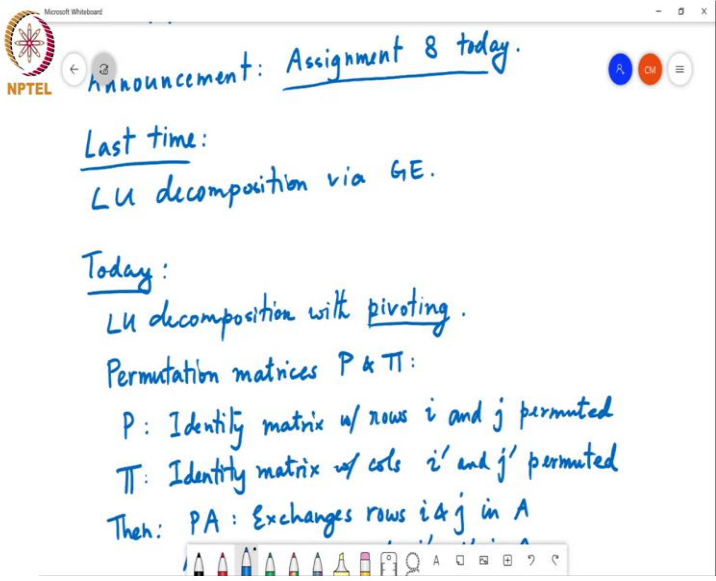
Matrix Theory
Professor Chandra R. Murthy
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore
LU Decomposition with Pivoting

(Refer Slide Time: 00:14)



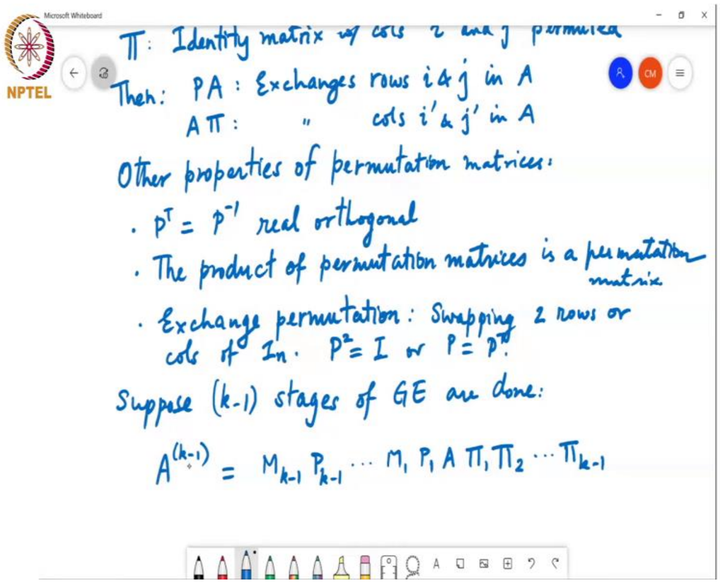
So, the last time we looked at LU decomposition by a gaussian elimination and we outlined the procedure and I also showed you an example and towards the end of the last class I told you about some numerical issues that arise in the LU decomposition.

Namely, that if there is a very small number that appears as a pivot element then inverting that when you compute it using a finite pressure precision machine can lead to incorrect answers.

And so that motivates us to look at LU decomposition with pivoting. So, pivoting is the process by which you try to stabilize the LU decomposition process and the way we do that is through the use of these permutation matrices. So, just very briefly a permutation matrix this, there are many types of permutation matrices but for the purposes of this discussion we will discuss about permutation matrices where a pair of rows or a pair of columns are getting exchanged.

So, we have notation P and pi, P is the identity matrix with two rows say row i and j being permuted are exchanged. And pi is the identity matrix with columns i dash and j dash being permuted or exchanged. Then if you define P and pi this way then if you consider P A for any matrix A it exchanges the ith and jth rows of A and A pi similarly exchanges the i dash and j dash columns of A.

(Refer Slide Time: 01:58)



So, I mean permutation matrices have many properties including these and also so I will just say. Other properties of permutation matrices. So, for example P transpose equals P inverse so it is a real orthogonal matrix and the product of permutation matrices is another permutation matrix. So, and these permutations that I discussed here are called exchange permutations which basically exchanges exactly two rows or two columns of the matrix it involves swapping two rows or columns of the n cross n identity matrix.

And such matrices have the property that P squared equals P A equals identity matrix. So, basically if you exchange two columns and then you exchange them back you get back the original matrix so if you apply P twice with an exchange permutation you will get the original matrix back.

So, P squared is the identity matrix or P equals P transpose because P so this means P is its own inverse but P inverse equals P transpose P transpose for any permutation matrix. So, P equals P transpose for such matrices so in other words if I take the n cross n identity matrix and exchange any two rows or any two columns I will still get a symmetric matrix.

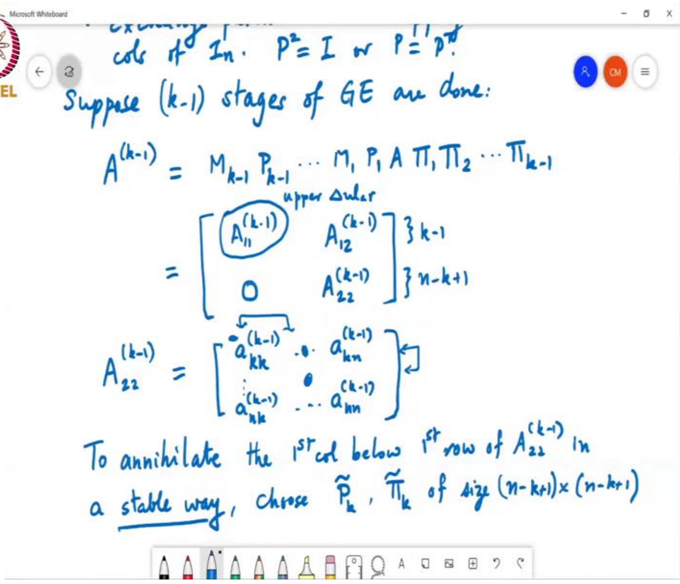
And that matrix is its own inverse that is special property of exchange permutations. So, now to connect this to the gaussian elimination based LU decomposition procedure so like the previous development suppose k minus 1 stages of the gaussian elimination have been done and we will describe how to proceed with the kth stage and from if you start with k equal to 1 2 up to n minus 1 then you get the entire LU decomposition.

So, suppose k minus 1 stages of gaussian elimination are done that is to say we have got this matrix Ak minus 1 which is equal to Mk minus 1 earlier I had Mk minus 1 all the way down to M1 times A is what I had as Ak minus 1 it was a pre-multiplication by these gauss transforms.

But, now I can, I will have a potential permutation that was done at the k minus 1 stage P k minus 1 all the way up to M1 P1 A pi 1 and then the second stage I may have to do another column exchange that is pi 2 up to pi k minus 1. So, these are the exchanges that I have done so far.

So, I have not told you how to do these exchanges but it will become clear as soon as I tell you what exchange will do with the Ak minus 1 matrix to get the A kth matrix.

(Refer Slide Time: 06:41)



Now, this matrix by construction has the form it has A11 k minus 1 at the top left and this is upper triangular and I have A12 k minus 1 and then 0 here and A22 k minus 1. And this is k minus 1 rows and this is n minus k plus 1 rows and similarly this is k minus 1 columns and n minus k plus 1 columns.

And now in order to describe what I want to do next I will consider the entries of A22 k minus 1. So, let us call them it starts with index k akk of k minus 1 akn because this is the first row of A22 k minus 1 is the kth row of this matrix Ak minus 1. So, this k minus

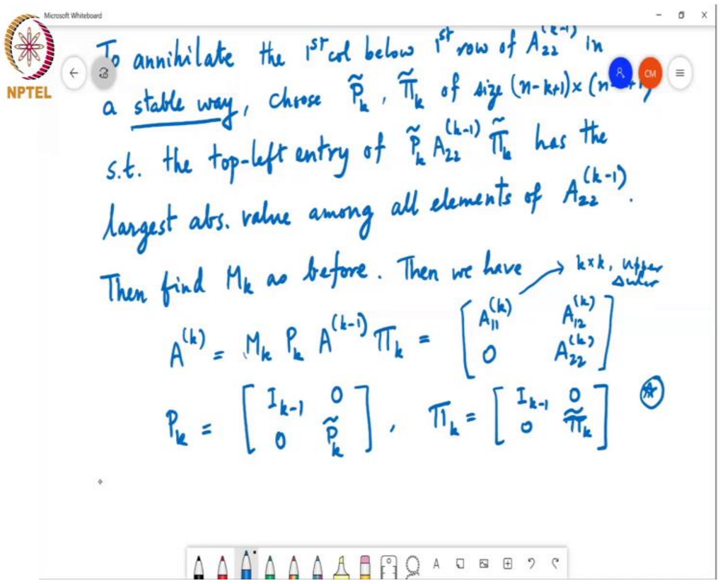
1 and ank of k minus 1 ann of k minus 1. So, this what we will call the entries of this matrix.

So, basically when I apply Mk what should happen is this entry will remain as it is and everything else below this will become zero. But, we want to do this in a stable way meaning that among all these entries here we want to get the largest magnitude entry and place it as the top left entry.

Because, we are going to be dividing all these entries by that entry and then doing row operations to make zeros appear in the below the, below the diagonal below this entry. So, to annihilate the first column below the first row of A22 k minus 1 in a stable way choose we will call it P tilde k and pi tilde k of size n minus k plus 1 cross n minus k plus 1.

So basically, the point is that suppose some entry over here is the largest entry then what you have to do is you have to exchange these two rows so that this entry goes up here and then you should exchange these two columns so that that entry ends up over here right. So, if I do one row exchange in one column exchange I can take whatever entry I find to be the biggest and make it appear as the top left entry.

(Refer Slide Time: 10:36)



So, you do this such that the top left entry of Pk tilde A22 k minus 1 pi k tilde has the largest absolute value among all elements of A22 k minus 1. Then we find Mk as before. So, we saw the last time that we will find Mk by, it will be a matrix with ones along the diagonal zeros above the diagonal and minus lk plus one comma k up to minus ln k and below the main diagonal of the kth column with lk chosen to be equal to Ai k of k minus 1 divided by akk of k minus 1.

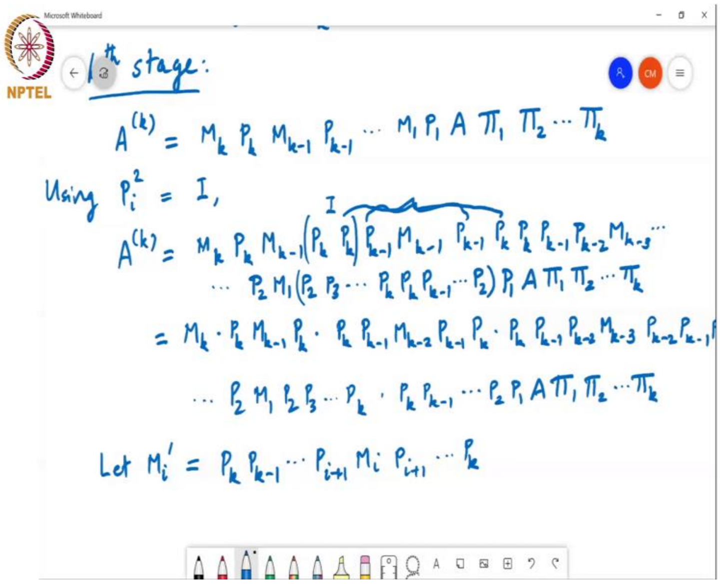
So, all this we discussed the previous time so then we will have Ak is equal to Mk Pk Ak minus 1 times pi k which will be of the form A11 k which is now k cross k A12 and upper triangular 0 and A22 k. And of course a Pk is related to this Pk tilde just by slapping on an identity matrix of size k minus 1 like this P k is equal to I k minus 1, 0 0 P tilde k.

And similarly pi k is I k minus 1, 0 0 by tilde k this just ensures that the first k minus 1 rows and columns of Ak minus 1 remain untouched. So, keep in mind that the exchanges are happening over entries columns and rows of Ak so this when I multiply this Pk times Ak minus 1 it is going to exchange rows of Ak minus 1 but it will exchange rows starting from the kth row up to the nth row it will not touch the first k minus 1 rows of Ak minus 1.

And similarly write multiplying by pi k will exchange columns starting from the kth column to the nth column it would not touch the first k minus 1 columns of Ak minus 1. So, we will come back to that point later so I will put a star over here. So, this is basically fine so we can execute this and at the end what will happen is that since Ak is of this upper triangular form if I do n minus 1 steps of this kind of thing I will get an upper triangular matrix out here.

So, but then I need to still show that this matrix which is the pre-multiplying matrix is of the form l which is a lower triangular matrix so that when I do when I pre multiply by l inverse then I get a is equal to L u where L is a lower triangular matrix and this u that I have over here is an upper triangular matrix. So, we still need to discuss how that matrix will end up becoming lower triangular.

(Refer Slide Time: 16:02)



So in the kth stage what we have is Ak is equal to I am just copying from here Mk Pk but Ak minus 1 itself is Mk minus 1 Pk minus 1 Ak minus 2 which is Mk minus 2 Pk minus 2 etcetera so this keeps going so I have Mk Pk and k minus 1 Pk minus 1 all the way down to M1 P1 A.

And then there is a pi k but when I substitute for Ak minus 1 I would have got Mk minus 1 pk minus 1 all the way down to M1 P1 A pi 1 pi 2 all those things will follow. So, this pi 1 by 2 up to pi k. So, this is the structure of Ak this is how it is obtained now because these P i is are exchange matrices we have that P i squared is equal to the identity matrix so using this we have Ak is equal to I will still keep Mk, Mk Pk and then Mk minus 1.

What I will do is before I write P k minus 1 I will write P k P k P k minus 1. So, P k P k is the identity matrix and then I have Mk minus 2 and so on all the way down to so maybe just to illustrate this thing I will just write one more term here so this will be clear.

So, there is P k P k times so this is the identity matrix, so this is the identity matrix P k minus 1 and then I have Mk minus 1 and then instead of writing m, P k minus 2, I will write P k minus 1 P k then I will multiply again by P k P k minus 1 and then I will write P k minus 2 and so on.

So, what I am doing here is notice that this has this structure P k minus 1 Mk minus 1 P k minus 1 and then I have P k sorry if I combine this together as well then I get this form P k P k minus 1 Mk minus 1 P k minus 1 P k. Here, I have Mk P k minus 1 P k and then the next term will be P k minus, P k P k minus 1 P k minus 2, Mk minus 3 and then there will be a P k minus 2 P k minus 1 P k so all of this things becomes like a symmetric product.

And so this will go all the way down to just before the last term I will have like a P 2 M1 then instead of writing M1 P 1 A, I will write it as P 2 P 3 up to P k times P k P k minus 1 P 2 P 1 A. And then I still have my pi 1 pi 2 up to pi k. So, I will just rewrite this so that it is clear how I am splitting this product. So, I will write this as Mk times these three times together P k Mk minus 1 P k times P k P k minus 1 Mk minus 2 P k minus 1 P k times just one more term for the sake of completeness P k P k minus 1 P k minus 2 Mk minus 3 P k minus 2 P k minus 1 P k and so on.

And down, all the way down to the last term will be P2 M1 times P2 P3 up to P k times P k P k minus 1 down to P 1 A and I still have pi 1 by 2 pi k. And so now what I will do is I will let Mi dash be equal to

Student: Sir.

Professor: Yeah.

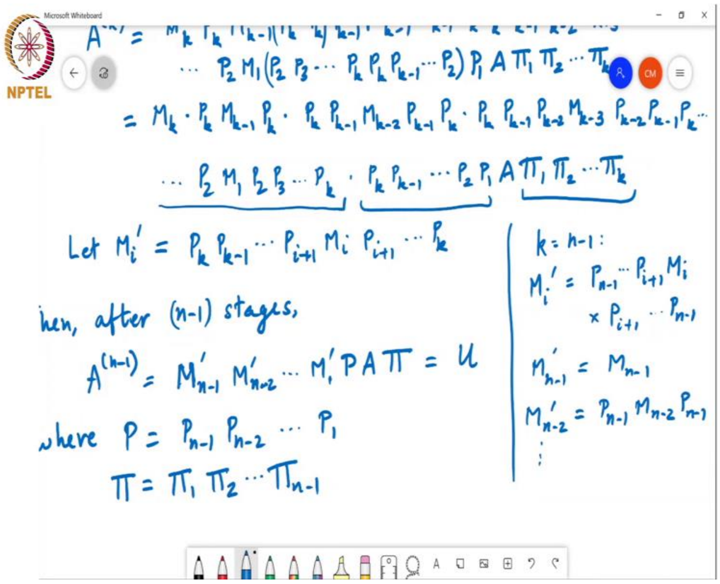
Student: In the first line where you wrote Ak after (())(22:19) how P k times P k is coming because at that place P k minus 1 is there so it maybe P k minus 1 times P k minus 1 that is.

Professor: No see I can insert a pk times pk wherever I want, it is the identity matrix. So, I am inserting it here.

Student: Okay, got it sir.

Professor: So, P k P k minus 1 up to P i plus 1 times Mi P i plus 1 up to P k. So, I will define this to be mi dash.

(Refer Slide Time: 23:13)



Then after n minus 1 stages what we have is that A n minus 1 which is an upper triangular matrix is equal to Mn minus 1 dash M n minus 2 dash all the way down to M1 dash times. So, I have all this together till here that will give me M1 dash and this product which will be P n minus 1 down to P 1 I will call that P and then this is A and then this product pi 1 through pi n minus 1 I will call that pi.

And this is an upper triangular matrix u. P is equal to P n minus 1 P n minus 2 all the way down to P 1 and pi equals pi 1.

Student: Sir.

Professor: Yes.

Student: Sir, in A n minus 1 should not and A M n minus one be multiplied because in Ak there was Mk being multiplied.

Professor: Can you say that again what is the question?

Student: In Ak sir the term is Mk P k Mk minus 1 P k so in An minus 1 there should be an A, Mn minus 1 there right?

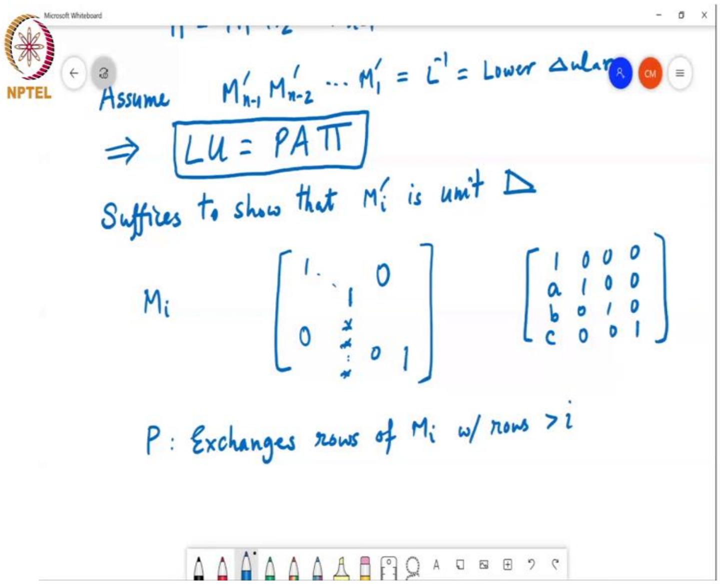
Professor: There will be an Mn minus 1 correct. So, you are right it will be Mn minus 1 and that is why M n minus 1 dash is defined to be so I think you can, I understand your confusion so if I want, so in the, so notice that the first so this thing is true for i being less than k. At i equals k these matrices would not be there I cannot go to so if I take, oops, yeah so let us maybe clarify this point, oops.

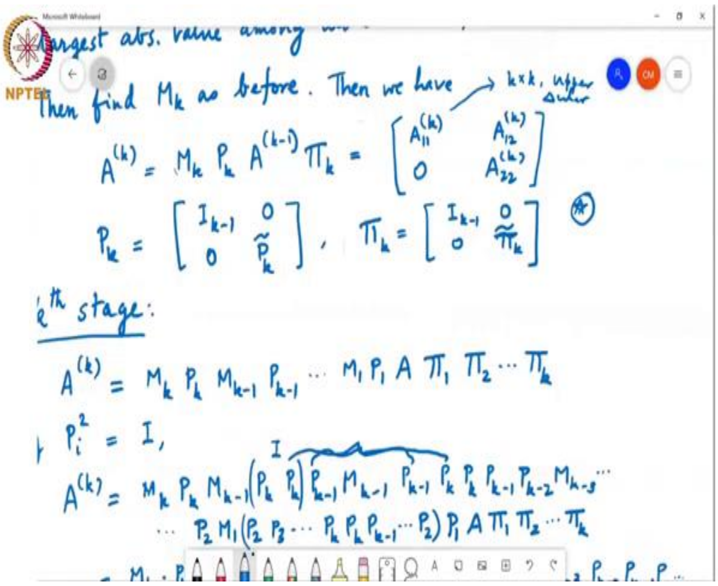
So, if I take if I take k equal to n minus 1 because I want A n minus 1 so then I will write over here k equal to n minus 1 then I have Mi dash is equal to P n minus 1 all the way up to P i plus 1 times Mi times P i plus 1 all the way up to P n minus 1. And so notice that if I take i equal to n minus 1 then I will get Mn minus 1 dash and I have Mn minus 1 here but this should be P n minus 1 plus 1 which is P n. but, there is no matrix P n that we are using in this process so these matrices will not be there for the M n minus 1 dash.

So, M n minus 1 dash is actually equal to M n minus 1 there is no matrices multiplying these. But, if I take M n minus 2 dash that is going to be equal to P n minus 1 M n minus 2 P M n minus 1 and so on. Does that clarify your question?

Student: Yes sir, thank you sir.

(Refer Slide Time: 28:28)





Professor: So, basically what we have then is so we have this product of all these matrices that times PA pi is an upper triangular matrix. So, for a moment let us assume M n minus 1 dash n minus 2 dash M1 dash is equal to L inverse and which is a lower triangular matrix.

Just for a moment imagine that this is true I will show you why this is true in a minute but assume that this is lower triangular. Then what we have is if I multiply by L inverse or if I multiply by L on both sides then I will have L times u is equal to P times A times pi. So, it is an LU decomposition not of A but it is a the LU decomposition the product of L and u gives you a row and column permuted version of A.

But, I still need to show that this product is going to be lower triangular. Now, the point is that if each of these matrices were lower triangular then of course their product would also in fact they are all matrices which we defined the original Mk s that we defined had were unit lower triangular.

So, we will end up showing that these matrices are also unit lower triangular. So, it is sufficient to show that Mi dash is unit lower triangular and the, but then this is simple because these permutation matrices like I mentioned at that star the permutation matrices only touch the rows and columns corresponding to so in the k minus 1th stage they will only exchange rows and columns corresponding to the A22 k part that is they are not touching the top k minus 1 cross k minus 1 entry.

So, in other words so for example if you take the original Mi if you remember the original Mi was of this form, was of this form where I had ones along the diagonal but in the k so Mi right so ith column so in the ith column I had a 1 here and then I had some let me just go back here in my notes yeah it was minus L k plus 1 or Li plus Li like that.

I plus 1 comma it is too difficult to write so I will just say here the sum star here and some star here whatever these entries were non-zero only below this thing everywhere else it was zeros but only these entries were non-zero. And now what I am doing in the kth stage is that I am applying an exchange matrix which is exchanging rows and columns corresponding to this part of the matrix.

So, if I exchange rows and columns corresponding to this part then let us see so what happens is that P matrix so let us see, so let me put it this way suppose I take this matrix with ones along the diagonal and then non-zeros only here. So, maybe it is even easier if I take a slightly more concrete thing so let me write it as 1 a b c and then 0 1 0 0 0 0 1 0 and 0 0 0 1.

And suppose I had exchanged some two rows let is say these two rows and then say, so I need to bring the largest element to the top left so what yeah so it will be, it would not be like this. So, I have to look at A22 of k minus 1 and I will be trying to bring the largest element over here.

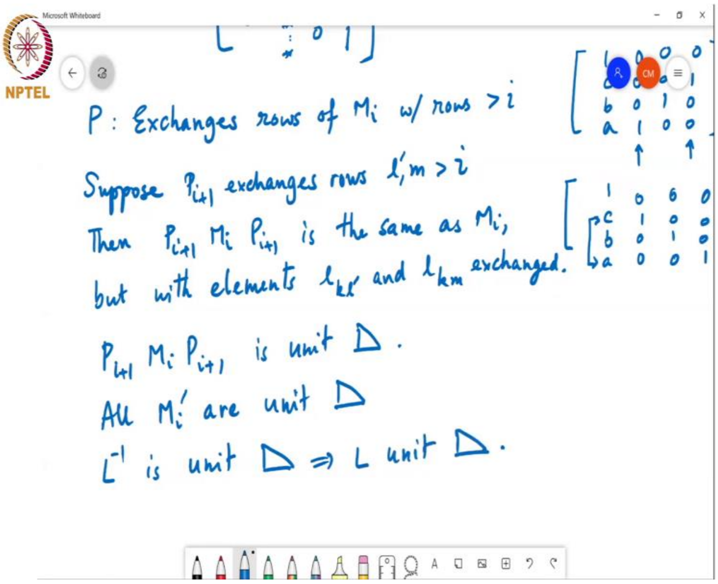
So, the exchange will be of the form where I would exchange maybe these two rows and then I have to exchange some pair of columns but of course I am trying to bring it bring the largest entry over here so maybe it is an exchange of these two columns. Now, if I exchange these two rows what will happen is I will get b okay let us write that b 0 1 0 out here and then a 1 0 0 here and then 1 0 0 0 here and c 0 0 1 here.

And now let is say I exchange these two columns then I will end up with yeah so I will put this here 1 if I exchange these two columns I will end up with no this is not how you go one second this is not correct.

Let me write it in a different way. So, the P matrix it exchanges rows of Mi with I think I understood the point is that Mi dash is yeah so I, yeah my example was what I was doing was not correct. So, maybe we will come back to that with rows greater than I and so basically let me let me put it like this. Suppose, I start with this matrix 0 0 0 a 1 0 0 b 0 1 0 and c 0 0 1.

Then remember that the Mi dash is obtained by doing P A, P M i times P i so if you go back here notice that the core structure is something like this P i plus 1 Mi P i plus 1 that kind of thing and P i plus 1 is going to exchange rows in Mi corresponding to some pair of rows which are greater than i. So basically, what you do is here in this structure what it will do is you pick two rows the it is the same matrix P i and P i plus 1 that is going on the left and right.

(Refer Slide Time: 37:32)



So, suppose you exchange these two columns then you will also be exchanging the first and third column you will be exchanging the first and third row.

If I do this let is see what happens when I exchange the first and third row I will get b 0 1 0 a 1 0 0 and 1 0 0 0 and c 0 0 1 here and now if I exchange these two columns then this column appears here. I will have so I still have the same just a second what am I doing wrong here?

I have to correct myself again anyway let us see so the exchange is going to be among rows of Mi with indices greater than i so this part is not going to get touched, it is going to be a pair of rows and columns among these. So, if I take let is say these two rows and I exchange them then let is see what I get I will get a matrix 1 0 0 0 c 0 0 1 and then b 0 1 0 and a 1 0 0.

And now, I have to exchange the same columns because it is P i plus 1 M i P i plus 1. So, I have to exchange the same second and fourth column of this matrix if I exchange these two then what I will get is the matrix 1 c b a 0 1 0 0 0 0 1 0 and 0 0 0 1 exactly. So, this is the idea so all that happened by this pi i plus, P i plus 1 Mi times P i is that the entries a and c got exchanged.

Otherwise, the structure is exactly retained so that is the basic idea here. So, for example suppose P i plus 1 exchanges rows l m which are both greater than I then P i plus 1 Mi times Pi plus 1 is the same as Mi but with elements l k, l k l, so, this is a bad notation because I had written l i j to be the entries of this matrix mi but anyway it is l k l so let me say this is l dash.

Then k l dash and l k m exchanged. Or in other words this matrix P i plus 1 M i P i plus 1 is unit lower triangular and that means that all M i dash are unit lower triangular and this means that l inverse is unit lower triangular which implies that l is also unit lower triangular.

So, that completes this discussion to say that this process even with pivoting gives you an LU decomposition not of A but of P A pi where P is a permutation, P and pi are permutation matrices.