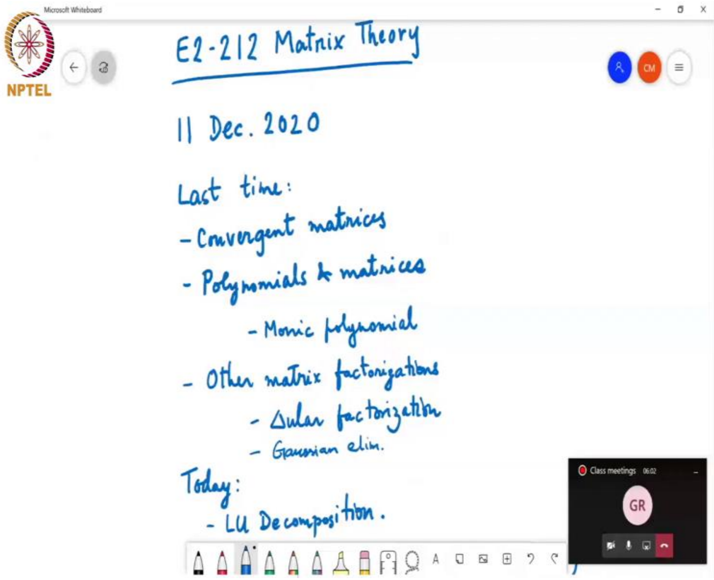
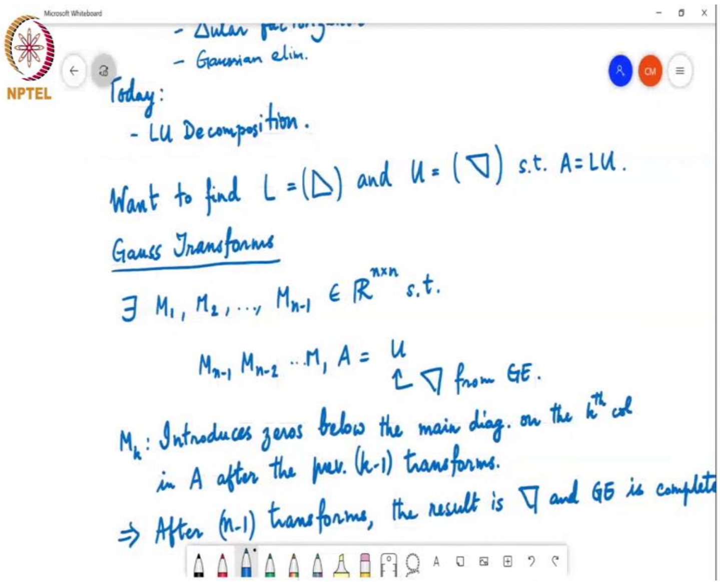
Matrix Theory
Professor Chandra R. Murthy
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore
LU Decomposition

(Refer Slide Time: 00:14)



Last time we looked at convergent matrices and also discussed about polynomials and matrices and in particular we defined the notion of harmonic polynomial and then we started looking at other matrix factorizations. In particular gaussian elimination and then at the end of the class we started discussing triangular factorization in particular the LU decomposition. So, today I will talk about this LU decomposition.

(Refer Slide Time: 00:44)



Our goal is to find a matrix L which is lower triangular and U which is upper triangular such that an n cross n matrix A can be written as the product of these two matrices L times U. And here A is an n cross n matrix and so are L and U both are n cross n matrices.

So, the basic unit of such a decomposition is these gauss transforms and essentially what this is going to do is to perform gaussian elimination. And this gaussian elimination is equivalent to a sequence of gauss transforms. That is to say that there exist matrices M1, M2 up to Mn minus 1 such that if you take the product of all these matrices times A you will get an upper triangular matrix and which is the same as the upper triangular matrix you would get if you had performed gaussian elimination on that matrix.

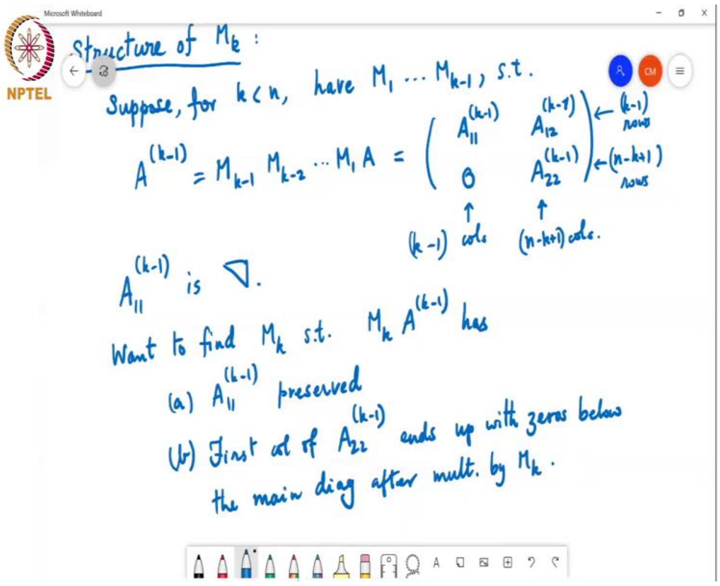
So, Mk here the kth matrix in this series of matrices is the one that introduces zeros below the main diagonal on the kth column of A after the previous k minus 1 transforms. And because of this when you do the first transform you are introducing zeros below the diagonal element of the first column.

When you do M2 when you multiply that with M2 you will enter, you will keep the first column intact but you will introduce zeros below the main diagonal of the second column and then the third column and so on. So, after n minus 1 transforms you have introduced zeros below all the first n minus 1 transforms and so the result is upper triangular and this gaussian elimination process is complete.

So, what I need to tell you now is how to determine what this M1, M2 etc is. And then if I want to write A as LU I need to do two more things one is I will have to take all these matrices M1 M2 up to M1 to the right hand side then I will get the inverse of the product of all those matrices times U and I need to show you that the inverse of the product of those matrices is lower triangular.

So, that you are effectively been able to write A as L times U. So, those are the steps that remain. So, we will start by first understanding what is the structure of this Mk matrix.

(Refer Slide Time: 03:18)



So, the in order to answer this suppose after so suppose you have already found M1 through Mk minus 1 and will discuss how to find Mk and if we know how to do that then we can start with M1 and then find M2 and 3 up to Mn minus 1. So, suppose for some k less than n we already have M1 through Mk minus 1.

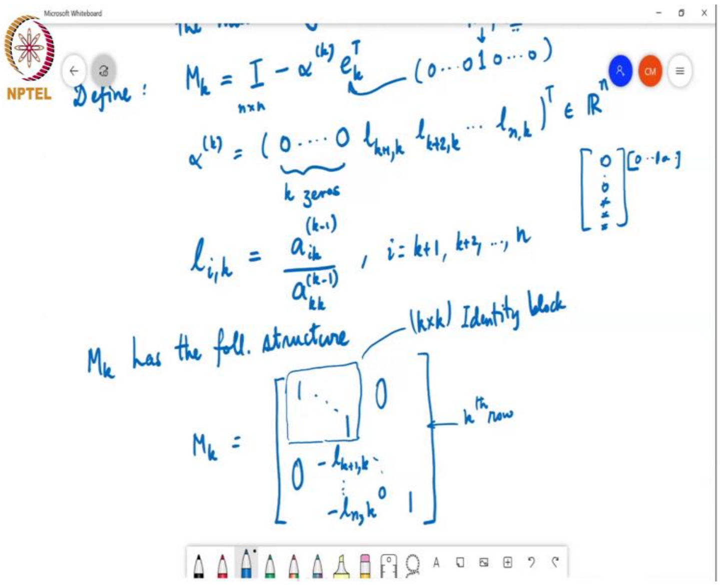
And these are such that if I take A, I will define Ak minus 1 to be Mk minus 1, Mk minus 2 all the way up to M1 times A has this structure because as I said in each step you are introducing zeros below the main diagonal of the successive column. So, this is A11 k minus 1 A12 k minus 1 this is 0 and this is A22 k minus 1.

So, this is the, this corresponds to the first k minus 1 rows and so this will have n minus k plus 1 rows and this is k columns and this will be n minus k plus 1 columns. This is the structure you have arrived at so assuming that you know how to find M1 through Mk minus 1 this the product of these matrices times A will have this structure.

Now, we want to and here in particular because we are assuming we have figured out the first k minus 1 matrices A11 k minus 1 is upper triangular. So, for the next stage our goal is to find Mk so we want to find Mk such that Mk times Ak minus 1 has two properties A11 k minus 1.

I do not want to disturb that I have already placed zeros where I want them so this is preserved. And b the first column of A22 k minus 1 as or ends up with zeros below the main diagonal. That is to say zeros below the first element of this matrix after multiplication by Mk. So, what is an Mk that will do this for me right.

(Refer Slide Time: 07:48)



So, this is going to be the Mk that will have these two properties. So, Mk is equal to the identity matrix minus this is of size n cross n alpha k e k transpose. So, here e k alpha k is an n by 1 vector so I will write what that is alpha k so it has k zeros followed by l k plus 1 comma k l k plus 2 comma k l n k transpose.

So, this is going to be a vector in R to the n and there are k zeros here. So, this is an n cross 1 vector, this is the transpose of an n cross 1 vector and e k. So, this vector here has zeros everywhere except in the kth position it has a one. So, what is the rank of alpha k e k transpose?

Student: 1

Professor: 1, correct. So, any matrix say B of the form u v transpose where u and v are n cross 1 vectors is always of rank 1. So, this kind of a matrix Mk is actually called a gauss transform. And so I also need to tell you how to choose these values so li k we will choose this to be ai k of k minus 1 divided by akk of k minus 1 for i equal to k plus 1 so all these indices.

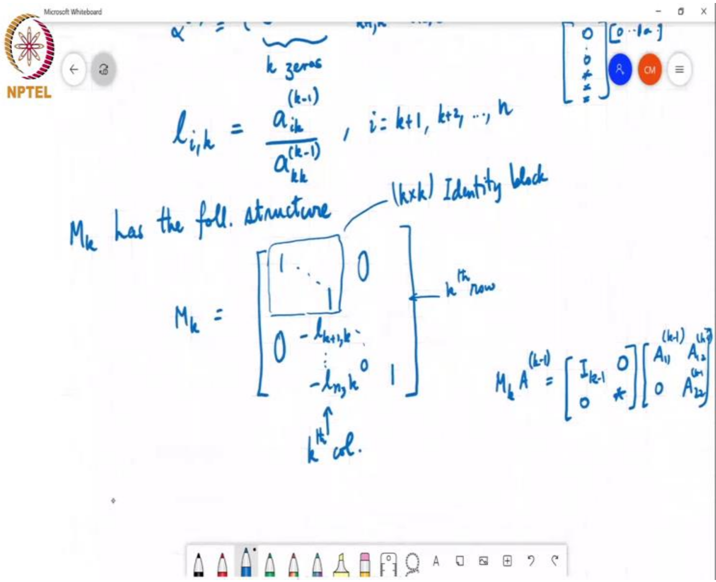
So, here I am assuming that this akk of k minus 1 is non-zero and this is, this plays a very significant role in gaussian elimination and this is called a pivot element. Now, if you go back and look at our gaussian elimination we defined the last time you will find that in gaussian elimination we are exactly doing using a quantity like this to multiply the rows of a matrix and add them to other rows.

So, this is actually doing exactly the same process as what we were doing in the gaussian elimination except it is putting it in a different way. So, with this so notice that this has non-zero entries only in the k L, k plus 1 to nth position and this is a vector like this so it looks like this with k zeros followed by something non zero over here and this is a vector e k has zeros in the first k positions for k minus 1 positions and a 1 there and then zeros everywhere else.

So, if I multiply these two together I will get a, I will get an n cross n matrix which has this whatever is in here is repeated in the kth column here in the product and everything else will be 0 because everything else is multiplying as 0. So, Mk has the following structure, it has ones along the diagonal and this is the kth column and here it has minus l k plus 1, not enough space.

Let me do this a little bigger and here it is minus l k plus 1 comma k and it is minus l k plus 2 comma k all the way up to minus l n comma k. And then 0 is everywhere else 0 here and 0 is here so this is the structure of Mk. And notice that it is this is a k cross k block this is the kth row k cross k, k cross k identity block and this is the kth column.

(Refer Slide Time: 14: 21)



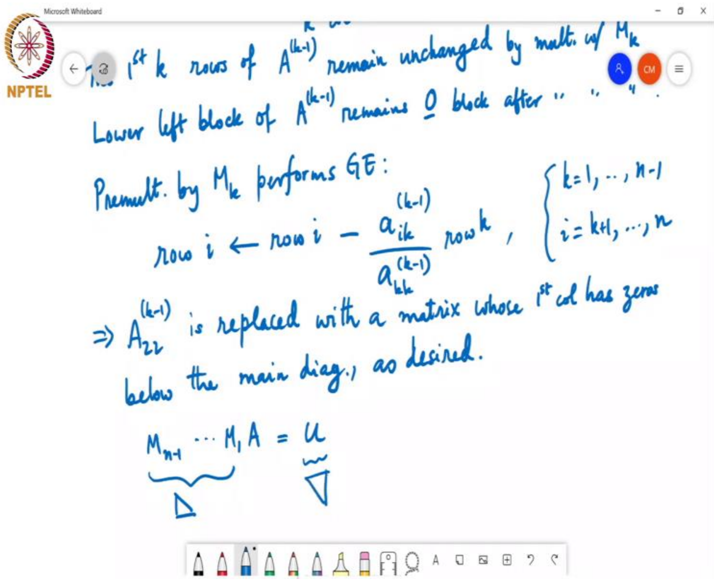
So, what this what this means is that if I do Mk times Ak minus 1 so this is like rough notes and write it over here. Mk Ak minus 1 it will be like multiplication of a k cross k identity 0 0 and then this has things over here and it has this form the diagonal entries are 1 and it has nonzero entries.

Actually, all of these are non-zero this entry let us fix this it is something here it is not, so it is this block here it is minus lk plus 1 comma k and then there is a 1 here and then ones along the diagonal the rest of the way and then all these entries over here. But, it has some structure like this and Ak minus 1 has the structure k11 k minus 1.

There is, these are not size matched so I will consider this to be the ik minus 1 matrix then it is easier to explain this is A12 k minus 1 and then this is 0 this is A22 k minus 1. So, if you do this then you see that this identity multiplies this A11 k minus 1. So, that it so this A11 k minus 1 will get preserved and this part will have some something the top right will have something because of this A22 k minus 1 and this bottom part will get multiplied by zero.

So, these zeros remain unchanged and this bottom part will get multiplied with this and I will get something and this is the part where I will end up introducing zeros below the main diagonal.

(Refer Slide Time: 16:22)



So, the first k rows of A k minus 1 remain unchanged by multiplication with Mk lower left block of A k minus 1 remains 0 after, remains a zero block after multiplication with entry. So, it has these two properties and m so basically Mk only affects this A22 k minus 1 block of this A k minus 1 matrix.

Now, these li k is are exactly the same as required by gaussian elimination to place zeros in these positions of the matrix and so pre-multiplication by Mk performs exactly the same row operations as gaussian elimination.

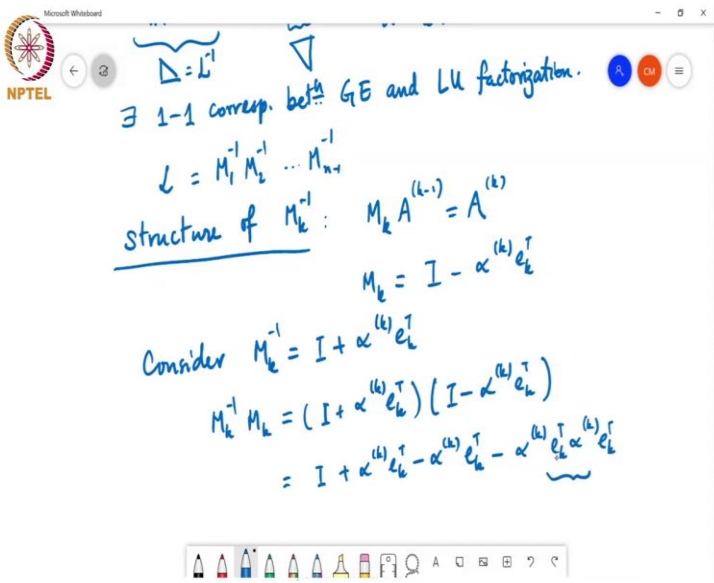
That is it will replace row i by row i minus ai k of k minus 1 divided by akk of k minus 1 times row k and this is for k equal to 1 up to n minus 1 and i equal to k plus 1 all the way up to n. And so specifically the if you look at the kth column what is happening is that the this operation is exactly cancelling off in the row kth, row case kth entry when you divide by this it gets (mi) the akk of k minus 1 cancels and you have a minus ai k 1 which cancels with the ai k of k minus 1 in the ith and you get 0 at that position.

And all other entries could potentially change. So, basically what this is, what this means is that a22 of k minus 1 is replaced with a matrix whose first column has zeros below the main diagonal. So, now we know how to construct these matrices M1 M2 up to Mn minus 1 and each of these matrices are upper triangular by construction.

And the product of a series of upper triangular matrices is upper triangular and so if we, so basically what we have is Mn minus 1 all the way up to M1 times a is equal to U and this product of all these matrices is lower triangular.

Each of these matrices M1 M2 up to Mn minus 1 they are all lower triangular by construction and this u ultimately after all these operations it will reduce this matrix A to an upper triangular form.

(Refer Slide Time: 21:04)



Now, if I define this matrix, this product to be L inverse then what I have is or L inverse A equal to u which implies this is lower triangular and the inverse of a lower triangular matrix is also lower triangular. And so this implies A is equal to LU and also by construction notice that each of these matrices is lower triangular with ones along the diagonal.

And so the for a lower triangular matrix the eigenvalues are the diagonal entries and so all its eigen values are equal to 1 and so it is non-singular can be inverted and so you have A is equal to LU.

So, basically there is a one-to-one correspondence between gaussian elimination and LU factorization. They are equivalent operations. Now, so at first glance it appears that in order to find the LU decomposition I need to take these n minus 1 matrices M1 to Mn minus 1 and I need to multiply them and then I need to invert that matrix to obtain this L.

But, it turns out that it is actually very easy to recover L because of the structure in these matrices. So, essentially first let us note that L is actually equal to the inverse of the product of these matrices and when you invert the multiplication order gets reversed so it is M1 inverse M2 inverse up to Mn minus 1 inverse.

And so we will find the structure of L by identifying the structure of these matrices. And then identify the structure of the product of these matrices and we will see that finding L is actually very easy. So, structure of Mk inverse so recall that Mk is the matrix such that Mk times Ak minus 1 is equal to Ak and Ak minus 1 is a matrix such that its first top left k minus 1 cross k minus 1 matrix is upper triangular.

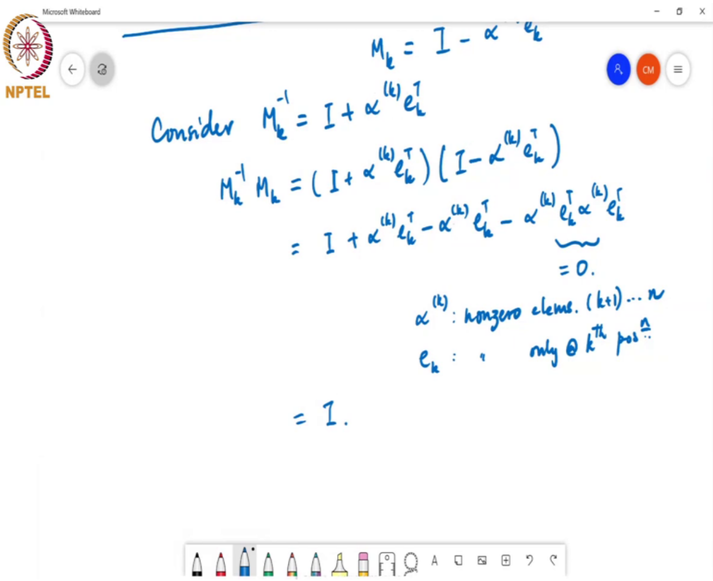
And below the upper triangular part you have zeros and this matrix is such that its top left k cross k matrix is upper triangular and it has zeros below the top left k cross k sub matrix. So, basically this Mk has the structure I minus alpha k e k transpose right that is that was our construction of Mk.

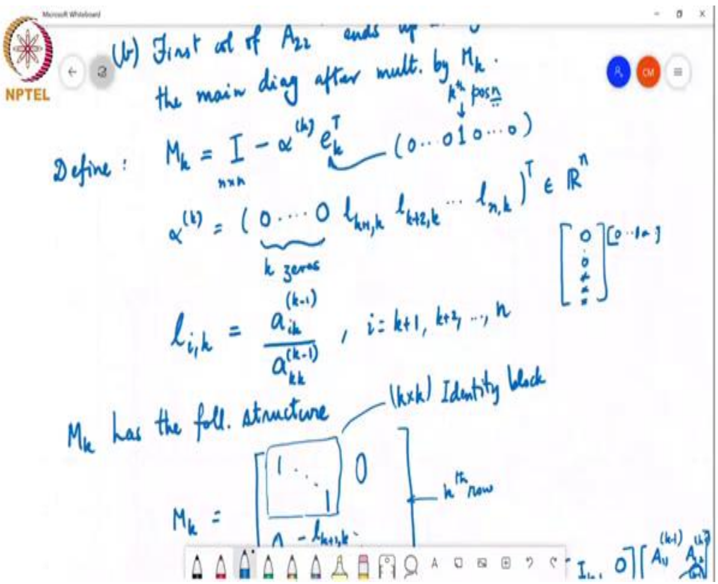
So, essentially what we see is that we get Ak from Ak minus 1 by taking Ak minus 1 minus something this, this times this matrix right. So, if we wanted to invert this operation and take Mk to the other side then it is sort of intuitive that maybe we need to do an addition operation.

So, the answer in fact that intuition is correct and so if you consider Mk equal to i plus alpha k e k transpose then if I multiply this so Mk minus Mk inverse to be i plus alpha k e k transpose. Then if I do if I multiply this with Mk then it is i plus alpha k e k transpose times i minus alpha k e k transpose which is equal to i plus alpha k e k transpose minus alpha k e k transpose minus alpha k e k transpose alpha k e k transpose.

Now, this e k transpose alpha k is actually the inner product between the vector e k and the vector alpha k. So, I claim that this is equal to 0 why is that true?

(Refer Slide Time: 26:40)





Student: Because, the matrix is constructed in such a way.

Professor: Yes.

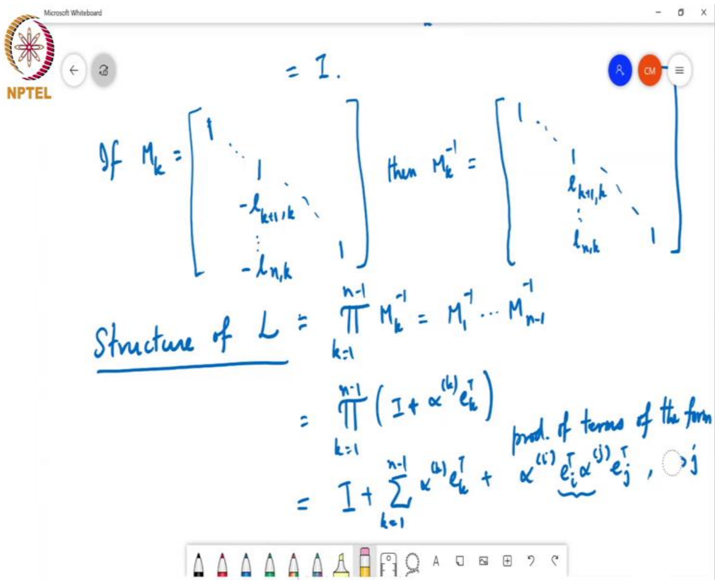
So, alpha k has non-zero entries in k plus 1 through n. Only those entries of alpha k are non-zero whereas e k as a 1 only in the kth position. So, this has a 1 in the kth position but the kth entry of alpha k is always equal to 0 only the k plus one to nth entries are non-zero.

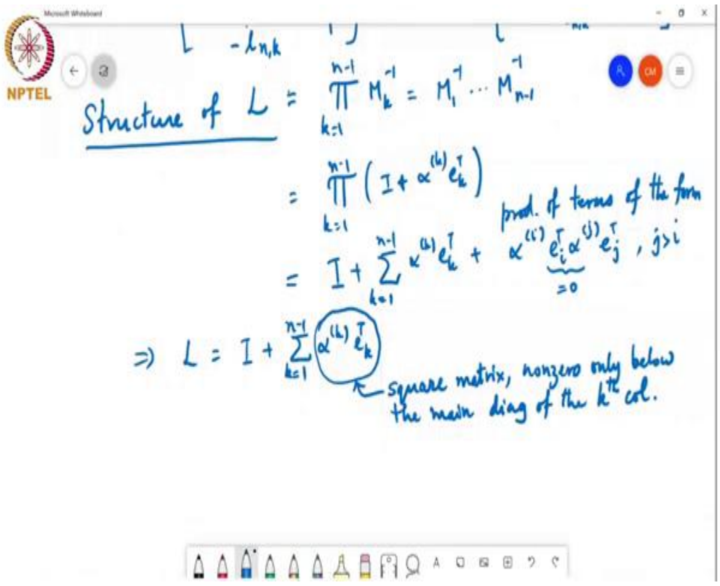
Let me go back here. See, alpha k had k zeros and then k plus 1 to n you have all these l k plus 1, l k plus 2 up to l k n. So, the these entries the so then e k has a 1 only in the kth position you see so if I take the inner product of this vector with this vector this 1 will multiply 0 and all these entries will multiply zeros here.

And so their inner product is 0. So, this matrix drops off and these two just cancel each other and so this is equal to the identity matrix. So, I plus alpha k e k transpose is the is the inverse of Mk and notice that this is actually the only thing we used here is that the inner product of these two is 0.

So, if I have a matrix A equal to I plus u v transpose and u is orthogonal to v then a inverse is equal to I minus u v transpose. So, this is generally true as long as these two vectors are orthogonal to each other. So, basically the finding the inverse of Mk is very easy all you have to do is to change the signs of this part which was i minus alpha e k transpose.

(Refer Slide Time: 29:11)





So, basically if Mk is of the form you have 1 1 then l k plus 1 comma k etc up to minus l n k and then zeros everywhere else. And then ones here on the diagonal then Mk inverse will be the same matrix but 1 1 and here I have plus l k plus 1 comma k l n k and then once along this diagonal.

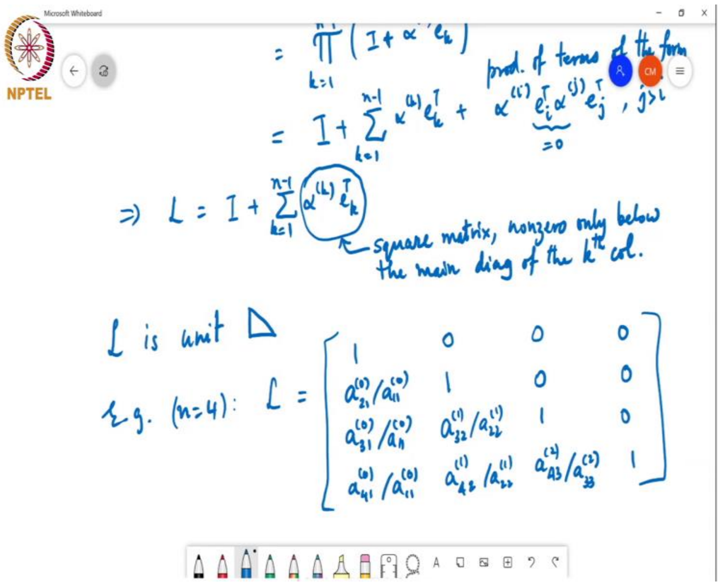
Now, so we now know how to find Mk inverse that is super easy given that we have already found what Mk is. So, if I now look at what is the structure of L, L is the product of these Mk's. So, this L is product k equal to 1 to n minus 1 Mk inverse which is M1 inverse Mn minus 1 inverse that is equal to the product of matrices of the form k equal to 1 to n minus 1 I plus alpha k e k transpose. So, it is I plus here because these are the inverse matrices.

And so if I just expand this out that is going to be equal to I will get an identity matrix when I multiply all the identity (matrix) matrices together plus I take one of these guys and multiply with the identity matrix that is the first, that is the next term k equal to 1 to n minus 1 alpha k e k transpose plus all these other terms which will look like of the form alpha i these are cross terms e i transpose alpha j e j transpose.

And this is for I greater than j. Then each of these if you look at the form of actually j greater than i if you look at the form of this this has a 1 only in the ith position and this will have non-zero entries from the j plus 1th entry position onwards. And so when I take this inner product it is always going to be equal to 0.

And so all the cross terms actually drop off and so L is actually equal to I plus sigma k equal to 1 to n minus 1 alpha k e k transpose that is it. So, basically each this is just the identity matrix which puts the, puts ones on the main diagonal of L and this is basically each of these is a square matrix which has non-zero entries only below the main diagonal of the kth column.

(Refer Slide Time: 34:17)



So, so basically L is unit lower triangular unit meaning it has once on the main diagonal and non-zero entries only below the main diagonal. So, for example for n equal to 4 4 cross 4 matrices L will be of the form 1 and here I will have a21 of 0 divided by a11 of 0 this is actually the entries of the a matrix itself a 0 is equal to a.

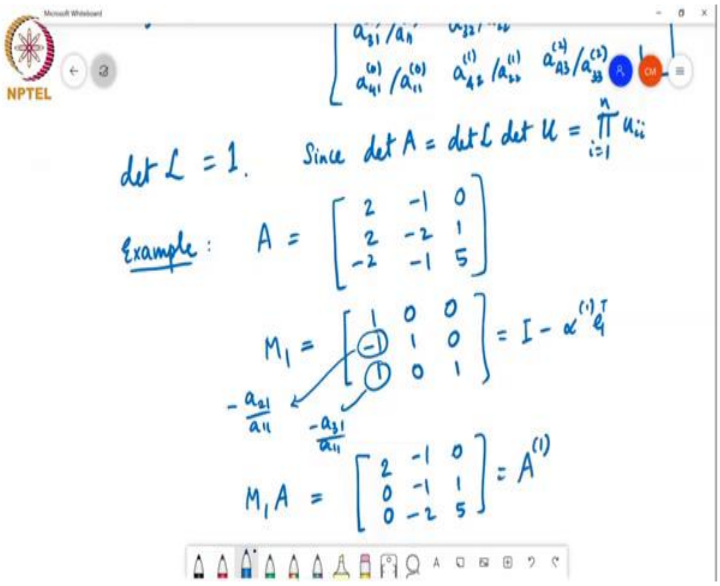
This is a31 of 0 divided by a11 of 0. So, I can just read off the first column of L from the matrix A a41 of 0 divided by a11 over 0 and then in the second column I have 0 here and a 1 here and a32 of 1. So, for this I need that matrix a1 divided by a22 of 1 a42 of 1 divided by a22 of 1 and this will have a 0 0 1 and then a43 of 2 divided by a33 of 2.

And then 0 0 0 0 1 in the last column. So, that is the structure of L. So, basically given this sequence of, so if you determine the sequence of gauss transforms you can form this matrix L without any further explicit computations. So, the inverses are accomplished by inverting a set of signs and multiplication is accomplished by just placing the non-zero elements of alpha k into the appropriate positions of L.

So, this so again so just to reiterate the sequence of gauss transforms we performed is exactly the same as gaussian elimination and therefore this LU decomposition is really actually a high level description of gaussian elimination there is no difference between the two.

And gaussian elimination itself is order 2 n cube over 3 floating point operations and that same thing carries over to the factorization I just discussed and in fact it is the lowest of any triangularization technique for square matrices without exploiting any further structure in the matrix. Now, since L is unit lower triangular the determinant of L is equal to 1.

(Refer Slide Time: 37:25)



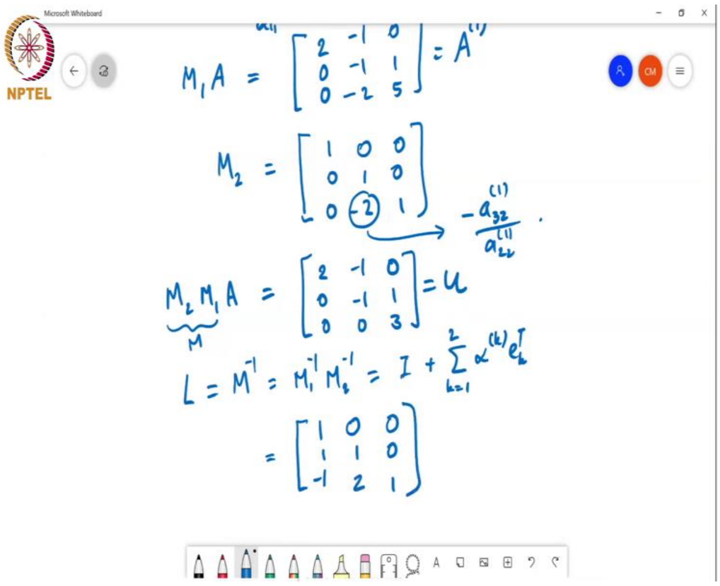
And so L is unit lower triangular. So, I have already written that so determinant of L equal to 1 and so basically A is equal to L U so since determinant of A equals determinant of L times determinant of u we have the determinant of A equals the determinant of u and u is upper triangular and so that is equal to the product of i equal to 1 to n uii.

So, the product of the diagonal elements of U will actually give you the determinant of A. So, let us maybe just illustrate this with an example. So, suppose my matrix A was the matrix 2 2 minus 2 minus 1 minus 2 minus 1 0 1 5 then the matrix M1 is going to be equal to this thing with I have ones on the diagonal it is an upper triangular lower triangular matrix with once on the diagonal and what is this entry it is minus of this a21 divided by a11 this ratio is 1 so this will be minus 1.

And this is, this entry is minus of this divided by this and so it is plus 1 and this will be 0. So, you are only placing non-zero entries below the first column so this is exactly that i minus alpha of 1 e 1 transpose. Just for your reference later I will just write what this is this is minus a21 over a11 and this is minus a31 over a11.

Now, if I did M1 times A I will get the matrix it will keep this entry as it is it will place zeros here and all the other entries you actually have to calculate them and you get minus 1 minus 1 minus 2 0 1 and 5. And this is what we are going to call the matrix A1.

(Refer Slide Time: 40:23)



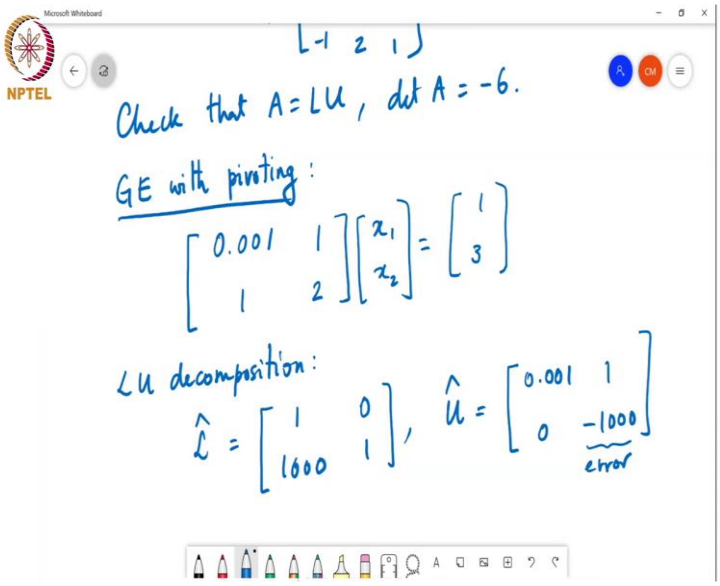
And now when you have this matrix A1, M2 is very easy you can immediately write it by inspection it is going to have non zeros below the main diagonal of the second column. So, only one entry here will be nonzero and then you have the identity structure. So, this is going to look like this 1 0 0 0 1 0 0 and this entry is going to be this divided by this with a negative sign.

So, this is 2 under the negative sign I have to write minus 2 here and this is 1. And again for the sake of completeness this is minus a32 of 1 divided by a22 of 1. And then if I calculate M2 M1 A that is going to be M2 times A1 which will give you 2 minus 1 0 0 minus 1 1 0 0 3 and this is my matrix u.

So, I have got it in the upper triangular form and L which is equal to M inverse so if I call this matrix M it is m inverse which is M1 inverse M2 inverse this will be equal to i plus sigma k equal to 1 to 2 alpha k e k transpose. And once again all I have to do is to invert the signs of these things and place them below the main diagonal.

So, that is equal to so I will have this 1 0 0 1 0 and 1 here and below this I just have to place the negative of this quantity so it will become 1 minus 1 and below this I have to place the negative of this quantity so it will become 2 so this is my matrix L.

(Refer Slide Time: 42:33)



So, you must check this that A is equal to L u so L is this matrix u is this matrix if you multiply these two together it will give you back the A matrix and determinant of A is equal to the product of the diagonal entries of u which is minus 6. So, you can easily check these for yourself so this is how the LU decomposition works.

So, this is the vanilla version of gaussian elimination or LU decomposition that I described but there are things one can do to make this more numerically stable and that is called gaussian elimination with pivoting.

So, we will start with a simple arithmetic a simple example which just to illustrate why we need to do this. So suppose, we are doing our computations in base 10 arithmetic but we can only store 3 digits of any numerical computation that is because all computers operate with finite precision arithmetic.

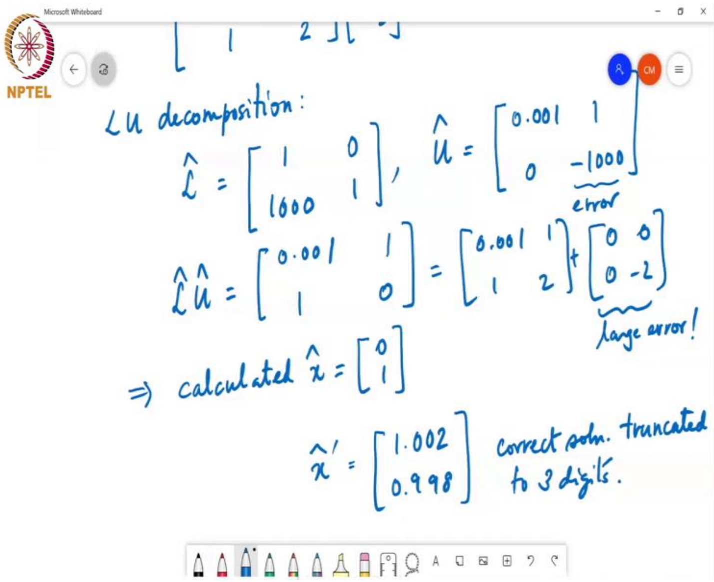
So, they are going to they are going to have to chop numbers that are too small that to be represented on the computer. So suppose, I had a matrix like a system of equations 0.001 1 1 and 2 this times x1 x2 is equal to 1 3. Suppose, I wanted to solve this you can already see that if I set x1 x2 equal to 1 this will be 1.001 and x this will become 3 and that almost solves this problem.

So, if I get an answer which looks close to 1 1 I know that I have solved the problem but if you try to do this using LU decomposition remember that LU decomposition is one way to solve these kind of problems and it gives you some computational advantage in very large dimensional systems.

Because, once you have the LU decomposition you can do a step of forward substitution followed by step of backward substitution, substitution to find x. So, if you work out the LU decomposition exactly as I worked out in that numerical example what you get is the following you will get L hat I will just call it L hat because this is the value you will get if you did this carefully but with finite 3 digit arithmetic so precision arithmetic.

So, this will be 1000 that is just the ratio of these two and then 0 1 and u hat will be equal to 0.001 1 0 and minus 1000. It turns out that this value is actually minus 1000 point 0 0 1 or something like that but because you are doing it in finite precision arithmetic this is, this has an error due to this round off or chopping that happens in finite precision arithmetic.

(Refer Slide Time: 46:30)

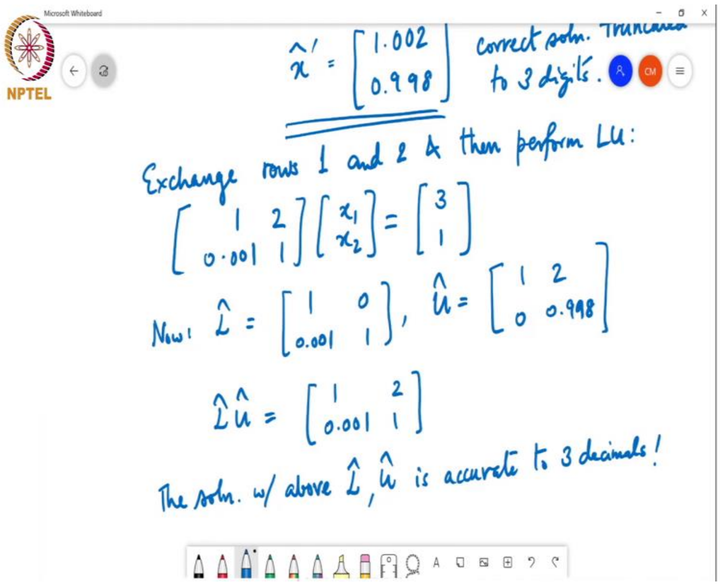


So, basically if you compute L hat u hat from this what you end up with is the matrix so that is easy you can just do it right very quickly. So, you get 0.001 then this times this I will get 1 this times this will give me 1. This times this actually gives me 0. So, L hat u hat is quite different from this matrix a this is actually equal to the matrix a 0.001 1 1 2 plus this error matrix 0 0 0 minus 2. So, you are actually making a large error.

So, if you use this L hat u hat and then you did your forward substitution followed by backward substitution what you will get is a x hat which will end up becoming 0 1 which is the answer truncated 3 digits but it is quite different from x hat dash which is equal to 1.002 and 0.998 which is the correct solution truncated to three digits.

So basically instead of directly doing the LU here the problem arose because when I when I tried to find this L I had to divide this by this and that gave me this large factor of thousand and in finite precision arithmetic these kind of large numbers mess up your calculations.

(Refer Slide Time: 48:26)



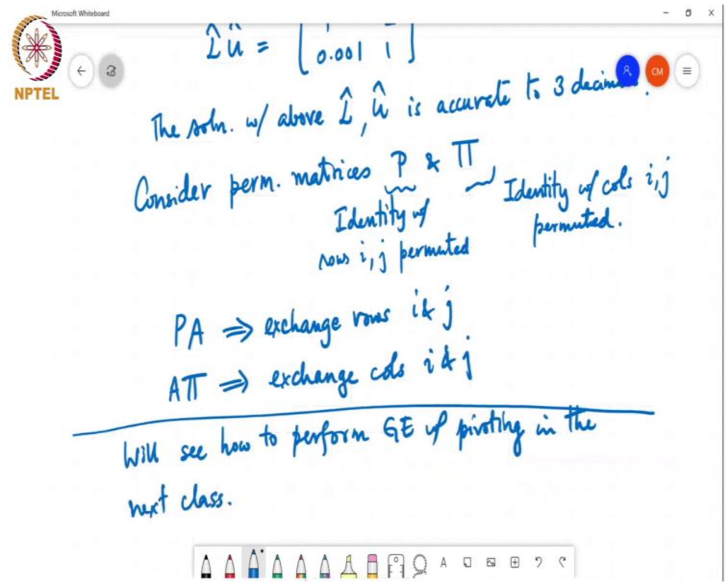
And so if I had exchanged rows 1 and 2 and then perform LU then what I will get is the once I exchange the rows I will get 1 2 0.001 and then 1 this times x1 x2 is equal to I have to exchange this side also so 3 1. So, when I exchange the rows everything here also it gets exchanged.

So, I start with this system of equations and if I now get now do the LU hat LU decomposition L hat will become equal to 1 0.001 which is the ratio of these two 0 1 and u hat will end up becoming equal to 1 2 0 0.998. And if I now compute L hat u hat that will be equal to 1 2 zero point need space 1 0.001 2 1 and if you use this L hat u hat the solution with the above L hat you had is accurate to three decimals.

Which is in other words it gives exactly the solution that I wrote earlier so the basic idea is to stabilize the gaussian elimination by exchanging rows and columns of when I exchange columns the entries of this vector get exchanged and I just have to undo that exchange after I have solved the problem.

So, we will stabilize the gaussian elimination by exchanging rows and columns such that the element with largest magnitude ends up in this top left corner of the matrix in the pivot position, upper left position. So, that that is going to be the core idea of gaussian elimination with pivoting.

(Refer Slide Time: 51:11)



So, basically we will end up with permutation matrices. Which I am going to denote by P and pi such that this P is an identity matrix with some rows i and j permuted. So, I could write it as P i j but just simplifying notation here and this is a also the identity matrix but with columns i and j permuted.

Then pre-multiplying by P so PA what it does is to exchange rows i and j and if I do A pi this gives me a matrix where I have exchanged columns i and j. So, that so basically what we will do is we will do this kind of row and column exchanges so that at each step we end up with the largest possible element as the pivot element and largest possible magnitude element as the pivot element.

And then we will use that to construct the LU decomposition. I will cover that in the next class so we will stop here for today. We will see how to use, so that is all I have for today. So, we will continue again on Monday.