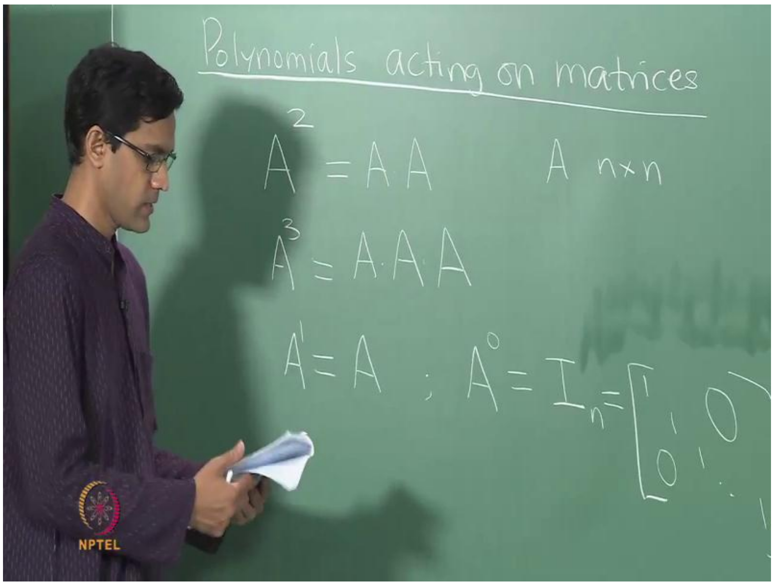
An Invitation to Mathematics
Prof. Sankaran Viswanath
Institute of Mathematical Sciences, Chennai

Unit
Matrices
Lecture - 32
Polynomials acting on matrices

(Refer Slide Time: 00:27)

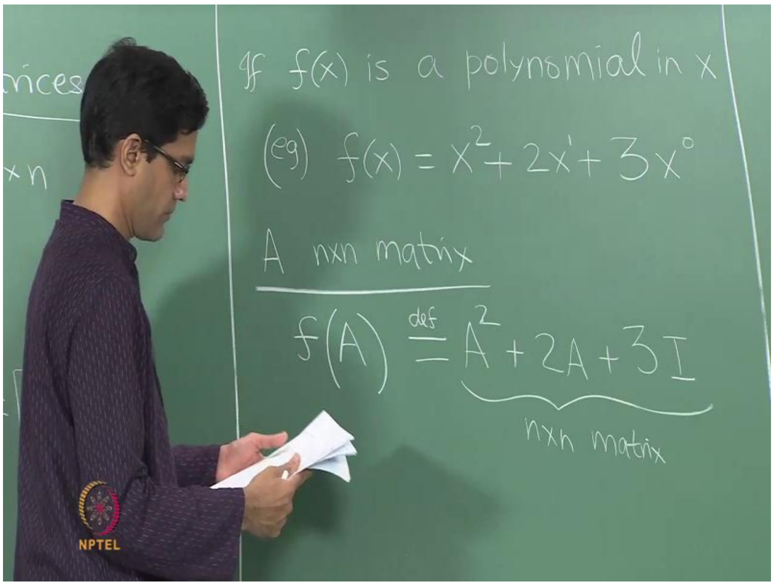


Today we will talk about Polynomials acting on matrices. So, observe last time when we talked about the applications of matrices and counting problems in graph theory, we had occasion to consider powers of a matrices. So, if you had a matrices A, then it was natural to look at the matrix A squared, which we just find to be A times A or the matrix A cubed, which is the product of matrix A with itself thrice. And this number counted, for an instance A squared would count the number of parts of length 2 in a graph, where provided A is the adjacency matrix of the graph, A cube would count parts of length 3 and so on.

So, powers sort of made in natural appearance in counting problems. Now, let us do the following, let us also define. So, A power 1 of course is just A, let us define A power 0 to be the identity matrix. So, what is this, if A is an n crossing n matrix? So, here I am thinking of A as been some n cross n matrix, then the identity matrix here is also of size n cross n. So, I denote it by I n, this just refers to the matrix with 1s along the diagonal, 0's everywhere else and having size enclosed.

So, the 0th power of matrix is defined to be the identity. Now, just like we consider powers of a matrix, it is natural to sort of go one step further and consider polynomials in which you plug in a matrix instead of x.

(Refer Slide Time: 01:50)

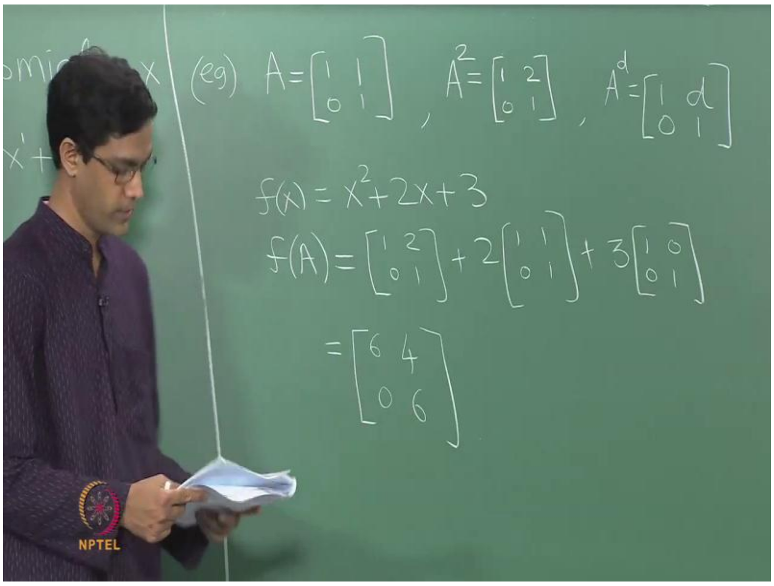


So, what is this mean? If f of x, so polynomial in the variable x for example, let say f of x is just x square plus 2 x plus 3, then here is what we would do. So, let us think of this 3 as just being 3 times x power 0, this is 2 times x power 1. Then, if A is a matrix, if A is any n cross n matrix, we can define the following, you plug in A in place of x. So, the quantity f of A is defined as follows.

So, here is a definition, we just read of the polynomial, but just replace all the x's with A's. So, this is now A squared plus 2 times the matrix A plus 3 times well x power 0, now that should be replaced by A power 0, but A power 0 as I said is just the identity matrix. So, the quantity f of A is just defined to be A square plus 2 A plus 3 I and observe, now that this is a matrix again, this is once more an n cross n matrix.

So, this sort of the general definition, if you have any polynomial with say x power d plus some multiple of x power d minus 1 and so on, wherever you see x power something you replace it by the corresponding power of A. So, that is how you make a polynomial act on a matrix. And let just study this to see, what sort of interesting things can come out of it.

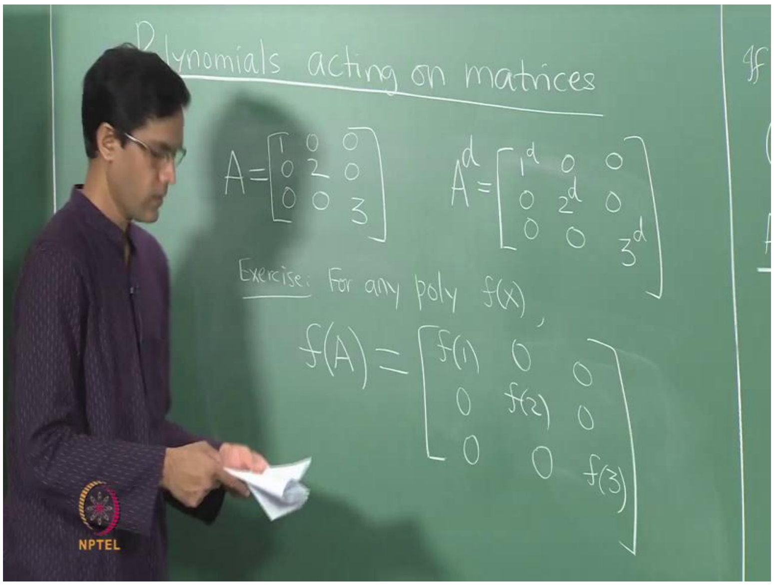
(Refer Slide Time: 03:42)



So, here is an example just to see, what this definition leads to suppose you have a matrix 1 1 0 1, just 2 cross 2, then what I am going to try is to apply make a polynomial on A. So, it will be useful to figure out, what powers of A look like. So, if you compute A square, you will find that it is 1 2 0 1, A cubed similarly will be well or more generally A power d will just become this, 1 d 0 1. So, here is something for you to check, check that A power d indeed leads to the matrix 1 d 0 1.

And so now, let us take the same polynomial f, f of x is x squared plus 2 x plus 3 and let us plug in this matrix A in place of x. So, as we said it is A squared, so A squared is 1 2 0 1 plus 2 times the matrix A, which is 1 1 0 1 plus 3 times the identity matrix, which is well 1 0 0 1. So, all we must do, you just add up these three matrices and check again that the answer here is 6 4 0 6. So, that is just an example to show, how this computation is carried out in general, it will produce some 2 cross 2 matrix. So, let us take another example, so one dimensional higher, let us look at the 3 cross 3 matrix.

(Refer Slide Time: 05:15)

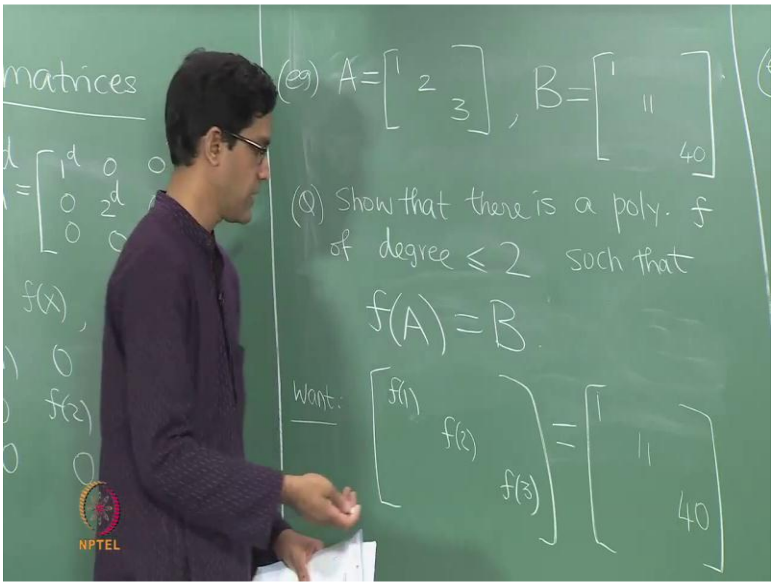


So, let us take the matrix A, you just have diagonal entries 1, 2 and 3 and 0's everywhere else, it is 3 cross 3. Now, observe if you raise this matrices to the dth power, keep multiplying it with itself d times, all it does is just the diagonal entries become 1 power d, which is well 1, 2 power d and 3 power d and the other 0's just remain as it is. So, it is very easy to take powers of diagonal matrices or to multiply diagonal matrices with each other.

And what is this in particular mean, it is very easy to apply polynomials to a diagonal matrix. So, if f is a polynomial, so here is a nice exercise, check that more generally this is something that I wrote out for the dth power of A. But, more generally for any polynomial f of x, check that if I apply f to a diagonal matrix to this particular diagonal matrix, all it will do is produce an another diagonal matrix with diagonal entries f of 1, f of 2 and f of 3.

And the main step of the proof of course is in using what we just said about, that A power d has this form. So, a typical polynomial is a just a sum of powers of A with some coefficients in front. So, if you put them all together, all it will produce is a matrix with diagonal entries f 1, f 2 and f 3.

(Refer Slide Time: 07:06)

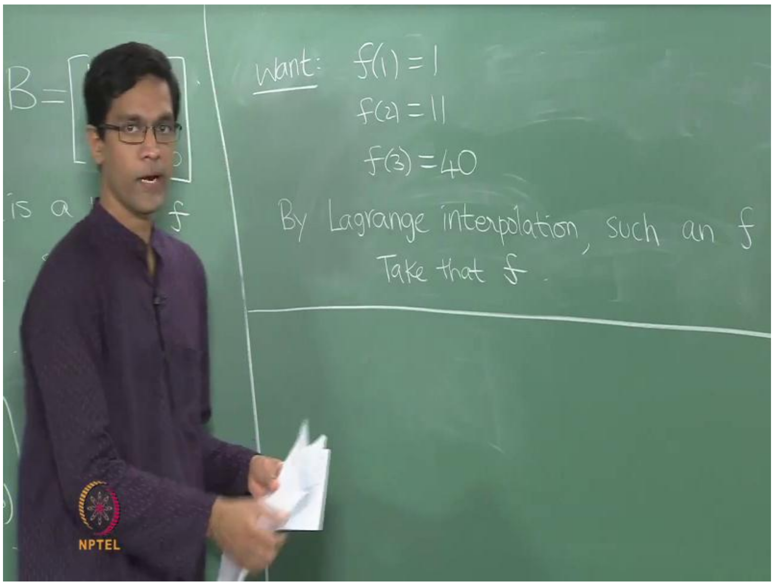


So, this raises or suggest a very interesting question, suppose you have here is the example again, if I take this very same matrix A 1, 2, 3 and B another matrix well, which can be pretty much any diagonal matrix. So, let me just take some random numbers. So, it does not matter, what numbers I pick here, I just picked a few three random numbers 1, 11 and 40, but I could just pick it to be any other matrix B. So, if for instance 11, 11 and 40 or any three other numbers that I can think of.

Now, here is the problem or question, show that there is a polynomial of degree 2, show that there is a polynomial f of degree at most 2, such that when you apply f to the matrix A, it produces the matrix B, so here is the problem. Show that by applying a polynomial of degree 2 to this diagonal matrix 1, 2, 3; you can make it into the matrix B; that is on the right hand side.

And the way to go about this is just for using the observation that we just made, that when you apply a polynomial to a diagonal matrix, all it does is just apply that polynomial to the three diagonal entries. So, let us try and solve this problem, what is that we want our polynomial f to satisfy. So, f of A remember is just, as we just said that diagonal matrix f 1, f 2, f 3, and so what we want is that this should be the matrix B that is given, should equal 1, 11 and 40. In other words, f must be a polynomial whose value at 1, f of 1 must be 1, f of 2 must be 11 and f of 3 must be 40. So, we just trying to solve the following three constrains.

(Refer Slide Time: 09:15)



The words what we want is a polynomial of degree at most 2, it satisfies f of 1 it is 1, f of 2 is 11 and f of 3 it is 40, but observe that exactly what Lagrange interpolation did for us. So, interpolation is exactly solving for this kind of polynomial, if you apply the interpolation procedure, what we would find is a polynomial of degree at most 2, which takes these prescribed values 1, 11 and 40 at these three prescribed points 1, 2 and 3.

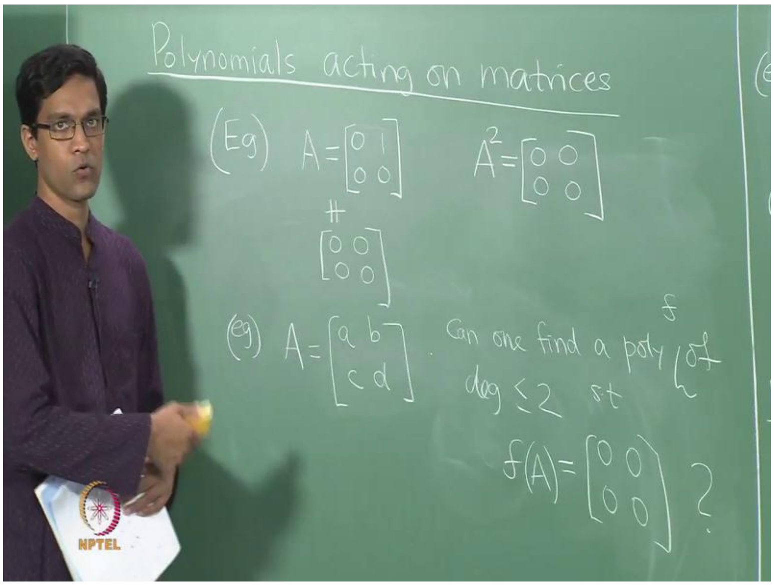
So, by Lagrange interpolation, you can actually write out the formula from Lagrange interpolation, it actually finds such an f. So, such an f exits and in fact you can also write a nice formula for it, such an f exists. So, you take that f; that is exactly the f you want, so take that polynomial alone. So, if you sort of just step back and think about, what we have just proved for a minute, what we have said is that no matter.

So, let us come back here, you fix this particular diagonal matrix A and observe that there is nothing secret about 1, 11 and 40, you could have pick any three numbers y 1, y 2, y 3 and Lagrange interpolation would have allowed us to find the polynomials which has those prescribed values. So, in other words given any diagonal matrix whatsoever with any three real entries on it is diagonals, there always exits a polynomial of degree at most 2, such that when you apply that f to A, you will get that given diagonal matrix.

In other words, every diagonal matrix can be obtained by applying polynomials to this particular diagonal matrix and again, thus nothing too secret about this diagonal matrix, any diagonal matrix with three distinct entries would have done the job. Because, for polynomial interpolation, recall what you need is x 1, x 2, x 3 to be three distinct numbers, y 1, y 2, y 3 could be you know equal, there is no ((Refer Time: 11:41)).

So, what we have really done is to show that, if you start with any diagonal matrix with three distinct entries, you can obtain every other diagonal matrix, but just by applying some polynomial to that diagonal matrix. So, I sort of encourage you to you know work out the few more examples, play with this a little bit more, may be with somewhat higher dimensions and so on.

(Refer Slide Time: 12:05)



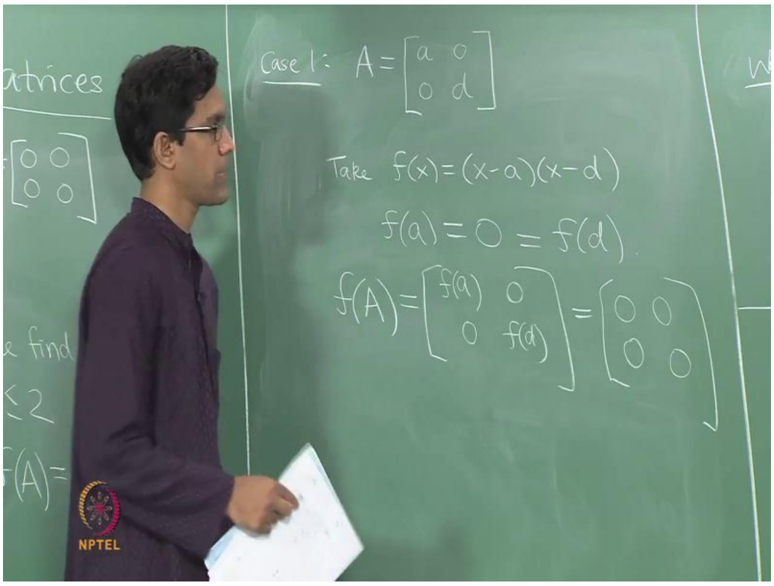
Like to do is again start out with the following example, take a 2 cross 2 matrix A, you just has 0 1 0 0. And again, think in terms of polynomials, so let us compute A squared for instance and here is what one finds, that squaring this matrix just produces a matrix of 0's. So, this is somewhat countering due to what one is used to, when we sort of takes squares of real numbers and so on.

If you have a real number, if A where instead a real number, which is not 0, so observe A here is not the zero matrix, so A is clearly a non zero matrix, but the square of A nevertheless could be 0. So, matrices in this sense behave very differently from how real numbers could behave, a matrix could be non zero, but it is square could be 0. So, again this sort of suggest an extension, which is you start with an arbitrary matrix and try and do something similar, try and apply some polynomial to it and try and make it 0.

So, the question really is, so again let us do it by example, sort of generalization of this here just squaring the matrix was enough to produce the 0. More generally, can you find polynomials which one apply to a matrix will produce the 0 matrix. So, let us just take, so I will take the matrix for now an arbitrary matrix of size 2, a, b, c, d and ask the following question, can we find a polynomial of degree at most 2, give it a name say polynomial f of degree at most 2, such that f apply to this matrix would give you the 0 matrix.

So, arbitrary 2 cross 2 matrix, what polynomial should one apply to it to get 0, well more importantly thus there exits such a polynomial. So, the preceeding discussion has answered this question in a special case.

(Refer Slide Time: 14:27)

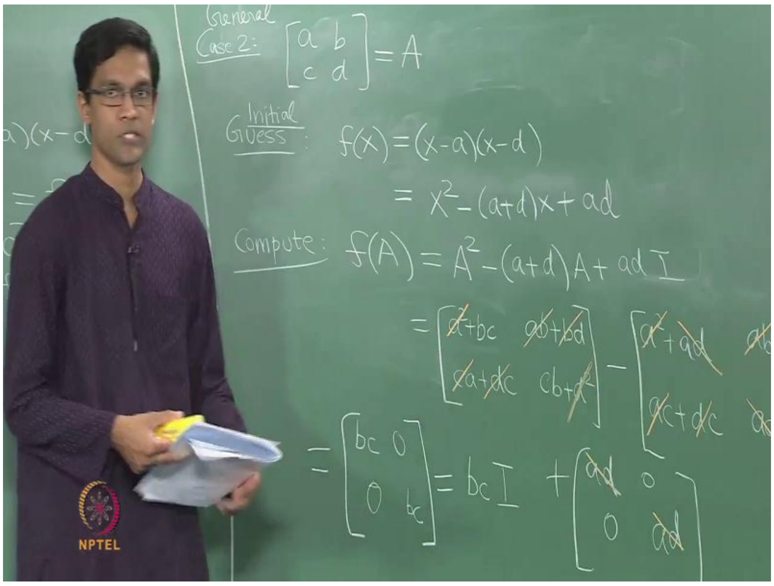


So, observe if A is diagonal then we know what do. So, case 1, if whatever the half diagonal entries b and c are 0. So, if A is 0 0 d, then we can sort of use the same sort of thing we did before, what polynomial should we apply to A in order to make both these entries 0's, well we just take the polynomial x minus a times x minus d. So, let us do the following, let us define the polynomial f of x to be x minus a times x minus d. So, that is a quadratic polynomial.

So, degrees 2 and what property does it have, well if you plug in x equals a or x equals b, it would just give you 0. So, it would make both these entries, it would kill both these entries. So, now, if you apply f to the matrix A as we just said before, what it will give is just f of a, f of d on the diagonals 0, such a way and since by design, these two diagonal entries are actually 0. So, what you do get this is 0 matrices.

So, this general problem that I just post at least in a special case, we know how to attack this if it for a diagonal matrix then we would just used the polynomial x minus a times x minus d. But of course, the question will that work, if you have may be of diagonal entries, so if have d and c being non 0 would be still be able to find the polynomial. So, that is the general case.

(Refer Slide Time: 16:20)



So, let us say case 2 is well, this is just the general case also assume case 1, if you wish. So, a, b, c, d; it is a matrix A. Question is can I find the polynomial f, so let us do the following as a first approximation, let us use the same polynomial that worked in the diagonal case. So, let us do the following, let us guess as our first guess initial guess, let us just take this same polynomial and see what happens.

So, this take x minus a x minus b and this is, so let me this expand out this x square minus a plus d x plus d and let us compute f of A and let see, what we get. So, let us compute f of A. So, what is f of A, it is a square first minus a plus 2 A, so I mean to square A, because ad times identify. So, squaring A of course, requires writing out the matrix A with itself and so out of carrying out the matrix multiplication in full.

So, let me just write out in the sound terms is looks like a square plus b c. So, you should certainly check that this calculation is okay, a b plus b d, c a plus d c and c b plus d squared. So, that just the square of A and the second term in which you multiply each entry of A with A plus d looks like, so it is as minus a square plus a d a b plus d b, a c plus d c and a d plus d square, and then the final term is ad times identity. So, that is just a d, a d with 0 some time. So, I just wrote this out in full. So, you can sort of check you matrix multiplication as well.

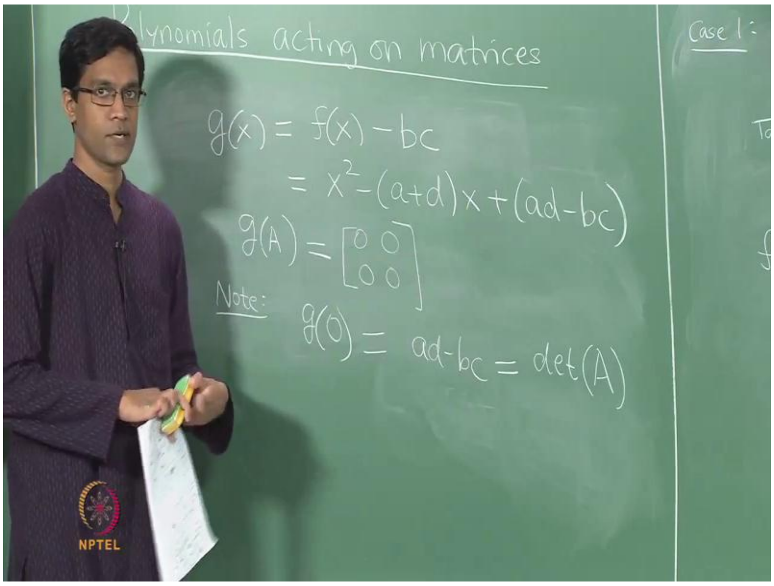
So, that is the complete answer, but now, there is lot of simplification that x place. So, what I need to do is for an instance, I take the top left entry a square plus b c minus a square plus a d plus a d. So, I need to take those three terms and combine them and see what I get. So, for instance here I already see some cancellations. So, a square for instance minus sign with in a square we canceled. So, I have minus a d here which will cancel with the plus a d on the other side.

So, similarly here I have a b and I have a b in the second term, which will canceled is send out, I have b d which will again cancel of the medium. So, similarly here c a plus d c; that is the same terms, they cancel each other of and here I have c d plus d square. So, the d square will cancel the d square and the a d cancels the a d. So, I just did this little quickly, you should check this out, somewhat more recently.

So, this finally, leads to the following answer and only left with two surviving terms there is a b c and a b c with 0 sums 1. So, which means that, well our initial guess not quite correct, because if I apply this polynomial f of x to the matrix A, I do not get the matrix 0, because I assumed of course, the d and c may be non 0. So, I would get the 0 matrix, if in at least 1 of b or c is 0, but here unfortunately it is not.

So, what we do well this is still not too bad, because observe that the answer that we obtain friendly is actually just b c times the identity matrix. So, the polynomial f applied to a, just produces the b c times the identity. So, it is easy to see, how to modify this, we just subtract the b c term as well.

(Refer Slide Time: 21:01)

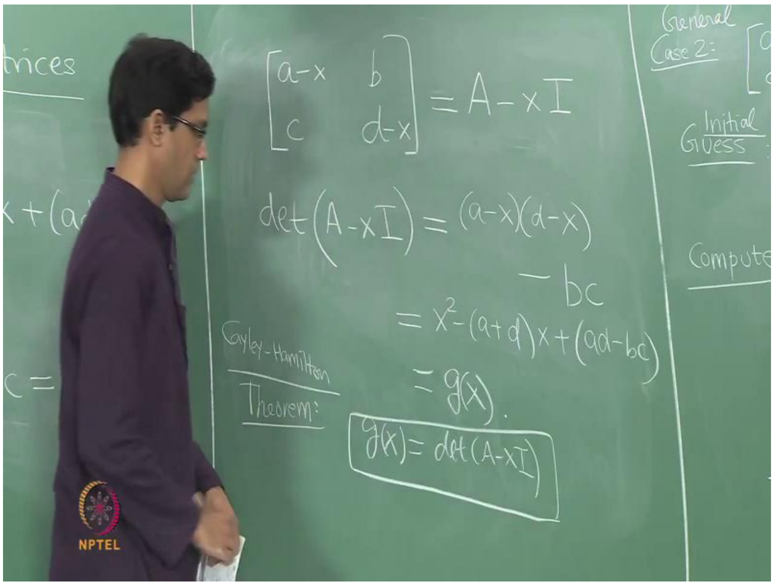


So, now, here is the final answer that suggested by what you just it, you take f of x, but from that you subtract of b c. So, subtracting a b c from f of x, means what, well this is x square minus a plus d x plus a d minus b c. And now, observe that our calculation said that, if you apply A, where if you apply g to a, then well the first the application of f to a produces b c times identity and that is going to be canceled of by the b c term that we know including. So, this is in fact, going to be 0's as required.

So, what we manage to here is to fine for the most general scenario, a polynomial which will make the given matrix 2 cross 2 matrix A into a 0. So, now, here are a couple of points birth noting. So, note that, if you take in a take this polynomial g of x that we just wrote out and look at it is constant term. So, meaning just look at what happens when you put at x equal to 0.

So, observe that if you take the polynomial g and you put x equal to 0, this just gives you a d minus b c. So, the constant term is a d minus b c that is of course, familiar to us that is, if you seen it to the, if you the contracts already it is what we called the determinate of the matrix. So, for the 2 cross 2 matrix determinant a d minus b c is exactly what seems to appear as a constant term of this matrix of this polynomial and in fact, the enter polynomial itself can be obtain in the following fashion.

(Refer Slide Time: 22:56)



So, observe in fact, here is the. So, remember here see as a b c d that was the 2 cross 2 matrix A. Now, here is the procedure which will allow us to get the entire polynomial g of x, you subtract x from both diagonal entries. So, I will take them the matrix A and done something to it. So, this is the matrix A from which I have done the following, I have subtracted x times the identity matrix, the effected has is of making both diagonal entries A minus x and d minus x.

So, I have taken this guy and now if you compute the determinant of this new matrix. So, already observe a b c d are real numbers and x is of course are not numbers any more that the symbols. But, recall we set before that the nice thing with having matrices or vectors and so on is that you can put any entries, you want as long as there is a sensible notion of addition and multiplication of those entries.

You do not always need those entries to be numbers, it is anything it could be other kinds of things, such as in variables like x and so on, which we know how to add a multiply to gather. So, polynomials are perfectly fine as entries of matrices and so now, if you compute the determinant of this resulting matrix here, matrix A minus x I, what we get is well the diagonal entries now are A minus x and d minus x.

So, A minus x times d minus x minus b c. So, that is the definition of the determine and observe this is exactly well I can switch this, I can call it x minus A, x minus d. So, this is exactly x square minus a plus d x plus a d minus b c, in other words is exactly the polynomial g that we constructed. So, this polynomial g is in fact, that determinant of A minus x I. So, that is the key fact that we just realized and this is one of the somewhat remarkable theorems concerning matrices are on polynomials applied to matrices. So, this is goes the name of the Cayley Hamilton theorem.

So, let me not write this out in full, this is the thing that we just set g of x equals determinant of, so I just say to in this context. So, g of x is just determinant of a minus x I. So, in general, what the Cayley Hamilton theorem has is no matter, what matrix A you take, it could be of any size n cross n for a instance, from that matrix you subtract of x times the corresponding identity matrix.

So, which means all diagonal entries must now be subtracted and x must be subtracted from the all the diagonal entries. And then you take the determinant of the in the resulting matrix, the polynomial that you show obtain the polynomial in x, will have the property that, when you apply the polynomial to the matrix A, it will kill the matrix, it will make the matrix 0. And we just have seen the 2 cross 2 proof of the Cayley Hamilton theorem. So, this is pretty much all we have to say about matrices.