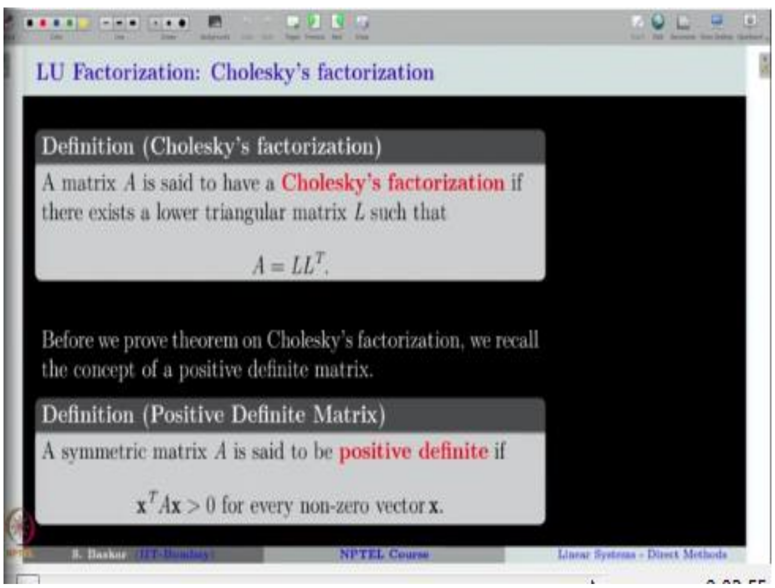
Numerical Analysis
Prof. Dr. S. Baskar
Department of Mathematics
Indian Institute of Technology - Bombay

Lecture-11
Linear System: LU Factorization (Cholesky)

Hi, we are learning direct methods for solving non-singular linear systems. Recall, the direct methods give exact solution to a linear system when there is no rounding error involved in it. We have learned Gaussian elimination method, Doolittle factorization method and Crout factorization method in our last class. In this class we will learn another LU factorization method called Cholesky's factorization.

(Refer Slide Time: 00:58)

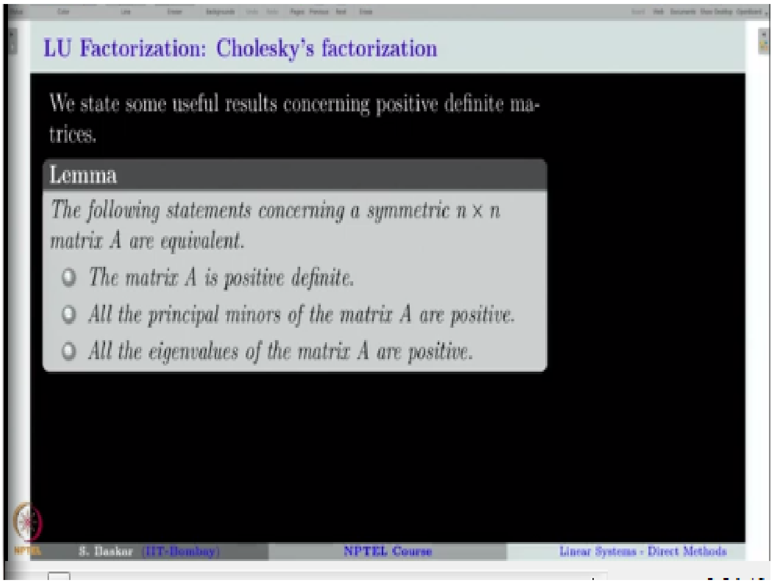


A matrix A is said to have a Cholesky's factorization if there exists a lower triangular matrix L such that A can be written as LL^T. Observe that if L is a lower triangular matrix then L^T is an upper triangular matrix therefore this form of factorization is indeed a LU factorization. Also you can observe that the right hand side matrix is a symmetric matrix because we have LL^T therefore if you take the transpose of that you again get LL^T.

Therefore Cholesky's factorization is possible only for symmetric matrices. Also we will see that Cholesky's factorization exist if the matrix A is a positive definite matrix. Let us recall from our linear algebra course what is mean by a positive definite matrix? A symmetric matrix A is said to be positive definite if $x^T Ax$ is positive for all non-zero vectors x. You can see from this definition that it is not that easy for us to check whether a given symmetric matrix is positive definite or not.

For that you have to take all the non-zero vectors and compute the expression $x^T Ax$ and check if it is a positive number that seems to be very difficult.

(Refer Slide Time: 02:56)

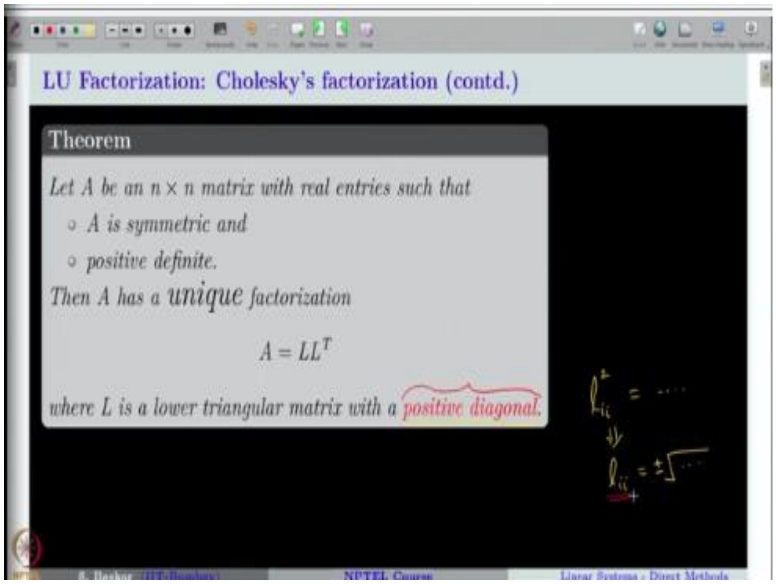


Therefore we have some equivalent properties which we would have studied in our linear algebra course. Here we will just recall those properties in the form of a lemma, suppose we have a symmetric $n \times n$ matrix A, then we say that A is positive definite. If and only if all it is principal minors are positive, that is one equivalent condition for positive definiteness. You can see that principal minors can be computed relatively in a easy way.

Therefore this condition is more handy for us to check whether a symmetric matrix is positive definite or not. Another equivalent condition is that a symmetric matrix A is positive definite if and only if all the Eigen values of A are positive. Again this can also be checked, well, finding Eigen values is little difficult computationally than finding principal minors however this can also be checked.

We will not get into the proof of this lemma because it is a part of linear algebra course but we will use these properties in our theorem on existence and uniqueness of Cholesky's factorization.

(Refer Slide Time: 04:17)



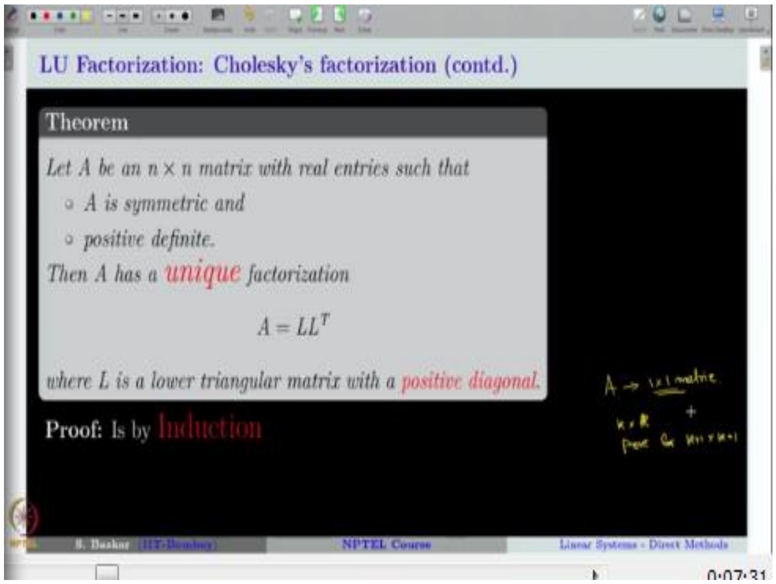
The theorem says that suppose you have an $n \times n$ matrix with real entries, note that we will always work with matrices of real entries. Therefore even if I do not tell this you should keep in mind that we always work with matrices with real entries. So, we have a matrix $n \times n$ such that it is a symmetric matrix and also it is positive definite. Then we can always find a lower triangular matrix L such that you can write $A = LL^T$.

Moreover if all the diagonal elements of this lower triangular matrix L are positive then we can say that that is the only lower triangular matrix such that $A = LL^T$. That is, we get a unique Cholesky's factorization. So, when we say Cholesky's factorization it means finding the lower triangular matrix L such that $A = LL^T$. We are imposing this condition that is positive diagonal elements is just to make sure that we have a unique way of obtaining the matrix L.

In our computation you will see that all the diagonal elements of the matrix L will appear as l_{ii} square some number. Therefore when you take square root you get l_{ii} = square root of that number therefore you may have plus or minus also. So, you have 2 choices here, to be very specific we will always make our mind that we will choose the positive sign for those diagonal elements.

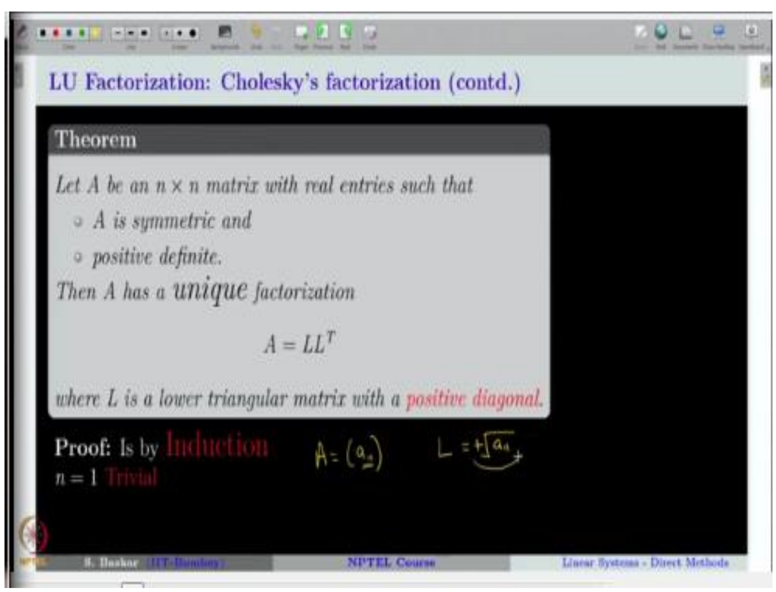
This is just to make the algorithm more precise without any ambiguity of what sign we have to choose. And in that way the theorem says that you will have a unique such matrix L that is what the statement is. Let us try to prove this theorem.

(Refer Slide Time: 06:46)



The proof is by induction. It means you first prove the theorem for the matrix A which is a 1×1 matrix. Then you assume that the factorization is possible for a $k \times k$ matrix and then prove it for $(k+1) \times (k+1)$ matrix. If you do so already you have proved it for 1×1 matrix therefore it is true for 2×2. Once it is true for 2×2 matrix then it is true for 3×3 matrix and so on, so that is the idea of induction.

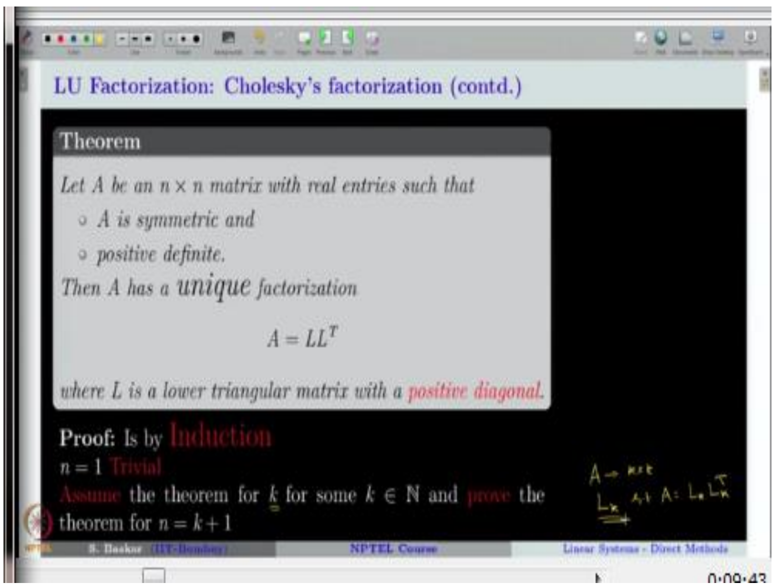
(Refer Slide Time: 07:32)



Therefore the first step is to prove that a unique factorization is possible when A is a 1×1 matrix. Let us take A as a_{11}, then obviously you can choose $L = \sqrt{a_{11}}$. Then you can see that A can be written as LL^T. Of course you can also choose $-\sqrt{a_{11}}$ but just because we made our mind to pick only the positive sign we will have $+\sqrt{a_{11}}$ here. Otherwise you can as well take minus here; there is nothing wrong in that.

Therefore the Cholesky's factorization is true when the matrix A is a 1×1 positive definite matrix. It means a_{11} should be positive, that is how we are getting a real matrix L. Remember, that L should be a lower triangular matrix with all its entries as real numbers, then only we will declare that Cholesky's factorization exists.

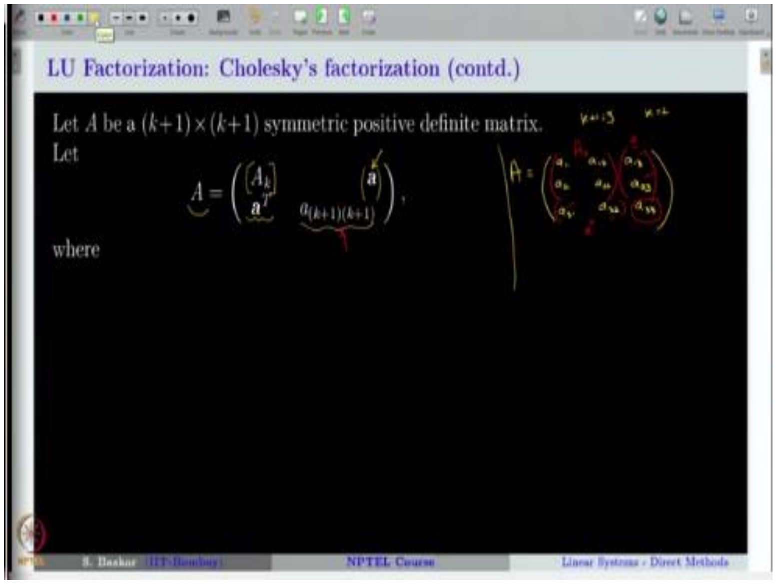
(Refer Slide Time: 08:45)



Now we will fix our induction hypothesis, as per that we will assume that the Cholesky's factorization holds for any $k \times k$ matrix for some natural number k. So, we are taking a k and then we are assuming that if A happens to be a $k \times k$ matrix then you can always find a L_k such that A can be written as $L_k L_k^T$, unique such L_k will exist with all it is diagonal elements as positive.

That is the assumption we are making, as per the induction method what we have to prove? If we have a matrix which is $(k+1) \times (k+1)$ matrix then we have to prove that the Cholesky's factorization exists.

(Refer Slide Time: 09:44)

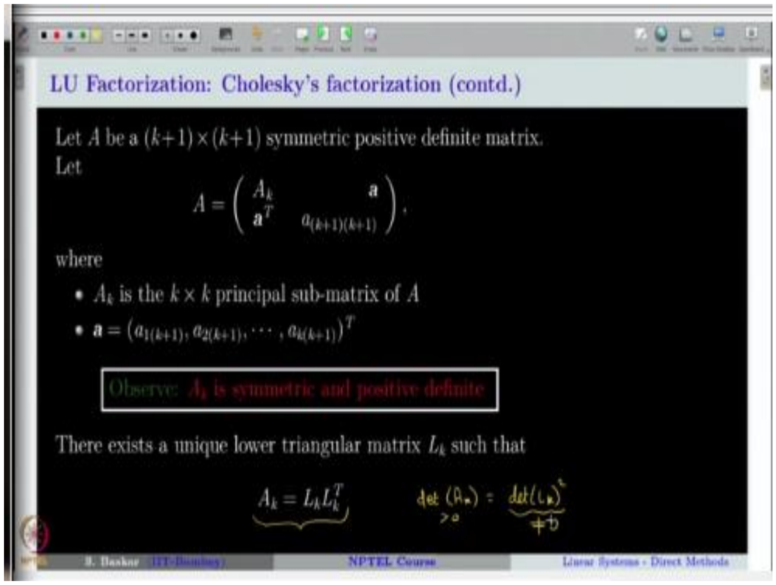


So, let us assume that A is a $(k+1) \times (k+1)$ symmetric positive definite matrix. Now we have to find the L such that $A = LL^T$. Remember, we have assumed that the Cholesky's factorization exist for any $k \times k$. Therefore what we will do is? We will write our matrix A as A_k which is the principal sub matrix of order k for A and then we will write the remaining column vector as a and it is transpose as a^T and then the last diagonal element like this.

Just to visualize let us take a 3×3 matrix as $\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$. So, here $k + 1 = 3$ we have taken in this example, therefore our k is 2. So, A_k should be the principal sub matrix of A of order 2, it means it should be this matrix, this is your A_2 here and then your column vector a is taken as this, this is your a. In all our discussions in this chapter we will always take a vector as a column vector.

Therefore its transpose that is the row vector will always be written as A^T. So, this is what is a transpose, remember it is a symmetric matrix. Therefore this and this will be the same and this and this will be the same, that is a_{31} will be equal to a_{13} and $a_{32} = a_{23}$ and then you have this element sitting here. So, this is how we are just splitting A into block wise.

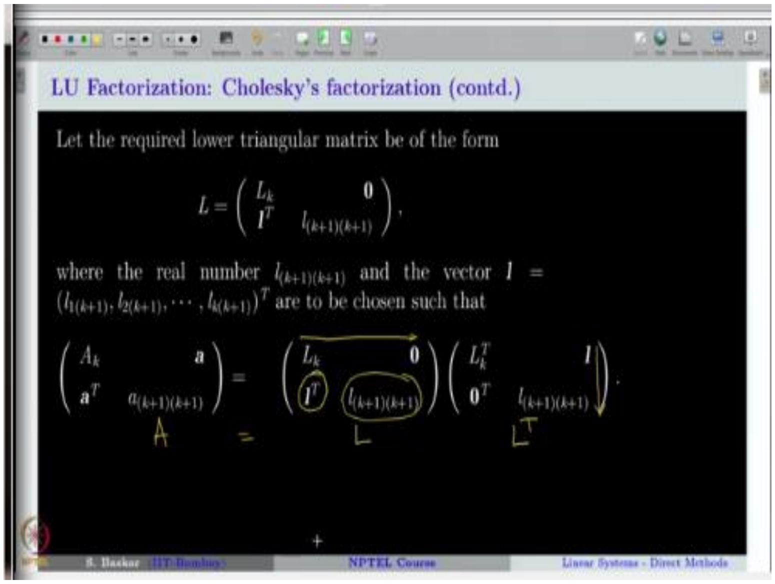
(Refer Slide Time: 12:25)



Where A is the $k \times k$ principal sub matrix of A and this vector a is the column vector at the last column removing the diagonal element that is $a_{(k+1)(k+1)}$. You can observe that A is a symmetric matrix therefore A_k is also a symmetric matrix. Also you can see that A_k is a positive definite matrix because A is positive definite by our lemma all it is principal minors are positive. In particular, the principal minors of A_k are also principal minors of A with lower orders. Therefore they all are also positive, in turn A_k is also a positive definite matrix, this is an observation.

Therefore by our induction hypothesis you can find the Cholesky's factorization for A_k, that is you can find a unique lower triangular matrix L_k such that $A_k = L_k L_k^T$, where all the diagonal elements of L_k are positive. This is the assumption as per our induction hypothesis. Now let us see how to construct the Cholesky's factorization for A with the help of the Cholesky's factorization of A_k.

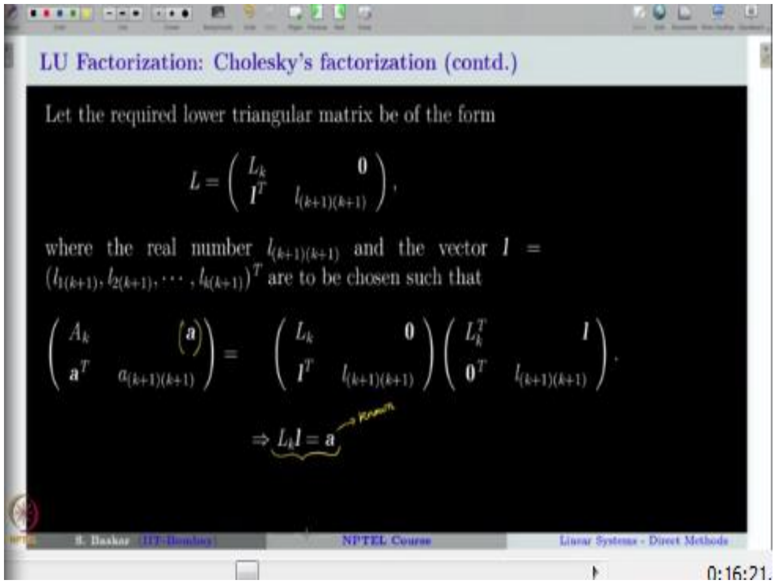
(Refer Slide Time: 14:07)



Let us propose that the Cholesky's factorization of A looks like this matrix, where L_k is coming from our Cholesky's factorization of A_k and then all these elements are 0 because it is a lower triangular matrix. And you have a vector l which is written in the row form therefore it is l^T and then this is a number $l_{(k+1)(k+1)}$. Now here we know this, this is known to us as per our induction hypothesis therefore we do not need to compute that.

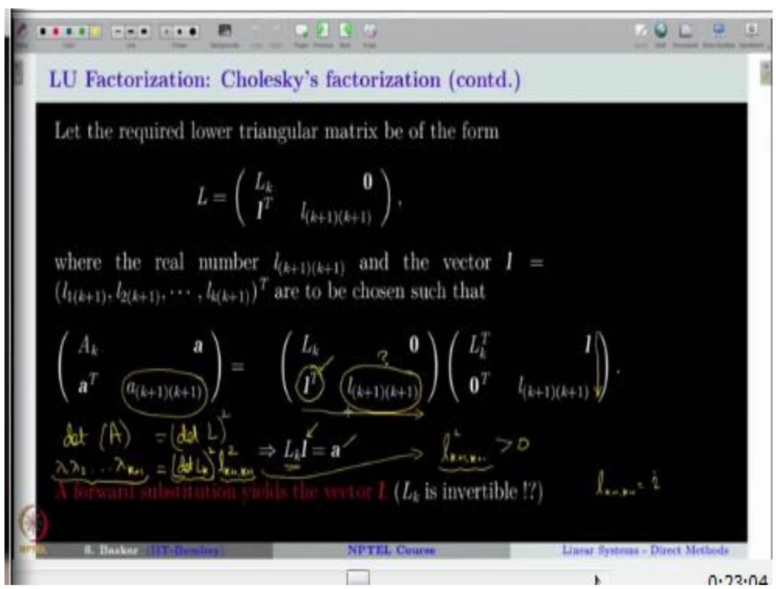
But we need to compute the vector l and the real number $l_{(k+1)(k+1)}$, how we have to find that? We should find these quantities in such a way that $A = LL^T$, so that is our Cholesky's factorization. So, we have to find this vector and this real number. How are we going to find this? Let us see, well, what we can do is? You take this the first block of the matrix L this is the row block and multiply it with the last column of L^T.

(Refer Slide Time: 15:51)



That will give you $L_k l$ is equal to we have to compare that with the corresponding entries of the left hand side matrix which happens to be the vector a. So, the vector a is known to us and therefore we got a lower triangular system with solution as the unknown vector l. Now how will you find it?

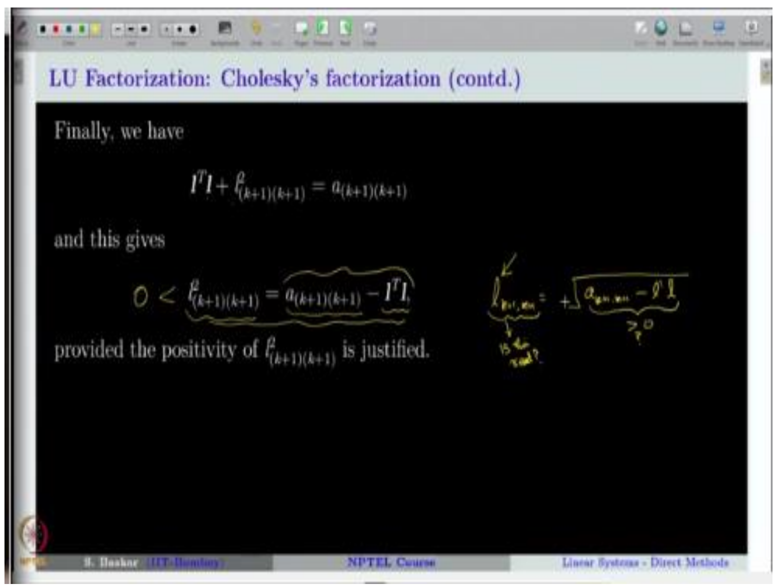
(Refer Slide Time: 16:22)



Well, you can just use the forward substitution to get the vector l because L_k is a lower triangular matrix, you do not need to go for any elimination process. You can simply use the forward substitution to get the vector l provided the matrix L_k is invertible. How do we know that the matrix L_k is invertible? Well, you can see that from the way we have constructed L_k. We have constructed L_k such that $A_k = L_k L_k^T$.

Now you take the determinant on both sides you have $det(A_k) = det(L_k)^2$. But this is positive because A is a positive definite matrix therefore this is surely non-zero that is what we can see from the way L_k was computed. Therefore L_k is an invertible matrix, so you will get a unique vector l such that $L_k l = a$ which is a known quantity. So, we obtained this vector l, now we need to only find this real number that is the only part left out for us.

(Refer Slide Time: 17:54)



Let us see how to find that. Again what we will do is, you take the last row of L and multiply it with the last column of L^T. That gives us the vector $l^T l + l^2_{(k+1)(k+1)}$ and that needs to be compared with the corresponding element of the matrix A and that is going to be $a_{(k+1)(k+1)}$. Now that gives us $l^2_{(k+1)(k+1)} = a_{(k+1)(k+1)} - l^T l$. Now from here you may take $l_{(k+1)(k+1)} = +\sqrt{a_{(k+1)(k+1)} - l^T l}$.

Of course you have to have plus or minus but we have already made our mind that we will take only the positive sign therefore I will take here only positive sign. But then the question is, is this real? In other words we have to first justify that is this greater than 0? That is the question because remember in order to say that the Cholesky's factorization exist we have to find a lower triangular matrix with all it is entries as real numbers.

Therefore this number should also be a real number but it need not be because if this number happens to be negative then you will have $l_{(k+1)(k+1)}$ as a imaginary number. Therefore you have to justify this, how will you justify it? Well, again go back to the form that we are writing, you take $det(A) = det(L)det(L^T)$ which is nothing but $(det\ (L))^2$.

What is $det(L)$? $det(L) = det(L_k)l_{(k+1)(k+1)}$ and that square means, this is square and square. Now determinant of A is nothing but the product of all it is Eigen values. Therefore you can say that the product of all the Eigen values of A is equal to, this is a positive number because we already know that L_k exists it means all it is entries are real that is already assumed.

Therefore this is a positive number and this is what we do not know whether it is positive or not. Now you can see that all the Eigen values of A are positive, why? Because A is a positive definite matrix, so we have stated one equivalent property of positive definite matrix in the last lemma that all it is eigenvalues are positive. Therefore the left hand side is the product of positive numbers therefore it is positive.

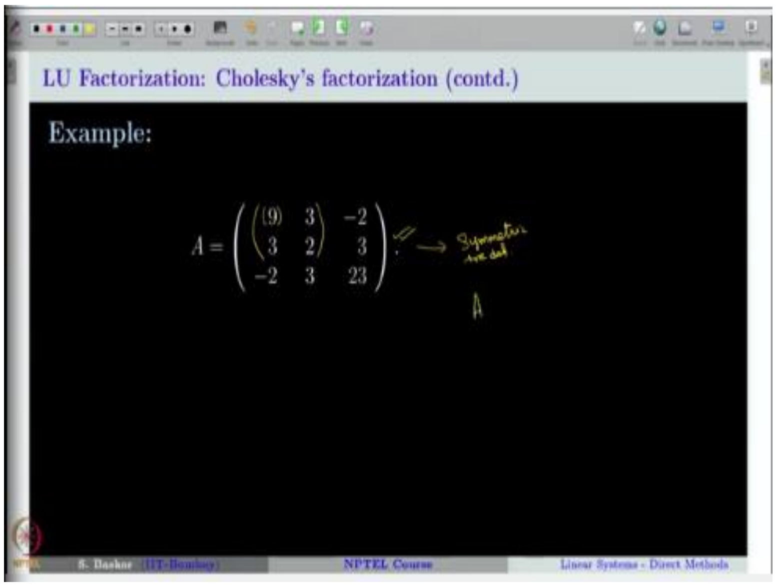
That shows that $l^2_{(k+1)(k+1)}$ is a positive number, remember just because you are squaring this does not mean it is positive. Say for instance if $l_{(k+1)(k+1)} = i$ then it is square is -1, therefore you just cannot directly say that this is a positive number. You have to justify this because we have not yet proved the existence of Cholesky's factorization. In fact that is what we are trying to justify that this is positive and that comes from this representation.

Therefore we have proved that this is positive and that implies that $l_{(k+1)(k+1)}$ is a real number and that proves the Cholesky's factorization exists. And also you can see the way we have constructed that the Cholesky's factorization is unique, why? Because from the induction hypothesis L_k is unique and then this system has a unique solution l and then finally this also is a unique representation.

Therefore with all this we can see that our Cholesky's factorization is unique, provided that all the diagonal elements sign are fixed as a unique sign either it should be positive or negative. You can see that at every stage of induction you are taking the square root for getting the diagonal element. There you have 2 choices; you may go for a plus sign or minus sign, so we made our mind to take only the plus sign in all the steps.

In that way we have a unique factorization that you have to keep in mind, there is nothing wrong in taking minus in all the steps also that gives a different Cholesky's factorization. Therefore as such Cholesky's factorization is not unique but if you fix the diagonal element's sign then it is unique.

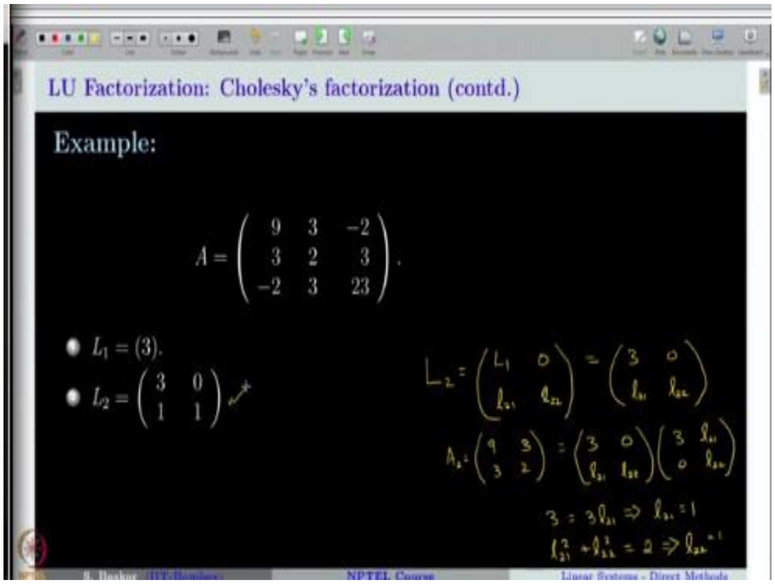
(Refer Slide Time: 24:10)



Let us take an example. Let us take the matrix A as given like this, you can observe directly that it is a symmetric matrix, also you can see that it is positive definite, how will you do that? One easy way out is to check all it is principal minors are positive. The principal minors of order 1 is 9 for this matrix, that is positive, the principal minor of order 2 is 9×2 - 9, that is again 9, that is also positive and you can also see the determinant of A is positive.

Therefore this is a symmetric and positive definite matrix. Therefore we can find a unique Cholesky's factorization with all the diagonal elements being positive, let us see how to compute that. There are many ways to compute but we will just follow the construction procedure we adopted in the previous theorems proof and try to construct the Cholesky's factorization for this matrix. For that we have to first compute L_1, what is A_1?

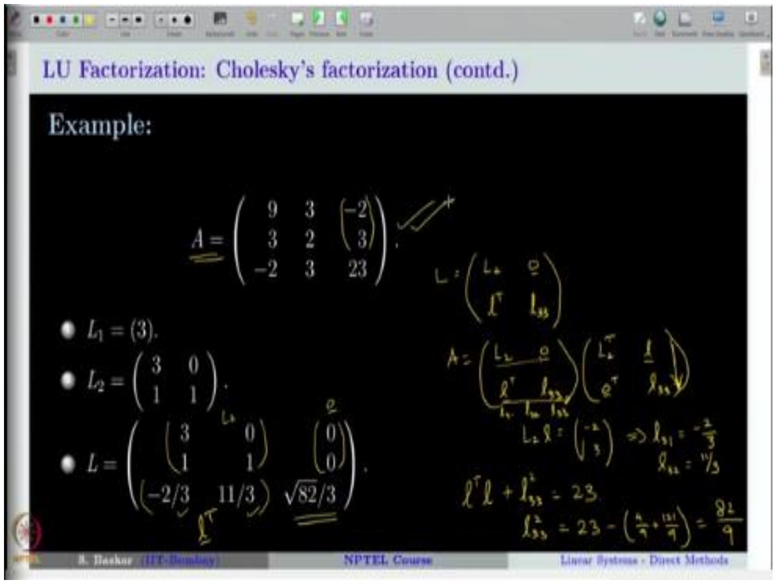
(Refer Slide Time: 25:42)



A_1 is nothing but 9 therefore L_1 can be taken as 3 or -3, again I am emphasizing we will always take the positive sign therefore we are taking L_1 as 3. So, with that we will go to find what is L_2? For that we will write $L_2 = \begin{pmatrix} L_1 & 0 \\ l_{21} & l_{22} \end{pmatrix}$, L_1 is 3 therefore it is $\begin{pmatrix} 3 & 0 \\ l_{21} & l_{22} \end{pmatrix}$, we have to find what is l_{21} and l_{22}? Such that A_2 which is nothing but $\begin{pmatrix} 9 & 3 \\ 3 & 2 \end{pmatrix}$ = $\begin{pmatrix} 3 & 0 \\ l_{21} & l_{22} \end{pmatrix} \begin{pmatrix} 3 & l_{11} \\ 0 & l_{22} \end{pmatrix}$.

And that is going to be equal to $3l_{21}$ = 3, that implies l_21} $= 1$ similarly $l_{21}^2 + l_{22}^2 = 2$ that will again imply that this is going to be $l_{22}^2 = 1$ or $l_{22} = 1$. Again remember whenever there is a diagonal element we will always get it in terms of the square of that element and then when you take the square root we will always take the positive sign and that gives us $L_2 = \begin{pmatrix} 3 & 0 \\ 1 & 1 \end{pmatrix}$.

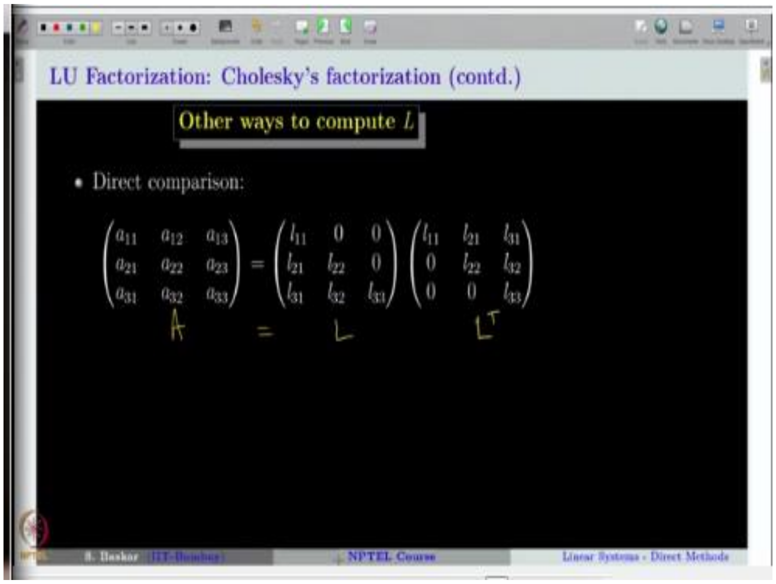
(Refer Slide Time: 27:42)



Let us now compute L_3 which is also the required Cholesky's factorization. For that we will take $L = \begin{pmatrix} L_2 & 0 \\ l^T & l_{33} \end{pmatrix}$. That gives us $A = \begin{pmatrix} L_2 & 0 \\ l^T & l_{33} \end{pmatrix}$ and then it is transpose, that is $\begin{pmatrix} L_2^T & l \\ 0^T & l_{33} \end{pmatrix}$. When you take the first block row with the last column here we get $L_2 l$ is equal to this vector$\begin{pmatrix} -2 \\ 3 \end{pmatrix}$, that gives us l_{31}. So, remember this is l_{31}, l_{32} and l_{33}. $l_{31} = -2/3$ and $l_{32} = 11/3$, so that is what we have here.

This is our L_2 and this is the 0 vector and we have l^T here, now we have to find the diagonal element l_{33}. For that we will multiply this with the last column of l^T that gives us $l^T l + l_{33}^2 = 23$. That gives us $l_{33}^2 = 23 - (4/9 + 121/9)$ and that 9 into 23 minus this is going to be 82 / 9. And therefore if you take the square root on both sides and choose the positive sign we will have $l_{33} = \frac{\sqrt{82}}{3}$. And that gives us the Cholesky's factorization for the matrix *A* given like this.

(Refer Slide Time: 30:11)

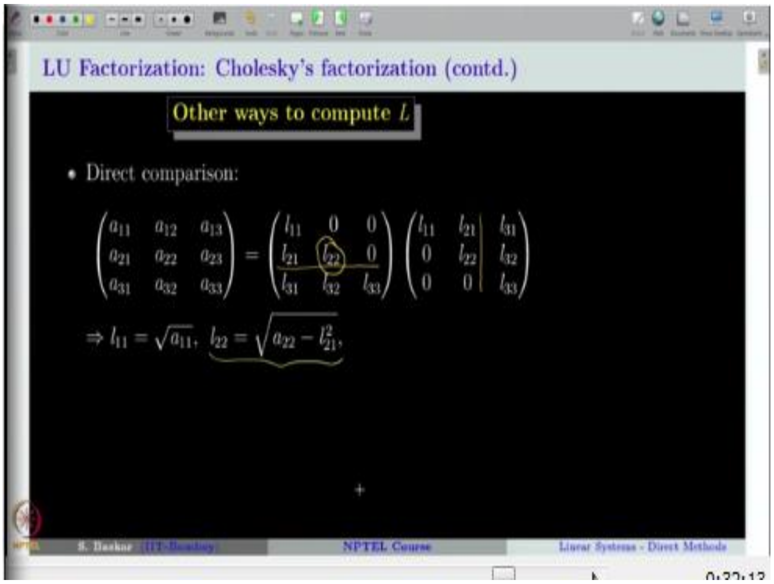


This is not the only way to compute Cholesky's factorization. Recall, in the Doolittle case we computed the Doolittle factorization by direct comparison. What we did? We wrote $A = LL^T$, of course in Doolittle all the diagonal elements of L are 1. So, you have to write such an L and then find the other entries of L as well as U. So, in fact this was U there in the Doolittle factorization but here in Cholesky's factorization we have to take it as L^T itself.

Now just like how we did with Doolittle factorization, what we did? The right hand side product of 2 matrices we just multiplied them and then compare the elements of the right hand side matrix with the corresponding elements of the left hand side matrix and we got all the elements of L and U there. The same idea can be followed in constructing Cholesky's factorization also, you need not go step by step as we did in the last computation.

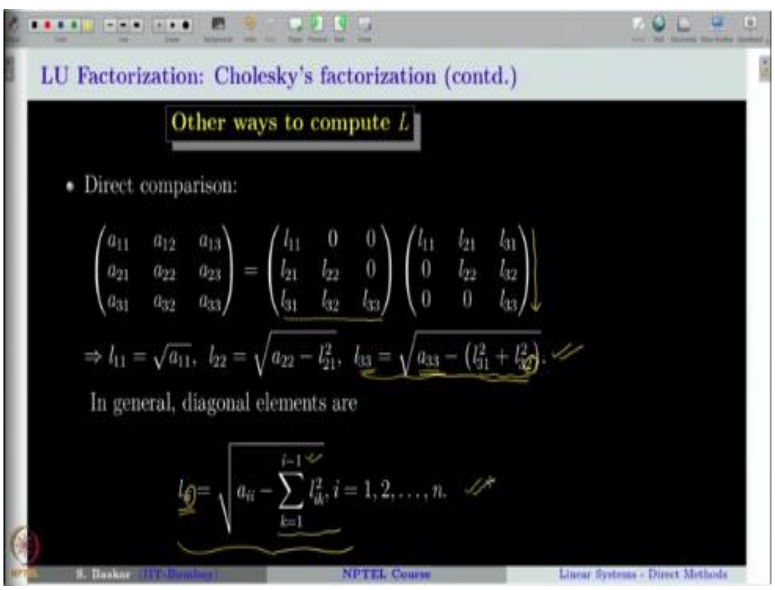
This is to make the proof of the theorem more rigorous; we used it in the form of an induction. Otherwise you can also go for a direct comparison calculation which will also lead to a very efficient algorithm.

(Refer Slide Time: 31:45)



Let us see how these expressions look like. Well, l_{11} can be directly obtained as $\sqrt{a_{11}}$, remember again we are fixing our sign as plus, that is why we got it otherwise you can also take $-\sqrt{a_{11}}$. Now l_{22} is obtained by multiplying this row of L with this column of L^T and that gives l_{22} as this.

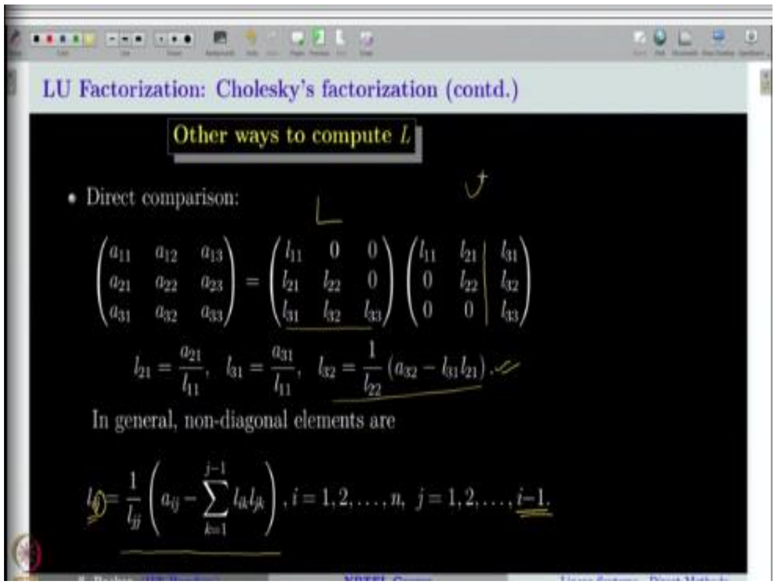
(Refer Slide Time: 32:14)



Similarly to get l_{33} you can multiply this with this and that gives you l_{33}. All the diagonal elements are therefore obtained. In general if your matrix is a $n \times n$ matrix, you can just look at these expressions and try to generalize how this expression will look like for a $n \times n$ matrix. That will be given as l_{ii} = root a_{ii} minus you have the sum starting from $k = 1$ and goes up to $i - 1$.

So, here it is 3 therefore it goes up to 2, similarly if you are computing the diagonal element at the ith row it goes up to $i-1$. Because i is already there that is what you are computing, $i+1$ onwards the entries are 0 because it is a lower triangular matrix. Therefore this will go only up to $i-1$. So, therefore the diagonal elements are given by this expression, how the non diagonal elements are obtained?

(Refer Slide Time: 33:26)



Again let us see, to get this element you can just multiply this row with the first column of L^T and that gives you l_{21}. l_{31} we have to find. For that you take this and again you do multiplication with the first column of L^T this gives you l_{31}. And how will you get l_{32}? Well, you multiply this with the second column and that gives you l_{32} again you can just observe the expression and try to generalize it for any $n \times n$ symmetric positive definite matrix A and that gives you this expression.

Remember, this has to go for each column other than the diagonal element. All the elements after diagonal elements are also not computed. Therefore j should go from 1, 2 up to $i-1$ and this has to be done for all the rows therefore the row index i should go from 1 to n. So, this way also you can compute Cholesky's factorization. So, since we are working with symmetric matrix.

Cholesky's factorization is more efficient than Gaussian elimination and the Doolittle or Crout factorization. Because we are making use of the property that A is a symmetric matrix therefore you are only computing L, you are not computing U explicitly. In that way you gain lot of computational time. Now how to compare 2 methods in terms of their computational time? Well, that can be done by counting the number of arithmetic operations involved in these methods.

In the next class we will compare Cholesky's factorization with Gaussian elimination method and see which method is more efficient in terms of the computational time. Well, that will be in terms of how many arithmetic operations are involved in them, this we will do in the next class. Thank you for your attention.