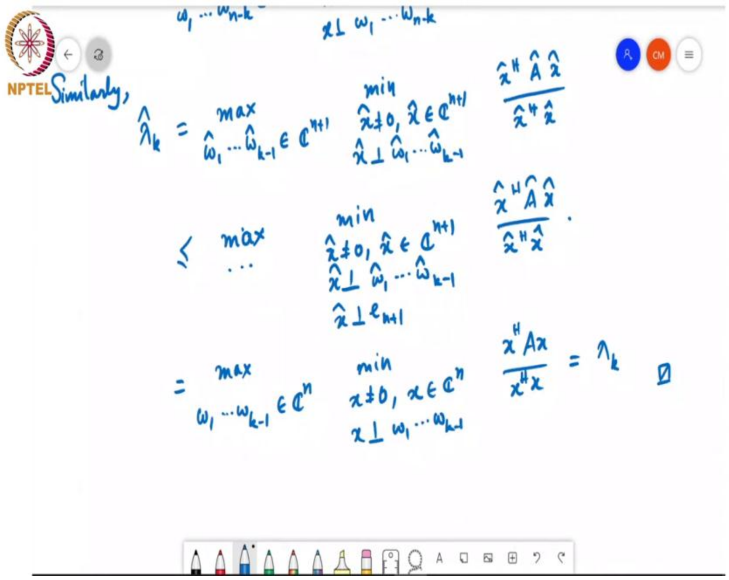
Matrix Theory
Professor Chandra R Murthy
Department of Electrical Communication Engineering
Indian Institute of Science Bangalore
Interlacing theorem II (Converse)

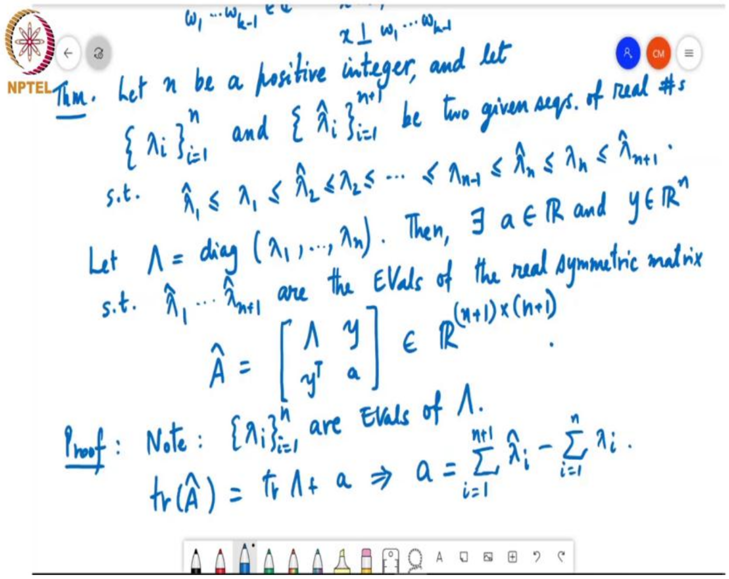
(Refer Slide Time: 0:14)



The previous two theorems showed that if you add a rank-one matrix the previous two theorems meaning the theorem we showed just now and the last theorem we proved in the previous class, they showed that if you add a rank-one matrix or if you border a Hermitian matrix then the eigenvalues of the matrix interlace.

So, now, the question is if you take two interlacing sets of real numbers, then can you realize these interlacing set of real numbers by a Hermitian matrix with a suitable modification. So, can I find matrices such that these set of interlacing numbers are such that one subset of numbers are the eigenvalues of one matrix, and then if you, for example, border that matrix with Y and an A you can get another matrix for which the other set of numbers are the eigenvalues. Then the answer is yes. And that is what is the essence of the following theorem.

(Refer Slide Time: 01:28)



So, it is kind of a Converse result to the theorem which has proved. So, let n be a positive integer and let lambda i i equal to 1 to n and lambda hat i, i equal to 1 to n plus 1 be two given sequences of real numbers such that they have this interlacing property. So, lambda hat 1 is less than or equal to lambda 1 less than or equal to lambda hat 2, lambda 2 lambda n minus 1 lambda hat n less than or equal to lambda n less than or equal to lambda hat n plus 1.

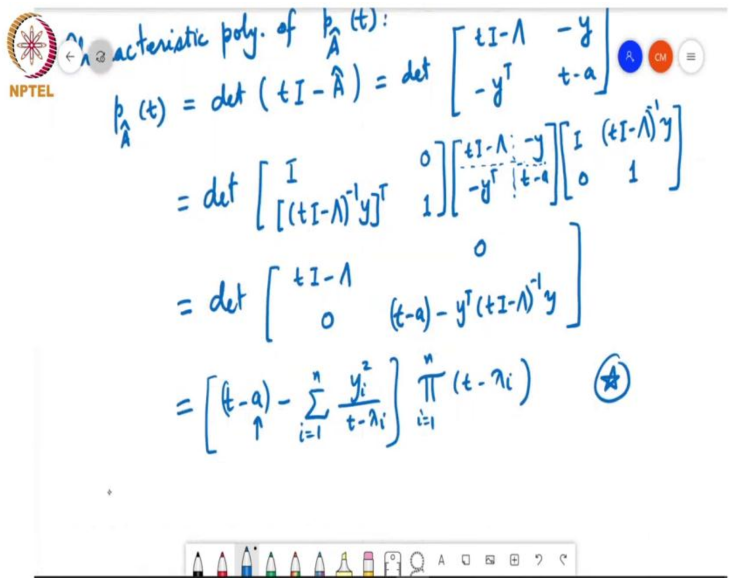
Now, let lambda equals a diagonal matrix with diagonal entries equal to lambda 1 up to lambda n then there exists a real number a and y in R to the n. Such in fact, it is enough to choose a real vector y such that lambda hat 1 up to lambda hat n plus 1 are the eigenvalues of the real symmetric matrix.

A hat which is equal to lambda y, y transpose and a, which is an R to the n plus 1 cross n plus 1. So, that is the statement of the theorem. So, we will show this. So, first of all, note that lambda i i equal to 1 to n are eigenvalues of lambda. So, we already have that first property that this matrix here has lambda i as its eigenvalues and what we are saying now is that the other set of eigenvalues lambda hat 1 lambda hat 2 up to lambda hat n plus 1 will be the eigenvalues of this matrix A hat if you choose y and an appropriately.

So the proof is essentially by to be constructive. So, we will show how to choose y and a such that lambda hat 1, lambda hat 2 etcetera, lambda hat n plus 1, are in fact, the eigenvalues of A hat. So, straight away, what you can do is you can look at trace of A hat and that is equal to trace of lambda plus a, and trace of lambda is just the summation of lambda i, i going from 1 to n and trace of A hat is a summation of the eigenvalues of A hat.

And what we want them to be is lambda hat 1, lambda hat 2 up to lambda hat n plus 1. And so, this just implies that a is equal to sigma i equal to 1 to n plus 1, lambda hat i minus sigma i equal to 1 to n lambda i. And so, we already figured out what a should be.

(Refer Slide Time: 06:23)



Now, let us look at the characteristic polynomial of pA hat of t. So pA hat of t is equal to the determinant of tI this is an n plus 1 cross n plus 1 identity matrix minus A hat, which is equal to just substituting for A hat, we can write it as the determinant of the matrix which has tI minus lambda, this is an n cross n identity matrix and then minus y minus y transpose and t minus a.

Now, what I can do is, I can do one small trick here, which is I can pre and post multiply by other matrices whose determinant is 1. And that will not change the value of this determinant. So, that is equal to the determinant of the identity matrix. And below that I will have tI minus lambda inverse times y transpose 0 and 1 here, this times this matrix tI minus lambda minus y minus y transpose t minus a, I will just draw some partitions here, so that the quantities do not get mixed up.

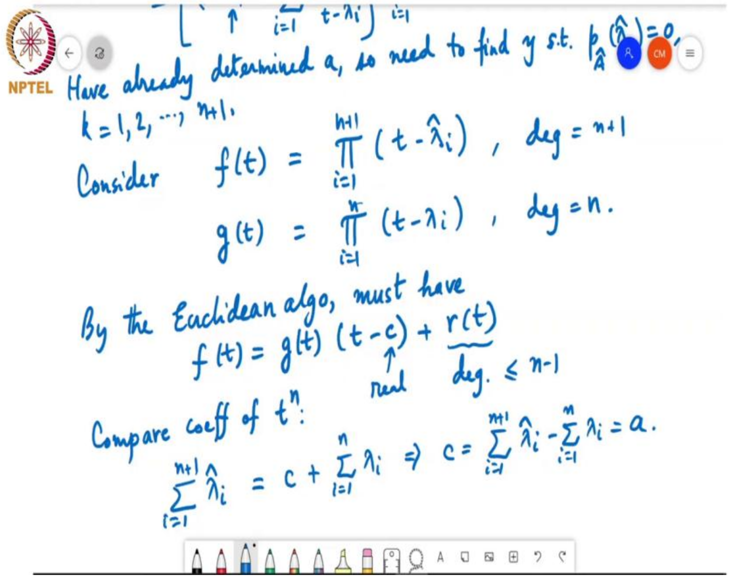
And the transpose of this matrix over here I, tI minus lambda inverse y, 0 and 1 here. So, these matrices have, these are lower triangular and an upper triangular matrix, and their determinant is equal to 1. So that does not change the value of this determinant. Now, if you carefully multiply these out, what you will find is that this reduces to the following form, it becomes determinant of tI minus lambda.

And over here, I will get t minus a minus y transpose tI minus lambda inverse y, and 0 here and 0 here. So, that is the reason I did all this so that I get a block diagonal matrix. In fact, this is a completely diagonal matrix. And since it is a diagonal matrix, I can now readily compute its determinant. And so that is just equal to the product of the diagonal entries, which is this term, t minus a minus sigma.

So, I will just expand this product here. i equal to 1 to n, y i squared divided by t minus lambda i. This is a diagonal matrix with lambda i's, t minus lambda i's on the diagonal, this inverse of this matrix because after all tI minus lambda is diagonal. So, tI minus lambda is also a diagonal matrix.

So, computing this inverse is super easy, you just invert all the diagonal entries. And so that is this value. So, this is the same as this times the product i equal to 1 to n, t minus lambda i. So, we will call this equation star, we will come back to it in a bit. So, now, we have already determined what a is. And so, what we need to do is to find y such that this pA hat of lambda k hat is equal to 0 for k equal to 1 to up to n plus 1.

(Refer Slide Time: 10:52)



So, we have already determined a, so need to find y such that p A hat of lambda k hat equals 0, k equal to 1, 2, up to n plus 1. So, this is a little bit challenging, so, let us see how to do that. So, consider two functions f of t, which is the product i equal to 1 to n plus 1, t minus lambda hat i.

And this is a degree n plus 1 and another polynomial g of t, which is equal to product i equal to 1 to n, t minus lambda i, this is a degree. So, what we really want is that, this characteristic polynomial should end up becoming and coming out in this form, so that then we know that lambda hat i, i equal to 1 to n plus 1 are the eigenvalues of this that the 0s of this polynomial.

So, want to and this thing here is your g of t. So, we have pA hat of t to be this quantity times g of t, and we want that to somehow end up in this form where f of t is equal to product i equal to 1 to n plus 1, t minus lambda hat i. Now, this is degree n plus1 and so, this is degree n. So, what you can do is you can write.

You can divide f of t by g of t and then you will get a quotient and the remainder, and the quotient will be of degree 1, and the remainder will be of degree at most n minus 1. So, by the Euclidean algorithm, we must have f of t equal to g of t times some t minus c plus r of t, where c is some real-valued quantity because all the coefficients here are real.

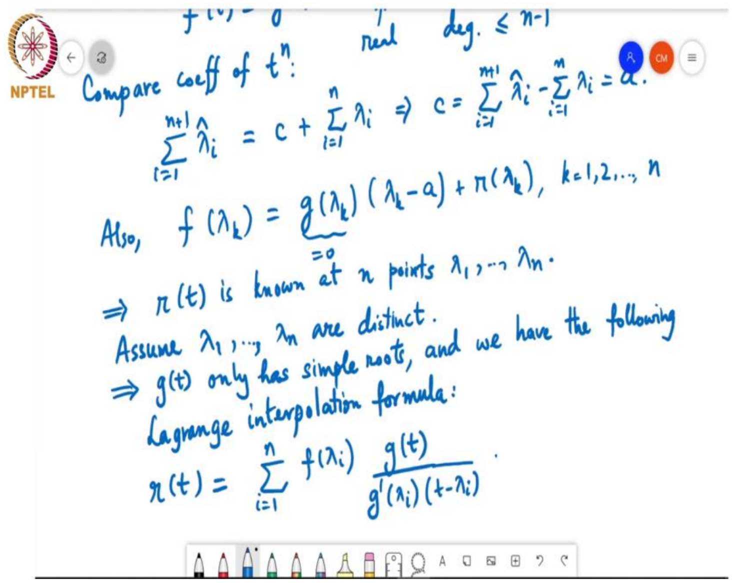
And r of t must be of degree at most n minus 1. This is the remainder polynomial and this is the quotient polynomial. So, what we can do is now, let us compare the coefficients of t power n on both sides. So, the coefficient of t power n plus 1 is just going to be 1 because the t power n plus

1 comes from this t power n here times this t here. But if you look at the coefficient of t power n, the coefficient of t power n here is just the summation of lambda hat i, because you will take n of these terms and 1 of these lambda hat i, and, or rather it is minus lambda hat i, but I will ignore the minus sign. I will consider a minus sign for the other thing also.

And so, the coefficient of t to the n is the summation of lambda hat i, i going from 1 to n plus 1. And if I look at this, the coefficient of t to the n is going to be either I can take all the t terms here, then it will multiply with minus c or I can take n minus 1 terms here and multiply with this t and then I will get a minus lambda i.

So, that what that means is that sigma i equal to 1 to n plus 1 lambda hat i is equal to c plus sigma i equal to 1 to n lambda i. And so, then this means that c is equal to summation i equal to 1 to n plus 1, lambda hat i minus sigma i equal to 1 to n lambda i, which is nothing but a. And so, this t minus c that we are looking at here, that is nothing but t minus a.

(Refer Slide Time: 16:00)



And further, if I compute f of lambda k, that is going to be equal to g of lambda k times lambda k minus a plus r of lambda k. And this is equal to 0 because there is a t minus lambda i term here product of t minus lambda i terms here.

So, g of lambda k for any k is equal to 0. And so this is equal to r of k. And this is true for k equal to 1, 2 up to n. So now, what that means is that if I compute f of lambda k for k equal to 1 to n, I then know what r of lambda k is at n different points. So r of t is known at n points. lambda 1 through lambda n.

So, f of lambda k I can calculate because it is just this polynomial here. And so, I can just substitute lambda 1, lambda 2, etcetera, I know what f of lambda k is. And by this thing, I know then what are of lambda 1, lambda 2 up to lambda n. So, this has a degree at most r of t has a degree at most n minus 1. And I know this know its value at n different points. And what that really means is that we know what r of t is.

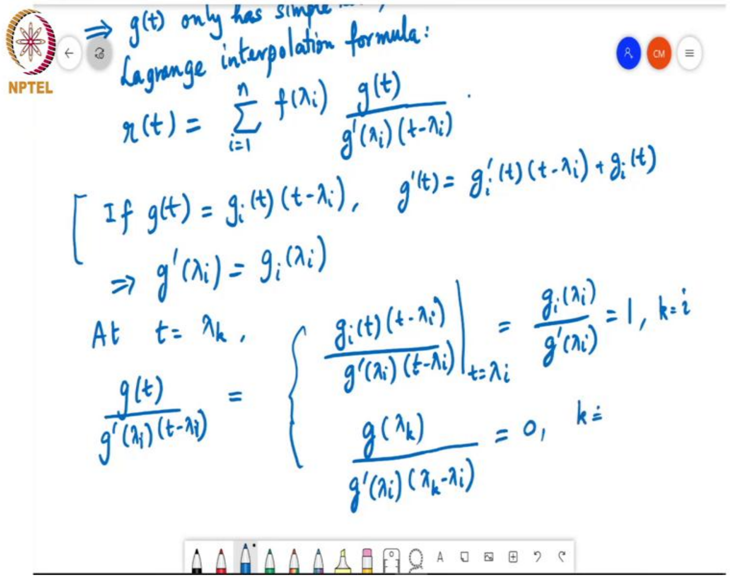
So, just for the moment, in order to proceed further, I will assume that this lambda 1 to lambda n are distinct. And then I will come back to the case where some of these eigenvalues are repeated, and I will deal with that case later. So, for the moment assume lambda 1, lambda n are distinct.

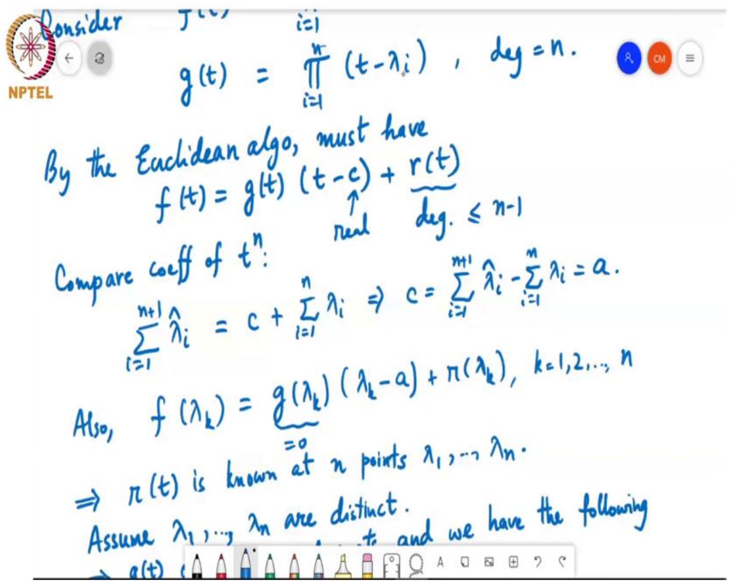
Then, what that means is that g of t is the product of all these t minus lambda i terms and each of these are going to be first-order terms, and so, g of t only has simple roots each lambda i occurs only once as a root. And so, g of t has only has simple roots and we have the following. See, for example, if you are given a first-degree polynomial with unknown coefficients, all you need is the value of the polynomial at two points and you can determine, what the polynomial is.

Similarly, if you are given a second-degree polynomial, all you need is the value of the polynomial at three points, and you can determine what the polynomials. And for the first-order polynomial case, I am sure you have seen this Lagrange interpolation formula, which tells you how to write out what the straight line that matches the two values that you have observed is, and we see that in linear regression in various problems many times, but this is a generalization of that.

So, we are looking for n minus 1-degree polynomial, such that its value matches with some given values r of lambda 1, r of lambda 2, up to r of lambda n, and at these n different points. And so that formula is directly in there is a direct formula to write out what r of t should be. And so this r of t is equal to the summation i equal to 1 to n, f of lambda i, times g of t divided by g dash of lambda i. That is the derivative of g of t evaluated at t equal to lambda i, times t minus lambda i. So, this is called the Lagrange interpolation formula.

(Refer Slide Time: 21:26)





So, we will take this on faith, but maybe I can indicate why this is correct. So, for example, if g of t. If I write g of t to be equal to gi of t times t minus lambda i, so all the other n minus 1 factors in g of t are in this gi of t, then g dash of t, I can write to be g i dash of t times t minus lambda i plus gi of t.

Which means that if I want to evaluate g dash of lambda i, that is equal to, now, if I substitute lambda i here, I get lambda i minus lambda i, so this term goes off to 0, and so I will be left with gi of lambda i. So, the derivatives are equal to gi of lambda i. It is a simple fact, but it is true.

And so, that means that at t equal to some lambda k, if I were to evaluate what happens to this part here g of t, divided by g dash of lambda i, times t minus lambda i, what I get is I will get either gi of t times. So, g of t is gi of t times, t minus lambda i, t minus lambda i over g dash of lambda i, times t minus lambda i, and I need to evaluate this at t equal to, so, I need to evaluate this as a t equal to lambda k.

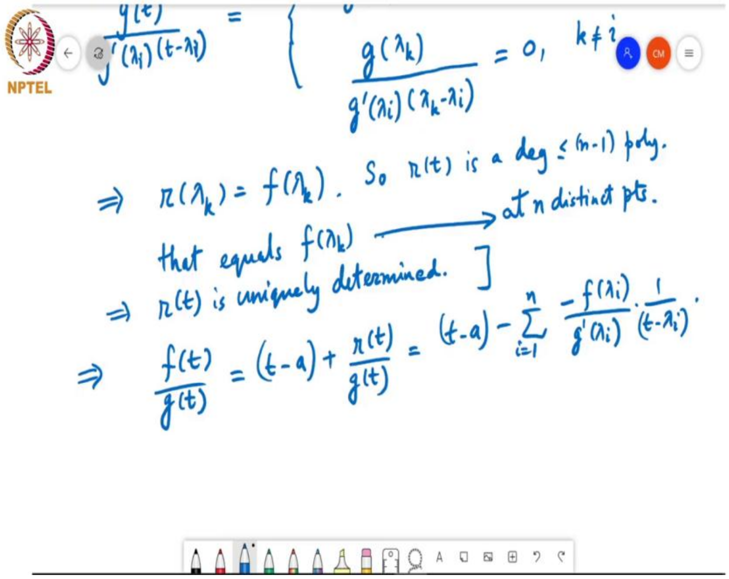
So, I will have to consider the k equal to i and k is not equal to i separately, so, I will consider k equal to i here. And so, then I will be substituting t equal to lambda i. And if I substitute t equals lambda i, these two terms cancel and so, I will have gi of lambda i over gi dash of lambda i, but gi dash of lambda i equals g of lambda i. So, this is equal to gi of lambda i over g dash of lambda i, which is equal to 1. So, this is for k equal to i. And the other cases when k is not equal to i, I do not need this factorization.

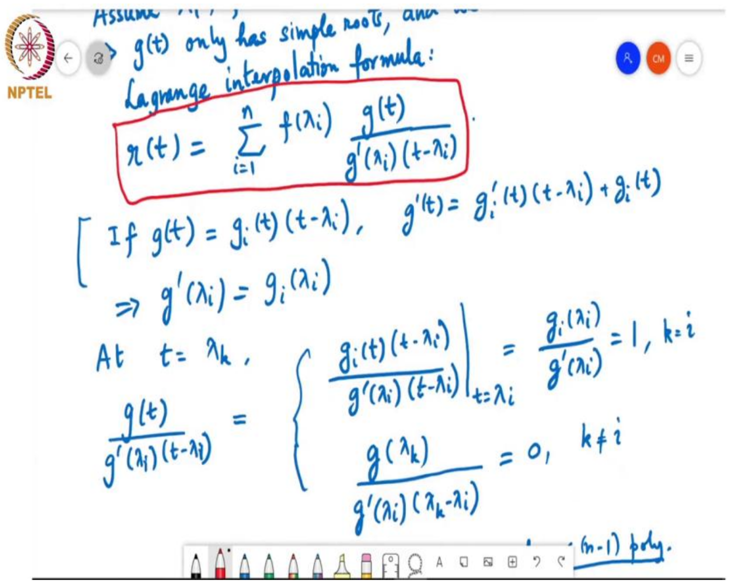
So, I can just write it as g of lambda k divided by g dash of lambda i, time t minus lambda i is lambda k minus lambda i but g of lambda k is equal to 0 for any lambda, right, because it has all these factors. g of t is the product of t minus lambda i. So, if I look at g of lambda k, that is always going to be equal to 0.

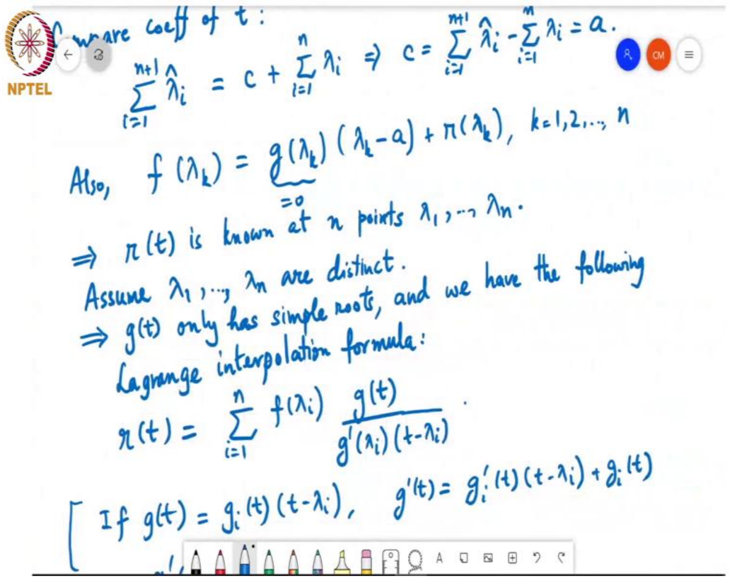
And g dash of lambda i is all the other factors, it is gi of lambda k. And if I have dropped the ith term from g, then this will be a nonzero quantity. And lambda k minus lambda i also nonzero, because I assumed the eigenvalues are distinct. And so, this is always going to be equal to 0 for k not equal to i. So, then, so now we know what happens to this.

So, if I look at what happens to r of lambda k, then I will have a summation i going from 1 to n, fi, f of lambda i times this thing evaluated t equal lambda k. But this is nonzero only for k equal to i. And so only the kth term in this summation will survive. And for k equals i, this quantity equals 1. And so, I am just left with f of lambda k.

(Refer Slide Time: 26:24)







r of lambda k is equal to f of lambda k. Which is what we wanted. We started out by saying that f of lambda k is something that is known, and we want r of lambda k to equal f of lambda k at this k at these endpoints, lambda 1, lambda 2 up to lambda n.

So, I will just write that here. So r of t is a degree less than or equal to n minus 1 polynomial that satisfies or I will put it this way that equals f of lambda k at lambda i, i equal to 1 to n by n distinct points. So, that means that r of t is unique and it is given by the formula we determined this is a consequence of the Lagrange interpolation, it is a consequence of polynomials.

So, then, what that means is that we now know what the form of r of t is. So, if we now consider what f of t divided by g of t is, this is equal to t minus a. So, f of t was t minus a, g of t plus r of t. So, plus r of t divided by g of t is equal to. Now, I will use this formula for r of t, which is summation f of lambda i, g of t divided by g dash of lambda i times t minus lambda i. So, I will write this as t minus a minus summation i equal to 1 to n. I wrote it with a minus. So, I have to write a minus f of lambda i over, g dash of lambda i, times one over t minus lambda i.

Student: Sir?

Professor: Yes.

Student: Sir, in the above point the one above rt is uniquely determined, should it be f of lambda k or f of lambda i?

Professor: It is at, okay. So, let me to avoid confusion and just remove this. So, at n distinct points. All I am trying to say is that r of t matches, certain given values at n distinct points. Happy? Is it clear?

Student: No sir, I will think about it. Sir, I will ask on teams.

Professor: Yeah, so r of t is some polynomial, we do not know what it is. But we know that r of lambda k equals f of lambda k at k lambda 1, for k equal to 1 to n. That is all that I am saying, it is a degree n minus 1 polynomial, up to at most n minus 1 polynomial, where its value at n distinct points is completely determined.

So, for example, if I have the real line, and I gave you 1, 2, 3 points, and I say, here is one value, here is another value. And here is another value. Now, if I asked you to fit, degree two polynomials, which is a quadratic, which matches with these, then it turns out that there is only one way you can do that, which is a quadratic, that somehow looks a bit like this.

I am not good at drawing these things. But it is a quadratic that looks like this, there is no other way you can fit a quadratic that matches these three points. If you had given me, if you allowed me to choose a third-order polynomial, then it you can choose many different third-order polynomials where it matches with these three values.

But if I have to choose a quadratic, this is the only way to do it, it is easier to think of it if you go to an even more trivial case, which is a straight line. So, if you give me two points, and then you say the value here is this, and the value here is this, then there is only one way I can fit a straight line through both these points, and this is the first-order polynomial.

And if you allow me to fit a quadratic through these, I can fit many different quadratics. This is in fact, a quadratic where the t squared coefficient equals 0, but of course, I can fit a quadratic, maybe like this, maybe like this, etcetera, etcetera. So, there are many ways to fit a quadratic through this, but there is only one way to fit a straight line where it meets these two points.

So, if you take if you are given that polynomial of degree at most n minus 1, and if you specify its value at n distinct points, then the polynomial is completely determined. I am just illustrating that the way we have chosen r of t by using this Lagrange interpolation formula is such that it is r of lambda k equals f of lambda k for lambda k, lambda 1 lambda are for k equal to 1, 2 up to n. So, I am not shown the uniqueness here.

That is a property of polynomials. But I am just saying that the way we have chosen is something that works in the sense that it matches f of lambda k, r of lambda k matches f of lambda k for k going from 1 to n. I hope that is a bit clearer.

Student: Yes, sir, yes sir.

Professor: Okay. So, now…

Student: Sir.

Professor: Yeah.

Student: I have a question. So, we need to determine a polynomial of degree at most n minus 1 that was r t, and we know the value at n minus 1 points…

Professor: n points.

Student: n points, yeah. So, we could write it in terms of general polynomial expression and we know the different values. So, the coefficients are unknown. So, we could express it in ax equal to be linear set of equations, and solve.

Professor: Absolutely, yes, that also works. This is a direct formula for the answer, read that this is the answer you would get.

Student: Okay, okay.

Professor: This Lagrange interpolation formula is the answer you would get if you did what you just suggested.

Student: Okay, sir. Sir, another side question that you told that if I am given two points, and then I can fit infinite number of quadratics through them. So, if I, in this way from the same linear set of linear equations, then I will find some vector in the null space. That is interpretation in this.

Professor: Absolutely. So, there is a very nice connection between polynomials and linear systems of equations. And, so what you said is correct.

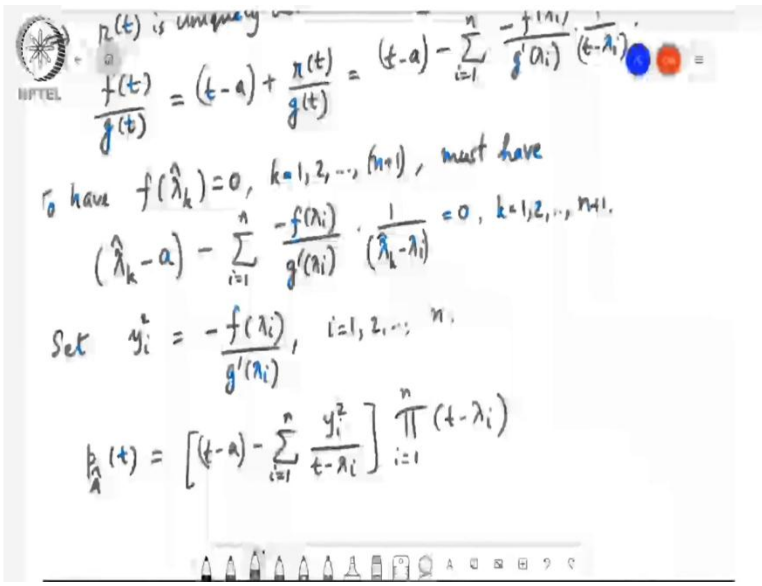
Student: n number of solutions if there is something in null space.

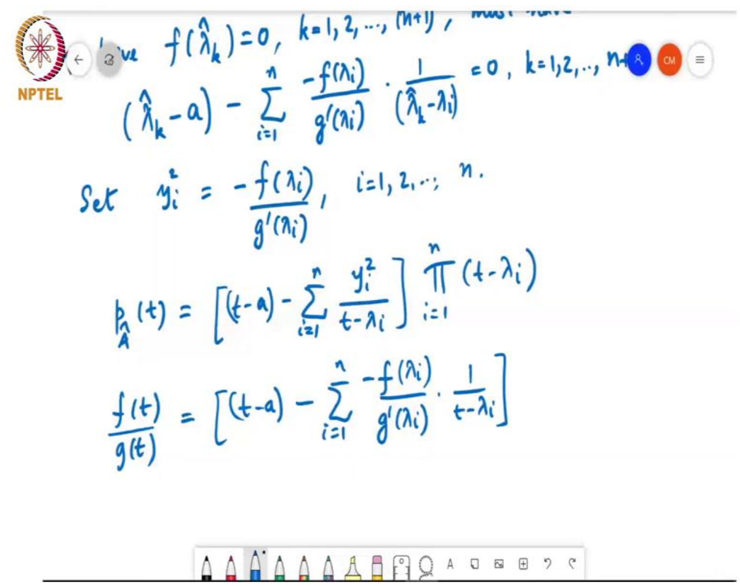
Professor: Exactly.

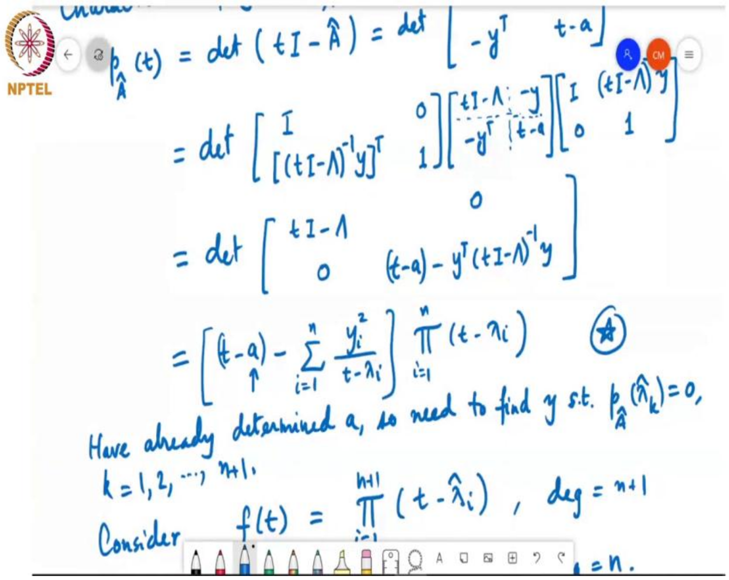
Student: Okay.

Professor: So maybe, time permitting, I will take a digression and discuss this kind of connections also. But we are closing in on the end of this course. And so, there are some more material that I need to cover. Let us see how it goes. But your intuition is correct.

(Refer Slide Time: 34:45)







Now, coming back to this, so what do we want, we want that f of lambda k hat equals 0 for k equal to 1, 2 up to n plus 1. So, to have f of lambda hat k equal to 0, k equal to 1, 2 up to n plus 1, we must have now substituting directly into this formula here, lambda hat k minus a minus sigma i equal to 1 to n minus f of lambda i divided by g dash of lambda i times 1 over lambda hat k minus lambda i equal to 0 for k equal to 1, 2 up to n plus 1.

Now, one small point is that what if I am dividing by lambda hat k minus lambda i. And so if lambda hat k equals one of these lambda i, then, of course, I am dividing by 0, that does not make sense. So, this is not a problem, because if lambda hat k equals lambda i, then I have f of lambda i sitting here and so this coefficient will also be equal to 0 because f of lambda hat K, and lambda hat k equals lambda i, so f of lambda hat k equals 0.

So, these two first cancel, and then you have 1 over g dash of lambda i. So, it there is no singularity at t equal to lambda hat k. So, now what we will do is we will set y i squared to be equal to minus f of lambda i over g dash of lambda i. And this is for i equal to 1, 2 up to n. Now, if we can show that this y, this quantity is greater than or equal to 0, then now I am telling you how to choose yi, then so then let us see.

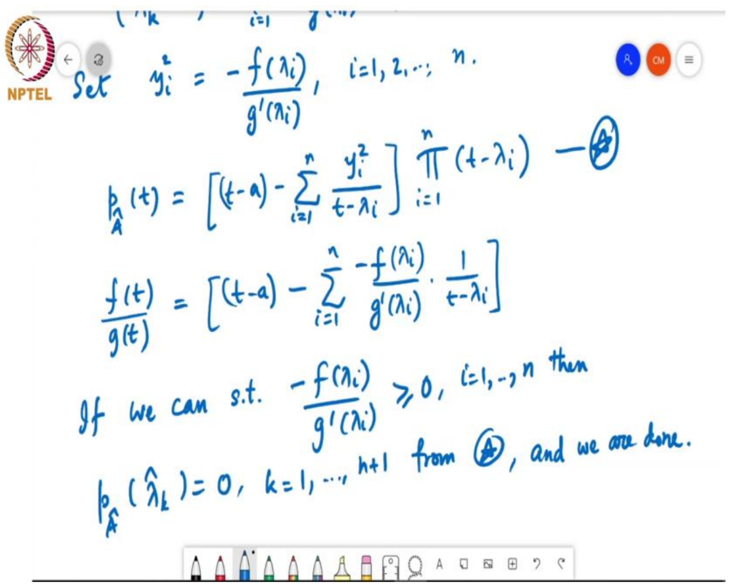
If I go up here. So, pA hat of t is equal to t minus a divided by minus yi squared over t minus lambda i times this product. So, now let us compare this against, so you need to keep this in mind or maybe what I will do is I will write that here, so that it will be clear. pA hat of t is equal to t minus a minus sigma i equal to 1 to n, y i squared over t minus lambda i times the product i equal to 1 to n,t minus lambda i, this is what we had earlier.

And what we have here is that f of t, over g of t is equal to t minus a, minus sigma i equal to 1 to n f of minus f of lambda i, over g dash of lambda i, times 1 over t minus lambda i, and g of t is exactly this polynomial here. Now, f of t is something that we are choosing such that it has lambda hat 1 through lambda hat n plus 1 as it 0.

So, if this polynomial exactly matches with this polynomial, we know that A hat, the characteristic polynomial of A hat has lambda hat 1 through lambda hat n as it 0s. So, if you look at this g of t is exactly this thing here. And so, all you need is to make this equal to this, and then your home right in the sense that these are exactly the same polynomials.

So, all you need now is to show that y i squared for me to be able to define this to be y i squared, this quantity should be non-negative, then I can define this to be y i squared for some real-valued quantity y i. And so, if I can show that.

(Refer Slide Time: 41:38)



So, if we can show that y i squared or I will put it this way, minus f of lambda i over g dash of lambda i is greater than or equal to 0 for i equal to 1 through n, then pA hat of lambda hat k equals 0 for k equal to 1 through n plus 1 from this equation star, which is the same as this.

So, that means that we have found the polynomial that we want, I mean, we found the y such that the characteristic polynomial of A hat is exactly the one that has lambda 1 one up to lambda hat n as its roots. So, now I still need to show that these y i squares are greater than or equal to 0, or that f of lambda high over g dash of lambda i is less than or equal to 0. So, that is just 1 or 2 more steps. But will do that in the next class because I have run out of time.

And then I need to extend this to the case where the eigenvalues could be repeated because I assumed that lambda 1, lambda 2, up to lambda n are distinct, the extension is going to turn out to be almost trivial because if for example, lambda 1 equals lambda 2, then lambda 2 hat equals lambda 1 because lambda 2 hat is supposed to interlace between lambda 1 and lambda 2.

And so that means that I can pull out some factors and deal with them separately, and then consider only the distinct eigenvalues that remain. So, will complete this proof in the next class. We will stop here for today.