

Professor: Okay. So, if the sum of all Eigen values is real, does that mean that the eigenvalues must occur as complex conjugate pairs?

Student: Yes, sir, then the sum of the conjugate pairs becomes real, if any one of them…

Professor: That goes the other way, right? If I take complex conjugate numbers and add them up, I will get A real number. But, it does not mean that the only way to get real numbers is by adding complex conjugate pairs.

Student: Sir can we say that the characteristic, the root of the characteristic polynomial is eigenvalue and the roots always occur in complex conjugate pairs?

Professor: That is correct. But why is that true?

Student: Sir, that I do not know.

Professor: That is because the characteristic polynomial coefficients are some strange combinations of the entries of the matrix A, right? And if A is real valued, the coefficients of the characteristic polynomial are all real valued. So, if I have an equation like so again, this is an aside. If I have an equation like say, a n lambda to the n plus a n minus 1 lambda to the n minus 1 plus dot dot dot plus a naught equal 0, this is the characteristic equation, all these coefficients come from the matrix A.

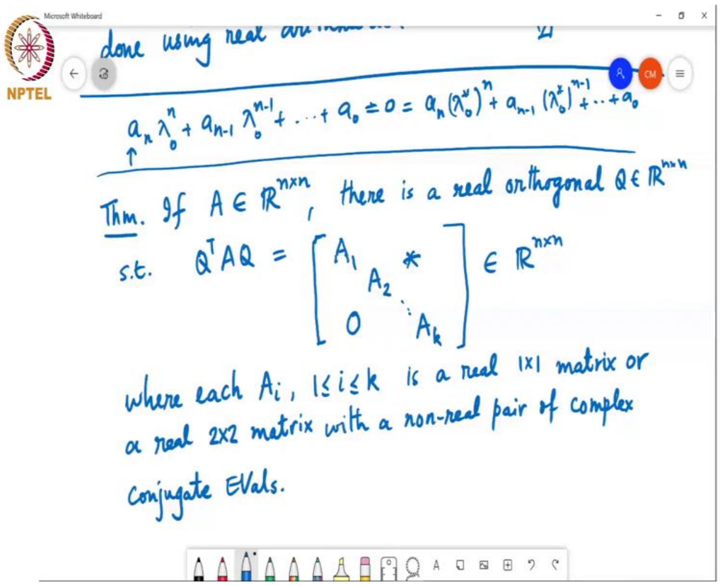
They are just some strange combinations of the entries. It is a multinomial combination of the entries of the matrix A and so all these are real valued. So, if there is a lambda naught for which this is true, if I just take the complex conjugate of this equation, then 0 the complex conjugate of 0 is 0. So, I have that a n lambda naught star to the n plus a n minus 1 lambda naught star to the n minus 1 plus et cetera plus a naught is also equal to 0.

So, basically, if lambda naught is 0 of the characteristic polynomial, then lambda naught star is also 0 of the characteristic polynomial. So, they always occur in complex conjugate pairs.

Of course, if lambda naught is real valued, lambda naught star is root of the characteristic polynomial, but it does not mean that it needs to be a, it has to be a repeated root it could be A solitary root.

But in that case saying that if lambda naught is a root, lambda naught star is also a root is not saying anything else anything new because lambda naught and lambda naught star are actually the same number.

(Refer Slide Time: 35:37)

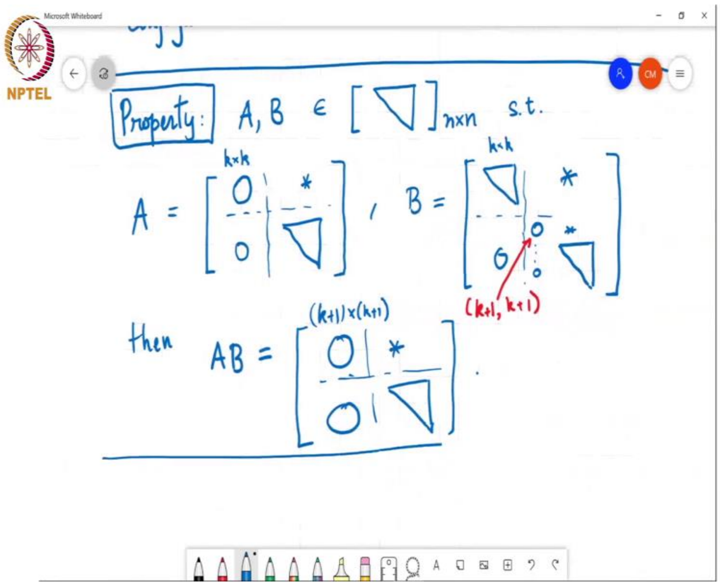


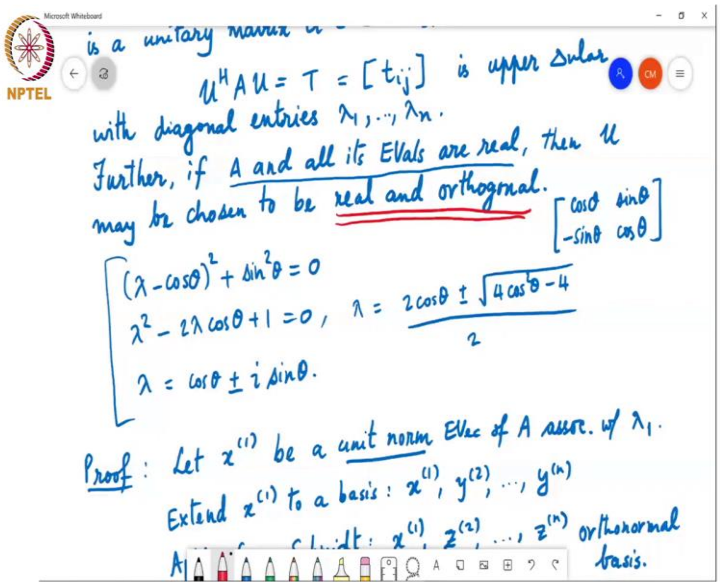
So, let us now discuss the next theorem, if A, then there is a real orthogonal matrix Q in R to the n cross n, also real valued matrix such that Q transpose AQ is equal to big matrix containing A1 A2 down to some Ak along the diagonal 0s here and arbitrary things above the diagonal which is also real valued.

Of course, here all these are real values, so when I take their product, it cannot suddenly become complex valued where each Ai is real 1 cross 1 matrix or the real 2 cross 2 matrix with a non-real pair of Eigen values. So, the only difference that it makes is that you will not necessarily get an upper triangular form, you will get a block upper triangular form, where these blocks are either 1 cross 1 blocks or 2 cross 2 blocks.

If they are 1 cross 1 blocks they correspond to real valued Eigen values of the matrix A and if they are 2 cross 2 blocks they correspond to non-real Eigen values of A which occur as complex conjugate pairs. I will not prove this theorem The proof is actually similar to the previous theorem and you can see the text. But we will discuss some consequences of the Schur's Triangularization theorem because we said it is a very useful result.

(Refer Slide Time: 38:40)





So, let us discuss some outcomes or some interesting things you can show by using this this theorem. First…

Student: Sir.

Professor: Yes go ahead please.

Student: Sir, in the theorem, we have told that if A have, A n Eigen values are real then u is real and, orthogonal or orthonormal matrix sir that you…? The example you have written is orthonormal only.

Professor: Yeah, so my notation is that an (orthonormal), so when I say orthonormal matrix, what I mean is I mean A matrix satisfying u Hermitian u is the identity matrix. When I say real orthogonal matrix, I mean a matrix Q such that Q transpose Q is the identity matrix and the entries of Q are real. So, there is a slight abuse of nomenclature here, but when I say real and orthogonal it actually means the real and orthonormal.

But I am using orthonormal to generally represent complex valued matrices. So, instead of saying complex orthonormal, and real orthonormal, I am saying orthonormal, which is for the complex case, and real orthogonal for the real, real case, but really the columns of this real and orthogonal matrix are all (())(40:20) now. And the columns are all orthogonal to each other.

Student: Yes sir, thank you sir.

Professor: So, there is one interesting property I will first discuss and we will show one cool result that you can show or establish by using Schur's Triangularization theorem. By the way, have any of you seen this, so suppose A and B are n cross n upper triangular matrices, such that they have some special structure. Both are upper triangular but A is of the form 0 and then of course below this will be 0, and then it has an upper triangular form here and this can be arbitrary, where this is a k cross k block of 0s.

And B is of the form, this k cross k block can be upper triangular and non-zero and of course below this is 0. And here below this, you will have 0 and of course below this, since it is upper triangular it has to always be 0. And here it can be arbitrary, but this part again upper triangular and this is arbitrary again. So, basically this here, is the k plus 1 comma k plus 1th element.

So, this the k plus 1, k plus 1th element is 0, here there is a k plus k block of 0s. Then if I consider the product AB, anybody wants to guess what will be the structure in this matrix? Multiply these together, basically when you multiply this with this you will get 0 and here you are multiplying column which has 0s down here with this matrix and so it will lead to a matrix which has 0, which is of size k plus 1 cross k plus 1 and of course below that it is all 0s.

And here it is arbitrary and here it is again upper triangular. So, products of upper triangular matrices is upper triangular, but this is a special extra structure that I am imposing as a consequence of which, this matrix has a k plus 1 cross k plus 1 block of 0s. So, you can verify this by direct multiplication by considering entries with you know B igs entries and A ijf entries here and see what happens when you take them, take the product, but this is true, this property holds true. So, now we will use this property in the result.