Advanced Matrix Theory and Linear Algebra for Engineers

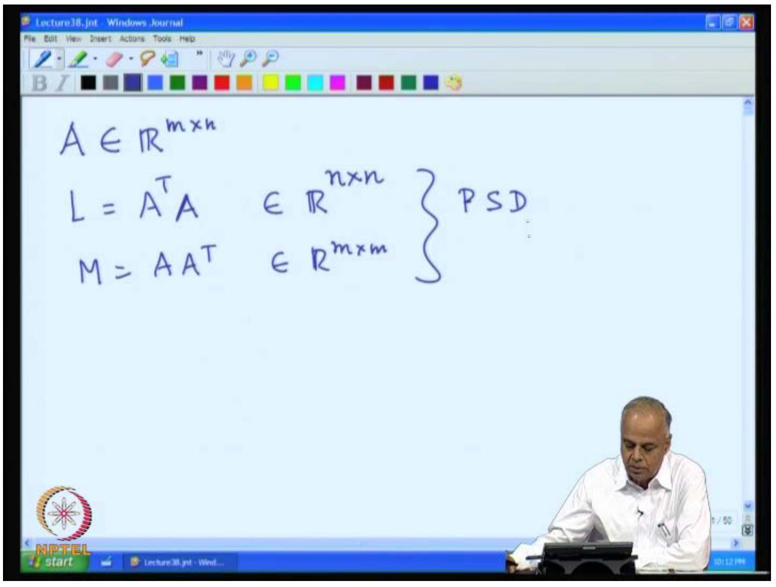
Prof. R. Vittal Rao
Centre for Electronics Design and Technology
Indian Institute of Science, Bangalore

Lecture No. # 38

Back To Linear Systems – Part 1

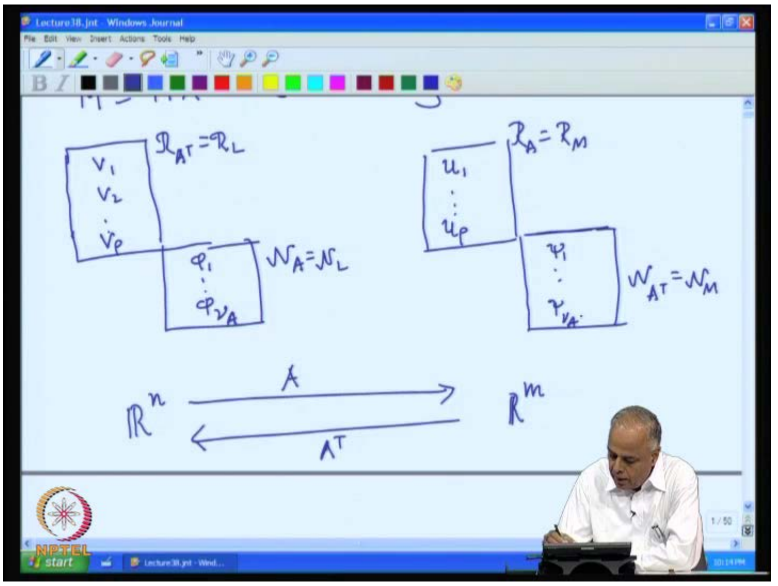
We have seen how to use the choice of our bases we make for the four subspaces connected with the matrix A to get the product decomposition and the sum decomposition of a matrix.

(Refer Slide Time: 00:32)



We had A, any real matrix m by n. And then, from that, we constructed the two square matrices, which was L – A transpose A, which is an n by n matrix. M – A A transpose, which is an m by m matrix and both were positive semi-definite matrices.

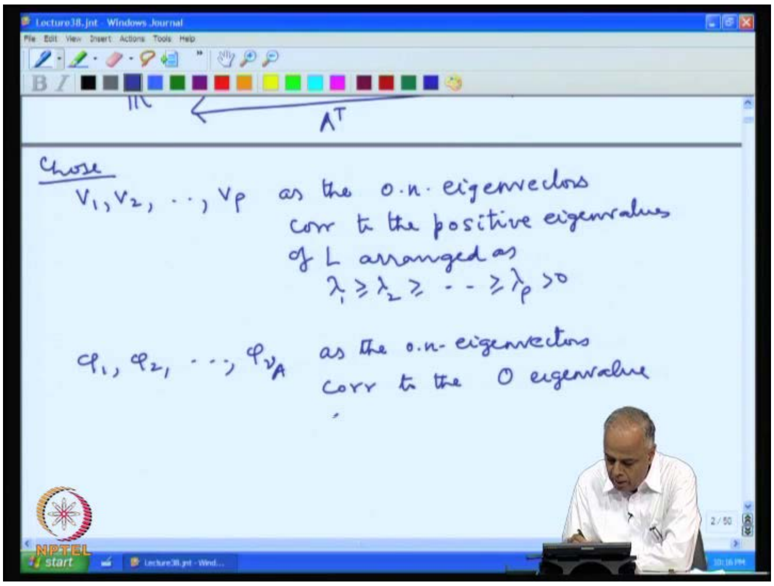
(Refer Slide Time: 01:09)



Now, using these two matrices, we found that the bases for the subspaces can be chosen as follows. Recall our decompositions. We had R n and the R m – two fundamental vector spaces A takes n component vectors to m components vectors; A transpose takes m component vectors to n component vectors. And, we had the decomposition of R n as two orthogonal complements of range of A transpose, which was the same as range of L; and, the null space of A, which was the same as null space of L. On this side, on the m side, we had the two orthogonal complements consisting of the range of A, which was the same as the range of M; and, the null space of A transpose, which was the same as null space of M.

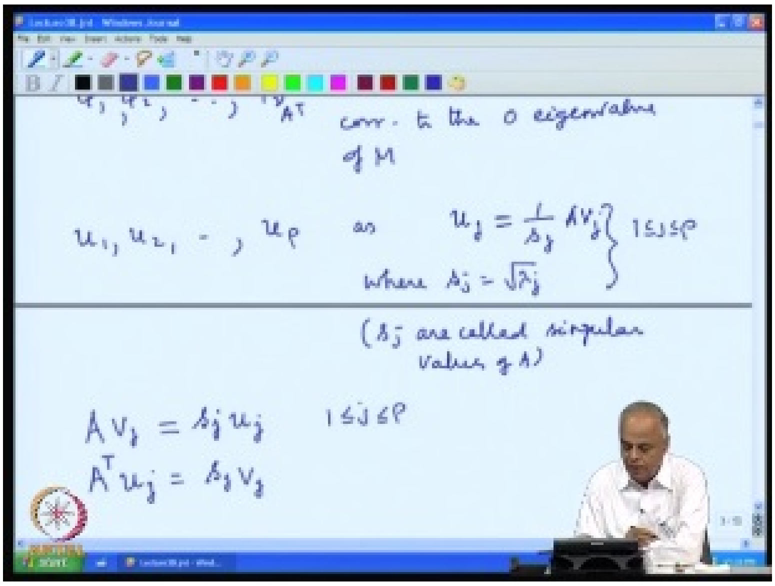
These gave rise to a series or a sequence of four orthonormal bases sets: one for the range of A transpose; the other for the null space of A; then, the third one for the range of A or the range of M; and, the forth one for the null space of A transpose. And, these bases we denoted them as V 1, V 2, V rho; we are assuming that the rank of A is rho and the nullity is of A is nu A. This is our standard notation. And, we denoted the bases for the range of L as v 1, v 2, v rho. The basis for the null space of A as p 1, p 2, p nu A; and, the orthonormal basis for the range of A as u 1, u 2, u rho; and, the orthonormal basis for the null space of A transpose as psi 1, psi 2, psi nuA transpose.

(Refer Slide Time: 03:30)



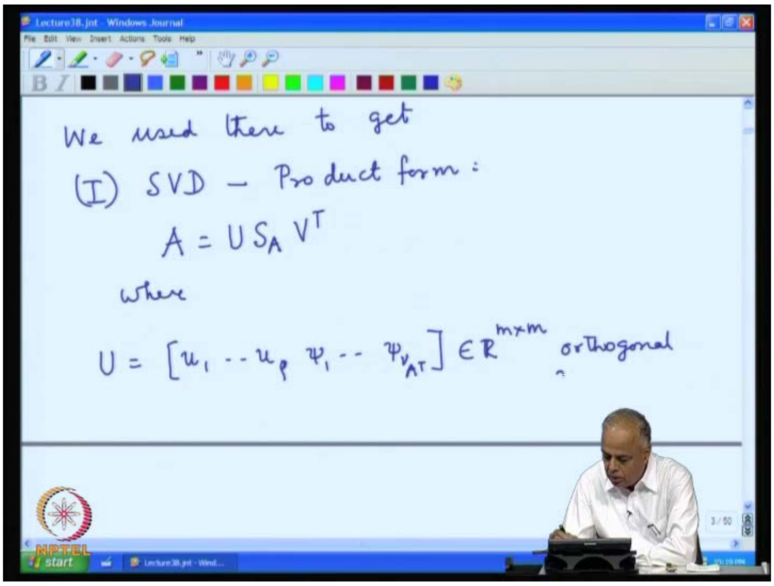
Now, it was the choice of basis that was important. We chose v 1, v 2, v rho – this was our choice as the orthonormal eigenvectors corresponding to the strictly positive eigenvalues of L arranged as lambda 1 greater than equal to lambda 2 greater than equal to lambda rho greater than 0. And, we obtained the phi 1, phi 2, phi nu A as the orthonormal eigenvectors corresponding to the 0 eigenvalue of L.

(Refer Slide Time: 04:38)



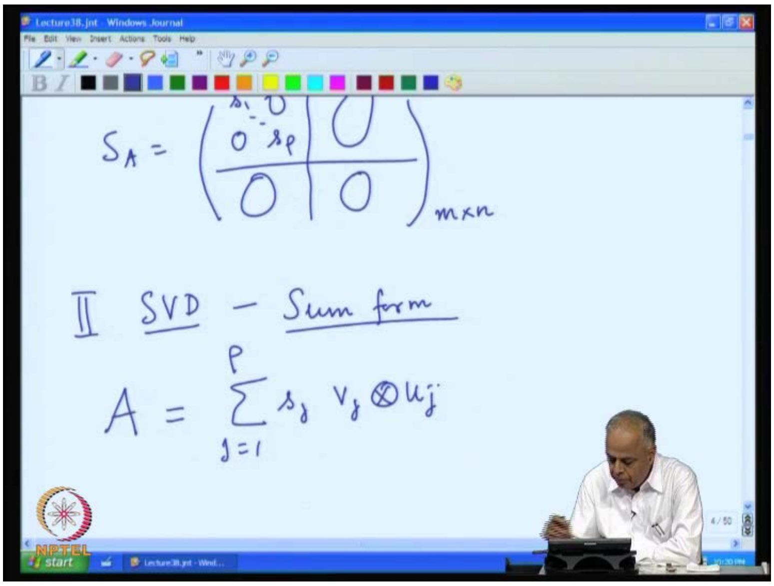
We then obtained psi 1, psi 2, psi nu A transpose as the orthonormal eigenvectors corresponding to the 0 eigenvalue of M. Now, it was in the choice of u 1, u 2, u rho bases, we linked the v basis and the u basis. We chose u 1, u 2, u rho as u j is equal to 1 by s j times Av j. We know the matrix A; we have chosen the vectors v 1, v 2, v rho; and, we construct u j as 1 by s j into Av j; where, s j is square root of lambda j – the positive square root of… this is for 1 less than j equal to rho. And, s j's were called the singular values of A. So, this was the choice of our basis and we chose the basis for the range of A in a way linking it with the basis for the range of A transpose. And, the linking gave us this following relation A v j – the jth vector in the basis for the range of A transpose under the transformation A goes to the jth vector for the range of A with the scaling factor s j. Similarly, the jth vector in the basis for range of A under the transformation A transpose goes to the jth vector in the range of A transpose with the scaling factor sj, the scaling factor being the same in both the directions. So, this was our fundamental choice of the basis.

(Refer Slide Time: 06:58)



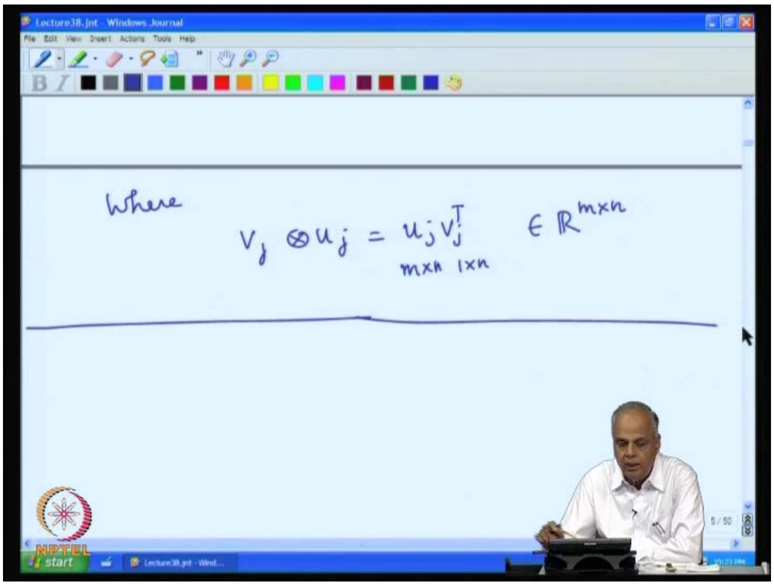
We used these to get 1 – SVD – the singular value decomposition in the product form. What was the product form? We got A as US A V transpose; where, U was the matrix, whose columns were the bases for the R m we chose – u rho, psi 1, psi nu A transpose. These were the bases for the space R m on the right side. And, these were the orthonormal bases we chose. This is an m by m matrix; and, since the columns are orthonormal, this becomes orthogonal matrix – orthogonal m by m matrix. Similarly, we was chosen, whose columns were the bases for the R m space that we chose. And therefore, this is an n cross n matrix and this is also an orthogonal matrix.

(Refer Slide Time: 08:20)



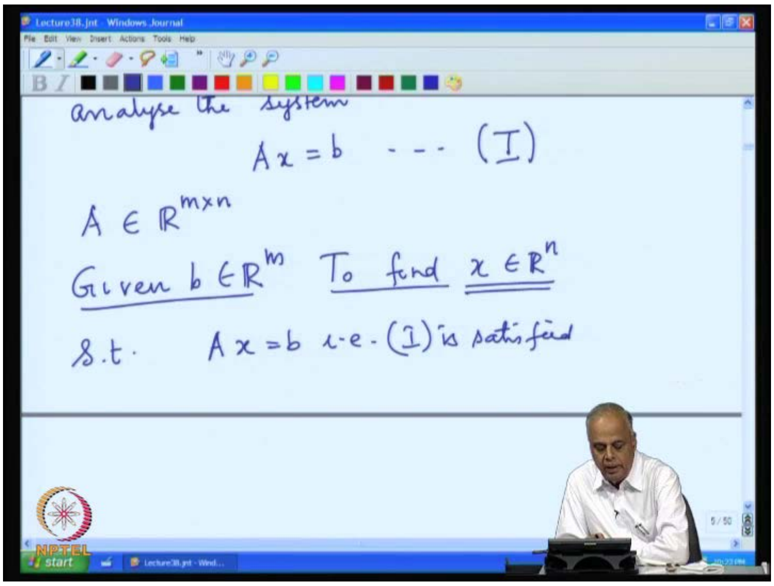
Then, the SA was essentially a diagonal block consisting of these singular values along the diagonals; and, all the rest of them are 0 and of course, it is an m by n matrix. So, all the zeroes are adjusted so that we get an m by n matrix. Thus, we got the product decomposition of the matrix. We also got the sum SVD as the sum decomposition. This was the product decomposition (Refer Slide Time: 08:54). Now, we get the SVD in the sum form, where we decompose the matrix A as the sum of rho one-rank matrices. The decomposition was A equal to summation j equal to 1 to rho s j v j tensor u j.

(Refer Slide Time: 09:19)



Where, v j – we use the notation (Refer Slide Time: 09:21) v j tensor u j – it is just the matrix u j v j transpose. So, since u j is m by 1, v j is n by 1, v j transpose is 1 by n, this is an m by n matrix and it is of rank one. And therefore, we have A as the sum of rho one-rank matrices. So, these were the decompositions we obtained in the last lectures using this particular choice of our basis. So, that is our important construction. The most important construction in all these is choosing (Refer Slide Time: 10:04) these right bases for these four subspaces.

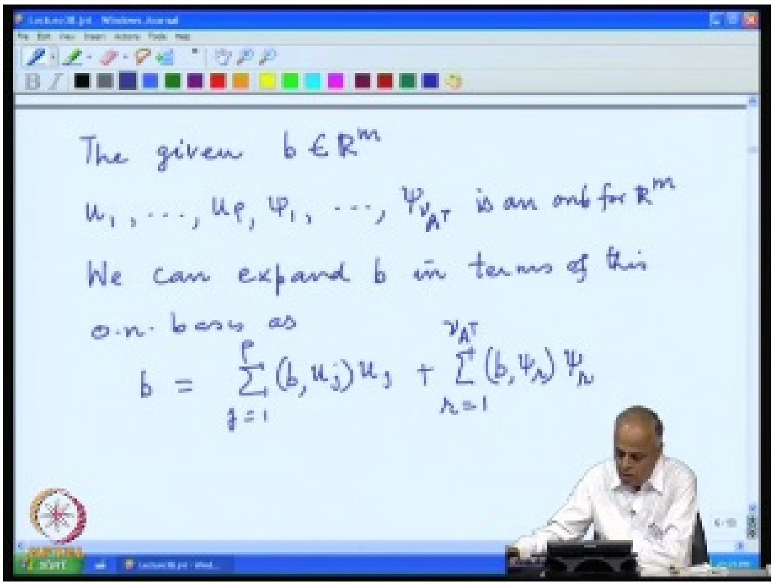
(Refer Slide Time: 10:20)



Now, how do we use these four bases for analyzing our system of equation? So, the use of these 4 orthonormal bases to analyze the system Ax equal to b. So, what is… Let us denote the system by 1. So, the problem is A is given – the matrix; it is a real m by n matrix. So, for given b in R m, we want to find… That is our problem. This is given; we want to find x in R n such that Ax equal to b; that is, 1 is satisfied. This is the problem of the system of equation. The matrix A is R m n. For any given b, which is in R m, we have to find an x in R m, such that A x is equal to b.

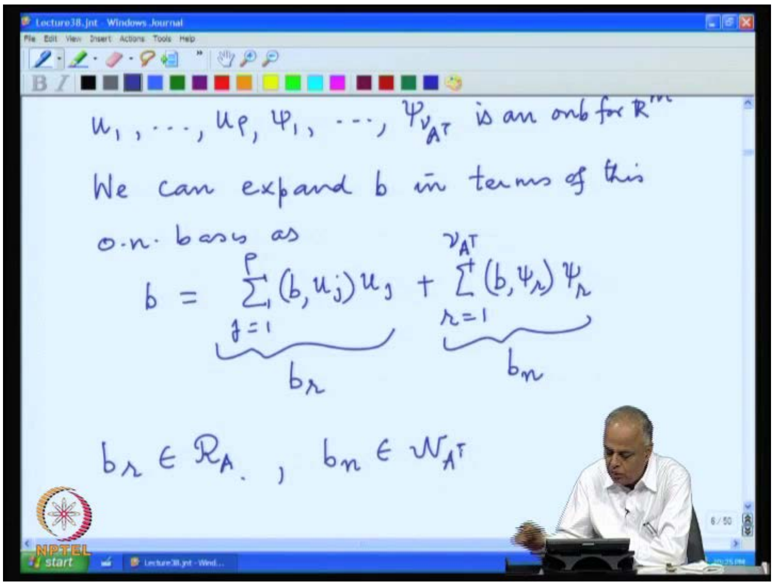
Let us see how we use these bases for answering questions regarding such a system. Now, we start with the quantity given. The given quantity is b; we must use the given information. The given information is matrix A and the vector b. Using these given information, we have to construct that unknown vector x in such a way that ax is equal to b. So, how do we do this?

(Refer Slide Time: 12:14)



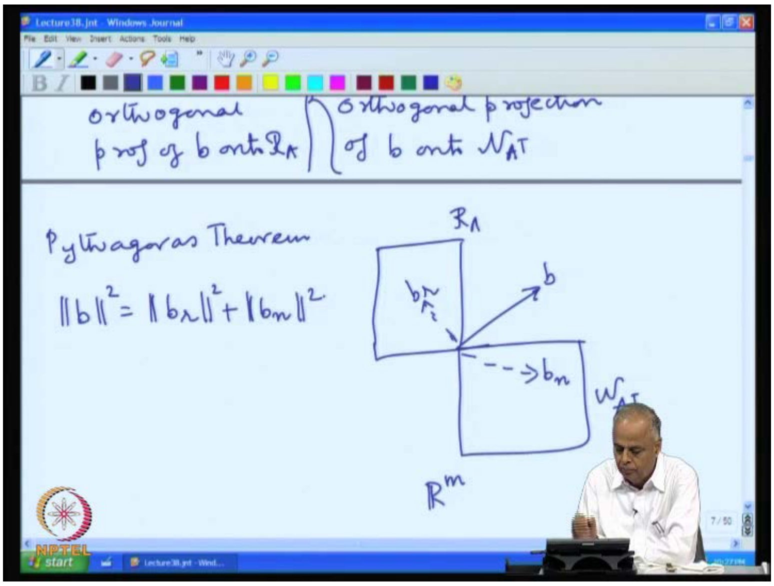
Let us start with the given information. So, the given b is in R m. And, for R m, we had chosen the bases u 1, u 2, u rho, psi 1, psi 2, psi nu A transpose, is an orthonormal basis for R m. Now, b is in R m. We have a basis for R m. Any vector in the space R m, now, can be expanded in terms of this orthonormal basis. So, we can expand b in terms of this orthonormal basis as b – first, let us look at the components in terms of the u vector – so, b comma u j into u j. Because we have an orthonormal basis, the component along u j direction will be precisely the inner product of b with u j plus the components along the side directions – r equal to 1 to nu A transpose b psi r psi r.

(Refer Slide Time: 13:45)



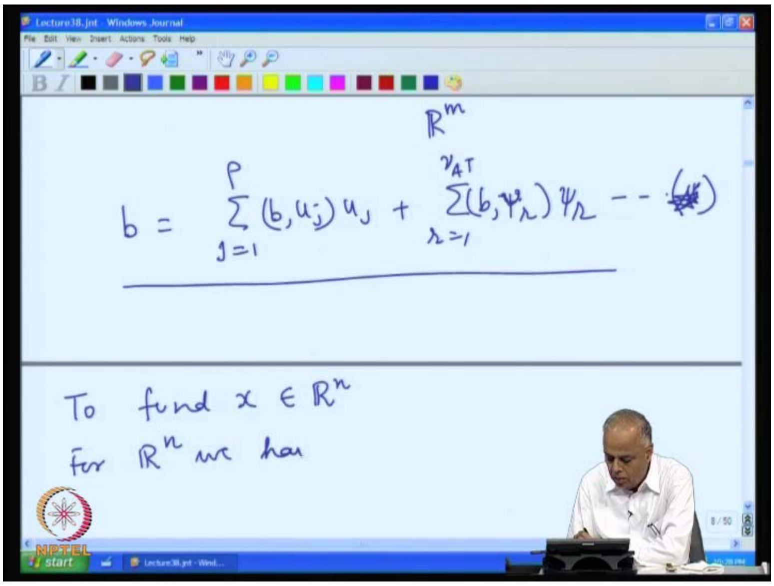
We shall denote this vector by b r and this vector by b n. Now, b r belongs to the range of A, because it is the linear combination of the vector u 1, u 2, u rho. The vectors u 1, u 2, u rho are all in the range of A. The range of A is a subspace. Any linear combinations of the vectors in the range of A will be again in the range of A. Similarly, b n belongs to the null space of A transpose. And, these are the projections.

(Refer Slide Time: 14:32)



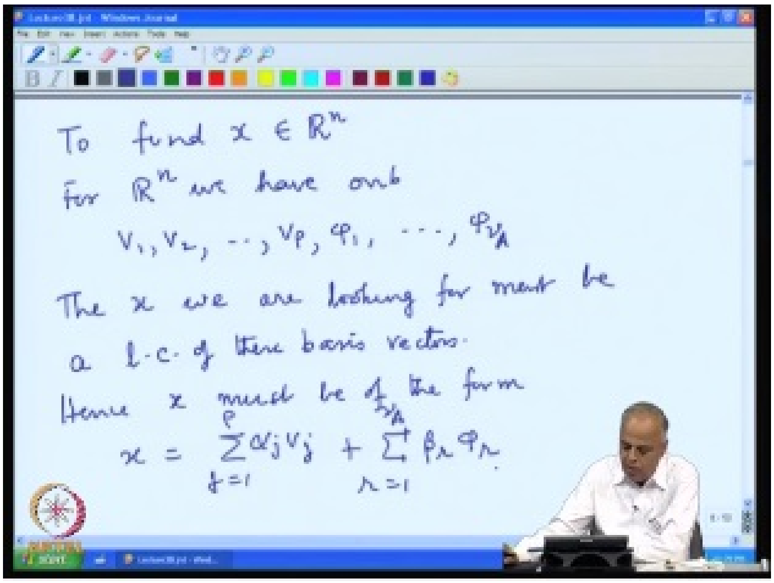
b r is the orthogonal projection of b onto range of A. And similarly, b n is the orthogonal projection of b onto the null space of A transpose. So, it (( )) down to something – it is like this. We have the R m. In R m, we have the null space of A transpose and the range of A; b is some vector. And, it has been orthogonally projected into these two subspaces – this projection is b r and this projection is b n. So, b has been projected orthogonally onto R a onto N A transpose. And, the Pythagoras theorem tells us that the length of b square is the same as length of b r square plus length of b n square.

(Refer Slide Time: 15:50)



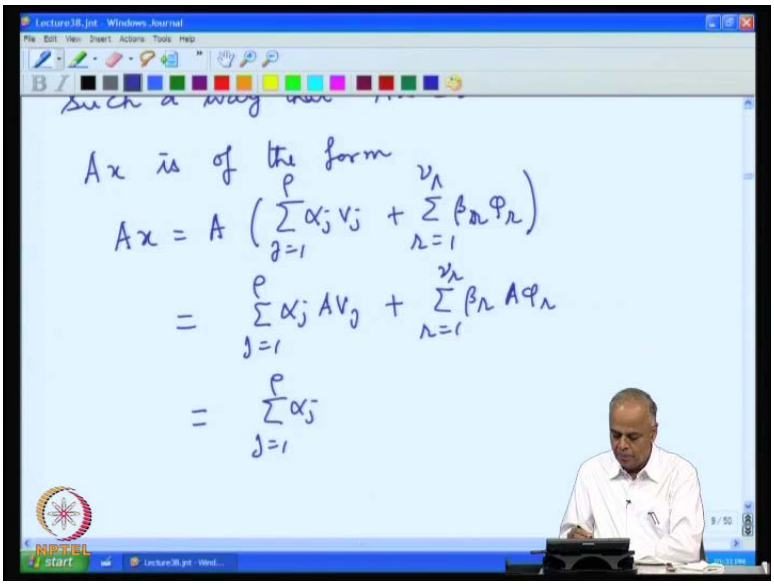
The first achievement that we have therefore is b has this expansion j equal to 1 into rho b, u j u j plus summation r equal to 1 to nu A transpose b, psi r psi r. Let us call this equation as star. So, we have the expansion of b. We know this b; our job is to find x. Where do we want to find x? We want to find x in R n.

(Refer Slide Time: 16:45)



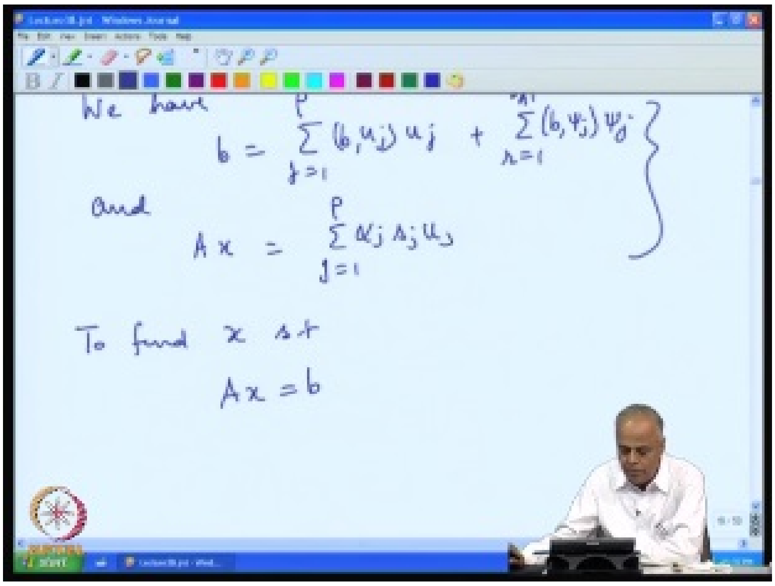
Now, for R n, we have orthonormal basis, which is v 1, v 2, v rho, phi 1, phi 2, phi nu A. And therefore, if you are looking for a vector in R n as our solution, we are looking for our solution as a vector x in R n. So, if it is going to be living in R n, it must be a linear combination of these basis vectors. So, the x we are looking for must be a linear combination of these basis vectors. Hence, x must be of the form x is equal to say j equal to 1 to rho alpha j v j. This takes care of the linear combination of the v's; r equal to 1 to nu A beta r phi r. So, the x that we are looking for must be of this form. So, the moment we know alpha 1, alpha 2, alpha rho, beta 1, beta 2, beta nu A, the vector x is known. So, the vector (Refer Slide Time: 18:09) x is known. The moment we know these numbers – alpha 1, alpha 2, alpha rho, beta 1, beta 2, beta nu A. So, our job is to find these numbers. We have to find these numbers, these scalars in such a way that Ax is equal to b. So, if x is this, what is A x? We have the representation for (Refer Slide Time: 19:02) x. It must be of this form.

(Refer Slide Time: 19:06)



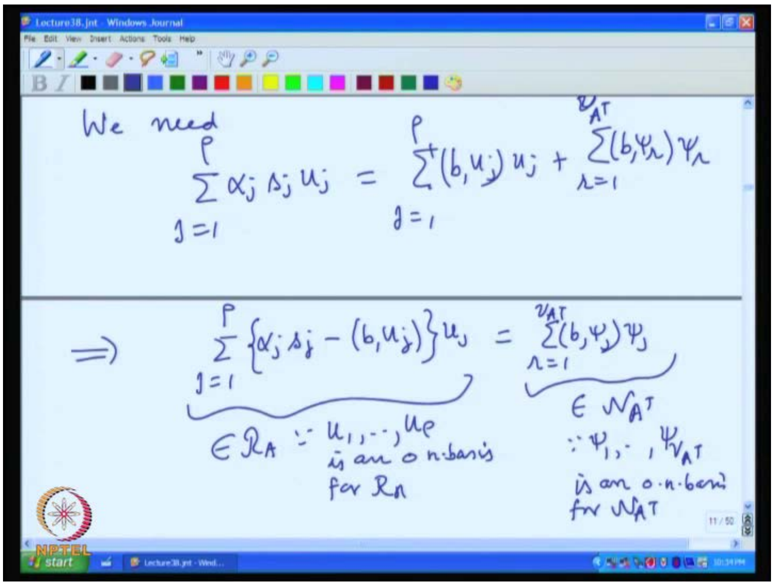
Therefore, Ax is of the form Ax equal to A of (( )) – any x that we are looking for is of this form is a linear combination of the basis vectors. Now, if we apply A, matrix multiplication is distributive. So, we have summation j equal to 1 to rho; alpha j's are scalars into Av j plus summation r equal to 1 to nu r beta r into A phi r. Now, what is this equal to? We have our choice of basis, was such that A v j – the v j basis went to the u j direction with the scaling factor s j. So, A v j was s j u j. So, the first sum becomes summation j equal to 1 to rho s j v j. In the second sum, the phi 1, phi 2, phi nu A; that is, all the phi r's are in the null space of A. So, A times this phi r will be the 0 vector. So, the second term will contribute only the 0 vector and hence, we get A x equal to alpha j s j u j.

(Refer Slide Time: 20:38)



Therefore, we have b expansion; we had expanded b as summation j equal to 1 to rho b, u j u j plus summation r equal to 1 to nu A transpose b, psi j, psi j. This is the first expansion we got looking at the known quantity. And now, we have Ax is equal to summation j equal to 1 to rho alpha j s j u j. Now, our problem is, we have to find x such that Ax equal to b. Therefore, we have the Ax, we have the b.

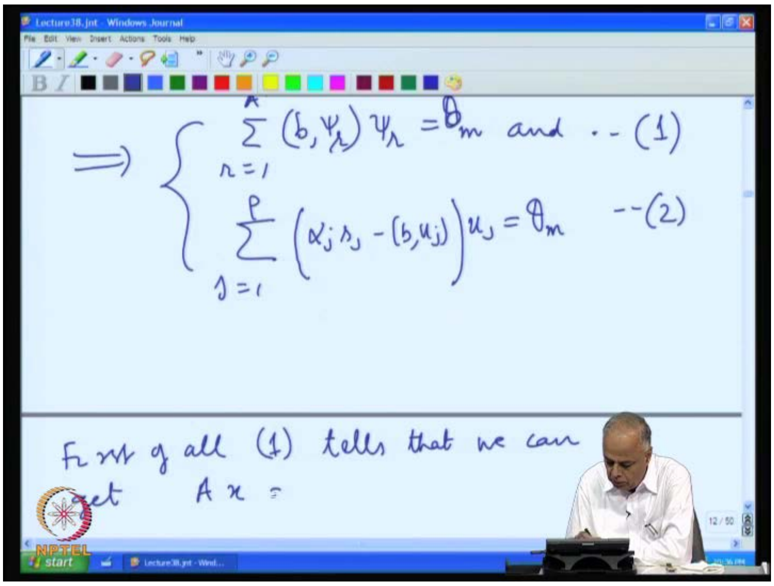
(Refer Slide Time: 21:27)



Therefore, we need summation j equal to 1 to rho alpha j s j u j must be equal to the b, which is j equal to 1 to rho b, u j u j plus summation r equal to 1 to nu A transpose b, psi r psi r. Now, we can write this as summation j equal to 1 to rho – we will take all the u j terms on one side – alpha j s j minus b, u j – this whole quantity into u j. On the right-hand side, we keep all the psi j terms. Now, the left-hand side denotes the quantities, which are linear combinations of u j. Now, the vectors u j are all in the range of A. If you look at the picture that we drew, the u 1, u 2, u rho were bases for the range of A. So, any linear combination of u 1, u 2, u rho will also be in the range of A. So, this is in the range of A, because u 1, u 2, u rho is an orthonormal basis for range of A.

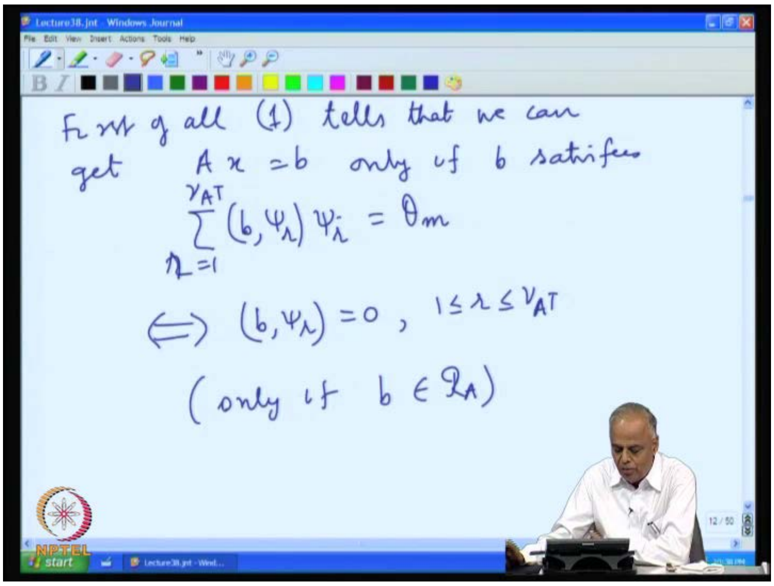
Similarly, this is in the (Refer Slide Time: 23:02) null space of A transpose, because psi 1, psi nu A transpose is an orthonormal basis for null space of A transpose. Now, the left-hand side is the vector in the range of A and the right-hand side is the vector in the null space of A transpose. But, the range of A and the null space of A transpose are orthogonal to each other. And therefore, the only vector common to them is the zero vector. And hence, both sides must be 0. See we have this range of A perpendicular to the null space of A transpose. So, this equal vector, which is common to both of them must only be the zero vector, because the only vector common to orthogonal complement is the zero vector.

(Refer Slide Time: 23:58)



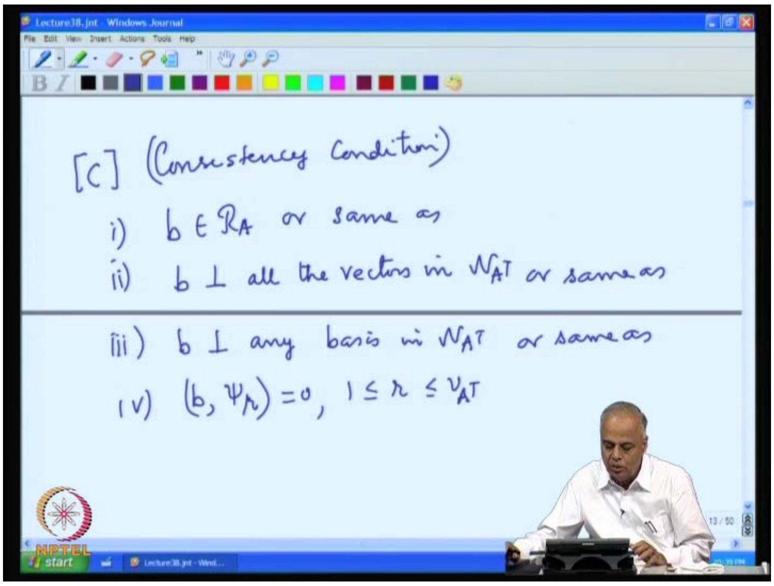
This implies the first – the right-hand side must be equal to 0 or this should be r; r must be equal to 0. And, summation j equal to 1 to rho alpha j s j minus b, u j – this whole thing u j must be equal to theta m; not 0, but the zero vector. Now, the given vector V is arbitrarily chosen from R m and therefore, the system demands, if you want to have a solution, if you want to have A x equal to b, you must have chosen b in such a way that the first sum is 0. Let us call this as 1 and 2.

(Refer Slide Time: 25:02)



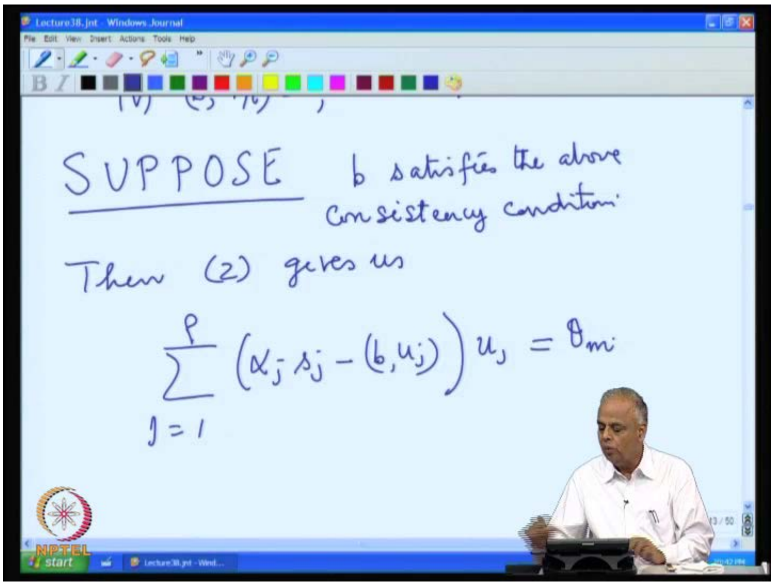
Now, first of all, 1 tells that we can get A x equal to b only if b satisfies summation r equal to 1 to nu A transpose b, psi r psi r equal to theta m. Now, psi r's are orthogonal vectors; they are linearly independent. And therefore, they are only called linear combination that will give zero vector is b, psi r, must all be equal to 0; 1 less than or equal to r less than or equal to nu A transpose. Therefore, the system Ax equal to b will have a solution only if b psi r is equal to 0. This means that the components of b along the null space of A transpose directions are all 0; or, the orthogonal projection of b on to N A transpose is 0. And hence, that is only if b belongs to the range of A, which is natural, because we want Ax equal to b. So, b better be in the range of A. This fact that it belongs to range of A is stated in many ways – b must be range of A; b must therefore be perpendicular to the null space of A transpose; and therefore, it must be perpendicular to any basis in particular to any orthonormal basis; and therefore, in particular to the orthonormal basis that we have chosen. This we will call as the consistency condition.

(Refer Slide Time: 27:09)



Consistency condition – various ways of stating is: b must be belongs to range of A – one; or, same as b must be perpendicular – orthogonal to all the vectors in null space of A transpose, which we can also create as b must be perpendicular to any basis in N A transpose; that is, all the vectors in any basis must be perpendicular; or the same as… The same fact is now b if perpendicular to the particular choice that we have got, b, psi r equal to 0 for 1 less than or equal to r less than or equal to nu A transpose. This is our particular choice and we want this to be equal to 0. So, the consistency condition. If this is not satisfied, then we cannot get Ax equal to b, because for Ax equal to b, if at all it has to be satisfied, we need both these to be satisfied. So, the first one demands that b belongs to the range of A, which is stated as b psi r is equal to 0.

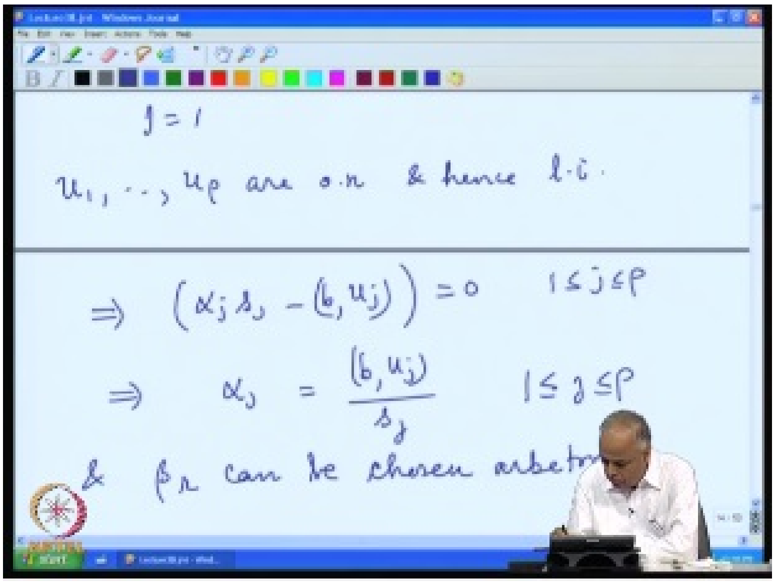
(Refer Slide Time: 28:45)



Now, suppose b may or may not. So, let us consider the easiest case first. Suppose b satisfies the above consistency condition. Let us go about what we have done. We had the known vector b; we could expand the known vector b in terms of the basis for the range of M as this (Refer Slide Time: 29:21) expansion. Then, we could expand the vector x in terms of the known basis and we got this representation; and, using that, we found that A x must be of this form (Refer Slide Time: 29:37). And therefore, using that Ax must be of this form and b is of the above form, we wanted Ax equal to b. And, that Ax equal to b requirement tells us that we need to have these two conditions: 1 (Refer Slide Time: 29:54) and 2. And, the first condition gave us the consistency criterion that b must belong to this.

Suppose b satisfies this consistency condition, then out of these two statements (Refer Slide Time: 30:07) we had, one has been taken care of. Now, we have to worry about 2. So, then, 2 gives us (Refer Slide Time: 30:18) summation j equal to 1 to rho alpha j s j minus b, u j – this whole quantity into u j is the zero vector.

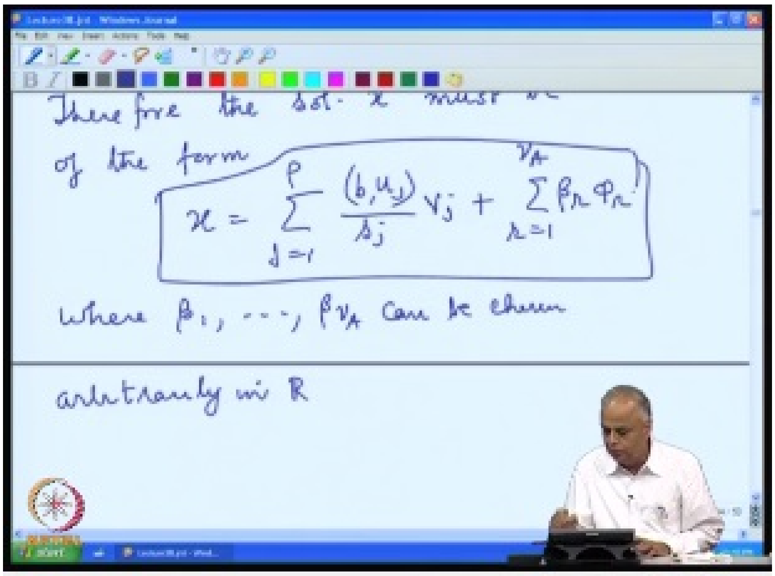
(Refer Slide Time: 30:40)



Now, u 1, u 2, u rho are orthonormal. And therefore, we have seen that any orthonormal set is linearly independent – and hence, linearly independent. And, here is a linear combination of the linearly independent vector 0. And therefore, the coefficients must be equal to 0 for 1 less than or equal to j less than or equal to rho. This means that alpha j must be equal to b, u j by s j. So, this says how to choose… Remember we were trying to find x. Finding an x means we must find alpha 1, alpha 2, alpha rho, beta 1, beta 2, beta rho.

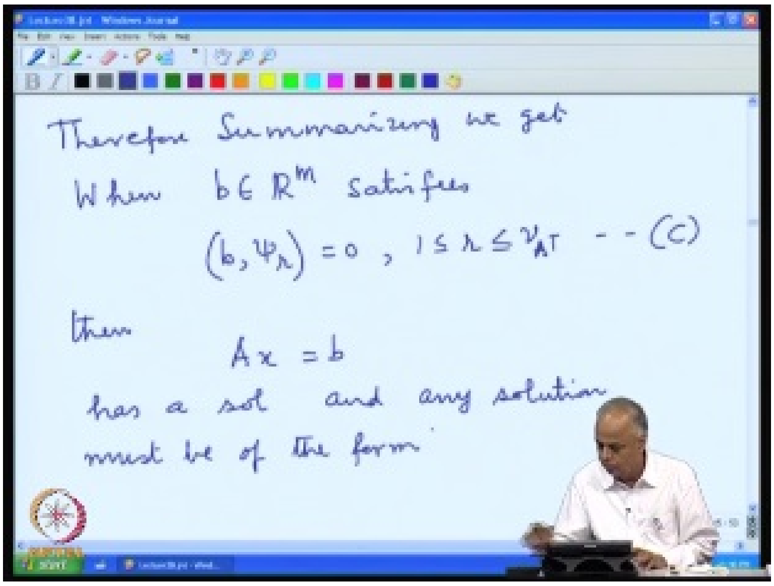
Now, all the analyses say that you have to choose the alphas such that alpha j is equal to b, u j by s j. And, this does not say anything about beta j. Whatever beta j's you choose, they are not going to affect your solution, because when we computed the A u j – if you remember, when you calculated the A u x, the beta j's came completely disappeared, because A phi was 0. So, whatever beta's we choose, that is not going to affect the system. And hence, beta's can be chosen in any way we want. And thus, we have… Alpha's have to be chosen as we have here (Refer Slide Time: 32:24) – alpha j's are beta, u j by s j; and, beta r can be chosen arbitrarily.

(Refer Slide Time: 32:39)



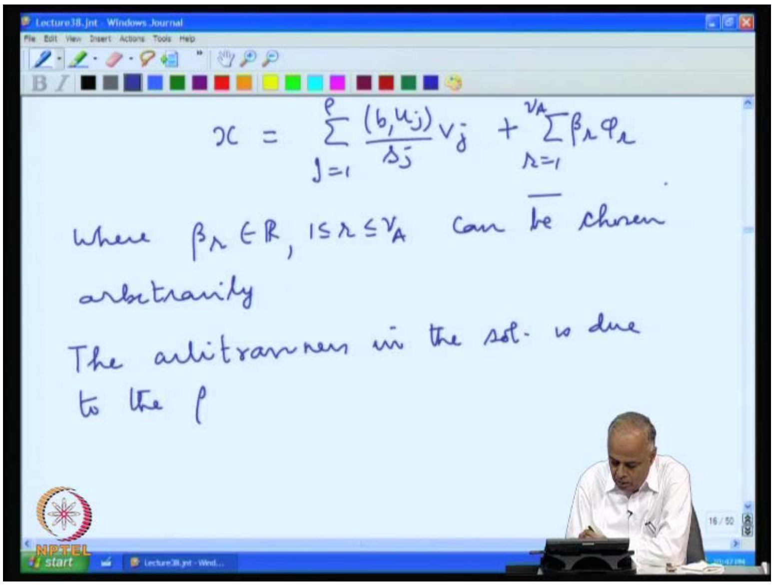
Therefore, the solution x must be of the form x is equal to summation j equal to 1 to rho alpha j must be equal to – b, u j by s j into v j plus beta R can be chosen any way you want – summation r is equal to 1 to nu A beta r phi r; where, beta 1, beta 2, beta nu A can be chosen arbitrarily in R. Thus, we have the solution. The moment b satisfies the consistency condition, we have the solution and we know that any solution must be of this form. So, any solution is of this form.

(Refer Slide Time: 33:51)



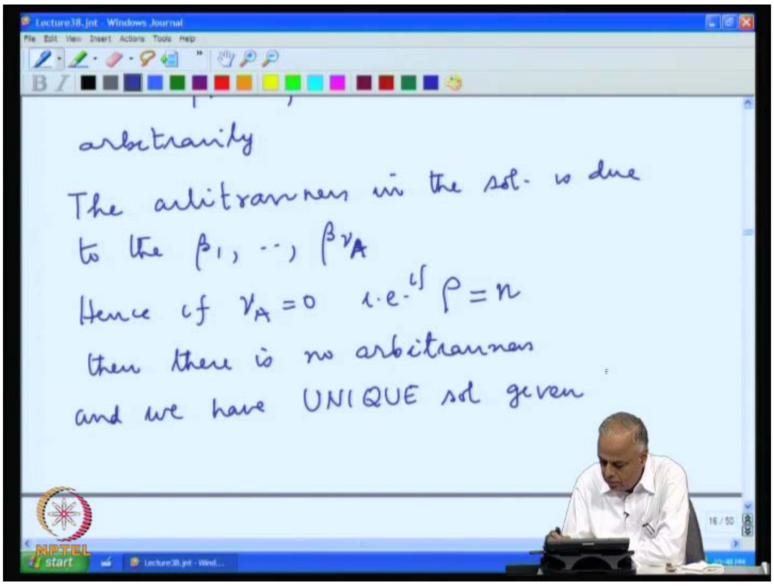
Now, let us therefore summarize. Therefore, summarizing we get – when b belonging to R m satisfies the consistency condition b, psi r equal to 0 for 1 less than or equal to r less than or equal to nu A transpose – consistency condition.

(Refer Slide Time: 34:52)



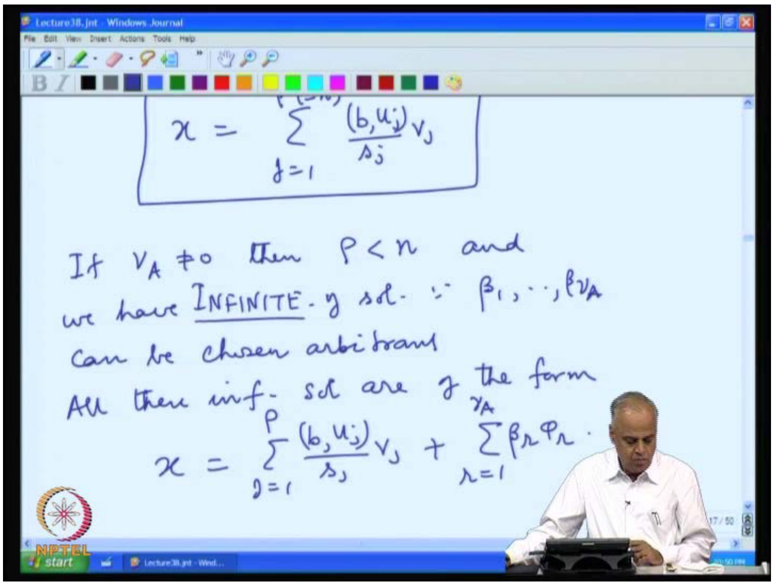
Then, Ax equal to b has a solution and any solution must be of the form x equal to summation j equal to 1 to rho b, u j by s j into v j plus summation r equal to 1 to nu A beta r phi r; where, the beta r belongs to R, 1 less than or equal to r less than or equal to nu A can be chosen arbitrarily. So, when b satisfies the consistency condition, we have the solution and any solution must be of this form. Now, where does the arbitrariness into the solution come? The solutions are arbitrary in nature in the sense the beta 1, beta 2, beta nu A can be chosen arbitrarily. Now, if the nu A is 0, then rho will be N and the second part will completely disappear. And therefore, the arbitrariness will be disappeared.

(Refer Slide Time: 36:06)



So, the arbitrariness in the solution is due to the b 1, b 2, b nu A. Hence, if nu A is 0, that is, if rank plus nullity is n – if nu A is 0; that means if rank is n, then there is no arbitrariness, because there are no beta’s there now; there is no arbitrariness.

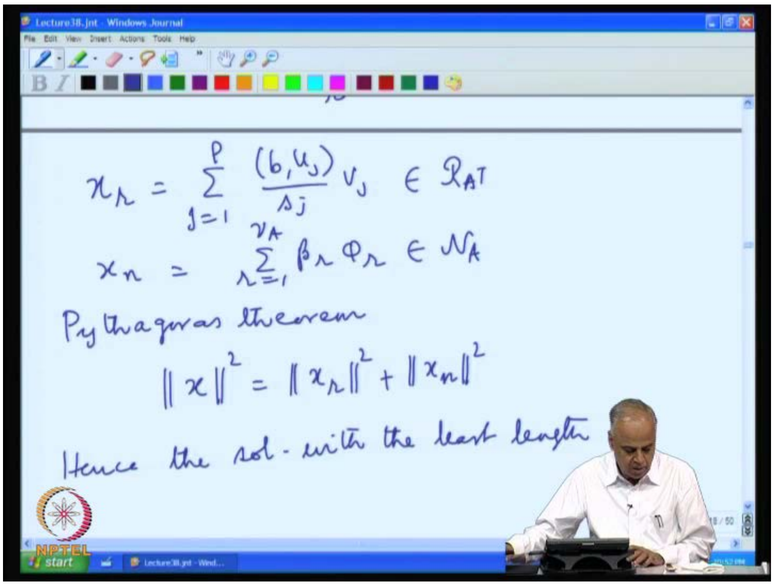
(Refer Slide Time: 37:08)



And, we have unique solution given by x is equal to summation j equal to 1 to rho; where now, rho is equal to n. We are taking the case rho equal to n. b, u j by s j into v j. This is the unique solution. If nu A – if the nullity of A is not equal to 0, then rho is less than n and we have infinite number of solutions, because beta 1, beta 2, beta nu A can be chosen arbitrarily. So, in this case, when we are having rho is less than n, we have infinite number of solutions. When rho was equal to n, we have unique solution; when rho is less than n, we have infinite number of solutions. Now, all these infinite solutions are of the form x is equal to summation j equal to 1 to rho b, u j by s j into v j plus summation r equal to 1 to nu A beta r phi r, because now the betas can be chosen arbitrarily.

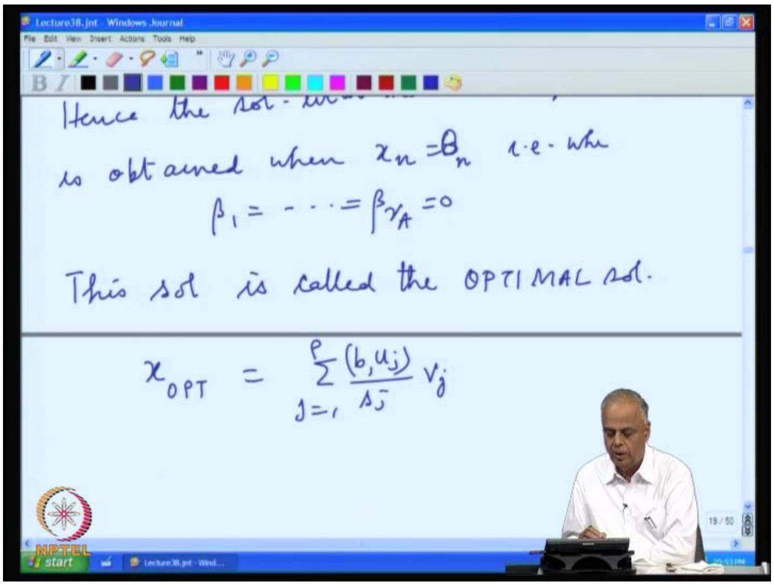
Now, if you look at this (Refer Slide Time: 38:50) part, these are all linear combinations of the vectors v 1, v 2, v rho. The vectors v 1, v 2, v rho form an orthonormal basis for the range of A transpose. And therefore, any linear combination will also be in the range of A transpose. So, that is the vector in the range of A transpose. And, the second sum is a linear combination of the vectors phi 1, phi 2, phi nu A. Phi 1, phi 2, phi nu A form an orthonormal basis for the null space of A. And therefore, any linear combination of them will belong to the null space of A. We will call this as x r and this as x n (Refer Slide Time: 39:37).

(Refer Slide Time: 39:39)



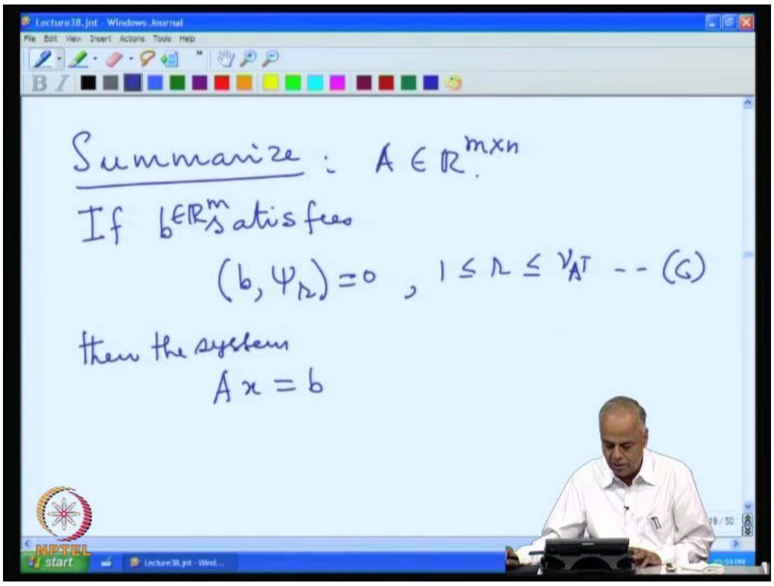
So, x r is the summation j equal to 1 to rho b, u j by s j into v j. This belongs to the range of A transpose. x n is summation r equal to 1 to nu A beta r phi r, which belongs to the null space of A. And, by Pythagoras theorem again, since the range of A transpose and the null space of A are orthogonal complements of each other, we get by the Pythagoras theorem, the length of x square is equal to the length of x r square plus the length of x n square. And therefore, if you look at the right-hand side, we find that the length of the solution will be minimum when x n is equal to 0; that is, when we choose all the betas to be zeroes.

(Refer Slide Time: 40:42)



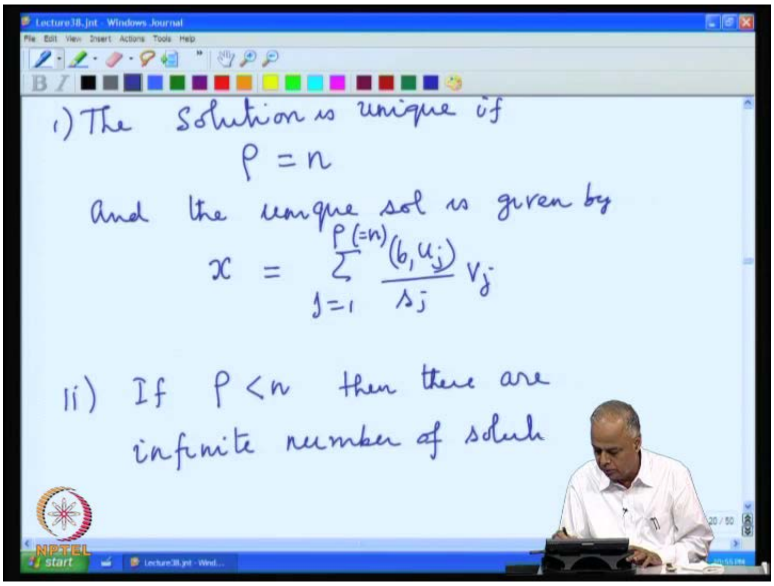
Hence, the solution with the least length is obtained when x n is equal to 0, 0 vector; that is, when all beta 1 equal to beta nu A equal to 0. So, this solution is called the optimal solution. And, we will denote it by x optimal and that is given by j equal to 1 to rho b, u j by s j v j; where now, rho is less than n. So, what is the conclusion?

(Refer Slide Time: 41:58)



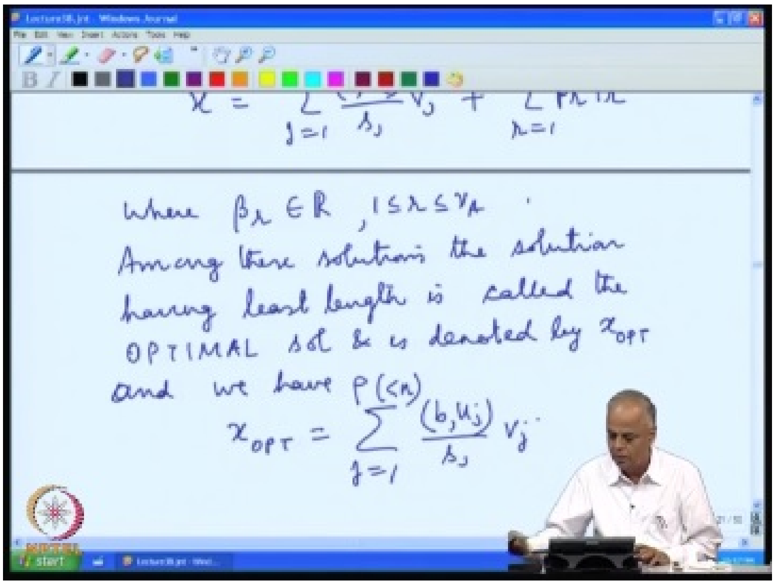
Let us again summarize. If you now summarize this whole process, we have – if b satisfies the consistency conditions… These are the consistency conditions. If b satisfies these consistency conditions, then A x equal to b; the system A x equal to… Of course, b is in R m. If b satisfies this consistent… A is a matrix, which is an R m n. Let us write all the notations in the summarization. If b satisfies the consistency condition; where the psi's, phi's, u's and v's are the same as we have in the figures that we have been drawing, the corresponding orthonormal basis – when b satisfies this consistency condition, the system has a solution.

(Refer Slide Time: 43:11)



One – the solution is unique if the rank of the matrix is equal to the number of columns and the unique solution is given by x is equal to summation j equal to 1 to rho – rho is same as n – b, u j by s j v j. If rho is less than n, then there are infinite number of solutions and they are all of the form (Refer Slide Time: 44:25) x is equal to summation j equal to 1 to rho b, u j by s j v j plus summation r equal to 1 to nu A beta r phi r; beta r can be chosen any way we want r.

(Refer Slide Time: 44:53)



And, among these solutions, the solution having least length is called the optimal solution. And, is denoted by x optimal and we have x optimal is equal to summation j equal to 1 to rho – now, less than n b, u j by s j into v j.

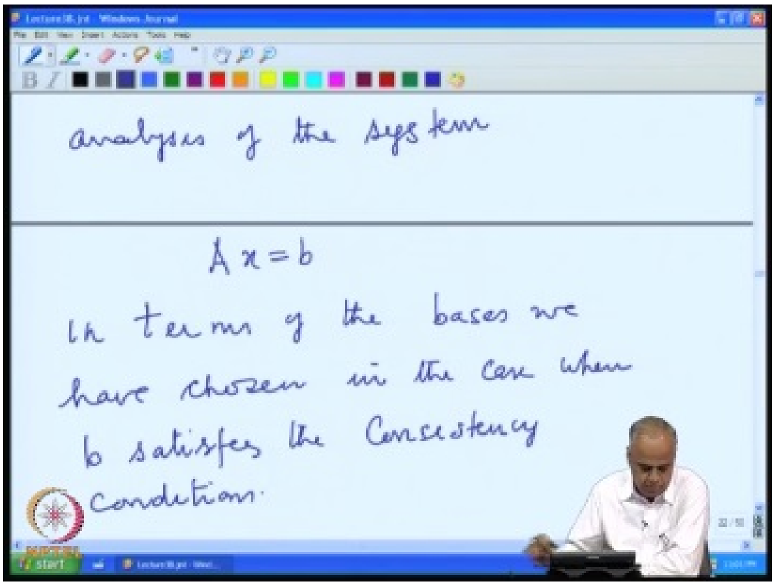
In the case when rho equal to n, there (Refer Slide Time: 45:50) is only one solution. Therefore, that is also the optimal solution. There is only one person; he is the lowest as well as the highest. So, among all the solutions, there is only one solution, and therefore, he is also the optimal solution. But, when rho is less than (Refer Slide Time: 46:06) n, there is this possibility that the infinite number of solutions can be found. Now, among these infinite solutions, which are all of this form, this gives us the (Refer Slide Time: 46:18) structure of all the solutions. Among all these solutions, there is the solution which has the least length and that is called the optimal solution. And, that optimal solution has this structure. So, if b…

We know what is the consistency condition in terms of the basis that we have chosen; and, we have also seen so far that these consistency conditions are satisfied; we know that the solution exists; and, we know when the solution is unique. This is the case (Refer Slide Time: 46:46) when the solution is unique. And, we know exactly how the unique solution looks like. This is the case (Refer Slide Time: 47:06) when rho is less than n. When rho is less than n, we have infinite number of solutions and we know how all these solutions look like; and then, from that, we know how to generate a representative solution namely, the solution with the least length or the optimal solution. All these are obtained in the terms of the bases we have chosen.

Now, if we look at the unique solution that we (Refer Slide Time: 47:37) got, all we have done is we have taken the jth component of b and scaled it down by a factor of 1 by s j and took that as the j th component of x. So, basically therefore, as we observed, A takes the jth basis vector on R n to the jth basis vector R m with a upward scaling factor of s j. So, when we want to find the solution, we are pulling back things. So, we are pulling back the jth component of b to the jth component of x. But, now, scaling down the factor, because the scaling up by A has been undone by the scaling down of s j.

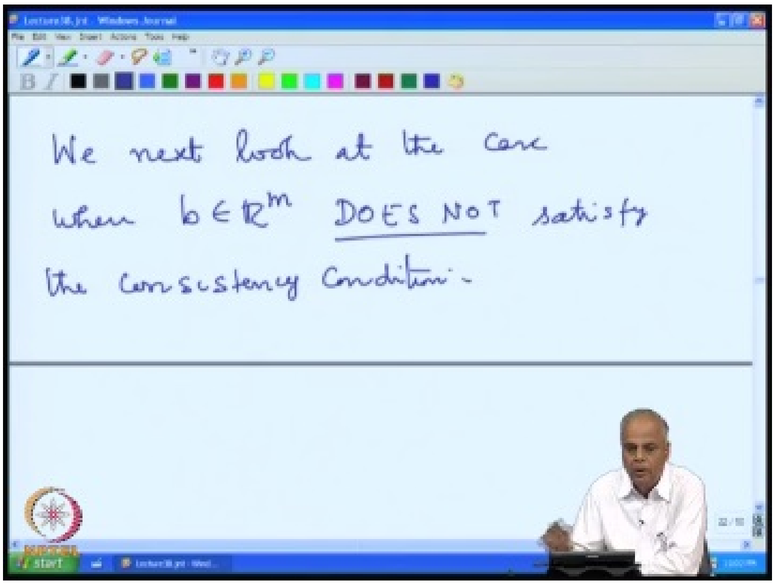
Similarly, even in the optimal solution case, we are looking at only that part of b, which belongs to the range of A and then we are pulling back again the components with the scaling factor. Basically, therefore, we are saying that since A has the fundamental effect of scaling up by s j while going from range of A to the range of A transpose, we have to do the scaling down while coming back to find the solutions. Thus, in the case of the b satisfying the consistency condition, we have all the answers; we know when the solution is unique; we know when it is unique what is the structure of the solution; and, we know when the solution is infinite; and, if it is infinite number of solutions, we know the structure of all the solutions; and, we also know how to choose a unique representative among these infinite solutions, is the optimal solution.

(Refer Slide Time: 49:43)



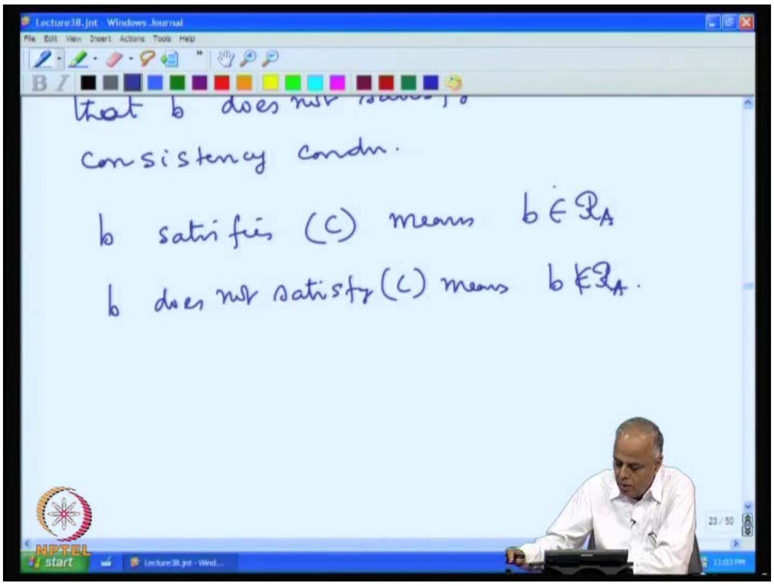
Thus, we have a complete analysis of the system A x equal to b in terms of the bases we have chosen in the case when b satisfies the consistency conditions.

(Refer Slide Time: 50:31)



Now, therefore, we next look at the case when b belonged R m does not satisfy the consistency condition.

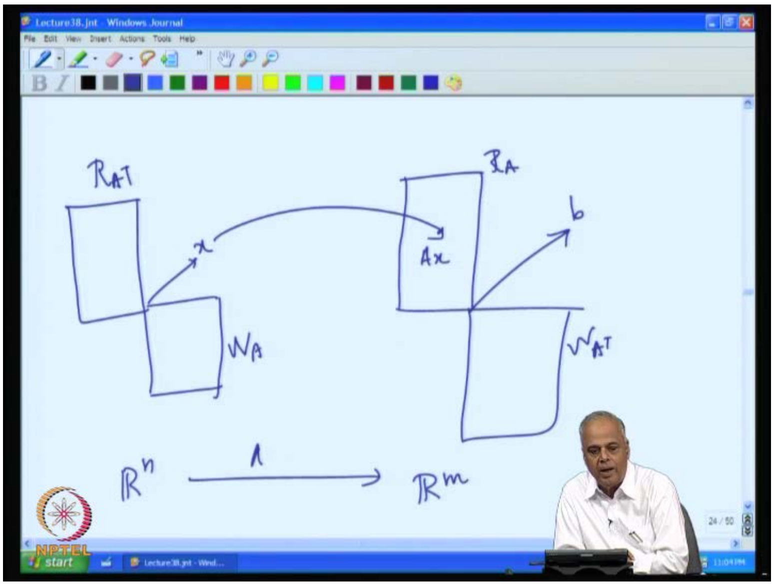
(Refer Slide Time: 51:08)



Now, what does it mean to say that b does not satisfy consistency condition? First, let us analyze that. What it means is the following. b satisfies means b satisfies the consistency condition c means b belongs to range of A. We put in many forms: b belongs to range of

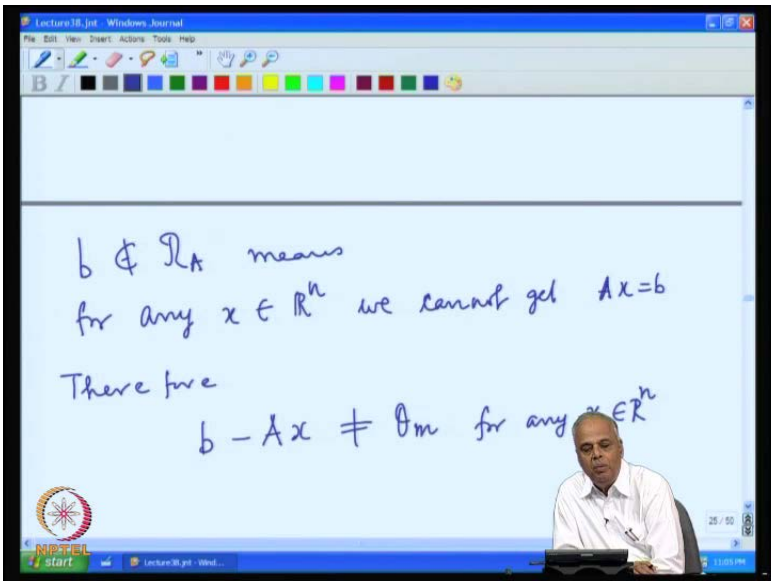
A; or, b is orthogonal to all the vectors in the null space of A transpose; or, b is orthogonal to the psi j vectors and so on and so forth. All of them mean the same thing; that means b belongs to range of A. But, if b does not satisfy c means b does not belong to range of A. Therefore, we have the following situation.

(Refer Slide Time: 52:21)



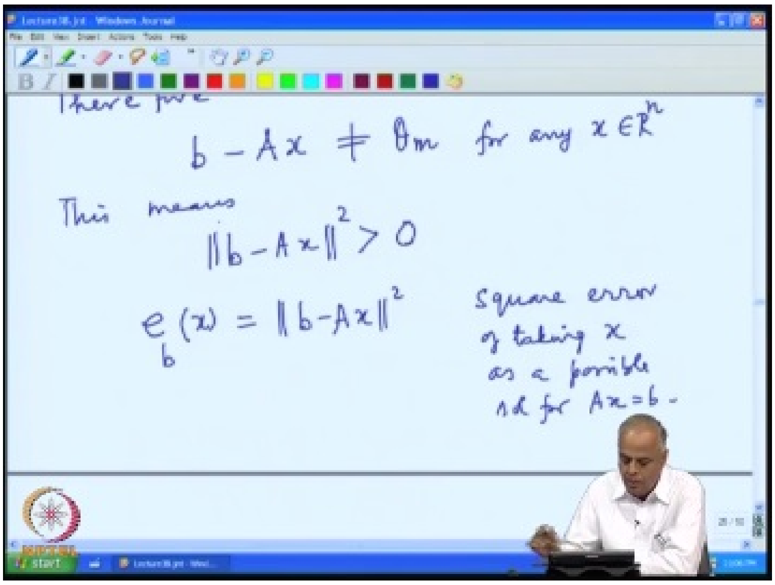
We have the R m side and we have the range of A; we have the null space of A transpose. So, the b is somewhere here falling outside the range of A; b does not belong to range of A. What does that mean? We have R n; A takes the vectors R n to R m. Therefore, if we now look at the R n side and if we take any vector x in R n, A x will always go and fall in the range of A. So, any vector in R n if we calculate A x, it will be in the range of A. But, since b is not in the range of A, b will not be equal to A x.

(Refer Slide Time: 53:14)



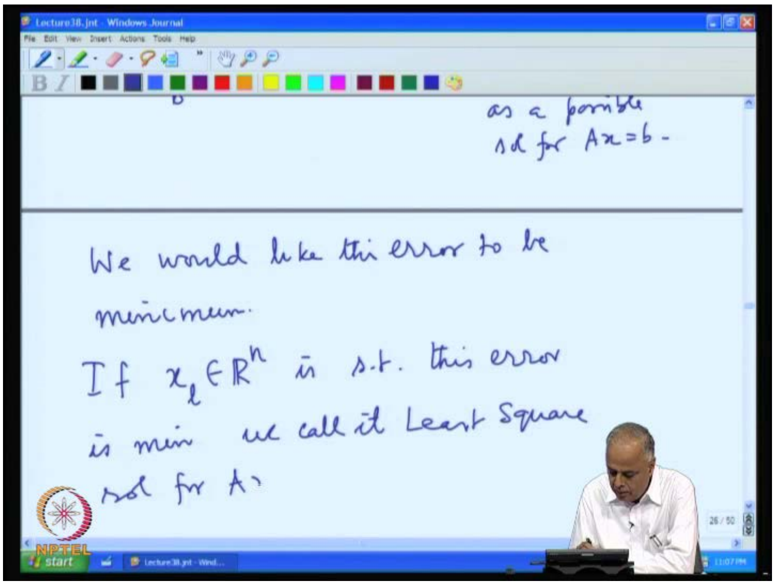
So, b does not belong to the range of A means for any x belonging to R n, we cannot get Ax equal to b. Therefore, b minus Ax will not be equal to the 0 vector for any x in R n, because the moment b becomes equal to A x, then b belongs to the range of A. So, b does not satisfy the consistency condition means b does not belong to the range of A; that can be translated into the fact that b minus Ax is not equal to theta m for any x in R n.

(Refer Slide Time: 54:17)



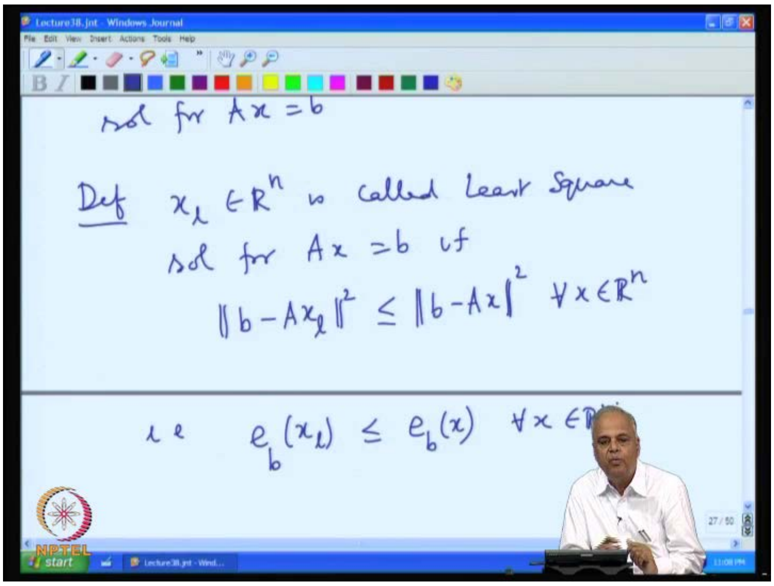
This means since it is not 0, its length squared will not be 0, and therefore, it will be strictly 0. And, this we will called as the error of… If we take x as the solution, if we had thought Ax equal to b; that is, if we had thought x as the solution of the system, we would get an error, because Ax is not equal to b; and, that error is measured by what is the difference I am getting. If I had thought Ax as the solution, I should have got b; I did not get b; so, the error is b minus Ax squared; and, that I will call as the square error e b x. So, e b x is called b minus Ax squared. And, it is called the square error of taking x as a possible solution for Ax equal to b.

(Refer Slide Time: 55:42)



Now, what we would like to do is – since we know nothing is going to go to b, we would like to get as close to b as possible. So, we would like this error to be minimum. If x belongs to R n, is such that this error is minimum – we will call it x l – we call it least square solution for Ax equal to b. Therefore, what is the notion of the least square solution?

(Refer Slide Time: 56:25)



Definition – x l belongs to R n is called the least square solution for Ax equal to b. If you look at the error taking x l as the solution, that will be less than or equal to if you take any x as the solution for every x in R n; that is, same as the error of taking x l as the solution will be always less than or equal to the error taking any other x. So, our only hope is to find least square solutions when b does not satisfy the consistency condition. We shall next see how to use the basis that we have chosen to find the structure of the least square solutions.