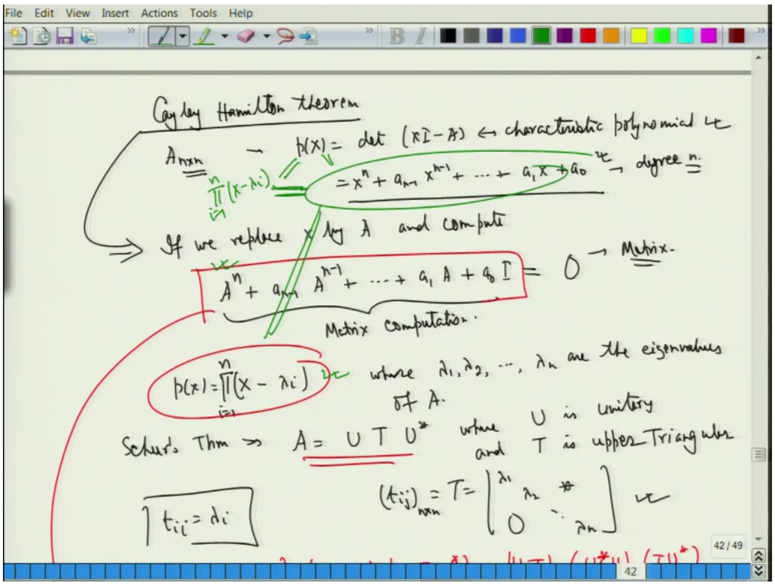
Linear Algebra
Prof. Arbind Kumar Lal
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur

Lecture – 61
Cayley Hamilton Theorem

(Refer Slide Time: 00:16)



Alright, so now, let us go to what are called, what is called the Cayley Hamilton theorem, theorem. So, we will be using the Schur Upper Triangularization theorem and the matrix multiplication nothing else. So, let us look at the definition of what is the Cayley Hamilton theorem? So, given A which is n cross n, we compute p lambda or p of X which was nothing, but determinant of X I minus A, the characteristic polynomial, polynomial, alright.

So, it is a polynomial of the type X to the power n plus a n minus 1 X to the power n minus 1; there will be plus or minus signs, so I am not putting out a n minus 1 anything plus a 1 of X plus a 0, this is what the polynomial is going to be. Since the matrix is n cross n; what we are saying is that, it will be a polynomial of degree n fine, that is more important for us, fine.

So, that is the characteristic polynomial we have been computed some of them for eigen values and so on, fine. What Cayley Hamilton theorem says that, this theorem implies that, we can replace x. So, if we replace X by A and compute A to the power n plus a n minus 1 A to the power n minus 1 plus so on plus a 1 of A plus a 0 of I; this is a matrix identity that we are looking at, we are looking at a matrix alright, matrix computation, alright.

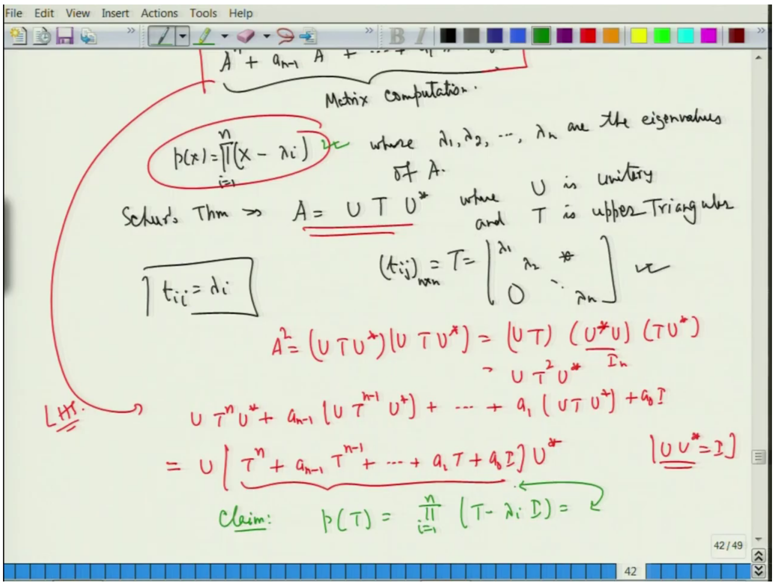
There it was looking at X as a unknown or X as a variable and we are looking at a polynomial in X with coefficients come as A I's, alright. Now, we are looking at powers of A and then computing something; it says that this is same as the 0 matrix. This is what the Cayley Hamilton theorem says that; if I look at this, then it is 0, fine. Proof is simple; let us go into the proof.

What we know is that, since it is a characteristic polynomial; so the roots of this are the eigen values. So, I can write p of x as X minus lambda i product i is equal to 1 to n. So, these are the eigen values, where lambda 1, lambda 2, lambda n are the eigen values of A, alright. So, what we have learnt? Schur's theorem says us Schur's theorem implies; I can write A as U T U star, where U is unitary and T is upper triangular

And not only that, it also tells me that; the diagonal entries of T. So, if I look at T, then T can be written as lambda 1, lambda 2, lambda n; 0 sorry upper triangular, so I am writing. So, 0 is here, I do not know what it is, is that ok? So, this is important that, what we are saying here is that; if I am writing T as t i j fine n cross n, then t i, i is equal to lambda i, I can rearrange my eigen vectors or I can rearrange the columns of U.

I can rearrange the columns of U in such a way that, I get lambda 1 as the first entry here, lambda 2 as the second entry and lambda n as the second entry, fine. Recall when I looked at diagonalization what we had done was; we wrote p and we showed that, A of u 1 is equal to d 1 u 1, A of u 2 was d 2 of u 2 and so on. So, I could rearrange them depending on the matrix p that I take, alright. So, there is just a rearrangement of permutation which will give me this part. So, now I want to understand, I know that A is this; what happens to A, A square and so on?

(Refer Slide Time: 04:37)



So, we if I look at A square; A square is nothing, but U T U star into U T U star, which is same as U times which is same as U T times U star U T U star which is. So, this part is identity, I n. So, this gives me U times T star U star, alright fine. So, therefore, I can write this matrix as the left hand side LHS as; I just multiply everything by U on the left and U star on the right, is that I do that?

So, what I do? I multiply here everything and then look at things, is that ok? So, A to the power n from here if I look at here what I get is U times T to the power n U star plus A to the power n minus 1 is a n minus 1; A to the power n minus 1 will be again U T to the power n minus 1 U star plus so on plus a 1 it will be U T U star plus a 0 of I this is what the left hand side is, fine.

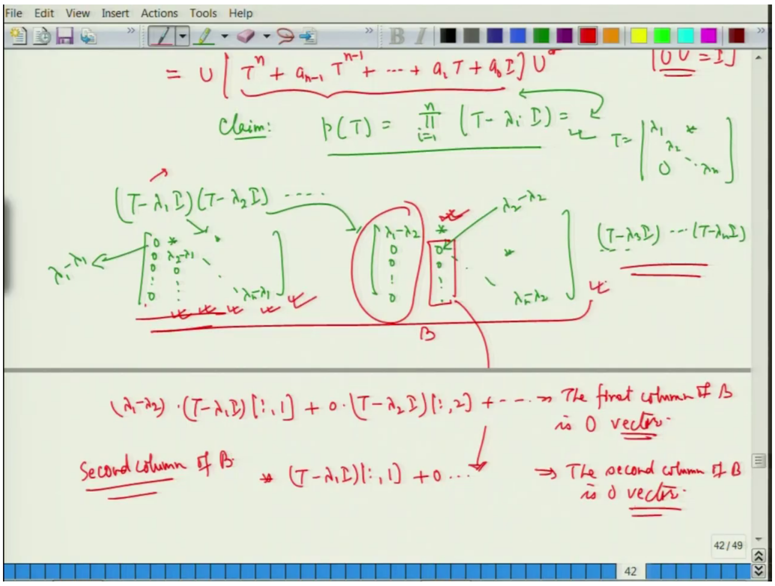
So, this is same as the left hand side is also equal to; I can take out U outside T to the power n plus a n minus 1 T to the power n minus 1 plus so on plus a 1 of T plus a 0 of I times U star, because U times U star is identity, U was an unitary matrix, alright. So, this is how unitary matrices are very helpful to us that, we can do a lot of things in place of going to the inverse; I can just use transpose or conjugate transpose and do the work, alright.

So, now, we are looking at the left hand side which is this. So, now, we need to compute this part for us; and from where we did we get it? I am looking at p of X as this, is that ok? So, p of X was equal to this. So, I would like to claim that, if I look at p of T as product i is equal to 1 to n T minus lambda i of I; then this is equal to this expression that I have, alright. Why I am saying this? Because let us go back and see; what we had was, this is equal to this part we are saying.

So, we are saying that, p X is equal to product i is equal to 1 to n X minus lambda i, fine. To get this expression in A, we replaced X by A, alright. So, in this expression I am replacing X by A; but this is same as this which is same as this is an identity, this is an identity in X. If something is an identity in X, I can replace X by any object that I want to make things meaningful. So, I can replace X here by A here, is that, that is more important for me; I can replace if this is an identity, there is an identity here going like this.

So, I can replace X by A and I have to replace the constant by constant times identity to make sense of it. So, this is what we have done and I am done nothing else. So, I have written this, p T as this, is that ok? So, now, let us compute this product i is equal to 1 to n T minus lambda i of I and we are done.

(Refer Slide Time: 08:05)



So, let us look at this product very nicely. What we have is T we wrote it as lambda 1 star, lambda 2 till lambda n, this is what we do. So, I want to compute what is T minus lambda 1 of I and let us compute T minus lambda 2 of I and so on, I want to compute this, is that ok?

So, what is T minus lambda 1 of I? Look at T minus lambda 1 of I; the first column is 0, alright. Second column if I look a; it is a star here, lambda 2 minus lambda 1, 0 here and there are some things here. So, this is lambda n minus lambda 1, this is a matrix this. What is this matrix?

This matrix if I look at the second matrix; second matrix is lambda 1 minus lambda 2, 0 here, 0 here, 0, a star here, 0 here, because this will increase lambda 2 minus lambda 2, this will increase lambda 1 minus lambda 1, 0 here and I have something here which will give me lambda n minus lambda 2 and a star here. And there are so many things here which are T minus lambda 3 of I so on till T minus lambda n of I, alright. So, I am just looking at this matrix product and nothing else.

So, be careful; do not get nervous, everything is nice here. So, if I multiply this matrix, this column with this matrix; what do I see is, it is lambda 1 minus lambda 2 times the first column. So, it is lambda 1 minus lambda 2 times.

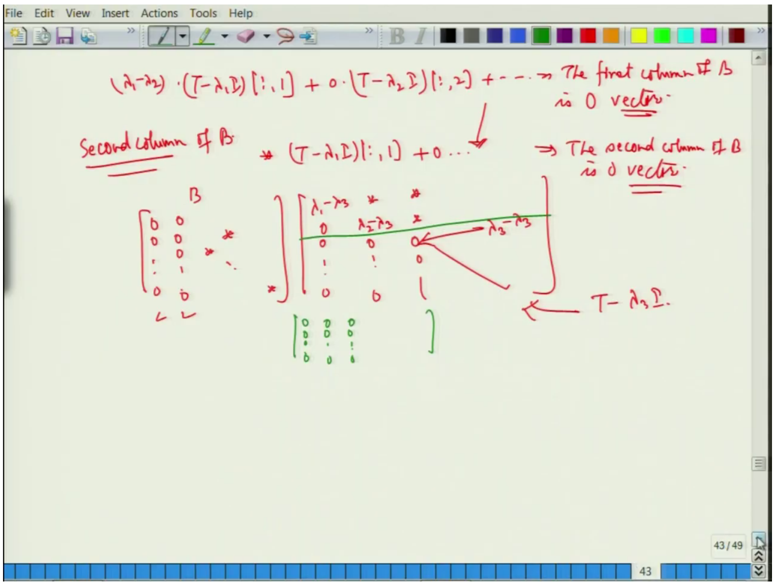
So, so if I write this as or let me write T minus lambda 1 of I itself; look at the first column of this plus T of T minus lambda 2 of I, the second column and so on, alright fine. I am just multiplying these two matrix; I am not bothered about this part, I am just trying to understand the product of the first two matrices.

So, it is lambda 1 minus lambda 2 times the first column which is all 0; now it is 0 times everything else. So, this will give me that, this implies the first column of this product whatever it is I say; if I write this matrix as B of B is 0 vector, alright. Here itself it is 0, I still get 0 here; is that I have not got anything extra, let us look at the second column of B, second column of B.

If you look at the second column of B, I am multiplying by a star. So, I am multiplying a star to T minus lambda 1 of I the first column and then again it is 0 times everything else; 0 times somethings, alright. Look at here it is 0 times, fine. So, I have got 0 times something here; a star is the first entry, but the rest everything is 0. So, I am multiplying the second column here, second column here to 0; second column, third column to 0, fourth column is being multiplied by 0, all the columns here are being multiplied by 0, is that fine.

So, therefore, this is also this implies that, the second column of B is 0 vector, alright fine. So, what I have shown is that, if I multiply the first two; what I get is, I get 0 in the first column, 0 vector in the second column. Earlier I had only 0 in the first column; after multiplication I get two part to be 0.

(Refer Slide Time: 11:50)



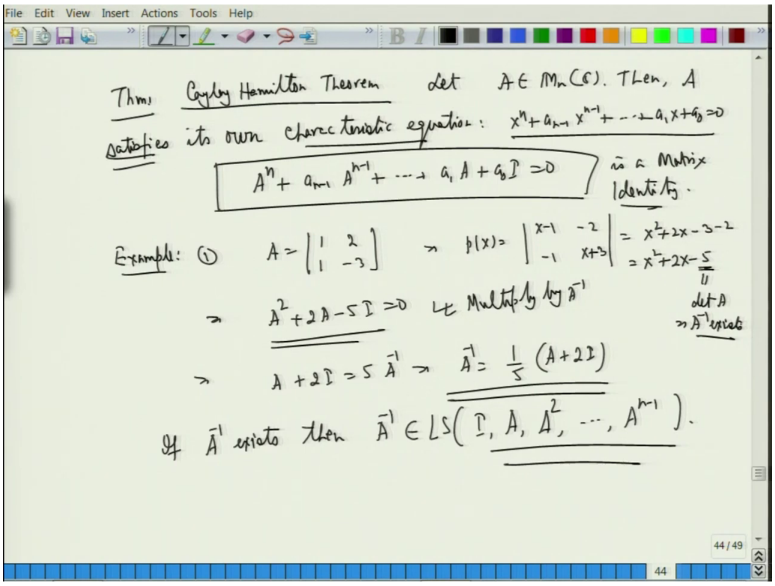
So, what we are getting is that, I have got in B. So, if I look at this B part, I have got 0 vector here, 0 vector here, fine. I do not know what is there for me here fine; there will be some entries here on the diagonal, I do not know.

Now, I want to look at T minus lambda 3. So, T minus lambda 3 means, it is lambda 3 T minus lambda 3. So, it is lambda 1 minus lambda 3, 0, 0, 0, a star here, lambda 2 minus lambda 3, 0, 0, 0, a star here, a star here; this is 0, because of lambda 3 minus lambda 3 and 0's here and so on, fine.

So, this corresponds to a matrix T minus lambda 3 of I, just multiply it again. Since the first two columns at 0 here fine; this part the top first two rows of this on the right hand side does not give me anything. So, what I will have is that, this product will give me 0 in the first column, 0 in the second column and 0 in the third column, is that ok ?

So, at each stage I am increasing the 0, the columns having 0 entries; and therefore finally I will get that the whole matrix is 0, is that ok? So, just try that out yourself, prove it yourself; I have given you the idea, I do not want to spend time there, just complete it yourself, fine. So, try that out. So, I would now like to write the statements here which are important for me.

(Refer Slide Time: 13:17)



So, theorem, the statement of the theorem let me write; Cayley Hamilton theorem I am writing it again, fine. So, let A belong to M n of c. Then, A satisfies its own characteristic equation. So, equation means, characteristic. So, what is characteristic equation? Characteristic equation is nothing, but X to the power n plus a n minus 1 X to the power n minus 1 plus a 1 of X plus a 0 is equal to 0, that is the characteristic equation.

It says that A satisfies its own characteristic equation; it means that, A to the power n plus a n minus 1 A to the power n minus 1 plus a 1 of A plus a 0 of I is 0. So, this is what we proved, is that fine.

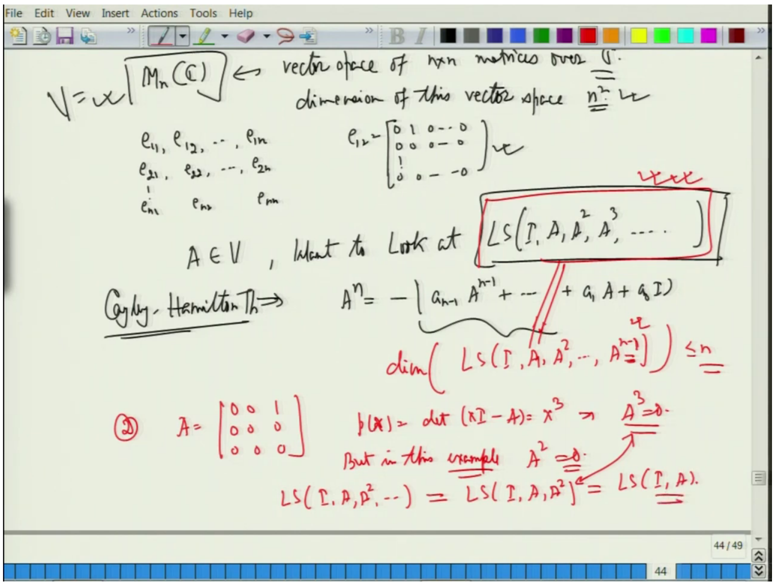
So, this is satisfies as a matrix identity that is more important, I am not equating this at X; what I am saying is that, I am not equating it at eigenvectors of A, we are saying that this is a, this is a matrix identity, alright. So, this is true for every X that you take in c n that is more important alright, I have given you the proof.

So, let us look at some of the some examples here, example 1. So, let us take A as 1, 2, 1, minus 3. So, p lambda or p of X will be equal to determinant of X minus 1, minus 2, minus 1, x plus 3; I hope I am correct, which is equal to x square plus 2 x minus 3 and minus 2 which is same as x square plus 2 x minus 5.

So, this implies that, A square plus 2 A minus 5 times identity is 0 fine, is that ok? This is what you get. Now, recall minus 5 is nothing, but determinant of A, alright. Just look at this. So, minus 5 is determinant of A; implies A inverse exists. So, I can multiply this by A inverse, multiply by A inverse; this will imply that A plus 2 times identity is equal to 5 times A inverse implies, A inverse is equal to 1 upon 5 times A plus 2 I, is that ok?

So, we are able to say that, A inverse whenever it exists. So, this part tells me that, as an application of this or the understanding of this it says that, if A inverse exists; then A inverse belongs to linear span of I, A, A square till A n minus 1, alright. So, we are able to say that A inverse is also in the linear span.

(Refer Slide Time: 16:46)



So, let us understand what I am trying to say here; what I am trying to say here is that, I have this matrix M n of C, which is a vector space, vector space of n cross n matrices over C complex numbers.

Dimension of this vector space is; what is the dimension of this vector space? Recall this is n square. Why it is n square? Because its basis are going to look like e 1 1, e 1 2, e 1 n; e 2 1, e 2 2, e 2 n; e n 1, e n 2, e n n. And what is e 1; e 1 2 is nothing, but the matrix 0, 1, 0, 0; 0, 0, 0, 0; 0, 0; 0 this is a matrix e 1 2, alright. So, recall that part; we had done it very early at some stage, fine.

So, this is the vector space. Since it is a vector space there is an A here in this; I would like to look at what is the linear span? So, given this vector space I write V as this A belongs to V.

So, I want to look at want to look at linear span of I, A, A square, A cube all the way till infinity, alright fine.

So, I want to look at the linear span of this. The Cayley Hamilton theorem says that, Cayley Hamilton theorem Hamilton implies alright theorem implies that; I can write A to the power n as minus of a n minus 1 A to the power n minus 1 plus so on plus a 1 of A plus a 0 of identity.

So, the whole space here; this linear span is nothing, but this is equal to linear span of I, A, A square till A to the power n minus 1. I do not have to go beyond n minus 1; even though I want to find linear span of I, A, A square till any power of A, I just have to restrict myself to understanding I, A, A square till A to the power n minus.

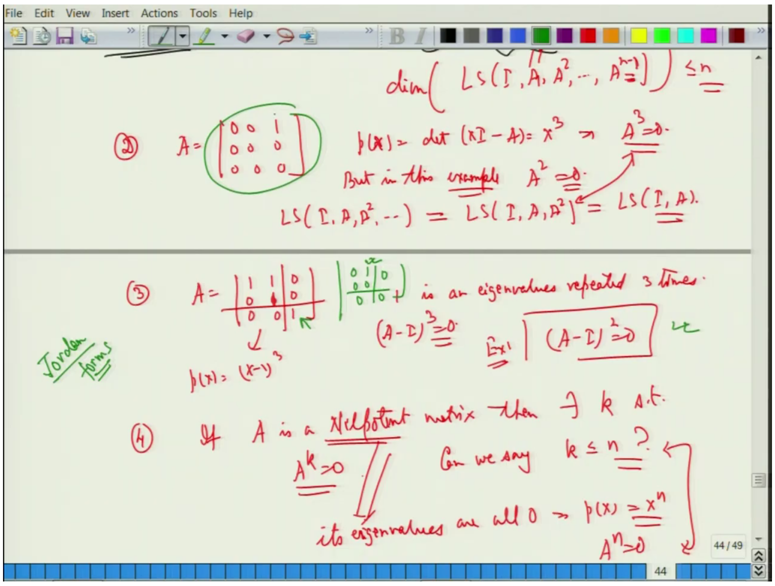
I do not need to go beyond that, is that ok, that is very very important for me. And that is a crucial thing in saying that, if I look at the dimension of this space; dimension of this space, it has how many vectors? 1 2 3 till n; so dimension of this is less than equal to n minus 1, fine that is important.

So, there are things that it tells me, it tells me that; this linear span of this which has infinite number of powers of A, there is no need to go to all the infinite powers, I just need to go till the power n minus 1 that is one thing it says. It also tells me that, its dimension is less than equal to n, not n minus 1; I wrote it wrongly, because there are n vectors here. There are examples where you can show that, this linear span, this dimension is strictly less than n.

For example: for the identity matrix its dimension is only one alright, the identity itself. So, try that out yourself, let us look at some more examples. Example 2, take a as 0, 0, 1; 0, 0, 0 and 0, 0, 0. Let us look at this A. Then you can see that p lambda or p of X for me is determinant of X I minus A will be equal to X cube. So, this implies A cube is 0 that you can check, this is the Cayley Hamilton theorem.

But what you can verify here is that, A square itself is 0; but in this case, in this example A square is 0, alright fine. So, therefore, what we are saying is that, in here linear span of I, A, A square and so on; Cayley Hamilton theorem says that, this is same as linear span of I, A, A square, because A cube is 0, fine. It has to go only till n minus 1; so I can go till A square only. But in this example, we see that this is same as linear span of I and A and nothing else, is that ok?

(Refer Slide Time: 21:25)



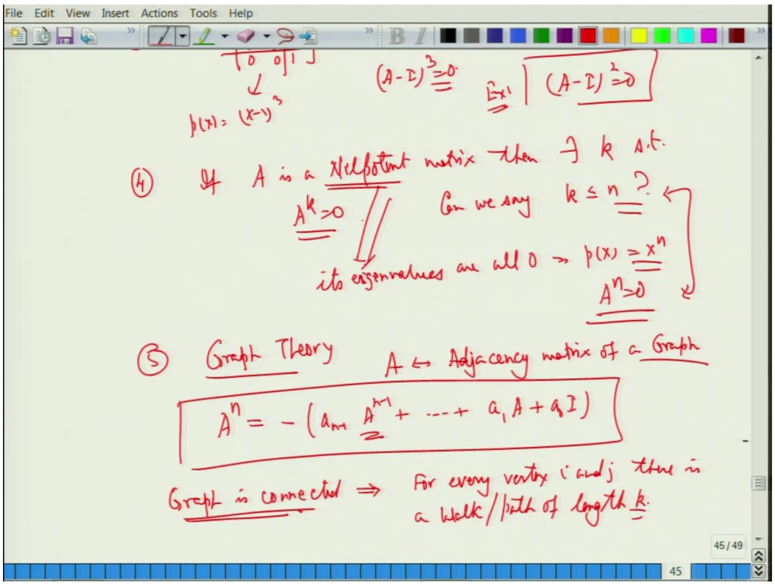
That is more important; anyway this is not invertible, but this is what you have. You can look at another example 3, where take A as which is invertible 1, 1, 0; 0s, 0, 1, 0; 0, 0, 1. If I look at this, I can write like this part sorry; not this one, but like this. So, I would like you to see here is that, 1 is an eigen value, 1 is an eigen value repeated thrice repeated, 3 times, fine.

So, by Cayley Hamilton theorem what I can say is that, A minus I whole cube will be 0 fine; because characteristic polynomial is X minus 1 whole cube, this is what p of X is, I can get this part. But I would like you to verify that, A minus I square itself is 0, alright So, try that out.

So, exercise try that out yourself, is that fine. Some more implications 4, what we have seen is that; if A is a Nilpotent matrix is a Nilpotent matrix, then there exists a k such that A to the power k is 0, fine. Can we say k has to be less than equal to n that is the question, alright? Initially we do not know what that is; but we would like to say that k is less than equal to n. How do I do that?

So, the idea is very simple that, since A is Nilpotent implies its eigenvalues are all 0; implies its p X, the characteristic polynomial is X to the power n and therefore, by Cayley Hamilton theorem I know that A to the power n is 0, is that ok? And therefore, I get that k can has to be less than or equal to n, fine.

(Refer Slide Time: 23:23)



Another important application of this is what is called graph theory, graph theory. So, take A to be the adjacency matrix of a graph, of a graph, alright. I just take this adjacency matrix of the graph. What do I see is, I will compute A to the power n; what we have seen that A to the power n is nothing, but a n minus 1 A to the power n minus 1 plus so on plus a 1 of A plus a 0 of I, fine.

So, look at this expression; if I know, alright. So, look at this, what we want to say is that, G is connected, the graph is connected; it means that, graph is connected means? So, this corresponds to saying that, for every vertex i and j alright, there is a walk or path of length k, alright. What this part tells me that; I just need to go till n minus 1.

So, if i and j are connected, then the path that I want to look at is only till n minus 1; the distance can be, maximum distance between any two vertices of a graph having n vertices will have n minus 1 that is what it says, is that ok. So, I have given you these ideas. In the next class, we look at what are called quadratic forms.

And I would like you to understand that, Nilpotent matrices they play a very important role; we have not done it, but in some sense we have shown here what are called, earlier also I had to used the word Jordan forms, alright fine. So, for matrix here, there is a notion what is called Jordan matrix and things like that, Jordan decomposition; I am not into that, because that is out of the syllabus.

But I would like you to see here that, in this example you had A minus I whole square is 0. So, if I remove I here, from this I can remove this part; then I am looking at the matrix 0 1 0 0; 0 here, 0 here, 0 this is the matrix I am looking at. This is a Nilpotent matrix and its order is 2, because of this extra 1, is that ok?

Similarly, in the previous example here; this is a Nilpotent matrix; this is what the example says. So, all the ideas; so whenever I have some matrix A which is not diagonalizable, somehow this Nilpotent matrix comes and play a very important role, whenever something is not diagonalizable. So, keep track of that, that is all for now.

Thank you.