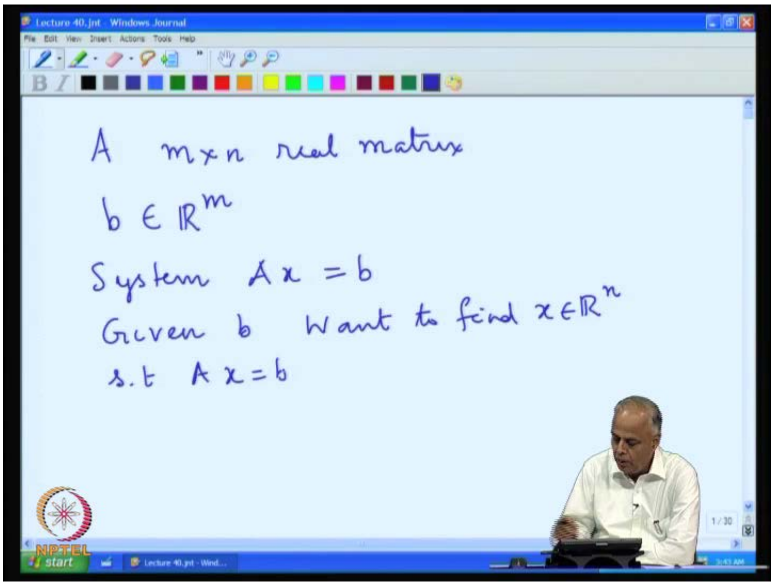
Advanced Matrix Theory and Linear Algebra for Engineers
Prof. R. Vittal Rao
Centre for Electronics Design and Technology
Indian Institute of Science, Bangalore

Module No. # 12
Lecture No. # 40
Epilogue

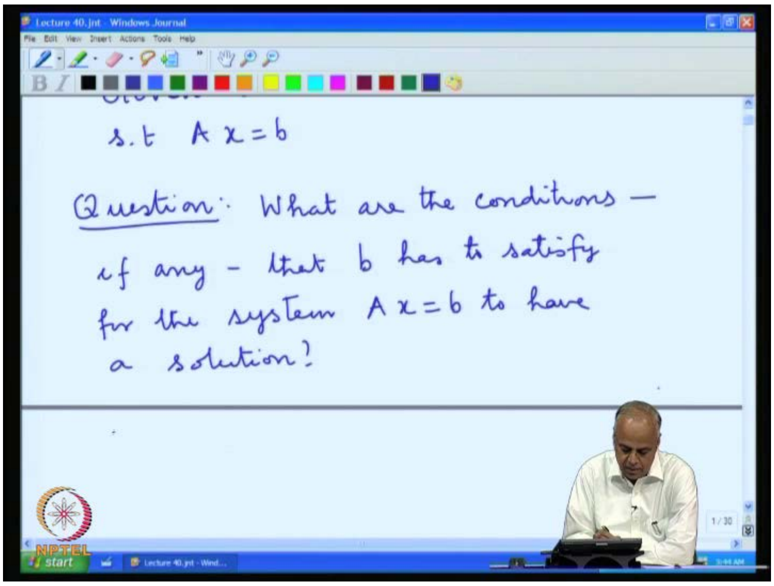
We have seen the choice of the basis for the various subspaces that we had considered connected with the matrix. We shall now go back to the various fundamental questions that we raised in the first two lectures and setting our goal for the course as finding the answers to these questions. Now, let us go back to these questions and see whether we have found all the answers.

(Refer Slide Time: 00:49)



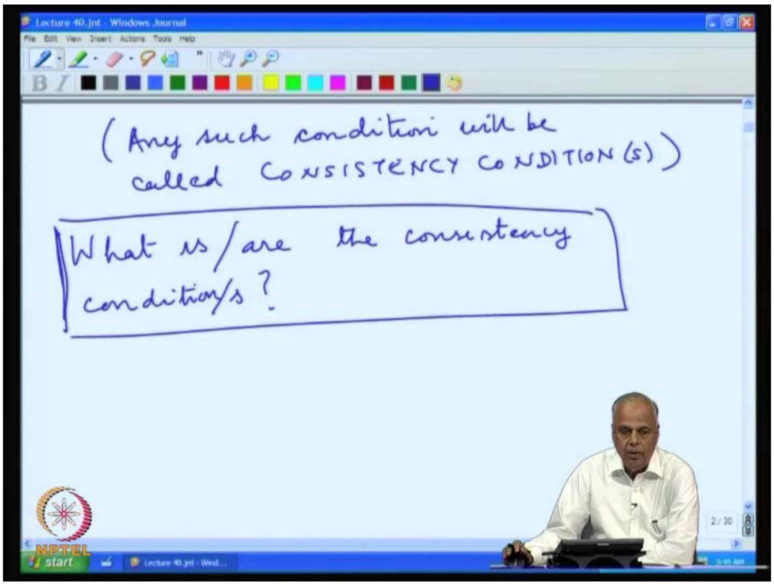
The first question that we looked at was suppose, a is an m by n real matrix. And b belongs to R m, and we look at the system A x equal to b. So, given b we want to find x n R n such that A x equal to b.

(Refer Slide Time: 00:49)



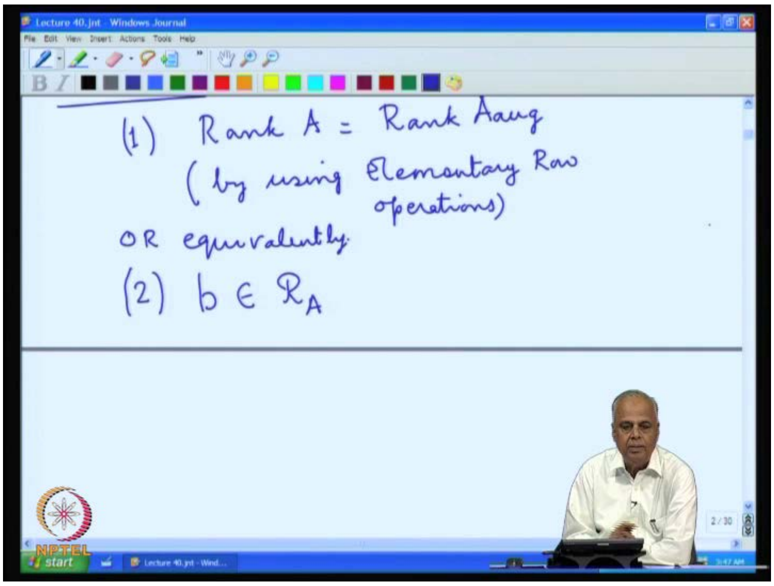
So, the first question was under what conditions that this possible. So, the question that we raised was what are or what is the conditions, if any that b has to satisfy for the system A x equal to b to have a solution.

(Refer Slide Time: 02:11)



And any such condition will be called consistency condition. This is what we called any such condition which, guarantees the existence of a solution is called consistency condition. So, what we are asking is what is or are the consistency conditions. So, the question is what is or are the consistency conditions? Condition or conditions whichever, way depending on how many condition we have to satisfy. So, this was the first fundamental question that we raised now during the course of our discussions we found the answers to these questions by interpreting it in various formats.

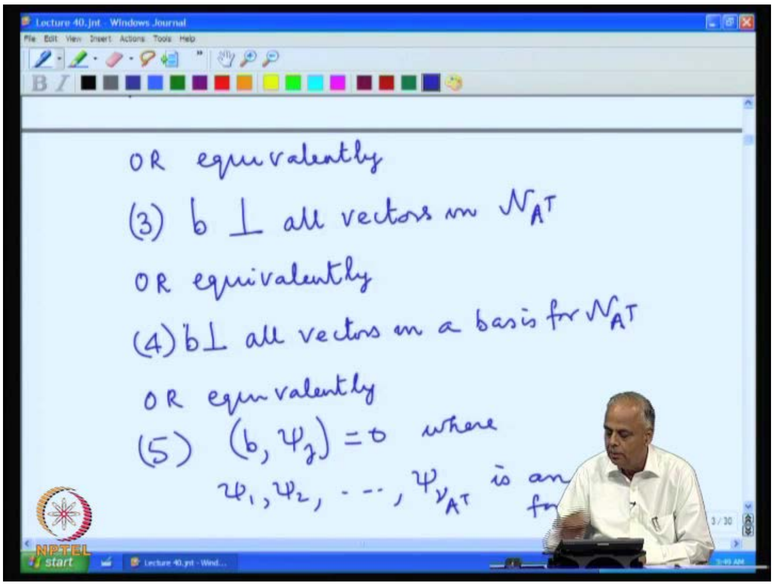
(Refer Slide Time: 03:17)



So, the answer to this question that we found was in many forms. The first form was that the rank of the matrix a must be equal to the rank of the augmented matrix, recall that the augmented matrix is the matrix obtained from a by appending as a last column the vector b. We got this conditions, when using by using elementary row operations this was the first way, first time we have got the consistency conditions for the system to have a solution.

Then we looked at them in the view of the transformations. And we observed that, if A x equal to b has to have solution b must be of the form A x and such vectors are called range vectors. So, the other way is that b belongs to the range of a. So, this was one or look at another manner the same thing equivalently, can be stated as b belongs to the range of A.

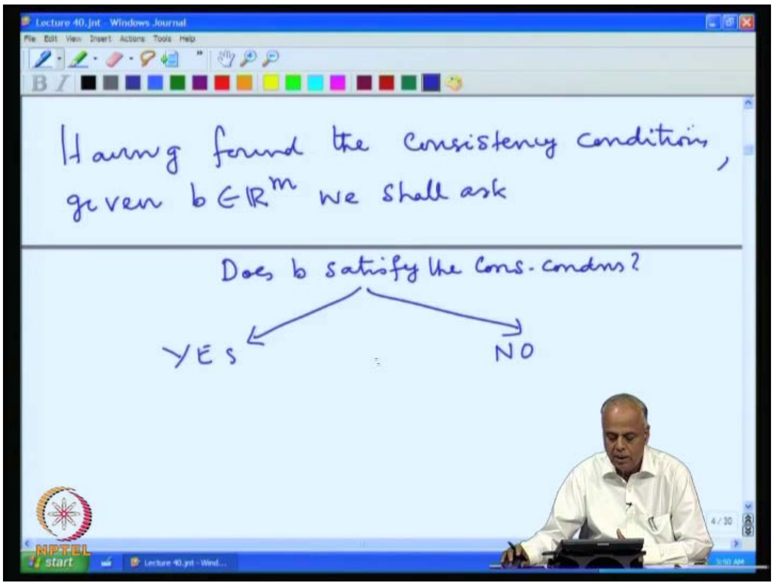
(Refer Slide Time: 04:47)



Or now we found that the range of A is the orthogonal compliment of the null space of A transpose. And therefore, b has to belong to the range of A simply means that b must be orthogonal to all the vectors in the null space of A transpose. So, the same can be written as b is orthogonal to all vectors in the rate, the null space of A transpose. Now, a vector is orthogonal to all the vectors in a subspace, if it is orthogonal to all the basis vectors and therefore, this can be stated equivalently as. B is orthogonal to all vectors in a basis or in A transpose, or in particular we can state this as by choosing the basis to be an ortho normal basis, which should be orthogonal to b psi j equal to 0 where, psi 1 psi 2 psi nu A transpose is an ortho normal basis for null space of A transpose.

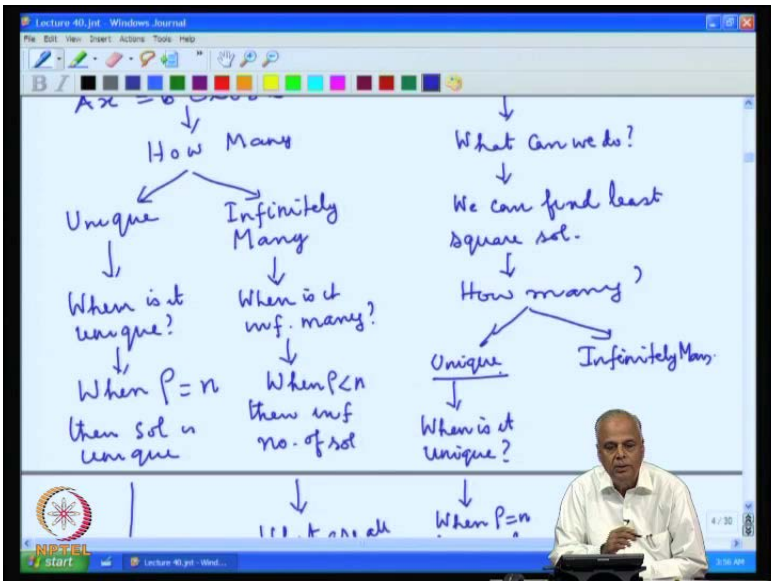
So, the answer to the first question what is or what are the consistency conditions? We found the answer in many formats all of them are equivalent and finally, we boil down to the format phi, which told us how to analyze these spaces.

(Refer Slide Time: 06:49)



Then having found the consistency conditions, then the first thing that we would do is given any b we shall ask does b satisfy the consistency conditions, that was our next question and we said we could get the answers to this in two forms either yes or no. Either b may satisfy the consistency conditions or not, if b satisfies the orthogonality to all the basis vectors then it satisfies the conditions, if it fails to be orthogonal to even one basis vector then it fails to satisfy the consistency conditions.

(Refer Slide Time: 06:49)



Now, in this case solution exists solution to system, A x equal to b exist then we ask the question how many we now know that we found the answer that either unique or infinitely many then we ask the ask the question when is it unique. And now we have the answer that when the rank of the matrix is equal to n the number of columns then solution is unique then finally, we ask the question in that case what is the solution? And now we know the answer the solution is given by our representation that we got ex or x, x is equal to summation j equal to rho 1 s j b u j V j, where u j V j s j are as we have found out the singular values the basis for the range of A transpose etcetera.

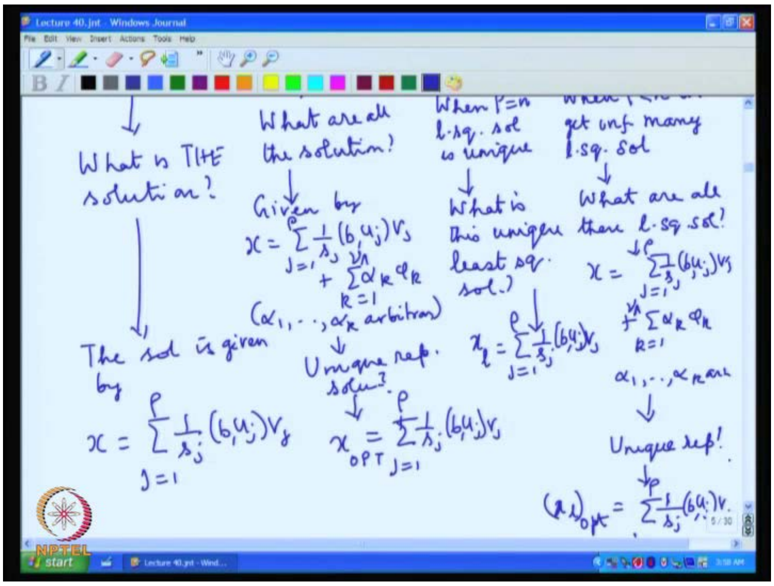
So, we have the answers for all the questions that we ask in this case the case where we have unique solution then similarly, we ask whether it is infinitely many when is it infinitely many. And uniqueness we got when rho equal to n. So, it should be the other case the rho cannot be greater than n. So, this can happen when rho is less than n then solution is then infinite number of solutions.

Then we asked what are all the solutions? And now we know the answer that they are all given by x equal to summation j equal to 1 to rho, 1 by s j b u j V j plus summation alpha k phi k, equal to 1 to nu a where alpha 1 alpha 2 alpha k arbitrary. Then we ask for unique representative solution is there anything like that and we found that that is given by x optimal, which is summation j equal to 1 to rho 1 by s j b u j V j. So, thus we have all the answers to all the questions in the case, b satisfies the consistency conditions. The next case that we have to deal with is what happens when the consistency conditions are? Not satisfied then we found.

So, we ask the question what can we do? First of all we get statement that there will be no solutions. So, first let is note that now there is no solution there is no solution. So, we said that if, there is no solution what can we do then? We observed that we can find least square solutions, which are very close to b and minimize the error we can get close to b as possible. And solution, which gives you the minimum error is called a least square solution. Then we, ask if we can find least square solution how many? Again we have the same situation as before either unique or infinitely many then we ask when is it unique.

Now, we know the answer that we get the unique least square solution when the rank is equal to n. So, when rank is equal to n least square solution is unique. And then once you know that you are in the case (( )) solution and that least square solution is unique we ask what is the least square solution? And we know the answers from our analysis that is given by x L equal to summation j equal to 1 to rho 1 by s j b u j V j. So, now we have all the answers in the case when we does not satisfy the consistency condition and we have unique solution the next. So, we have this situation covered the next series of questions that we raise where in the case of infinitely many solutions when.

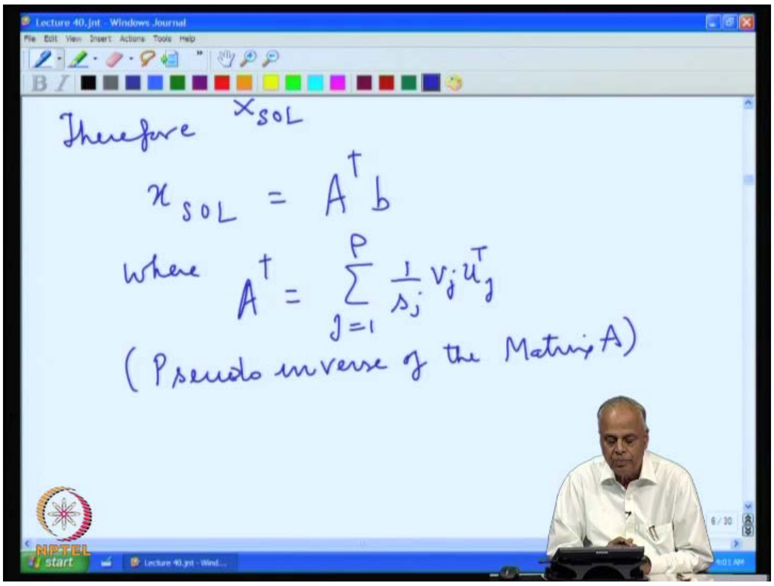
(Refer Slide Time: 06:49)



So, the question therefore, is when do we get infinitely many solutions. Then we know that when rho equal to n we get unique the other two cases are rho greater than n, and rho lesser than n, but the rho cannot be greater than n as we said before. So, we get when rho is less than n we get infinitely many least square solutions what are all the solutions? What are all these? Least square solutions that was the next question, that we asked. Now, we know how to get all these solutions they are given by x equal to summation j equal to 1 to rho 1 by s j b u j V j plus summation, k equal to 1 to nu a alpha k phi k where alpha 1 alpha 2 alpha k arbitrary.

So, then we ask what about a unique representative? And we had the answer we had x L optimal the least square solution optimal, which is given by j equal to 1 to rho 1 by s j b u j V j. So, thus we have the answers to the entire series of questions that we raised about a linear system of equations. We have the consistency situation the non consistency situation, consistency unique situation, consistency non unique, inconsistent unique least square, inconsistent infinitely least square in both the cases consistent and inconsistent when we have infinite solution, we have a corresponding optimal solution. And thus we have.

(Refer Slide Time: 17:06)

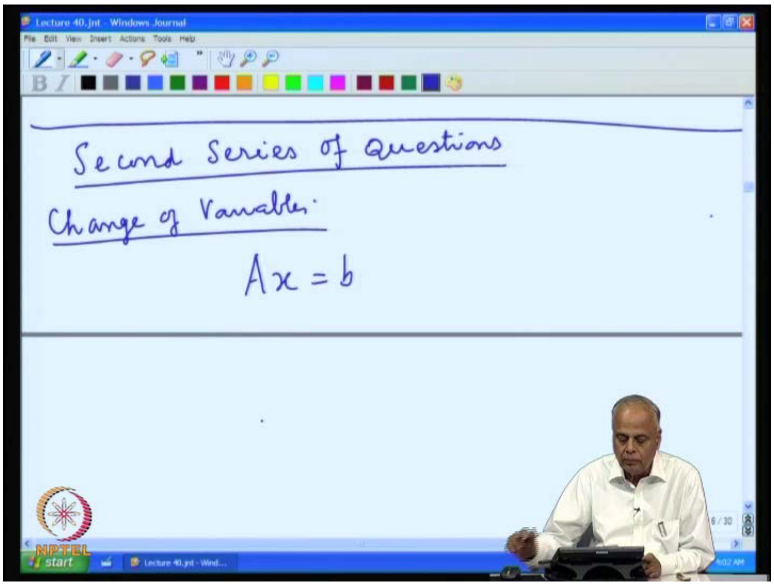


Therefore finally, our answer in all these cases, we call the answer we got this one this one we call this as X sol in all these cases, the final version of the solution the best possible solution, that you can get for the system a x equal to b when I say best possible I mean it is either a solution. And, if you cannot find a solution, least square solution and, if there are many solutions it has the least length. So, least error solution is called x. So, and we have X sol is equal to as dagger b where A dagger was the matrix summation j equal to rho 1 by s j V j u j transpose and this we call as pseudo in.

You call the pseudo inverse of the matrix this was the first set of basic questions, that we raised about a matrix A it is about the non homogenous system of equations consistency conditions, about the existence of solutions uniqueness of solutions.

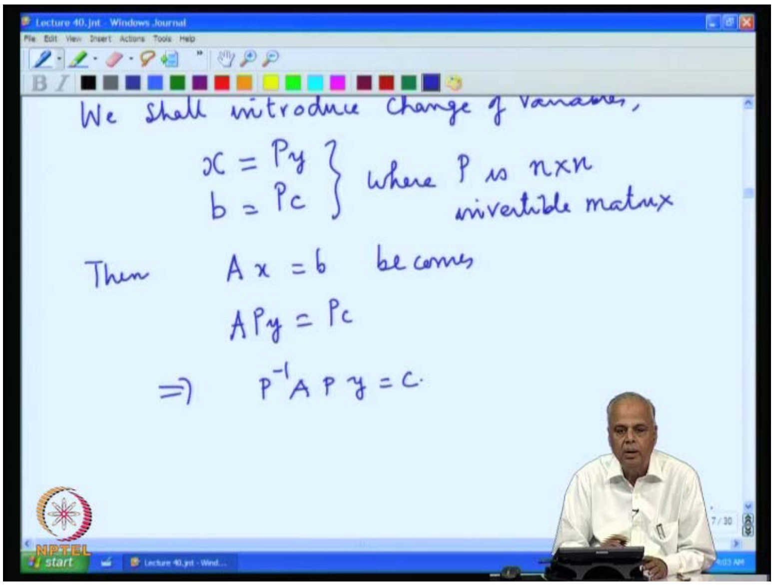
When it is non existence what can we do? And whenever we do that approximation will it be unique infinitely many. And what is the most representative solution? So, we have the answer for all these questions and we put them all in a single package. Whatever case you are in just compute, A dagger b and that will be your final answer whichever of the four cases, we discussed in all these cases the final answer is going to be A dagger b where A dagger is this. Of course the summation will depend on what the rank is, if rho equal to n A dagger will have n terms if rho is less than n A dagger will have less than terms, that was as I said the first series of questions.

(Refer Slide Time: 19:32)



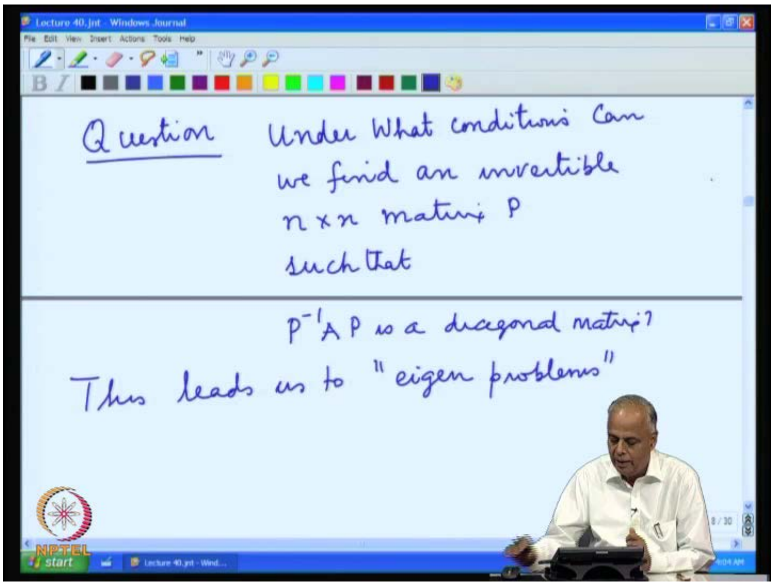
Then we raised a second series of questions. Now, that we have seen that during the course we have found that answers to all the questions, in the first series of questions the second series of questions, since that we raised were the following. This involved the change of variables for example, We had A x equal to b.

(Refer Slide Time: 20:10)



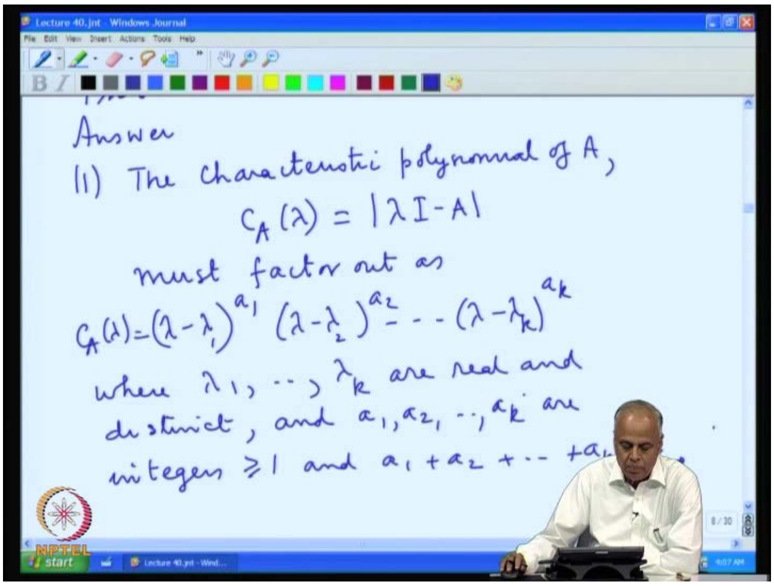
We said we shall introduce change of variables, where we call x as some P y b as some P c, where P is n by m invertible matrix we want invertible, because from the x variable to y variable and x variable to y variable you must be able to translate that problem. Then A x equal to b becomes A P y equal to p c. And therefore, p inverse A P y becomes c. So, if we can make P inverse A P as diagonal then we get a diagonal system and we can easily solve it.

(Refer Slide Time: 21:23)



So, the question was that we raised was under what conditions, can we find a invertible under n by n matrix P such that P inverse A P is a diagonal matrix then this lead us this leads us to the Eigen problems, we found this is connected with the notion of Eigen values and Eigen vectors.

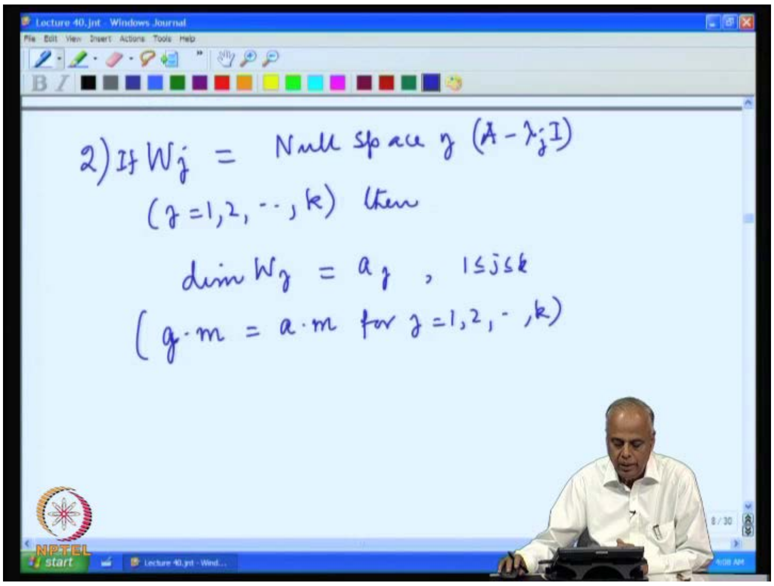
(Refer Slide Time: 22:40)



So, what was the answer that we got? So, we first said we introduce a notion of characteristic polynomial. So, the answer characteristic polynomial of a, which is defined as C A lambda equal to the determinant of lambda I minus A must factor out as lambda minus lambda 1 into a 1, lambda minus lambda 1 to the power of a 2 lambda minus lambda k to the power a k, Where lambda 1 lambda 2 lambda k are real and distinct we call that we are asking for a matrix since, we are having real we always think of real matrix p lambda 1 lambda 2 lambda k we are seeking as real and distinct.

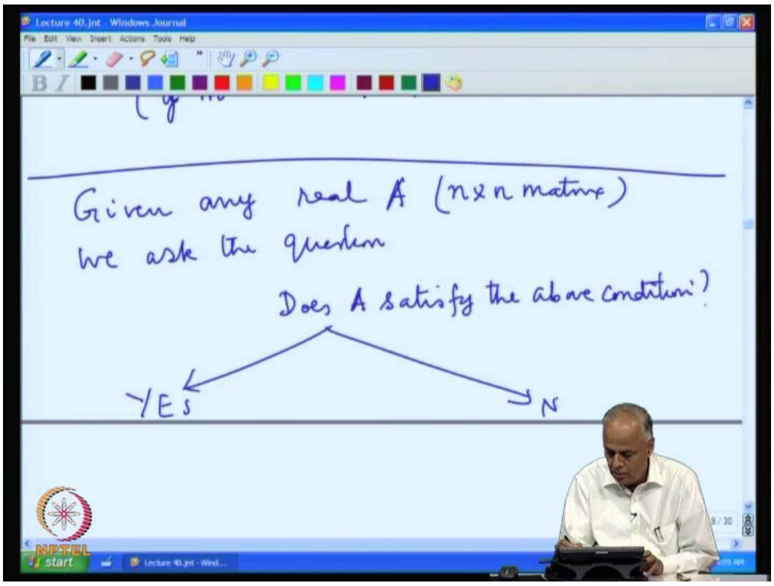
This k this distinct lambda 1 lambda 2 lambda k. And a 1 a 2 a k are integers greater than or equal to 1 such that a 1 plus a 2 plus a k equal to. So, the first requirement is that the characteristic polynomial of a factors out like this then, we call lambda 1 lambda 2 lambda k as the distinct Eigen values of the matrix a and a 1 a 2 a k were referred to as the algebraic multiplicities of this Eigen values then that was the first condition once that is then. So, the conditions that we are asking for under what conditions consists of many parts the first part is the characteristic polynomial must factor out as pointed above.

(Refer Slide Time: 25:03)



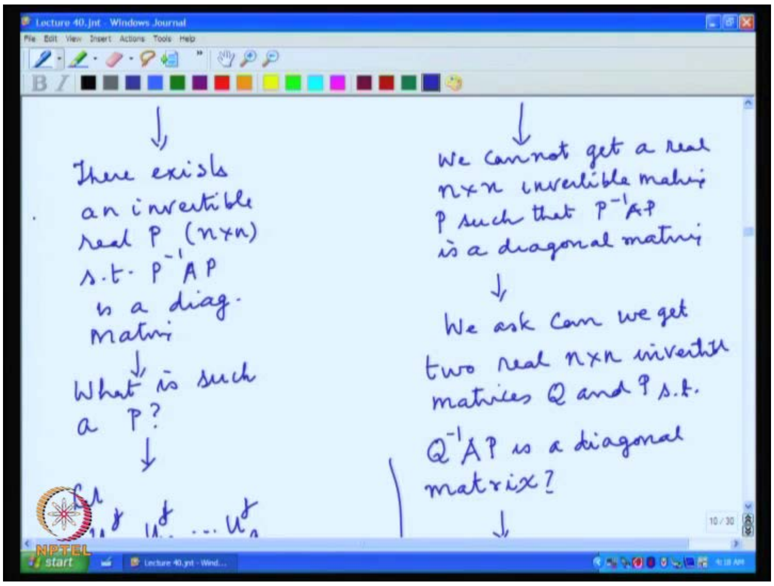
Then, if you define W j to be the null space of A minus lambda j I for j equal to 1 2 k these we call as the Eigen spaces corresponding to the Eigen value lambda j, if then dimension of W j, which we call as the geometric multiplicity this is what is known as? Algebraic multiplicity is equal to algebraic multiplicity this must happen for 1 less than j less than k or j equal to 1 2 k for every Eigen value the geometric multiplicity must be equal to the algebraic multiplicity. These are the two conditions that has to satisfy in order that the matrix p exist.

(Refer Slide Time: 26:18)



So, therefore, given any real p the real A n by n matrix we ask the question does A satisfy the above conditions. And again the answers can be yes or no.

(Refer Slide Time: 26:18)



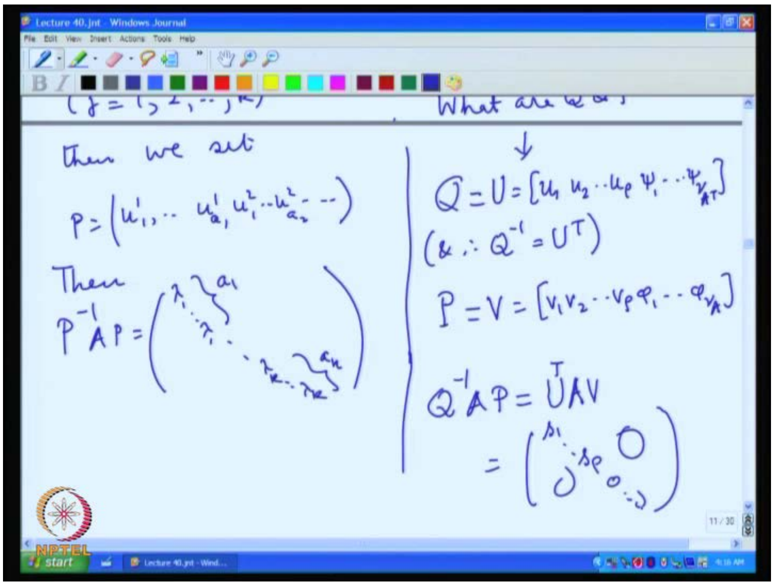
In the case of yes, we asked therefore, since we satisfied condition there exists a invertible real P n by n matrix such that P inverse A P is a diagonal matrix. So, our question was under what conditions there exists? A P is a diagonal matrix we have got these two conditions. Now, we ask whether matrix a satisfies these two conditions and the answer is yes therefore, there must be A P what is that p? What we asked what is such A P? And answer s was as follows let u 1 1 u 1 2 etcetera u 1 a 1 be a basis for w one. So, in general let u j 1 u j 2 u j a j be a basis for w j.

See a dimension of by a condition two we said that the condition is satisfied by condition two the dimension of W j is a j and therefore, a basis for W j will have a j vectors. Now, if the u a j vectors are denoted by a j 1 a j 2 u j a j and this we do for each j then we set, P equal to the matrix whose first column is u 1 1 and go on and u 1 a 1 then u 2 one etcetera u 2 a 2 etcetera. In other words P is obtained by putting these basis vectors along the columns. And p inverse A P will be a diagonal matrix where lambda 1 will occur a 1 times and so on lambda k will occur a k times.

So, what is the case? When p satisfies a satisfies the condition for the existence of A P. Now, what do you do when? It is a (( )) there are two alternatives we say that one lays to the Jordan canonical form, but we looked at another alternative and this we took this alternative in this of course,, we did not take the alternative course, which reached to the Jordan canonical form. On the contrary we took an alternative which lead to the singular value decomposition, because it also allows us to look at a rectangular matrix. So, what we did was the following instead of? Remember this question arose, because of the requisition for the change of variables, if you look at the change of variables we use the same matrix p for changing the variable x as well as changing the variable b.

Can you not use two different change of variables like x equal to p y and b equal p c or x equal to Q y and V is equal to r c some change of variables. So, we asked, if you change that then instead of asking p inverse A P to be diagonal we will ask Q inverse A P to be diagonal. So, the question is modified as can we cannot, if their conditions are not satisfied we cannot get a real n by n invertible matrix P such that, P inverse A P is a diagonal matrix. So, what we do? So, we asked can we get two real n by n invertible matrices Q and p Q is used to transform the variable b and p will be used to transform the variable x such that Q inverse A P is a diagonal matrix and the answer we got was yes.

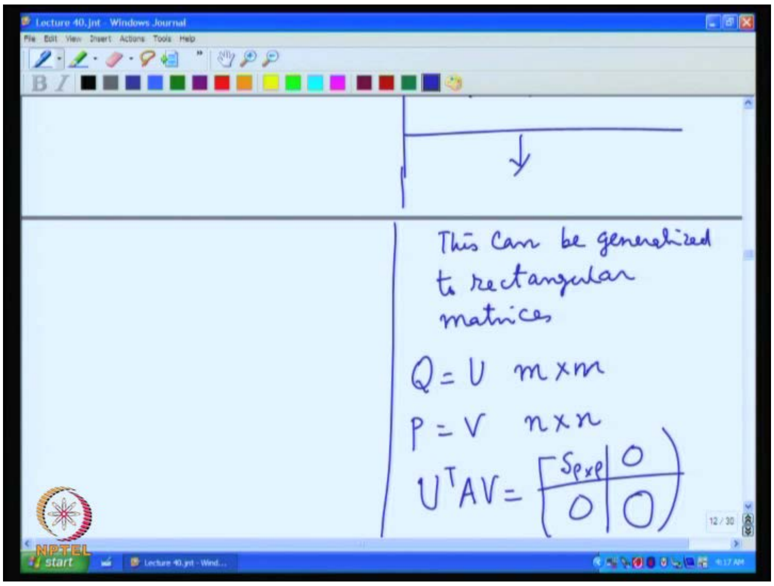
(Refer Slide Time: 26:18)



And therefore, what are Q and p? The answer we had was you choose Q to be the matrix U, which was u 1 u 2 u rho psi 1 psi 2 u A transpose remember, he u 1 u 2 u rho formed a basis for range of A psi 1 psi 1 psi u A transpose formed the basis for, the range of A transpose since we are in the square matrix state n equal n. So, rho plus U A transpose be n. And therefore, given n by n matrix and not only that this is an orthogonal matrix and therefore, and Q inverse is u transpose because u was an orthogonal matrix and we choose P to be the matrix V which consists of this V 1 V 2 V rho and then phi 1 phi 2 phi nu a.

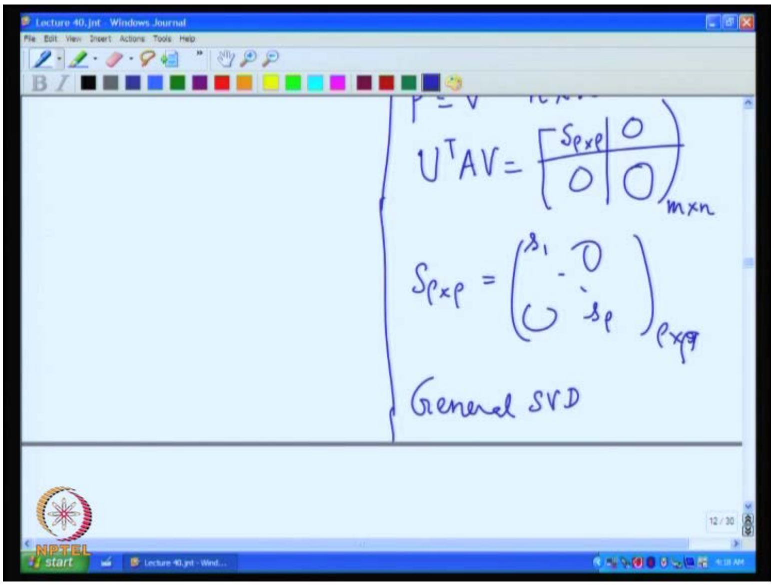
This is again an n by n matrix and then we have Q inverse A P is the same as u a V u transpose A V. And this we got as the singular value see s 1 s 2 s rho 0 0 0 and this was the singular value decomposition this was the SVD.

(Refer Slide Time: 26:18)



Then we ask the question that? This can be generalized our analysis shows that, this can be generalized to rectangular matrices and what will be our Q now Q will be our U, but now an m by m matrix and P will be our, which will be an n by n matrix. And U transpose A V will be our no longer a diagonal matrix, which have the s rho cross rho and the 0 matrix and all other. There will be a leading diagonal block of s rho cross rho, which is a diagonal matrix with the singular values along the diagonal.

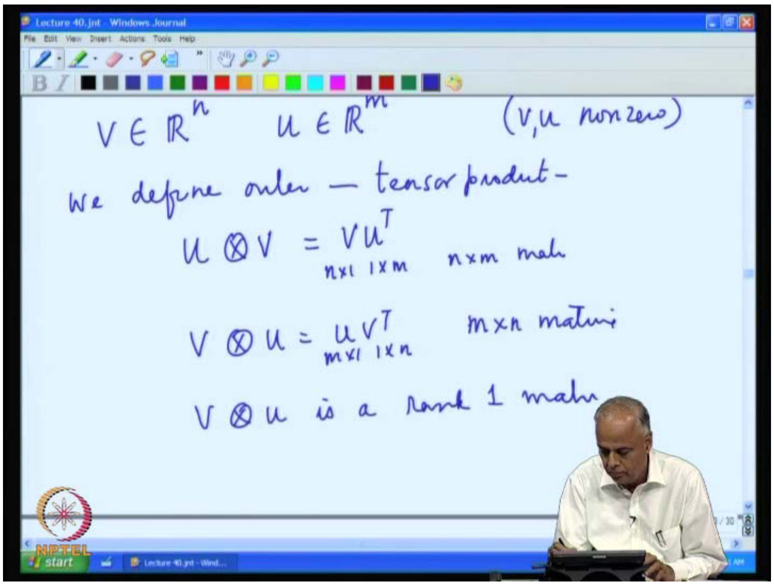
(Refer Slide Time: 26:18)



So, s rho cross rho is simply a rho by rho matrix, which has s 1 s 2 s rho or all others are 0. So, you have just put a diagonal block consisting of s 1 s 2 s rho and then adjust the other 0s. So, that you get an n by n matrix. So, this is the general SVD. So, once again we have now found the answers to all our questions in the second series about diagonalizibility, under what condition is diagonalizable? We got the answer as eventually the algebraic multiplicity must be equal to geometric multiplicity before that the characteristic polynomial must factor out and so on.

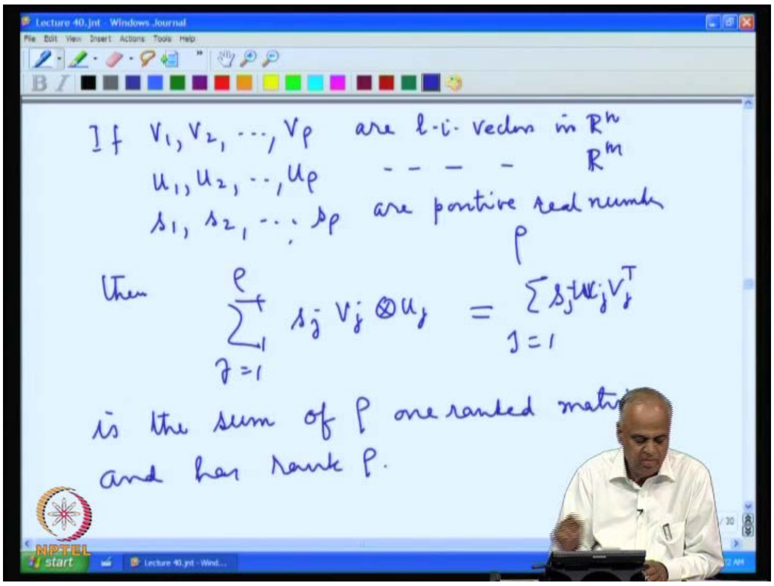
And when it was satisfied we knew how to find the diagonalising matrix p and when it was not satisfied, we can split it in to two transformations, Q and P such that Q inverse A P is a diagonal matrix. And the Q and P were found in this format. And then we said that this could be also generalized to rectangular matrices. And we can get the general singular value decomposition given for rectangular matrix. So, that we have the complete set of answers for the entire set of questions in the second series of questions.

(Refer Slide Time: 37:05)



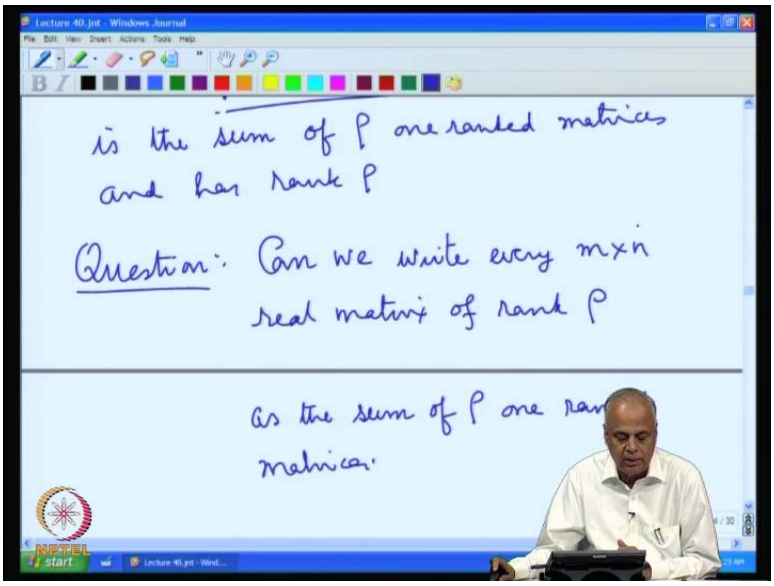
The third series of questions, that we raised was the following suppose I have a vector V in R n. And A vector u in R m both non 0 vectors V u non 0. So, suppose we have two non 0 vectors then we define the outer product or the tensor product sometimes called the tensor product. We shall define it as follows u tensor to be the matrix V u transpose. So, in that case what will we get? We get u is V is n by 1 and u is m by 1. So, we get an n by m matrix, if we take V tensor u then it will be u V transpose and this will be m by one and this will be one by n. So, we get an m by n matrix. So, if we take a matrix u, which is in R m a matrix V in R n and take u V transpose it will be a n by n matrix, then V tensor u is a rank 1 matrix we then said we can take such rank one matrices and superpose them.

(Refer Slide Time: 38:53)



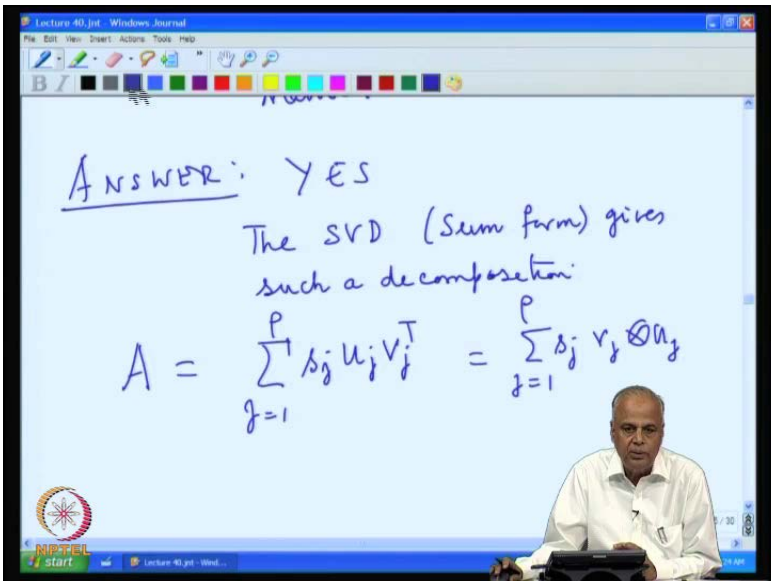
So, the next idea we said was if V 1 V 2 V rho are linearly independent vectors in R n. U 1 u 2 u rho are linearly independent vectors in R m. And then s 1 s 2 s rho are positive scalars positive real numbers then, if we superpose all this s j V j tensor u j, which is the same as summation j equal to 1 to rho s j u j V j transpose is the sum of rho 1 rank matrices is and as rank rho.

(Refer Slide Time: 40:10)



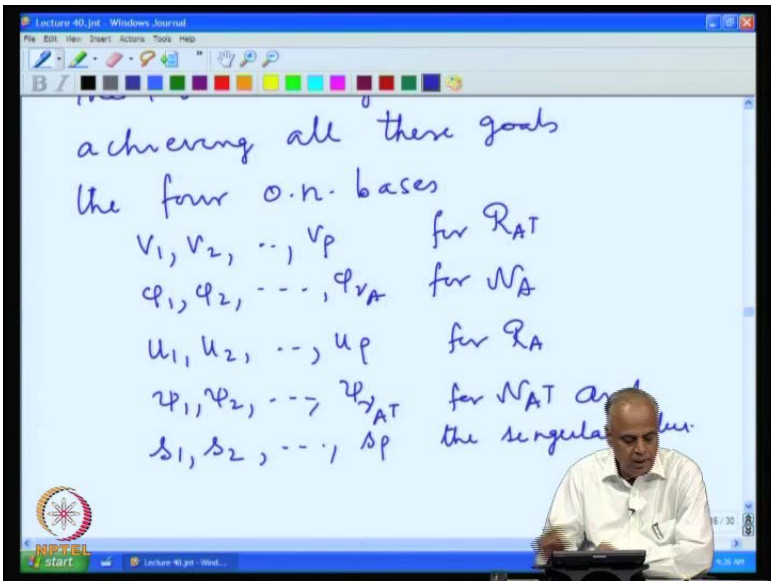
So, the question that we asked was can we write every m by n real matrix of rank rho as the sum of rho 1 ranked matrices? See the construction that, we have here shows that we can generate a lot of rho ranked matrices, by putting together these one ranked matrices. The question was whether we could take and exhaust all the rho ranked matrices that is whether all rho ranked matrices can be generated this way, which means can we write every rho ranked n by m matrix as the sum of rho 1 ranked matrices. This is the fundamental question of decomposing a rho ranked matrices matrix into rho 1 matrices and the answer we got was yes.

(Refer Slide Time: 41:29)



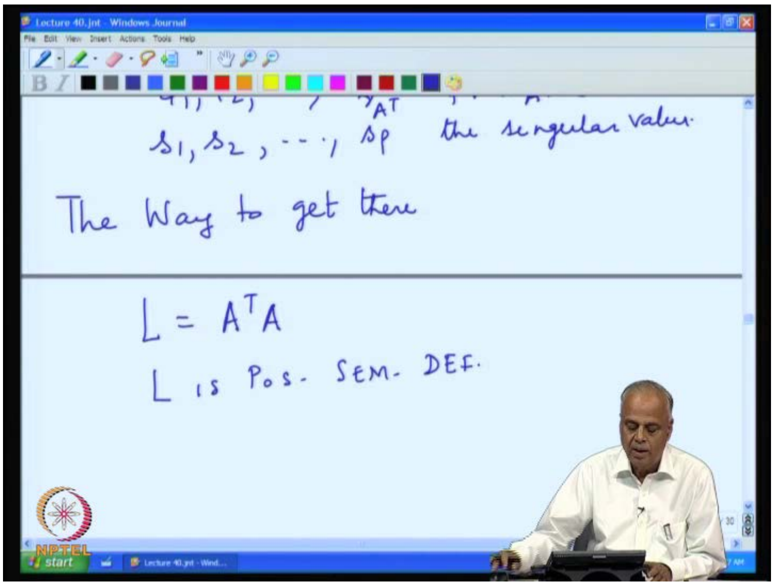
And the answer precisely the yes the SVD we got the SVD into formats and the sum form of the SVD gives such a decomposition, out of the decomposition A equal to summation j equal to one to rho s j u j V j transpose, which can also be in our tensor notation s j V j tensor u j. So, therefore, we have the answer to the decomposition also. So, thus we have answers for all the questions we raised in the first two lectures and set them as our goal for this course and since we have now all the answers we have achieved our goal.

(Refer Slide Time: 42:36)



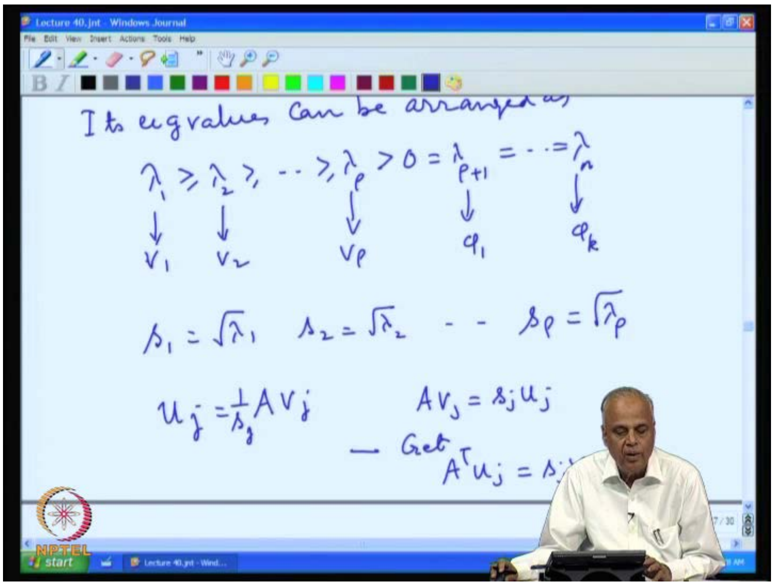
All our goal was achieved the main ingredient for gradients for achieving all these goals or all these answers are what? Everywhere the same u j V j will appear the four ortho normal basis write then down V 1 V 2 V rho for the range of A transpose phi one phi two phi nu a for the null space of A u 1 u 2 u rho for the range of A psi one psi two psi nu A transpose for the null space of A transpose and s 1 s 2 together that these four basis, we have s 1 s 2 s rho the singular values. So, these are the main ingredients for finding all the answers therefore, given a matrix these are the most important things that we have to compute with such a matrix.

(Refer Slide Time: 44:18)



Where do we get all these from where should we look. So, the way to get this let us recall was the following, we first construct L equal to A transpose A and then L is positive semi definite.

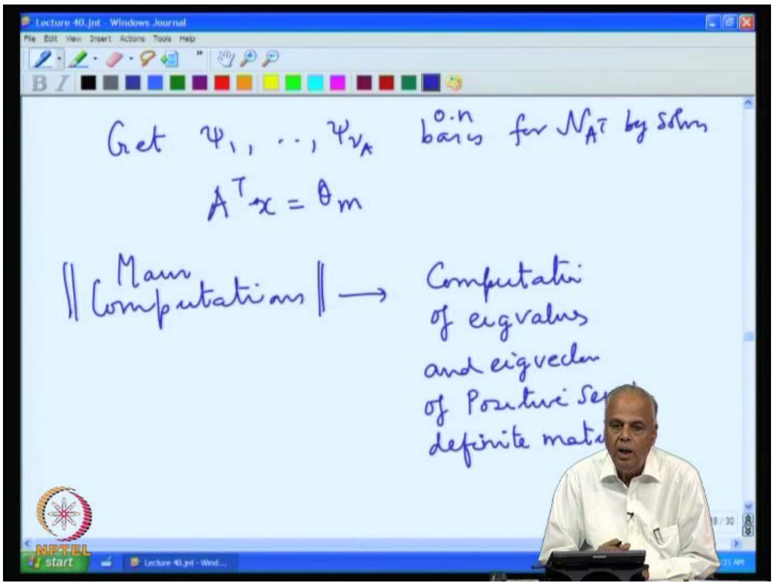
(Refer Slide Time: 44:40)



And it is Eigen values can be arranged as lambda 1 greater than or equal to lambda 2 greater than or equal to lambda 2 rho greater than 0, which is equal to the remaining Eigen values. When we dealt with positive semi definite Eigen matrices we found that their Eigen values can be arranged from this format. And corresponding to this we will get the Eigen vectors ortho normal Eigen vectors, which we call as V 1 V 2 V rho and these are the Eigen vectors corresponding to the null space.

And then we define s 1 to the square root of lambda 1 s 2 to be the square root of lambda 2 s k s rho to be the square root of lambda rho these were the singular values whenever, we say root we mean the positive square root. Then we define u j to be a V j divided by s j. So, that a V j become s j and we can find that from this we can also get A transpose u j is s j V j and.

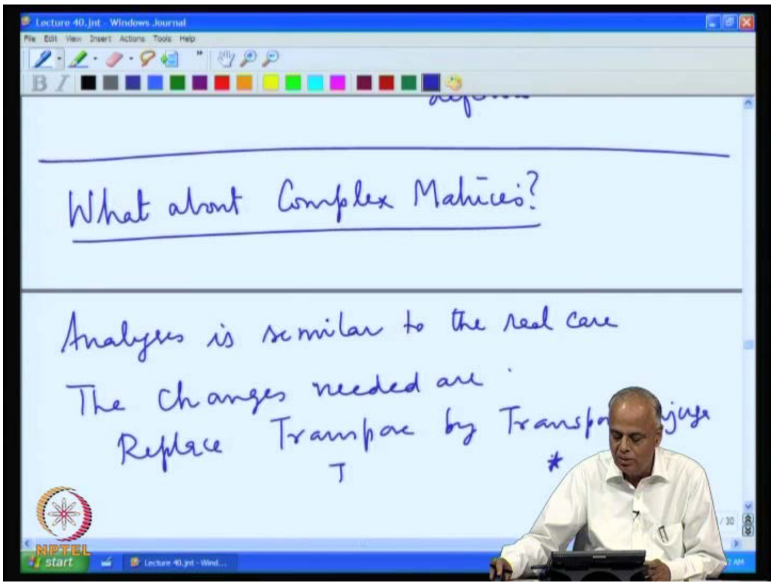
(Refer Slide Time: 46:28)



So, we got the u we got the V the phi's we got the lambda's and finally, get psi one psi two psi nu a basis ortho normal basis for A transpose for null space of A transpose by solving A transpose u. X equal to theta m and this can be done by elementary rho operations, thus the basic ingredients is therefore, or mainly these u and the V the phi's and the psi can be solved by the homogenous equations using elementary raw operations, these are the easiest, but the question is about the V 1 b 2 V rho u 1 u 2 u rho lambda 1 lambda 2 lambda rho s 1 s 2 s rho can be found the moment you get the lambda 1 lambda 2 lambda rho.

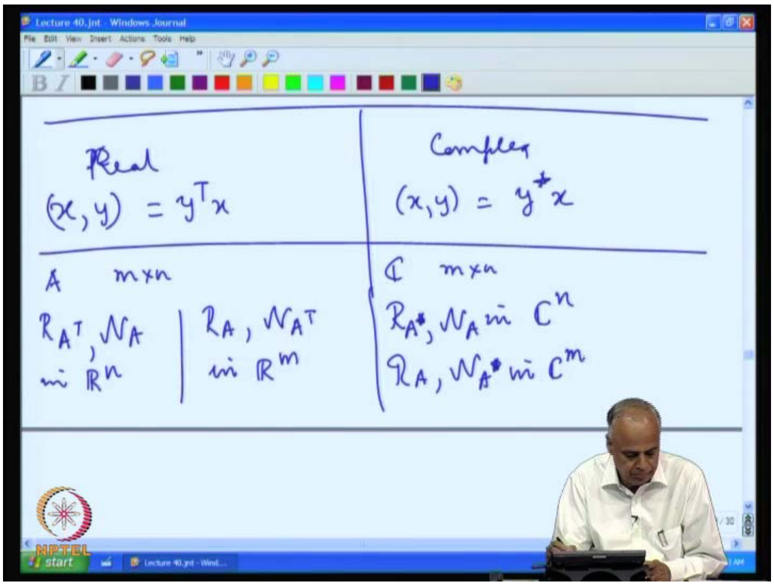
But we have seen that we can even get u 1 u 2 u rho the moment we get s rho 1 s 2 s rho and V 1 V 2 V rho, because the u can be found from the s v and the given matrix. So, the (( )) of the matrices are four is finding lambda 1 lambda 2 lambda rho V 1 V 2 V rho. Lambda 1 lambda 2 lambda rho are the positive Eigen values associated with the positive semi definite matrix L. And V 1 V 2 V rho are the corresponding ortho normal Eigen vectors. So, when we learn computations the main algorithms that would come into the picture are the competition of Eigen values and Eigen vectors of positive semi definite matrices, these are the most fundamental algorithms in computational linear algebra.

(Refer Slide Time: 48:55)



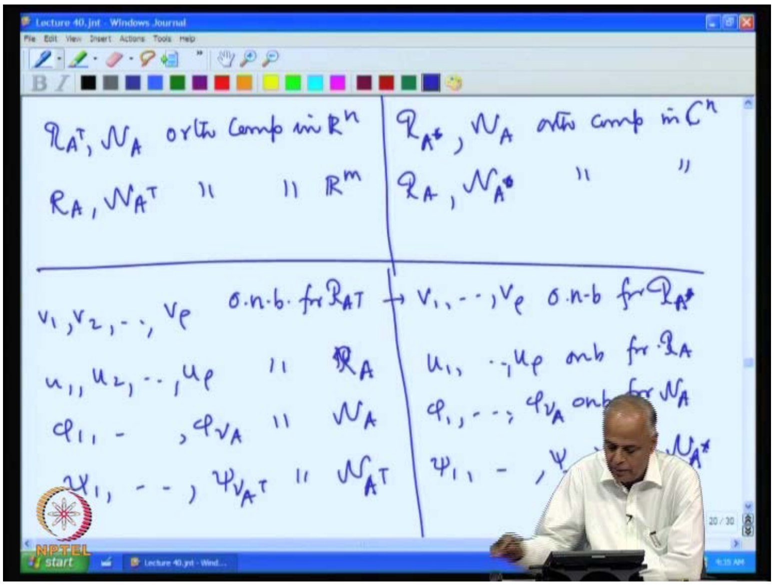
So, now we have seen a complete analysis of a real n by n matrix we will just ask a passing question what about complex matrices? Complex matrices the analysis will always be almost similar only change, that we will have to make is wherever we use transpose we simply take transpose conjugate. So, the analysis is similar to the real case, the change is needed are the following replace transpose by transpose conjugate. So, transpose, if you write it as T transpose conjugate is written as star what do we mean by this?

(Refer Slide Time: 50:11)



For example, the real case and the complex case. The real case we define the inner product of two vectors as y transpose x in the complex case, we define the inner product of two vectors y transpose conjugate, which is y star x. And then, if you have a matrix a we look at the range of A transpose and the null space of A in R n. And the range of A and the null space of A transpose in R m. In the complex case we look at range star and the null space of A in C n and the range of A and the null space of A star n C m.

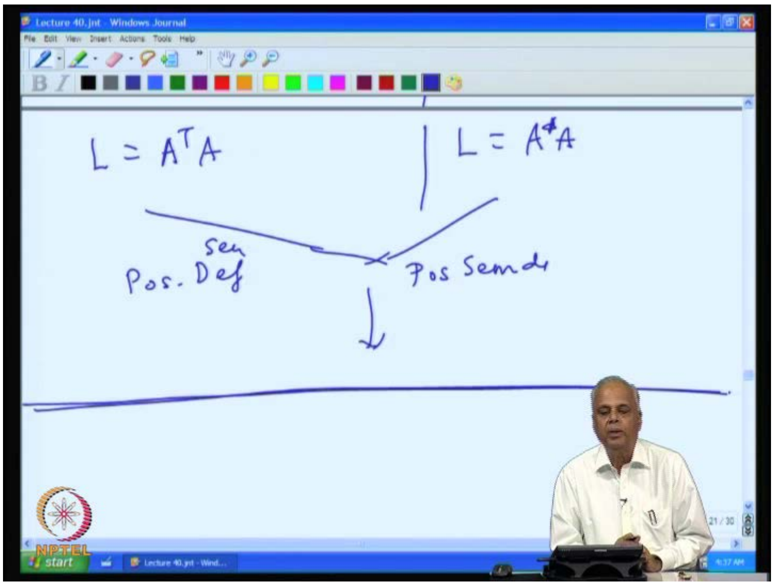
(Refer Slide Time: 51:18)



Then in the complex in the real case, we found that the range of A transpose null space of A are orthogonal compliments in R n. And range of A and null space of A transpose orthogonal compliments in R m. In the complex case the range of A star and the null space of A will be orthogonal compliments in C n and the range of A and the null space of A star will be orthogonal compliments in C m. And we would have got u 1 u 2 u rho ortho normal basis for range of A transpose v1 v 2 v rho and u 1 u 2 u rho ortho normal basis for the null space range of A.

In this case, we will get v 1 v 2 in the complex case v 1 v 2 v rho ortho normal basis for the range of A star and u 1 u 2 u rho ortho normal basis for range of A. And phi 1 phi 2 phi nu a ortho normal basis for null space here we will get phi 1 phi 2 phi nu a be ortho normal basis for null space of A. And psi 1 psi 2 psi nu a be A transpose null space of n be A transpose now so, it will be psi nu A transpose will be ortho normal basis for null space of A star.

(Refer Slide Time: 53:17)



Now, where do we get all these things? Again previously consider L in the real case, we consider L equal to A transpose A now we consider L equal to A star A. Now, this will be positive from this stage onwards it is all same this will be positive definite semi definite and this will also be positive semi definite and therefore, computations will proceed as before find the Eigen value the Eigen values arranged as lambda 1 lambda 2 lambda rho correspondingly V 1 V 2 V rho as singular values and so on.

So, the necessary changes that we have to make in the complex cases are remember we have a complex in the complex case, in the beginning itself by defining the inner product the geometry of orthogonality the is the conjugation involved. And that conjugation reflects right through our computation all the way up to defining L and the basis and then we get a positive semi definite matrix. So, all these analysis can be verbatim carried out to complex matrix now and everyone of the answers to the analog as to what we got the one route, which we not adopted in this course is when the matrix is square matrix and it is not diagonalisible.

What is that we can do? Still getting p inverse A P only not trying to resort to two matrices Q and P, but resorting to the same transformation P can be make P inverse e A P very close to the diagonal matrix this is the Jordan canonical form theory in the case of complex and real matrices. And in the case when this is not possible even you have general feel.

What we can do these are questions for at a your advanced course in linear algebra we have to consider a general field and analyze all these problems, but for a fundamental use as an engineer most of these things that we have discussed will either appear in solutions of system or other optimization problem or in image processing problems or in signal processing problems. So, these are the fundamental things that one will have to at least minimum compute with respect to a matrix. I hope these points will be referred to repeatedly and the basic necessities of these will be understood clearly.