

Advanced Matrix Theory and Linear Algebra for Engineers

Prof. R. Vittal Rao

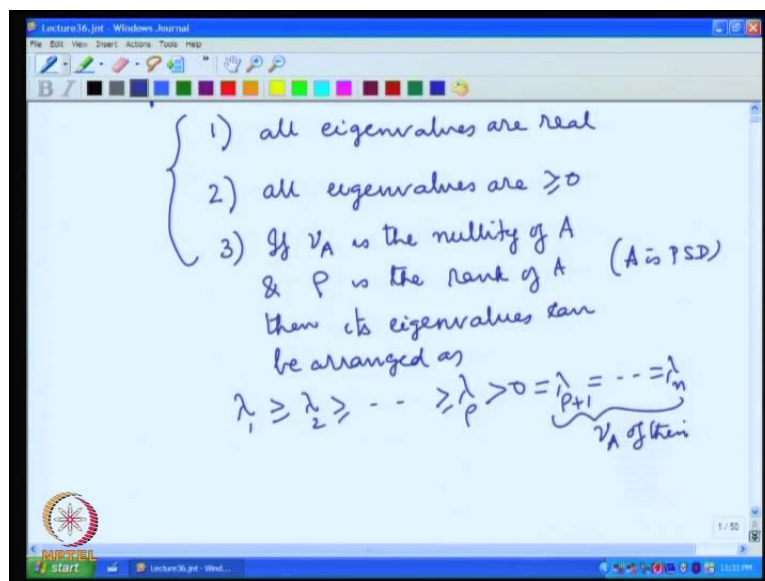
Department of Electronics Design and Technology

Indian Institute of Science, Bangalore

Lecture No. #36

Singular Value Decomposition (SVD) – Part 1

(Refer Slide Time: 00:26)



In the last lecture, we looked at the properties of a positive semi definite matrix. Now use all this ideas as we mentioned at the end of the last lecture; we should now use all the ideas that we have developed so far in this course to analyze a general given matrix. So recall the properties of a positive semi definite matrix; one - all Eigen values are real and this comes from the fact that a positive semi definite matrix is always Hermitian; the second point is that all Eigen values are greater than that or equal to zero. And this is that if ν_A is the nullity of A and ρ is the rank of A , where A is positive definite; A is positive semi definite. Then its Eigen values can be arranged as in a decreasing order Eigen value than the next smaller Eigen value and so on. These are all greater than zero; ρ of them are positive Eigen values and the remaining or all there are ν_A of them; these are all zero Eigen values.

(Refer Slide Time: 02:18)

4) Corresponding to them we have o.n. eigenvectors

v_1, v_2, \dots, v_p corresp. to $\lambda_1, \dots, \lambda_p$
(pos. eigenvalues)

$\phi_1, \phi_2, \dots, \phi_{n-p}$ corresp. to the eigenvalue 0

5) v_1, v_2, \dots, v_p Provide an o.n.b. for the range of A.

And then corresponding to this we have orthonormal Eigen vectors. We have v_1, v_2, \dots, v_p corresponding to the Eigen values $\lambda_1, \lambda_2, \dots, \lambda_p$ the positive Eigen values. So, we have the v orthonormal Eigen vectors corresponding to the positive Eigen values and then $\phi_1, \phi_2, \dots, \phi_{n-p}$ corresponding to the Eigen value zero. So, the Eigen vectors can be found and we also found that v_1, v_2, \dots, v_p provide an orthonormal basis for the range of A. So, these are all the fundamental properties of a positive semi definite matrix.

(Refer Slide Time: 03:40)

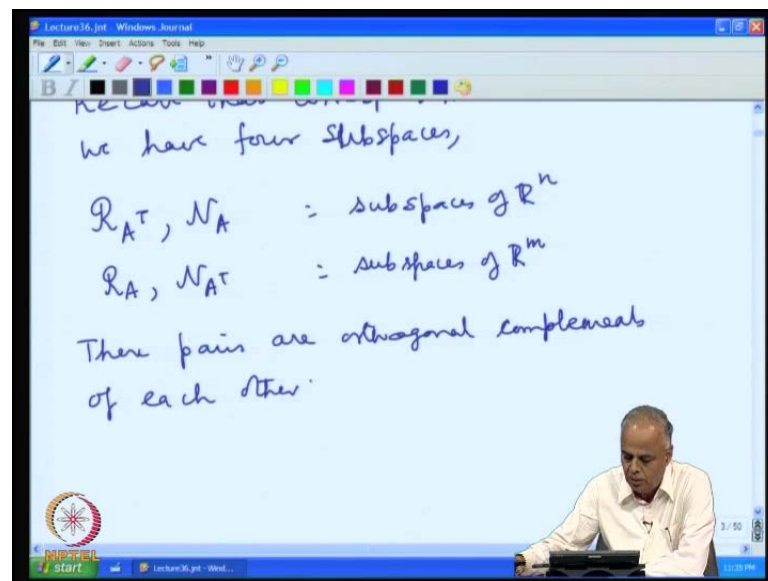
Analysis of a general $m \times n$ matrix

w.l.o.g. we shall look at $A \in \mathbb{R}^{m \times n}$

Recall that corresp. to A we have four subspaces,

Now, we shall look at a general matrix and see how we are going to use these properties to analyze a general matrix. So, let us now start with the analysis of a general m by n matrix. Now without loss of generality, we shall look at A as the real m by n matrix and we will point out what are the minor changes that we have to make whenever we deal with a complex m by n matrix.

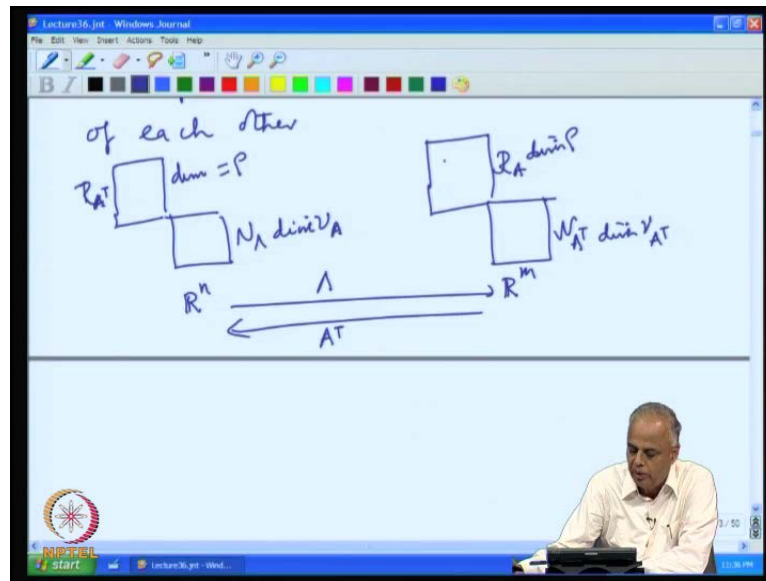
(Refer Slide Time: 04:38)



So first, we look at the real m by n matrices. Recall that corresponding to A , we have four fundamental sub spaces. Two of them namely range of A transpose and the null space of A sub spaces of \mathbb{R}^n , and the range of A and the null space of A transpose sub spaces of \mathbb{R}^m and these pairs are orthogonal complements of each other **orthogonal complements of each other**. What this means is, we have the space \mathbb{R}^n ; A maps vectors n component vectors to m component vectors and A transpose take m component vectors n component vectors and these space is... In this space, we have this two orthogonal complements. This is the range of A transpose; this is the null space of A . On this side we have the two orthogonal complements; the range of A and the null space of A .

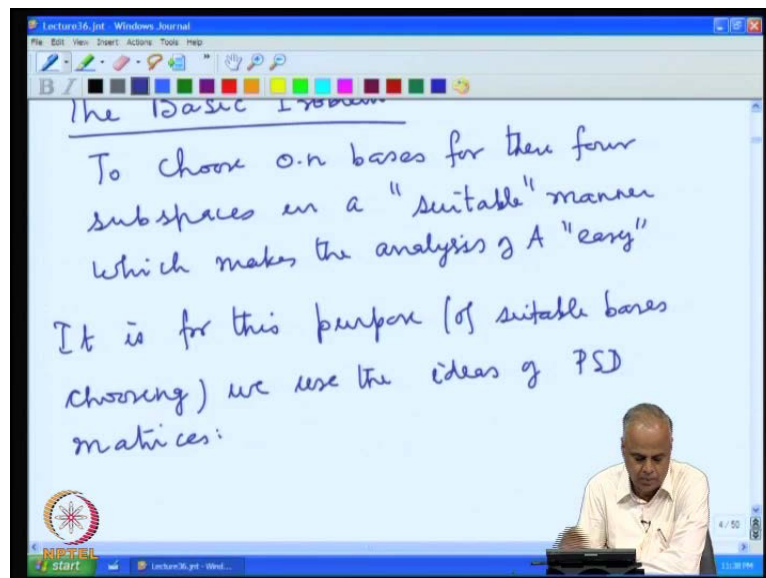
The fundamental problem is choosing suitable basis for this. Notice that the dimension of this is the dimension of the null space of dimension of the range of a transpose which is the rank of A transpose which is the same as the rank; will denote the rank of the matrix by ρ . Similarly, the dimension here is the nullity; here the dimension is again the rank; here the dimension in the dimension nullity of A transpose.

(Refer Slide Time: 05:28)



So, these are all the dimensions of these four spaces. The two spaces the range A transpose in the range A at the same dimension namely the rank of the matrix; null space of a dimension ν_A ; the null space of A transpose is the nullity of ν_{A^T} .

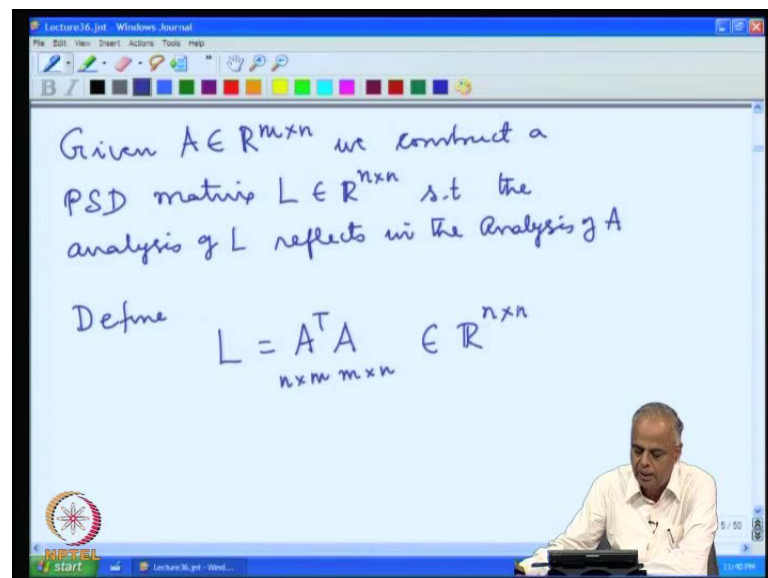
(Refer Slide Time: 06:50)



So the basic problem is to choose orthonormal bases for these four sub spaces in a suitable manner which makes the analysis of the matrix A easy. We shall see the meanings of the word suitable easy etcetera as we go along. This is the fundamental problem; the fundamental problem is to choose the sub spaces these four fundamental

sub spaces we have got them. So, we have split the two vectors spaces \mathbb{R}^n and \mathbb{R}^m into two parts each. Now in each part we are going to do the sampling namely get the basis. When we get the basis, we want it always be orthonormal. So, the computations become easy and we want to choose this basis in such a way that it makes our computations and analysis easy. So, it is for this purpose of choosing the suitable basis of suitable of suitable basis choosing. We use the ideas of positive semi definite matrices. How do we do this? That is the question. So, what we do is even though the matrix given matrix may be rectangular or it may be square we do not know whether it is hermitian it may or may not be hermitian; it may or may not be positive semi definite. What we will do is starting with the given matrix A we shall construct another matrix which is positive semi definite in such a way the analysis of the positive semi definite matrix that we construct will reflect in the analysis of the given matrix A .

(Refer Slide Time: 09:40)



So given A in $\mathbb{R}^m \times \mathbb{R}^n$, we construct a positive semi definite matrix which will call us L which is $n \times n$ such that the analysis of L reflects in the analysis of A . Now, the analysis of L will be easy because it is positive semi definite and we have seen all the properties of positive semi definite matrices. Now, how do we define this matrix L ?

(Refer Slide Time: 10:40)

$L = A^T A$
 $n \times m \quad m \times n$

1) $L \in \mathbb{R}^{n \times n}$

2) $x \in \mathbb{R}^n \Rightarrow (Lx, x) = (A^T A x, x)$
 $= (A x, A x)$
 $= \|Ax\|^2 \geq 0$

$\Rightarrow (Lx, x) \geq 0 \quad \forall x \in \mathbb{R}^n$

So, we define L to be A transpose A . Notice A is m by n and A transpose is n by m . So, the product is n by m . So, L belongs to n by m . So, first property of L is that L belongs to n cross m . So, L is an n cross n matrix. So, even though the matrix original matrix A may not be rectangular we have construct may be rectangular and may not be square, we still constructed a square matrix out of it which is L and which is an n by n square matrix. So, the moment we have a square matrix we look at its properties now. The second property that we look at is the following. Suppose, x belongs to \mathbb{R}^n then we have Lx comma x is equal to A transpose Ax comma x which is equal to (Ax, Ax) . Remember, when you move A transpose to the second factor you will go with the another transpose. So, it will become A transpose transpose, which is A which is equal to the length of Ax square which is greater than or equal to zero.

(Refer Slide Time: 12:10)

The screenshot shows a digital whiteboard with the following handwritten text:

- $\Rightarrow (Lx, x) \geq 0 \quad \forall x \in \mathbb{R}^n$
- $\Rightarrow L$ is a Positive semi definite matrix $\in \mathbb{R}^{n \times n}$
- 3) Analogously we can define $M = AA^T$. This is a PSD matrix $\in \mathbb{R}^{m \times m}$.

The video feed shows a lecturer in a white shirt sitting at a desk. The NPTEL logo is visible in the bottom left corner.

So therefore, Ax comma **sorry** therefore, Lx, x is greater than or equal to zero for every x in \mathbb{R}^n . This is precisely the meaning of the fact or this is precisely saying that L is a positive semi definite matrix. So, L is a positive semi definite matrix belonging to \mathbb{R}^n . So given an n by n m by n matrix, we can always construct a positive semi definite matrix which is n by m . Analogously, we can define m to be AA^T and this is a positive semi definite matrix in $\mathbb{R}^{m \times m}$. So given the rectangular m by n matrix, we have constructed two positive semi definite matrices. One of them is n by n ; the other one is m by n . We shall study the properties of L analogously will get the properties of m .

(Refer Slide Time: 13:38)

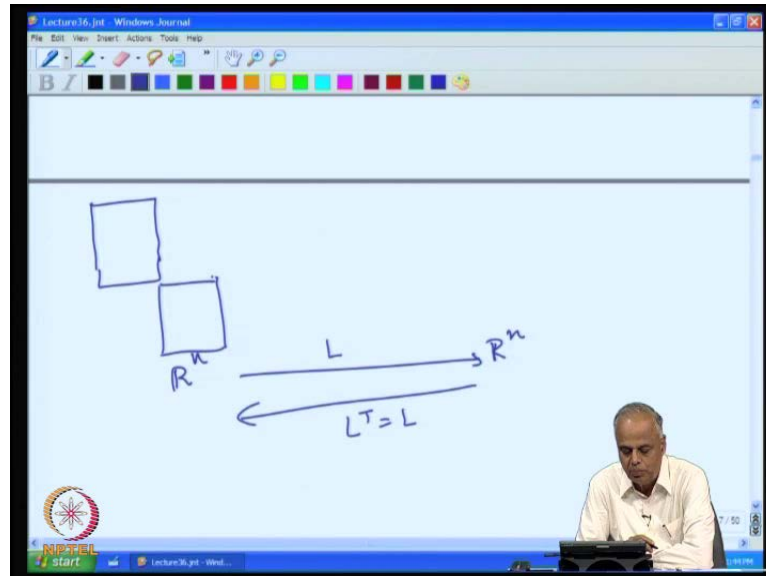
The screenshot shows a digital whiteboard with the following handwritten text and diagram:

- Given $L \in \mathbb{R}^{n \times n}$
- A diagram showing a linear transformation: $\mathbb{R}^n \xrightarrow{L} \mathbb{R}^n$

The video feed shows the same lecturer as in the previous slide. The NPTEL logo is visible in the bottom left corner.

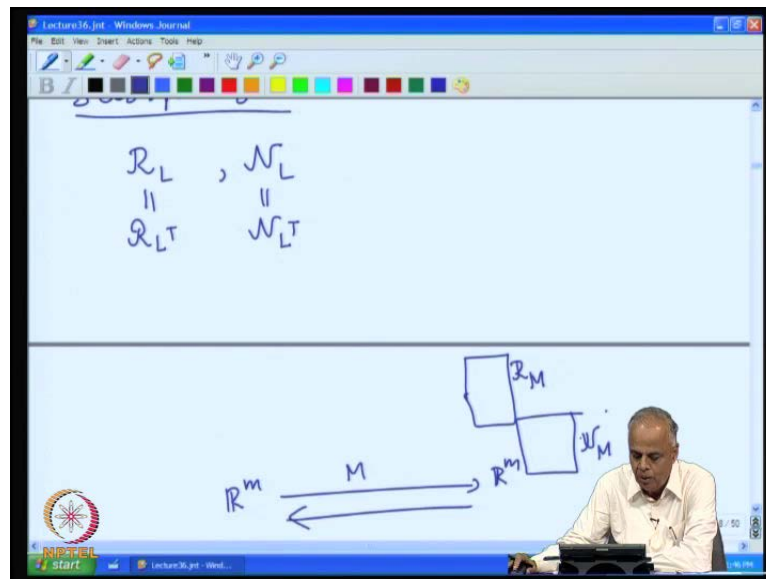
Now given L , what does we do? This is now, we have constructed this L ; now take this matrix L , what does this do? If an n by n matrix, so it maps, it takes \mathbb{R}^n vectors to \mathbb{R}^n vectors. Now what is L transpose?

(Refer Slide Time: 14:20)



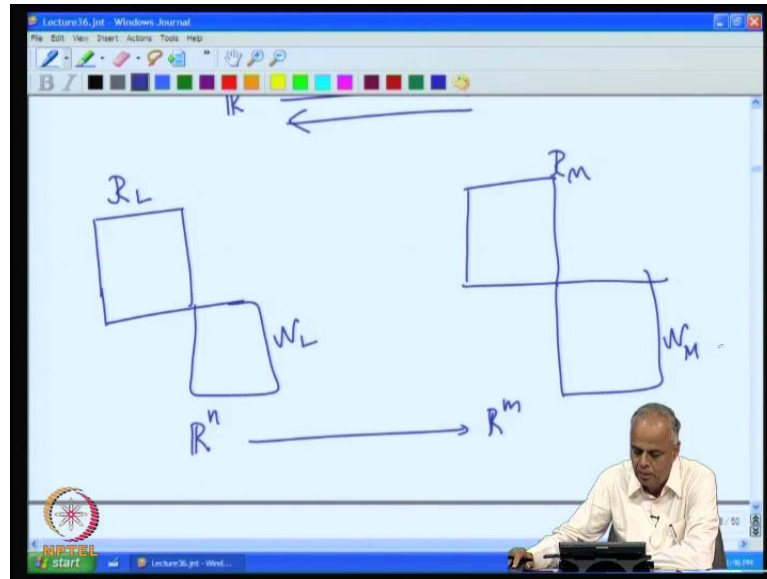
Since it is positive definite it we make sure that it is symmetric and by very definition, we see that L transpose is equal to L and therefore, L is symmetric real symmetric. So, L transpose is L ; so it is symmetric matrix. So L transpose also maps \mathbb{R}^n to \mathbb{R}^n . So, the reverse map \mathbb{R}^n to \mathbb{R}^n L transpose is the same as L so they we do not get anything new. Now corresponding to L , we must have a decomposition of \mathbb{R}^n . Remember, that the moment we have a matrix, we have the subspaces; what are the subspaces for L ?

(Refer Slide Time: 14:58)



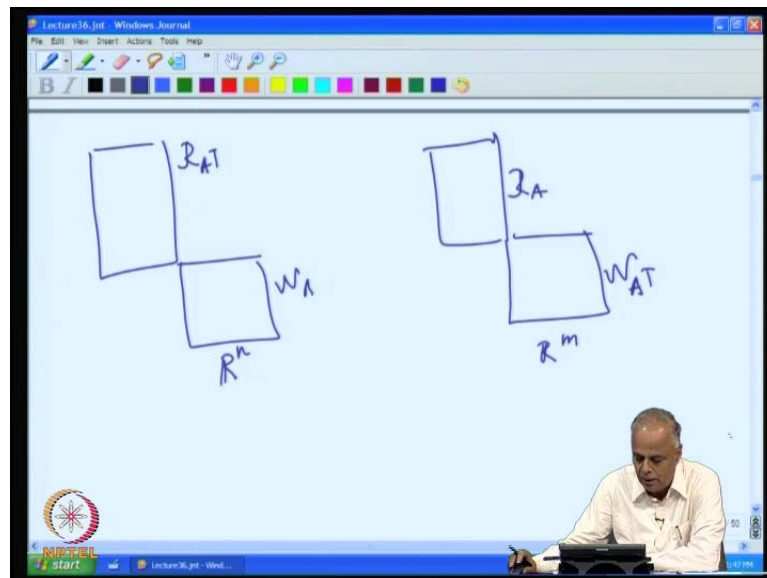
The subspaces of L or one range of L null space of L than we have to say range of L transpose, but L transpose is L so that is the same as this so nothing new. Similarly, L transpose is L ; therefore, null space of L transpose is the same as null space of L . So, there are two basic subspaces of \mathbb{R}^n ; namely the range of L and the null space of L . So, these are the two subspaces we get. Actually, I will write or I will transpose which is equal to \mathbb{R}^n . We know that the null space of any matrix is orthogonal complement of the range of the L transpose. But in this case, the transpose is the original matrix; therefore, the null space of L is the orthogonal complement of the null space of m or similarly, we get for the matrix m on this side the decomposition, because it is an n by n matrix; it decomposes \mathbb{R}^m by \mathbb{R}^m both this space sides are the same. It is the range of m transpose which is same as m everything and this is the null space of m transpose, which is the same as null space of m .

(Refer Slide Time: 16:32)



So, we have one decomposition of R^n given by L and one give decomposition of R^m given by m . So now, if you look at R^n and R^m , we have the decomposition on this side given by the range of L and the null space of L and the decomposition into orthogonal complements on this guide given by range of m and the null space of m .

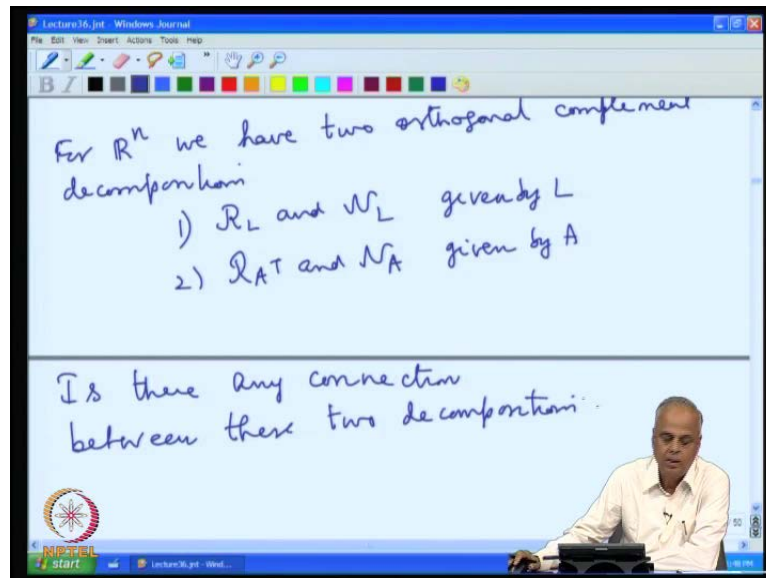
(Refer Slide Time: 16:56)



On the other hand, A also gives a decomposition and the other hand, we have R^n and R^m we have the decomposition given by A which is range of A transpose and null space of A on this side and range of A **range of A** and null space of A transpose on this side.

Now, let us look at this two pairs of decomposition and focus on \mathbb{R}^n first. On the \mathbb{R}^n side, we have one decomposition given by L ; namely R_L and N_L . Also another decomposition given by A ; namely R_{A^T} and N_A so on.

(Refer Slide Time: 17:44)



For \mathbb{R}^n we have two orthogonal complement decompositions. One of them is the range of L and N_L the null space of L given by L and the other is the range of A^T and the null space of A given by A . So, we have this two decomposition of the same space; one arising out of the matrix L and the other arising out of the matrix A , but the matrix L is connected with the matrix A , because it is defined from the matrix A starting from the matrix A we defined L as a transpose A ; so, we expect there must be some connection between these two decomposition.

(Refer Slide Time: 19:10)

To look at some properties of \mathcal{N}_L & \mathcal{R}_L

(I) $x \in \mathbb{R}^n$
 $x \in \mathcal{N}_L$ } $\Rightarrow Lx = \mathbf{0}_n \in \mathbb{R}^n$

$\Rightarrow (Lx, x) = (\mathbf{0}_n, x) = 0$
 $\Rightarrow (A^T A x, x) = 0$ ($\because L = A^T A$)
 $\Rightarrow (Ax, Ax) = 0$
 $\Rightarrow \|Ax\|^2 = 0$

So is there any connection between these two decompositions? We shall now investigate this question. So, to do thus to look at this we shall first look at some properties of the null space of L and the range of L and the range of L . So, we shall now look at the null space and the range of L . So first, the null space of L ; suppose, we have a vector x which is the null space of L . Remember, L is a n by n matrix so the null space of L is a part of \mathbb{R}^n . So, x is in \mathbb{R}^n . So x belongs to \mathbb{R}^n and x belongs to null space of L implies Lx since x is n by n x is n by 1 ; L is n by n ; Lx also belong to \mathbb{R}^n and since it is the null space, it must be the zero vector of that space. So, Lx is equal to $\mathbf{0}_n$. If Lx is equal to $\mathbf{0}_n$ that says Lx comma x the inner product of Lx with x is the same as the inner product of $\mathbf{0}_n$ with x which is zero. So that says, A transpose Ax comma x is zero because, L is define to be A transpose A .

Now in an inner product if you move the A 's from one side to the other side it gets added up with the transpose. So, remove it to the second factor, it becomes A transpose transpose; that is it becomes Ax , Ax is equal to zero. That says the length of Ax square is zero. Now, note that A is an m by n matrix; x is an n by 1 vector. So Ax is a vector in \mathbb{R}^m . Its length is zero and therefore, Ax must be the zero vector of the \mathbb{R}^m space; which means, x is in the null space of A .

(Refer Slide Time: 21:20)

$$\Rightarrow Ax = \theta_m$$
$$\Rightarrow x \in \mathcal{N}_A$$
$$\therefore \boxed{\mathcal{N}_L \subseteq \mathcal{N}_A} \quad \dots (1)$$

On the other hand

$$x \in \mathcal{N}_A \Rightarrow Ax = \theta_m$$
$$\Rightarrow A^T Ax = A^T \theta_m = \theta_n$$
$$\Rightarrow Lx = \theta_n$$

So, what we have seen is that whenever you have a vector in the null space of L, it must also be in the null space of A. So therefore, null space of L is contained in null space of A. So, we have started first look at some connection between these two pieces then, null spaces of L and A and the range of L and A A transpose. So, this is the first property; let us call it as 1. On the other hand, x belongs to null space of A implies Ax since A is m by n and x is n by 1 that is the zero vector in the m space. Now, if I multiply both sides by A transpose, I get A transpose theta m which is A is transpose is n by m; this is m by 1 so it will be the zero vector of the n space.

(Refer Slide Time: 22:46)

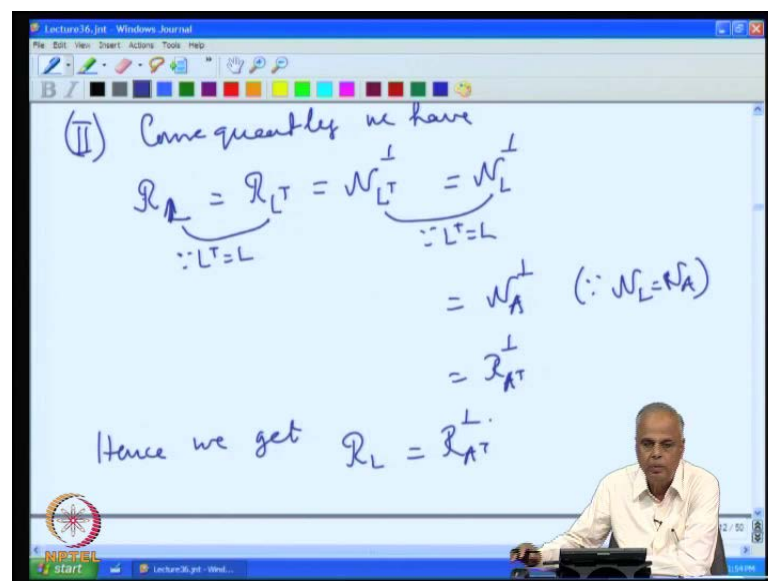
$$\Rightarrow x \in \mathcal{N}_L$$
$$\Rightarrow \boxed{\mathcal{N}_A \subseteq \mathcal{N}_L} \quad \dots (2)$$

By (1) & (2) we get

$$\mathcal{N}_L = \mathcal{N}_A$$

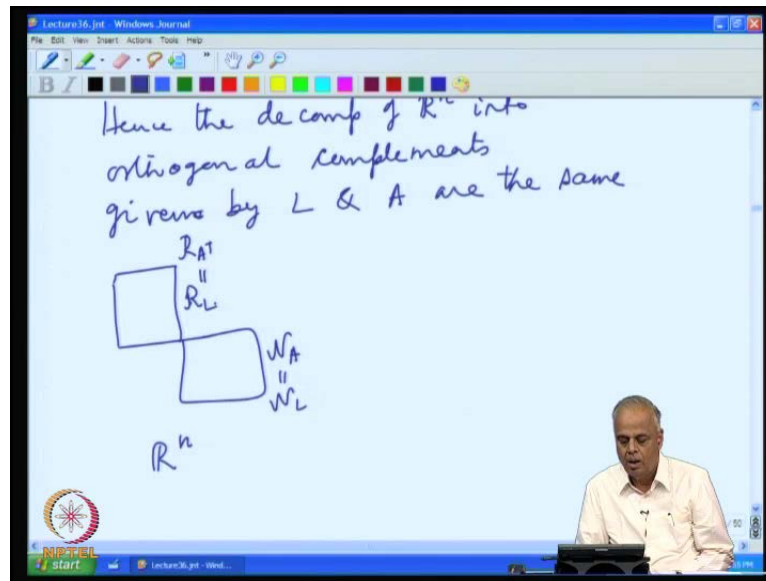
So that says, Lx is theta n because A transpose A is L . So that says, x is in the null space of L . So consequently, we have the null space of A is part of null space of L because, anything in the null space of A is also in the null space of A . Compare 1 and 2, we get **we get** null space of L is the same as null space of A . So now, if you look at this picture, the two decompositions on the R^n side this and this are the same. If these two are the same their corresponding orthogonal complements must be same. So basically, these two decompositions collapse into the same decomposition. So, we shall now put this together.

(Refer Slide Time: 23:48)



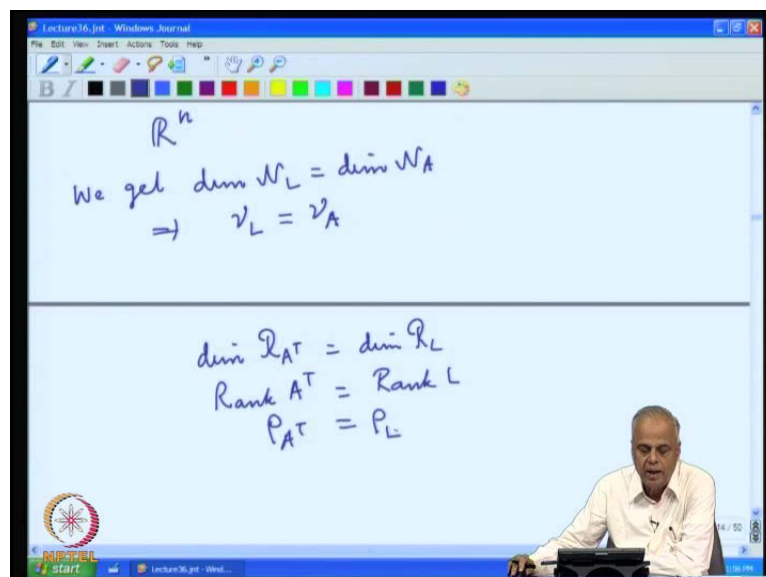
So, we have the null space of L is equal to null space of A . Consequently, we have the range of A range of L is the same as range of L transpose, but for any matrix the range of the transpose is the orthogonal complement of the null space, but the null space of L transpose is the same as null space of L perpendicular because, L is symmetric real symmetric L transpose is equal to L . This is what we have used here **this is what we use here** because L transpose is equal to L and because L transpose is equal to L . Now, $N L$ perp is the same as $N A$ perp, because we are just seen that $N L$ is the same as $N A$, but $N A$ perp is the range of A transpose perp.

(Refer Slide Time: 24:54)



So therefore, hence we get range of L is the same as range of A transpose perp. So hence, the decomposition of \mathbb{R}^n into orthogonal complements given by **given by** L and A are the same. So, we have this picture on the \mathbb{R}^n site. We have we call it null space of A is the same as null space of L and we have the range of A transpose which is the same as the range of n

(Refer Slide Time: 25:42)



Consequently, we get dimension of $N L$ is the same as dimension of $N A$, but the dimension of $N L$ is the nullity of L and the dimension of $N A$ is the nullity of A . So, nullity of L is the same of nullity of A .

(Refer Slide Time: 26:34)

The image shows a whiteboard with the following handwritten mathematical derivations:

$$\begin{aligned} \dim R_{A^T} &= \dim R_L \\ \Rightarrow \text{Rank } A^T &= \text{Rank } L \\ \Rightarrow P_{A^T} &= P_L \\ \Rightarrow P_A &= P_L \end{aligned}$$

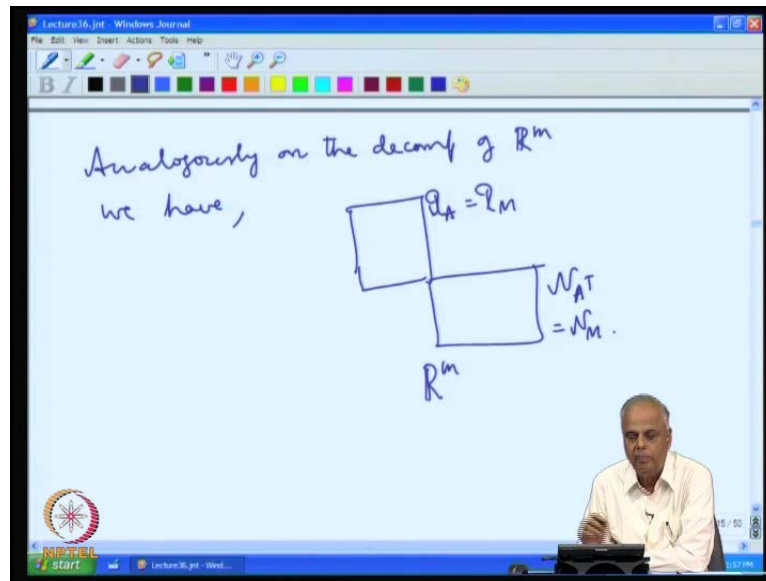
Below this, the following equations are boxed:

$$\left\| \begin{aligned} V_A &= V_L \\ P_A &= P_L (= P_{A^T}) \end{aligned} \right\|$$

Similarly, because $R A$ transpose is equal to $R A R L$, we have dimension of the range of A transpose. If the dimension of the range of L the dimension of range A transpose is the rank of A transpose and this is the rank of L , but rank of A , which we denote by ρA transpose, but the rank of A transpose is the same as rank of A which we have seen. So, the rank of A is equal to rank of L .

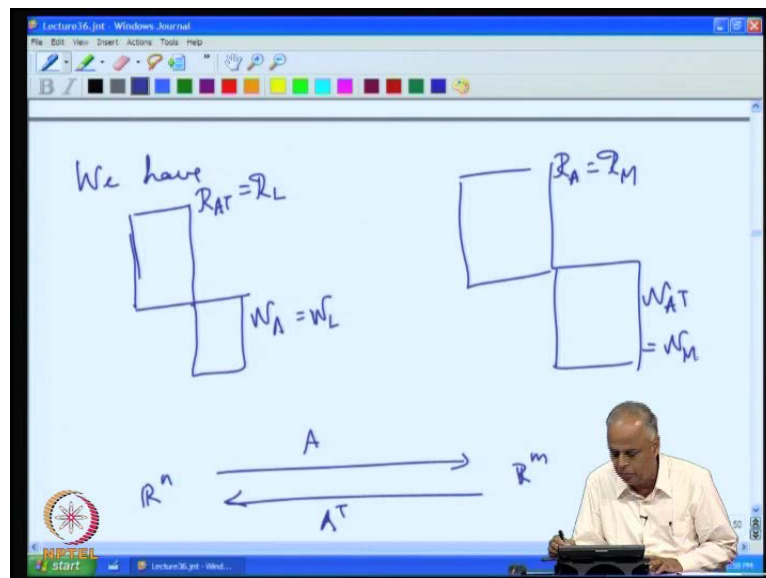
So, the L at the same rank as A and same nullity as A ; so, what are the conclusions that we have? The nullity of A is the nullity of L ; the rank of L at the rank of A is equal to rank of L is also equal to rank of A transpose. So, these are the properties that the fact that the decomposition given by the A and the by given by L of $R n$ is the same.

(Refer Slide Time: 27:14)



Analogously on the decomposition of R^m we have on the R^m site, we have the null space the range of A which will be same as range of m now. Remember, m is AA^T transpose and the null space of A transpose will be the same as null space of m .

(Refer Slide Time: 28:02)

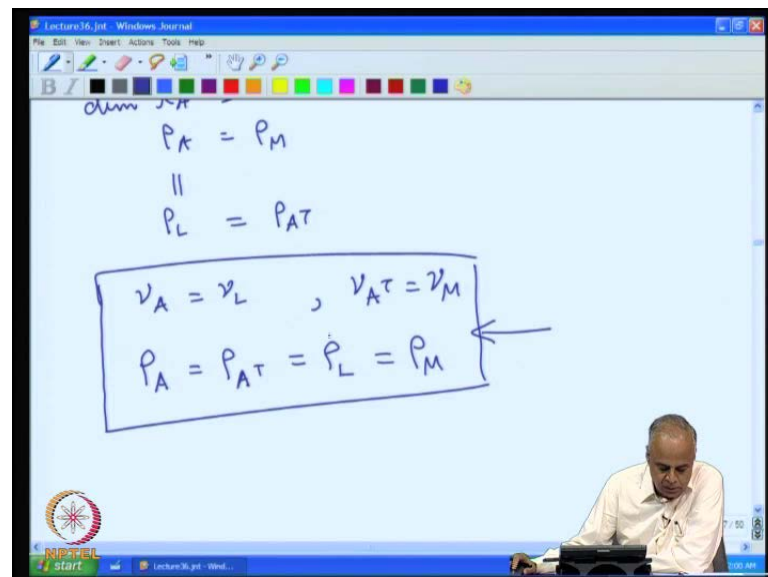


So putting all these pictures together, we have R^n R^m the two spaces; A goes this side; A transpose goes this side and on the R^n site we have the decomposition range of A transpose which is the same as range of L . Null space of A transpose which is same as null space of **null space of** A which is the same as null space of L . On this side, we have

the range of A which is the same as range of m and null space of A transpose which is the same as null space of m. Again you looking at the right hand side, we see that the dimension of the null space of A transpose is the same as the dimension of the null space of m which gives as the nullity of A transpose is equal to nullity of m.

Similarly, if we look at the decomposition of the R m, we get dimension of range of A is the same as dimension of range of m. This is the rank of A, which is the same as rank of m. So, we have seen that the rank of A the same of rank of L, which is the same as rank of A transpose.

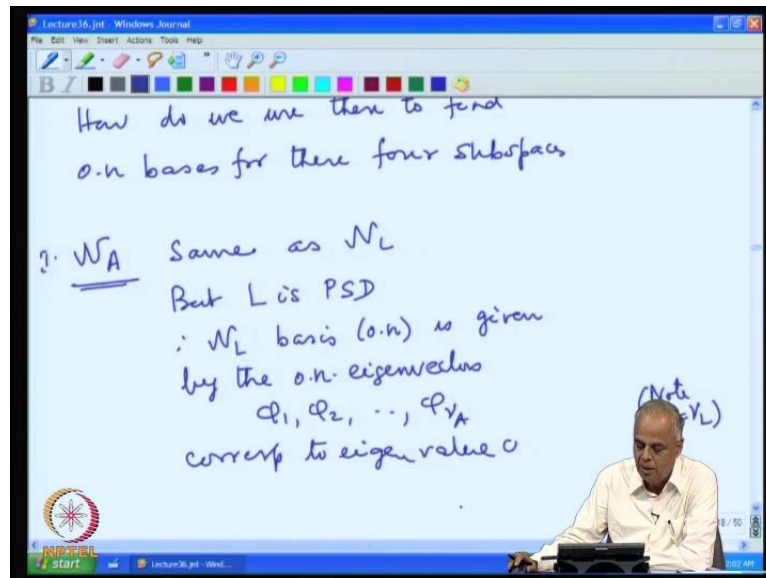
(Refer Slide Time: 28:44)



So we therefore have the following situations. The nullity of A is nullity of L; the nullity of A transpose is equal to the nullity of m. The rank of A is the same as rank of A transpose with same of rank of L the same of rank of m. All these matrices AA transpose L and m share the same rank, whereas the nullities will depend on A and A transpose; L will have the same nullity as A and m will have the same nullity of A transpose. Could these are the four fundamental relations. The relation between the sub spaces is described in this picture. The range of A transpose is same of range of L; null space of A is the same as null space of L; the range of A is the same as range of m; the null space of A transpose is the same as null space of m.

This give rise to the following relationship between the nullities and the rank of these various matrices; important think to notice that AA^T L m all share the same rank that number associate with the matrix, this is the very important number.

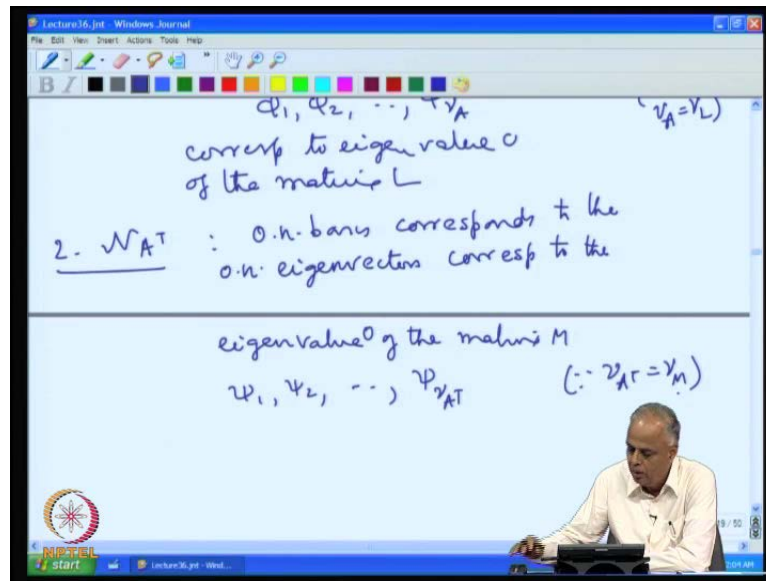
(Refer Slide Time: 31:06)



Now, let us see how we are going to exploit to find the basis for the four sub spaces now. So, how do we use these to find orthonormal basis for these four spaces that is the main question. Now, let us look at first the null space of A same as null space of L , but L is positive semi definite and therefore, null space of L basis orthonormal is given by the orthonormal Eigen vectors; let us call them ϕ_1, ϕ_2 the nullities ν_A is the same as ν_L . So, instead of writing ν_L , we will write ν_A . Note $\nu_A = \nu_L$.

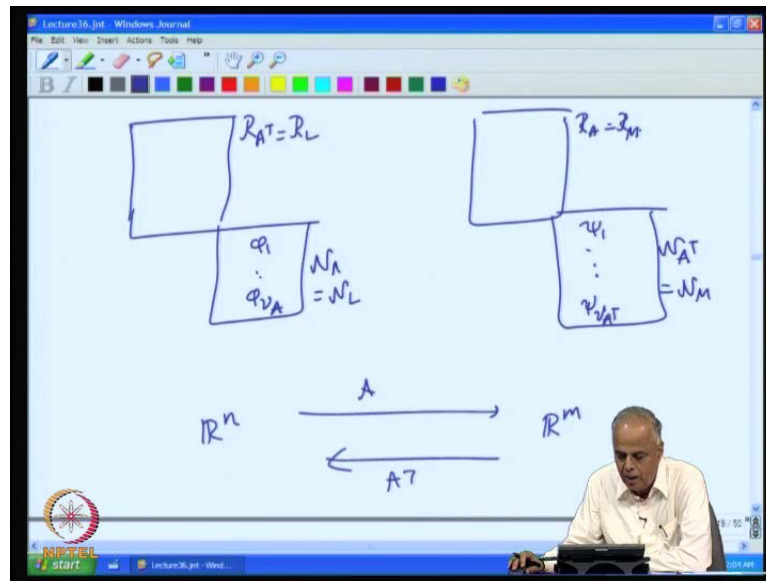
Orthonormal Eigen vectors corresponding to Eigen value zero. Recall that we have seen that the null space basis for a positive semi definite matrix corresponds to the Eigen vector sub zero. So, we can find the orthonormal basis for N_A through the positive semi definite matrix L and from its Eigen vectors corresponding to the Eigen value zero, we can find the orthonormal basis for the null space of A . So, we have finish one-fourth of our job; namely finding the Eigen was orthonormal basis for N_A we are able to get.

(Refer Slide Time: 33:04)



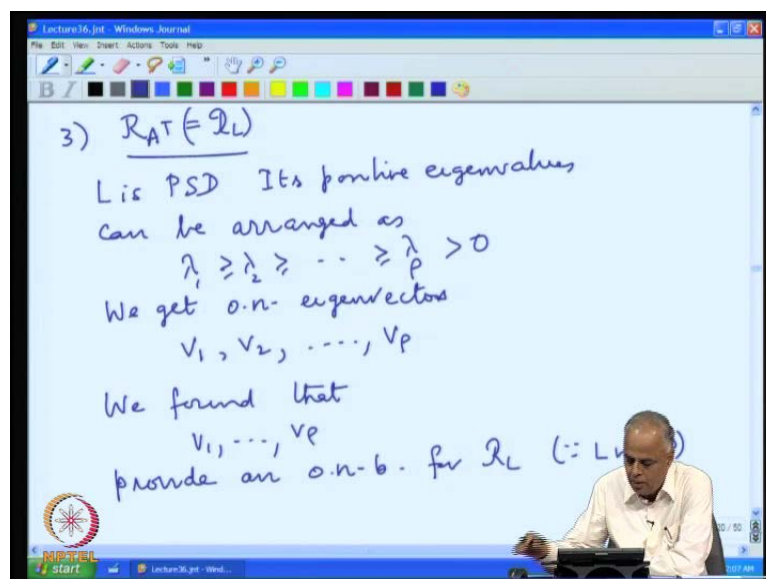
Similarly, we look at $N A$ transpose orthonormal basis corresponds to the orthonormal Eigen vectors corresponding to the Eigen value zero of the matrix m ; here we should write here it was L of the matrix L . So, we have now for we have now seen out of this four sub spaces the this the null space part here can be founded from the zero Eigen vector of L , and the null space part on this side can be found out from the zero vector **correspond** zero Eigen value corresponding to the matrix m . So, let us denote this basis by ψ_1, ψ_2 and it will have $nu A$ transpose vectors because, $nu A$ transpose is equal to $nu m$.

(Refer Slide Time: 34:20)



So now, you have this picture. Again let us keep on writing this picture until we complete this whole thing with the basis; we have \mathbb{R}^n ; we have \mathbb{R}^m ; we have A going this way; A transpose going this way and this is the null space of A which is the same as null space of L . Now, we have found a basis for this and then, this is the null space of A transpose which is the same as null space of m , we have found a basis for this. So now, our job is to find the basis for these two fellows. This is the range of A transpose which is the range of L ; this is range of A which is the range of m . So, these are two things that we have to find.

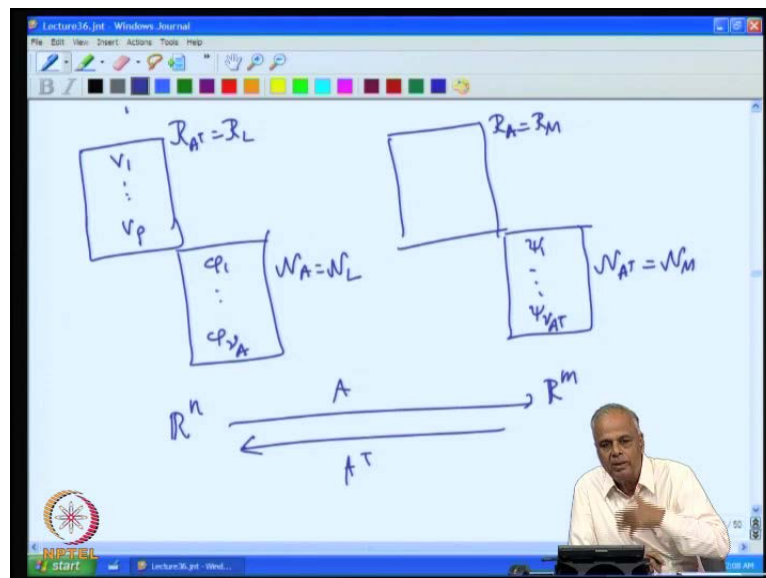
(Refer Slide Time: 35:14)



Now, let us look at the range of A transpose which is the same as the range of L . Now, L is positive semi definite; its positive Eigen values can be arranged as λ_1 greater than or equal to λ_2 and how many of them will be there that will exactly be the rank of the matrix A . We will not write ρ sub A because, ρ sub A is the same as ρ sub A transpose the same as ρ L is the same as ρ m . This all of them share the same rank; we will not distinguish and simply write; all these are the Eigen values greater than zero and the zero Eigen value has been taken care of here then finding the null space and we get we know that corresponding to the positive Eigen values, we get orthonormal Eigen vectors when we analyzed positive semi definite matrices, we found that corresponding to the positive Eigen values we get orthonormal Eigen vectors V_1 corresponding to λ_1 ; V_2 corresponding to λ_2 ; V_ρ corresponding to λ_ρ .

So we found that at the end of the last lecture, we found that the V_1, V_2, V_ρ which are the orthonormal Eigen vectors corresponding to the positive Eigen values of a positive semi definite matrix give rise to a basis for the range of that positive semi definite matrix. So, we found that V_1, V_2, V_ρ provide an orthonormal basis for range of L because, L is positive semi definite, but the range of L is the same as range of A transpose.

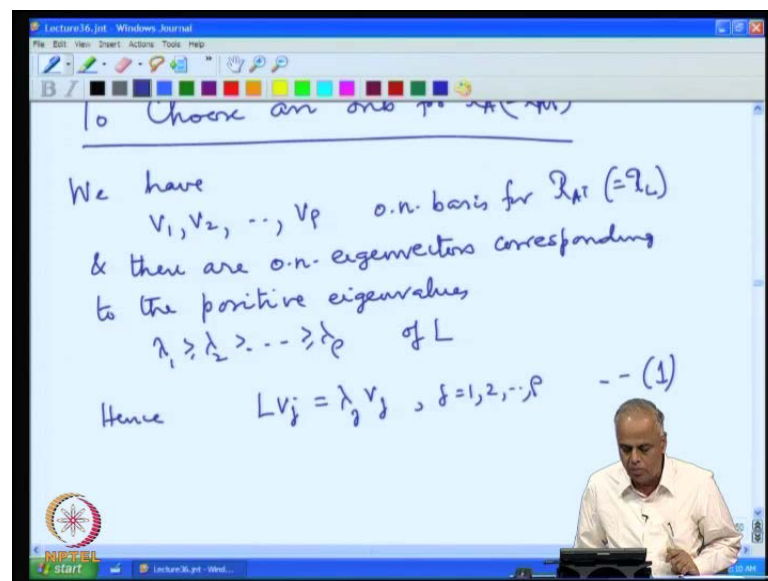
(Refer Slide Time: 37:35)



So now, we have the picture which is progressing L in the sense that we had R^n ; we have R^m and A goes this way and A transpose goes this way and we have the range of A transpose which is the same as range of L , and we have the null space of A which is the same as null space of L , and for this we found the basis.

This comes from the Eigen value zero of L and now we found the orthonormal basis V_1, V_2, \dots, V_ρ which come from the positive Eigen values L and the corresponding orthonormal Eigen vectors and on this side, we had the null space of A transpose which is the same as null space of m we found orthonormal basis which corresponds to the zero Eigen value of m and the Eigen vectors orthonormal Eigen vectors corresponding to the zero Eigen values of m . Now, the only thing that remains is finding an orthonormal basis for the range of A or the range of m . Now, we will show a clever method of choosing this in order that this is where the cleverness of choosing the basis come will choose the basis for the range of A in such a way that from then on our analysis becomes easy. Now, how do you choose this basis?

(Refer Slide Time: 39:12)



So, we have now the last part is to choose an orthonormal basis for range of A which is the same as range of m . Once we choose that, we have the four subspaces for orthogonal subspaces two on the one side, and two on the other side. We have the orthonormal basis from them and then we shall analyze everything in terms of this orthonormal basis. Now, we have V_1, V_2, \dots, V_ρ orthonormal basis for range of A transpose which is the same as

range of L , and these are orthonormal Eigen vectors corresponding to the positive Eigen values $\lambda_1 > \lambda_2 \geq \dots \geq \lambda_r$ of L and therefore, hence $L v_j$ is equal to $\lambda_j v_j$ for j equal to $1, 2, \dots, r$. Let us call this equation as 1. So, we have this r Eigen vectors of L , which form an orthonormal basis for the range of A^T .

(Refer Slide Time: 41:04)

We look at
 $AV_j \in \mathbb{R}^m$
 $m \times n \quad n \times 1$

 $AV_j \in \mathcal{R}_A, j=1,2,\dots,r$

Define $W_j = AV_j, j=1,2,\dots,r$

Then $W_j \in \mathcal{R}_A$

Now, we define we look at what happens to v_j under A so look at AV_j . Now, v_j is in \mathbb{R}^n so it is n by 1 ; A is m by n so that belongs to \mathbb{R}^m . So certainly, this a vector in \mathbb{R}^m . Secondly, it is of the form A of something so it belongs to the range of A . So, AV_j are all in range of A ; j equal to $1, 2, \dots, r$. Any vector of the form A of some vector must be in the range of A so these vectors are AV_j . So, if we define W_j equal to AV_j ; j equal to $1, 2, \dots, r$ then W_j all belong to range of A . So, we have captured some vectors in the range of A and if you are very lucky, they may even form an orthonormal basis for the range of A . Now, let us see whether they do this.

(Refer Slide Time: 42:22)

Will w_j form a basis for $\text{range}(A)$?
" " " an onb for $\text{range}(A)$?

We look at these w_j

We have

$$(w_j, w_r) = (AV_j, AV_r)$$
$$= (V_j, A^T AV_r)$$
$$= (V_j, LV_r)$$

Now, will w_j form a basis for range of A ? Will w_j form an orthonormal basis for range of A ? If so, our search for the basis for the range of A is over; so let us look at w_j . So, we look at these w_j . We have w_j comma w_r the inner product of w_j with w_r ; w_j by definition was AV_j ; w_j is AV_j . Similarly, w_r is AV_r . So therefore, w_j, w_r inner product is the same as $AV_j AV_r$ inner product. Once again, we observe that in the real matrix case when we move the matrix from the inner product from one factor to another factor it goes with the transpose. So, it becomes $A^T AV_r$. Now, $A^T A$ by definition is L so that is LV_r . Now recall that this v_1, v_2, v_r are Eigen vectors corresponding to the positive Eigen values and therefore, LV_r will be equal to $\lambda_r v_r$.

(Refer Slide Time: 44:14)

The screenshot shows a whiteboard with the following content:

$$\begin{aligned} &= (v_j, \lambda_r v_r) \\ &= \lambda_r (v_j, v_r) \\ &= \begin{cases} 0 & \text{if } j \neq r \\ \lambda_j & \text{if } j = r \end{cases} \end{aligned}$$

What this says is that w_1, w_2, \dots, w_p are orthogonal to each other

and $\|w_j\|^2 = \lambda_j$

The whiteboard is part of a software application titled "Lecture16.int - Windows Journal". The interface includes a menu bar (File, Edit, View, Insert, Actions, Tools, Help), a toolbar with various drawing tools, and a color palette. In the bottom right corner, there is a small inset video of a man in a white shirt sitting at a desk. The NPTEL logo is visible in the bottom left corner of the whiteboard area.

So, this will be equal to $v_j \lambda_r v_r$. Now, when we pull out a constant λ_r from the second factor of the inner product everything is real. The Eigen values are all real because positive semi definite so we can pull it out as $v_j \lambda_r v_r$. So, $w_j \lambda_r v_r$ is the same as $\lambda_r v_j \lambda_r v_r$, but this v_1, v_2, v_j are orthonormal and therefore, v_j, v_r will be equal to zero if j is not equal to r and 1 if j equal to r . So therefore, this will be zero if j is not equal to r and when j equal to r , $v_j v_r$ is 1 because, every vector has length one and multiplied by λ_r . So, it is equal to λ_r if j equal to r ; λ_r or λ_j which ever we want to write pull it.

So therefore, this fact says w_j, w_1, w_2, w_r are orthogonal to each other if we take any two of them **any two of them** there are orthogonal, but they do not have length 1.

(Refer Slide Time: 46:32)

What this says is
 w_1, w_2, \dots, w_p are orthogonal to each other
and $\|w_j\|^2 = \lambda_j \Rightarrow \|w_j\| = \sqrt{\lambda_j}$

Define $u_j = \frac{w_j}{\|w_j\|} = \frac{w_j}{\sqrt{\lambda_j}}$ when we take Positive sq. root of λ_j

So, what this says is, that w_1, w_2, w_3 are orthogonal to each other and the length of each one of them is equal to squared is λ_j because $w_j^T w_j$ is λ_j in our calculate. Put r equal to j you get $w_j^T w_j$ and that is equal to λ_j . So, w_j length of w_j squared is λ_j . If we have a set of orthogonal vectors and if you divide each one of them by its corresponding length you automatically get an orthonormal set of vectors.

(Refer Slide Time: 47:34)

Define $u_j = \frac{w_j}{\|w_j\|} = \frac{w_j}{\sqrt{\lambda_j}}$ when we take Positive sq. root of λ_j

We denote by $\delta_j = \sqrt{\lambda_j}$ Singular Values of A

$u_j = \frac{w_j}{\delta_j}$

u_1, u_2, \dots, u_p are o.n. vectors in \mathbb{R}^n
But \mathbb{R}^n has dimension p

So now, we define u_j to be equal to W_j by its length; its length is W_j by λ_j square root because, the length of j square W_j squared is λ_j . So, the length of W_j is square root of λ_j . Now recall that these λ_j are the positive Eigen value that we have talking about and therefore, there is no problem about the taking the square root of positive quantities. Now, the question is which square root we take, as a convention we take the non negative or the positive square root.

So, where we take positive square root of λ_j , we have to take the positive square root because we want it to be length of W_j . So now, we denote by s_j square root of λ_j and these are called the singular values of A . The s_1, s_2, \dots, s_ρ are called the singular values of A .

(Refer Slide Time: 48:36)

u_1, u_2, \dots, u_ρ are o.n.-vectors in \mathbb{R}^n
 But \mathbb{R}^n has dimension n
 $\Rightarrow u_1, u_2, \dots, u_\rho$ form an o.n.-basis for \mathbb{R}^n .

Note: $u_j = \frac{W_j}{s_j} = \frac{AV_j}{s_j}$
 $\Rightarrow AV_j = s_j u_j \quad j=1, 2, \dots, \rho$

So now, what do we have we have u_j therefore, is W_j by s_j then u_1, u_2, \dots, u_ρ are orthonormal vectors in range of A , but range of A has dimension ρ . So, if we have a ρ dimensional space and we have ρ orthonormal vectors; that means, they form an orthonormal basis **form an orthonormal basis** for range of A . So, we also got the orthonormal basis for range of A . Note we have u_j is equal to W_j by s_j ; by W_j by definition was AV_j and therefore, we get AV_j is equal to $s_j u_j$; part j equal to 1, 2, \dots, ρ .

(Refer Slide Time: 49:16)

We have $u_j \in \mathbb{R}^m$

$$A^T u_j = \frac{A^T A v_j}{\lambda_j} = \frac{L v_j}{\lambda_j} = \frac{\lambda_j v_j}{\lambda_j}$$

$$= \sqrt{\lambda_j} v_j \quad (\because \lambda_j = \sqrt{\lambda_j}^2)$$

$$= \lambda_j v_j$$

$A^T u_j = \lambda_j v_j$

On the other hand, we have u_j belongs to \mathbb{R}^m ; therefore, we can take A transpose u_j ; A transpose u_j is the same as a transpose AV_j by s_j because, u_j is AV_j by s_j . Now A transpose A is L ; A transpose A is V_j is L . So, A transpose A is V_j is L , but V_j is an Eigen vector corresponding to the Eigen value. So, it is $\lambda_j V_j$ by s_j , but s_j is square root of λ_j . So, it is just square root of λ_j V_j because s_j is equal to square root of λ_j , but that is the same as $s_j V_j$. So therefore, we have a transpose u_j is $s_j V_j$. So, what is the structure now?

(Refer Slide Time: 50:28)

$R_A^T = R_L$

$R_A = R_M$

$W_A = W_L$

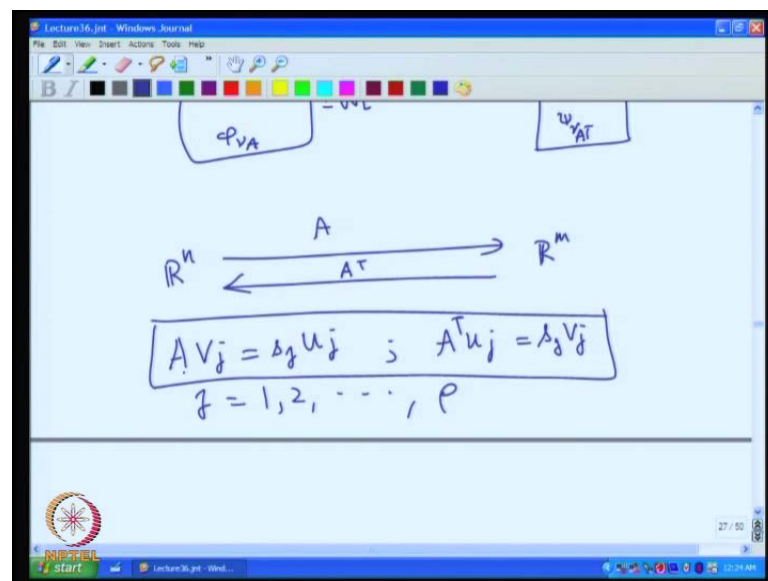
$W_A^T = W_M$

$\mathbb{R}^n \xrightarrow{A} \mathbb{R}^m$

$\mathbb{R}^m \xrightarrow{A^T} \mathbb{R}^n$

So, we have this \mathbb{R}^n ; we have the m dimensional space; A is a mapping from \mathbb{R}^n to \mathbb{R}^m ; A transpose takes \mathbb{R}^m vectors to \mathbb{R}^n . We have the subspaces here range of A transpose which is the same as range of A ; null space of A which is the same as null space of A transpose. We have the orthonormal basis V_1, V_2, \dots, V_p ; we have the orthonormal basis u_1, u_2, \dots, u_p and then, we have on this side the range of A which is same as range of A transpose; null space of A transpose which is the same as null space of A . We have the orthonormal basis u_1, u_2, \dots, u_p and $\psi_1, \psi_2, \dots, \psi_{m-p}$.

(Refer Slide Time: 51:24)



And the fundamental relation is that AV_j is $s_j u_j$ and $A^T u_j$ is $s_j V_j$; what this means is that the basis vector V_1 here; the basis vector V_1 on this side goes to the same direction as the basis vector u_1 with the scaling factor s_1 because, AV_j has the scaling factor s_j . So, V_1 goes to the u_1 direction with the scaling factor s_j ; V_2 goes to the u_2 direction with the scaling factor s_2 ; and V_p goes to the u_p direction with scaling factor s_p . So, basis directions are match to basis direction expect that there is a certain dilation taking place in each one of these basis direction.

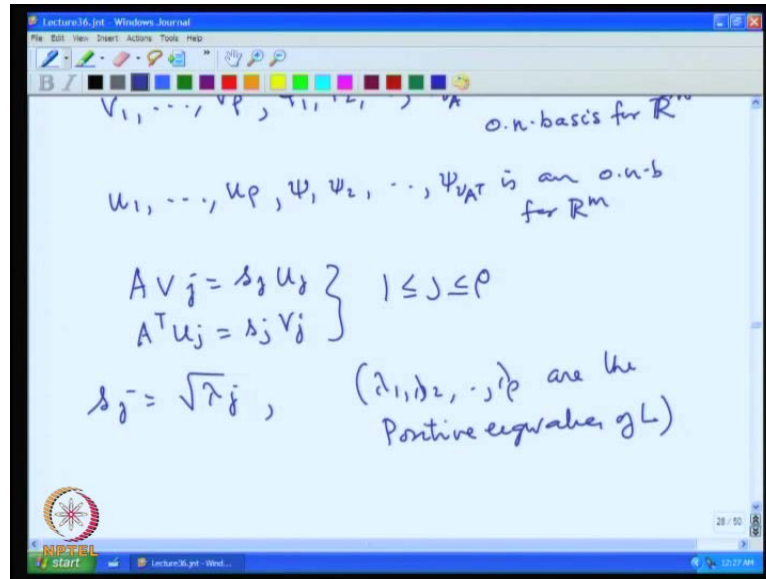
So, V_j is put to the u_j direction with the scaling s_j ; V_1 is put to the u_1 direction with the scaling s_1 ; V_2 is put to the u_2 direction with the scaling s_2 and so on. So, this is for j equal to 1, 2 up to p ; these connection between these. So, now we have constructed a basis for the range of A transpose and the range of A in such a way that the basis direction are connected that is the first basis vector goes to the same at the direction

of the first basis vector on \mathbb{R}^m side; second basis vector of range of A transpose go to the same direction as the second basis of the range of A and so on and so forth. So, the direction the basis let us map to basis expect there is a scaling factor. So, in our to get the orthonormalization we have to do this scaling factor. So, thus we have using the positive definiteness of the matrices L and m we have constructed, we have been able to get the orthonormal basis for these four subspaces.

Now, all of them come out have the Eigen vectors as the orthonormal Eigen vectors corresponding to the positive semi definite matrix L on this side and the positive semi definite matrix on the m on that side. So, all the competitions therefore are reduced to the competition of a positive semi definite matrix and therefore, whether the given matrix is rectangular or square; whether it is even if it is square whether it is hermitian or not, we can always construct a positive definite matrices semi definite matrices L and m starting from the given matrix A from which we can construct the orthonormal basis for all the four subspaces that we want and the orthonormal basis for the range of A transpose in the range of A or constructed in such a way they are link to each other.

The moment we know the orthonormal basis V_1, V_2, V_ρ for the range of A transpose, we can extract the orthonormal basis u_1, u_2, u_ρ from this relation AV_j equal to $s_j u_j$. So, thus we have chosen certain basis in a **in a** connected manner and we shall now see how we are going to and use all this basis to analyze a given matrix. As a feel for it the idea is we are trying to solve a system of equation Ax equal to b ; the vector b lies on the \mathbb{R}^m side; since it lies on the \mathbb{R}^m side that is the known vector. It lies on the r m side; it can be expanded in terms of the orthonormal basis u_1, u_2, u_ρ ψ_1, ψ_2, \dots νA transpose. The unknown vector x we have trying to find out lies on this side and therefore, it can be expanded in terms of this orthonormal basis V_1, V_2, V_ρ so and $\phi_1 \phi_2$.

(Refer Slide Time: 56:10)



So, we make that the observation, we have $V_1, V_2, \dots, V_\rho, \psi_1, \psi_2, \dots, \psi_{n-\rho}$ is an orthonormal basis for the space \mathbb{R}^n and $u_1, u_2, \dots, u_\rho, \psi_1, \psi_2, \dots, \psi_{m-\rho}$ is an orthonormal basis for \mathbb{R}^m and the fundamental relation is $AV_j = s_j u_j$ and $A^T u_j = s_j V_j$ for $1 \leq j \leq \rho$ and $s_j = \sqrt{\lambda_j}$ and $(\lambda_1, \lambda_2, \dots, \lambda_\rho)$ are the positive Eigen values of L . Now, we shall see how we use this four basis to analyze a given general matrix? And this analysis, we will begin in the next lecture.