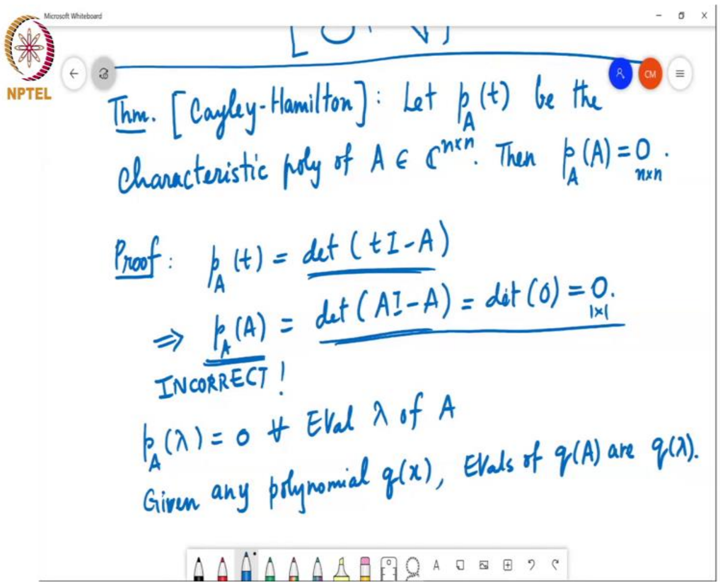
Matrix Theory
Professor Chandra R Murthy
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore
Lecture 41
Cayley-Hamilton Theorem

(Refer Slide Time: 0:15)



Sure, unlike the Cayley Hamilton unlike the Schur's triangularization theorem you have seen or in your undergraduate curriculum. So, this is the Cayley Hamilton theorem. Has anybody not heard of the Cayley Hamilton theorem? So, let p A of t be the characteristic polynomial of A. Then what does Cayley Hamilton theorem say?

Student: Yes it (())(0:59) characteristic question.

Professor: So, A of A is the 0 where the 0 is actually a 0 matrix. This is a polynomial but I am evaluating the polynomial at t taking the value equal to an n cross n matrix A. So, I am replacing t with an n cross n matrix. And I am evaluating the characteristic polynomial and I get all 0 matrix. So, this is 0 of size n cross n. So, that is what the Cayley Hamilton theorem says that any matrix satisfies its characteristic polynomial.

So, again it is a this is again one of those very cool results of linear algebra that is completely non intuitive to me at least. I cannot give you an intuitive reason why a matrix should satisfy its own characteristic polynomial. So, let us so have you guys seen a proof of this theorem. So, p A of t is by definition determinant of t I minus A. So, that implies that if I wanted to find p A of A that is equal to the determinant of AI minus A which is equal to the determinant of the all 0 matrix which is equal to 0. Is that correct?

Student: Yes sir.

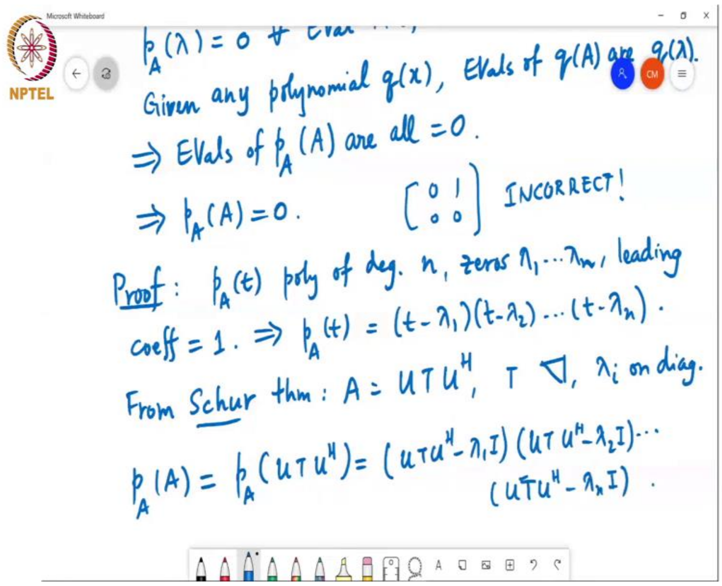
Professor: This is an incorrect proof you know you can see that p.

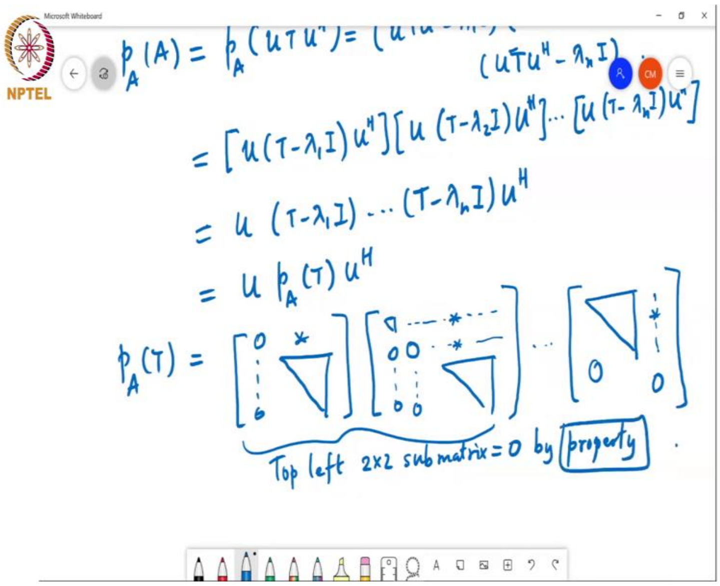
Student: Determinant real numbers it is not a matrix.

Professor: That is a very good point. So, obviously I already observed that p A of A is actually a matrix of size n cross n. But if I do it this way this thing is giving me a scalar value this is 1 cross 1. So, some the I mean there is no meaning to saying that p A of A is equal to determinant of A I minus A. So, what went wrong was I should have actually first evaluated this and I get a polynomial out of this.

And then substitute A instead of t in that polynomial I cannot directly substitute replace t with A inside the expression for the determinant. So, this is an incorrect proof. Here let me make another attempt. So, we know that p A of lambda equals 0 for every eigenvalue lambda of A. That is that we know because by definition an eigenvalue is 0 of a characteristic of the characteristic polynomial. So, we also know that given any polynomial q of x the eigenvalues of q of A are what q of lambda. So, I just evaluate that polynomial at the Eigen values and I get the Eigen values of q of A.

(Refer Slide Time: 5:15)





So, that means that if I wanted to find the eigenvalues of p A of A that would be p A of lambda which is equal to 0. Because this is true for every eigenvalue of A are all equal to 0. So, that means that whatever this p A of A is it is a matrix whose Eigen values are all equal to 0. So, that implies p A of A equals 0.

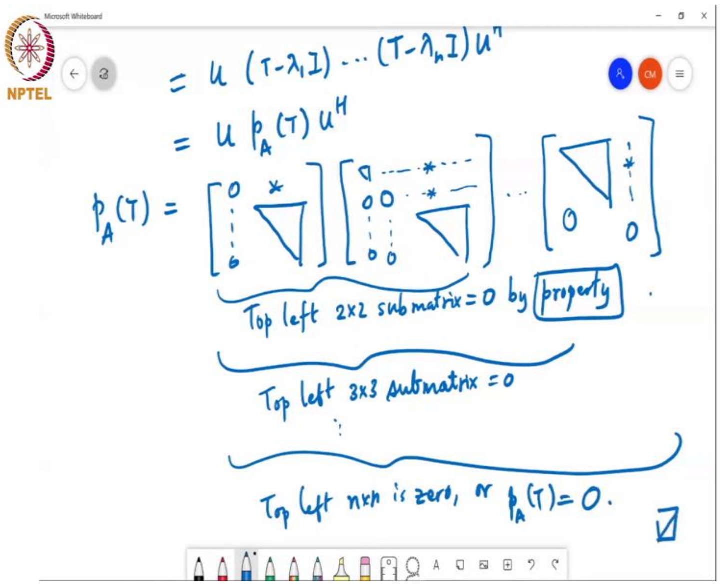
So, this is the proof so none of your object so the fallacy of this proof is returning even if the eigenvalues are 0 it does not mean that the matrix should be equal to 0. We already have seen this example 0 1 0 0 there are many other examples. Where the matrix is not 0 but its eigenvalues are all equal to 0. So, this is also incorrect. So, now we will actually now this is the correct proof.

So, basically p A of t is a polynomial of degree n with 0's lambda 1 through lambda n. And also, the leading coefficient of this p A of t is 1 that is the coefficient of t power n is 1. So, that implies I can write p A of t in the form t minus lambda 1 into t minus lambda 2 t minus lambda n. So, this is the form of the characteristic polynomial. Now, from Schur theorem so I said we will see some uses of Schur theorem.

So, this is the use we are using the theorem we can write A equal to u T u Hermitian I am just taking the u to the other side where T is upper triangular with lambda i on diagonal. So, then if I evaluate p A of A because I have written this polynomial out in the expanded form like this. I can now actually substitute for A in this in this polynomial equation. So, I will write it as p A of u T u Hermitian which is equal to u T u Hermitian minus lambda 1 I times u T u Hermitian minus lambda 2 I et cetera up to u T u Hermitian minus lambda n I.

So, that is equal to I can write it I as u u Hermitian and pull out a u on the left and a u Hermitian on the right. And write it as u T minus lambda 1 I u Hermitian times u T minus lambda 2 I u Hermitian u T minus lambda n I u Hermitian. And this u Hermitian u is all the identity matrix. So, I will be left with u on the left and a u Hermitian on the right and all the middle u's will actually disappear. They are just the identity matrix. So, I can write this as u T minus lambda 1 I T minus lambda n I u Hermitian Which is equal to u times p A of T. So, this thing is actually the characteristic polynomial of T as well because T is similar to A and u Hermitian. And so clearly p A of A will be 0 if and only if this p A of T is equal to 0.

(Refer Slide Time: 10:48)



Now, p A of T is equal to this matrix if I look at t minus lambda 1 I lambda 1 is the 1 comma 1 the element of T. And so, this will have a 0 on the top. And it may have something here and this is upper triangular. And next matrix has something here is the 1 comma 1 the element but the second element will be 0. And then this will have 0's all down the first column. And over here there could be arbitrary things. And then there are 0's down here. And this the rest of this row can be arbitrary but this part will be upper triangular and so on.

And the last one will have some structure like this and then 0 at the bottom then 0's down here. But now this form here is exactly the form we studied in the property where this element is 0. And so, when I multiply these two together the first two elements will become 0. So, if I multiply these two. The top 2 cross 2 top left 2 cross 2 sub matrix is 0 in by the property we just discussed. And if I take the first three matrices together after that then you will see that the top left 3 cross 3 sub matrix equals 0 and so on. So, if you multiply all of these together the top left n cross n is 0 or p A of T is the 0 matrix. And so that proves the result. So, I am a bit over time so we will stop here.