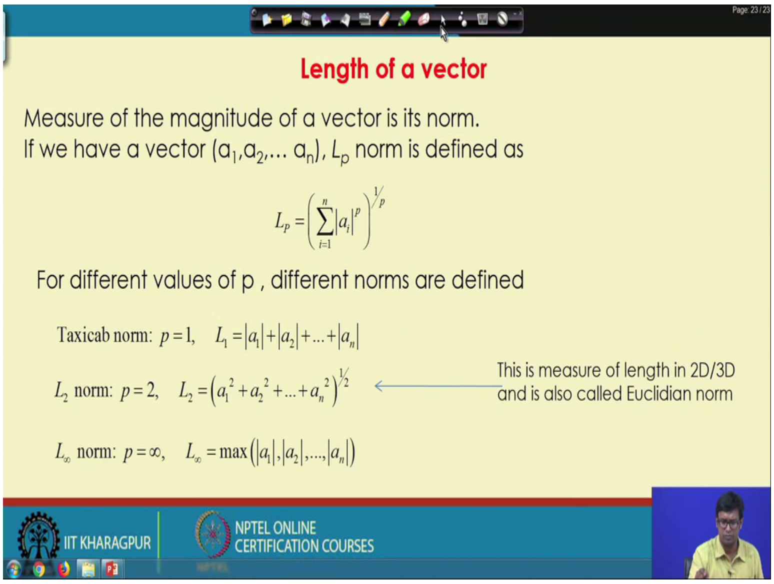
Matrix Solvers
Prof. Somnath Roy
Department of Mechanical Engineering
Indian Institute of Technology, Kharagpur

Lecture – 22
Orthogonality between the Subspaces

Welcome. In last class we took an example of rectangular matrix and looked into different subspaces associated with it and, observed that null space and row space, a perpendicular to each other. In a sense, any vector belonging to leave now space is perpendicular to any vector in the row space, and same for column space and left null space. Also made a comment that if a vector right hand side vector b has some component along left null space, it will be away from it will coming out of the column space. And therefore, Ax is equal to b cannot be solved.

So, then this the these are due to the orthogonality between the spaces, null space, row space, left null space and column space. We will explore it in more detail and we will start looking into what is an orthogonal, what are orthogonal vectors, what is orthogonality, and look into the property of orthogonality between the fundamental subspaces.

(Refer Slide Time: 01:24)

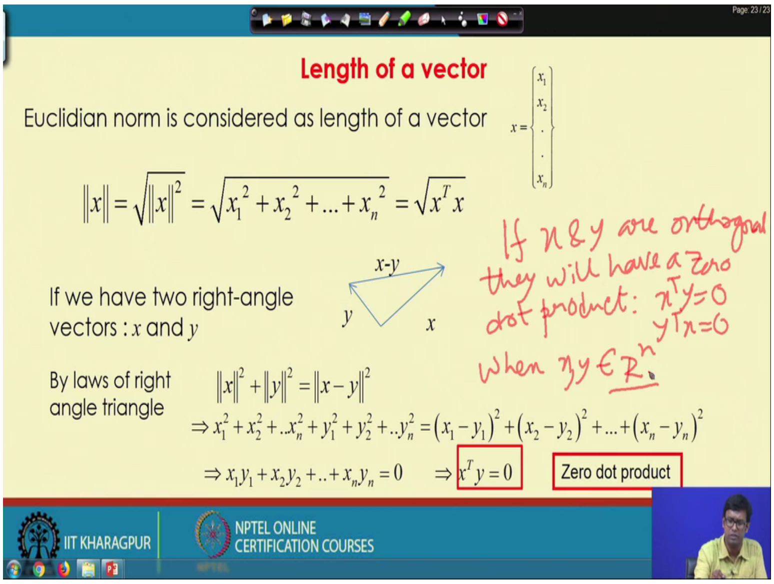


So, length first idea which is little away from what we are doing right now on matrices and the properties of matrices; we will look into the definition of length of a vector. And this is given the length of a vector is usually measured by the term norm, when measure of the magnitude of a vector is norm. If you have a vector a 1 to an which is the vector belongs to R n L P norm of the vector or pth norm of the vector; is defined as L P is equal to absolute magnitude of each if each of the component raise to the power p, and then some all the components and 1 by pth root of that.

So, for different values of p different norms and defined; for taxicab norm is defined if p is equal to 1 which is L 1, which is basically magnitude of each of the components of a and summation of that. L 2 norm is defined for p is equal to 2 which is L 2 is root over of a 1 square plus a 2 square up to n square. And L infinity norm is defined for p is equal to infinity. So, all of these components are raised to the power infinity. And then we take 1 by infinite root of that. So, the only the largest component stays there because, 1 by infinity all the components raise to the power infinity, the largest components becomes very large.

And then only this stays when I take 1 by infinite root of that. So, it becomes the maximum of magnitude of all the components of the matrix. This norm definition of this norms are important and later we will see when we look into iterative solvers and try to find out that convergence will use the norms for finding out convergence. The L 2 norm is important because, if we think of a 2D or 3D vector, it is a length of the vector, and it is also called it is a measure of length it is also called Euclidean norm.

(Refer Slide Time: 03:45)



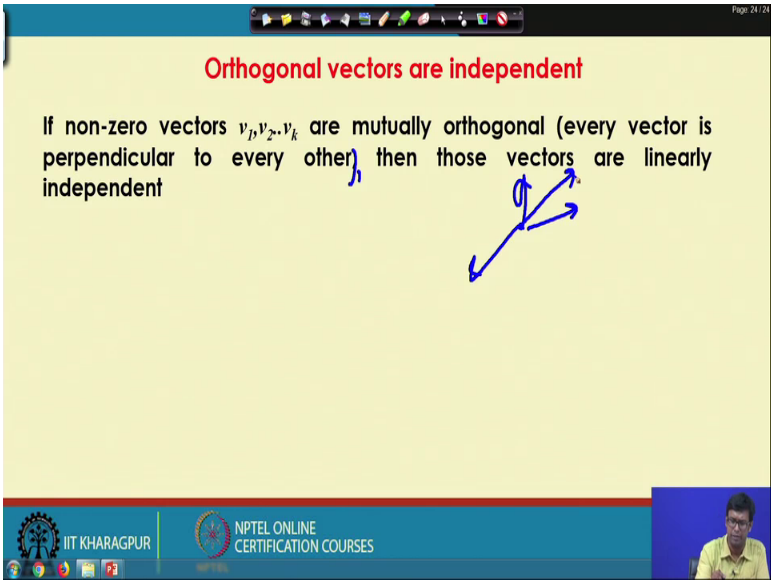
Euclidean norm is considered as length of the vector if I have a vector x 1 x 2 x n Euclidian. So, 2D, 3D it is visible x 1's root of or x 1 square plus x 2 square or root of or x 1 square plus x 2 square plus x square; use me the length of the diagonal. These are this can be visualized but, in higher dimensional vector this is found out in same way and they are considered as a length of the factor. And we can see L 2 norm is given as this form, 2 lines inside which x raise is root over of L 2 squared norm which is x 1 square plus 2 square plus x n square is root over of x transpose x when x is a vector.

If we have a right angle triangle the equivalent of Pythagoras theorem, there is a law of right angle triangles still stays that that the square length of squared of the length of 2 sum of the square of the length of the 2 sides is equal to square of the length of the hypotenuse. So, we have vectors x and y, and the hypotenuse becomes x minus y, x minus y is not the vector; it my this is not the x minus y vector; x minus y vector any vector should go through origin this is translated to get the vector difference.

Anyway so, we can write by laws of right angle triangle x squared, plus y squared is equal to x minus y square. And we can expand them, and while and then by rearranging and cancelling terms from which side, we can write x 1 y 1 plus x 2 y 2 plus x n y n is equal to 0, or x transpose y is equal to 0. So, if x and y are orthogonal, we get x transpose y is equal to 0 or which is basically equivalent of dot product of the vectors x and y, the vectors we have 0 dot product.

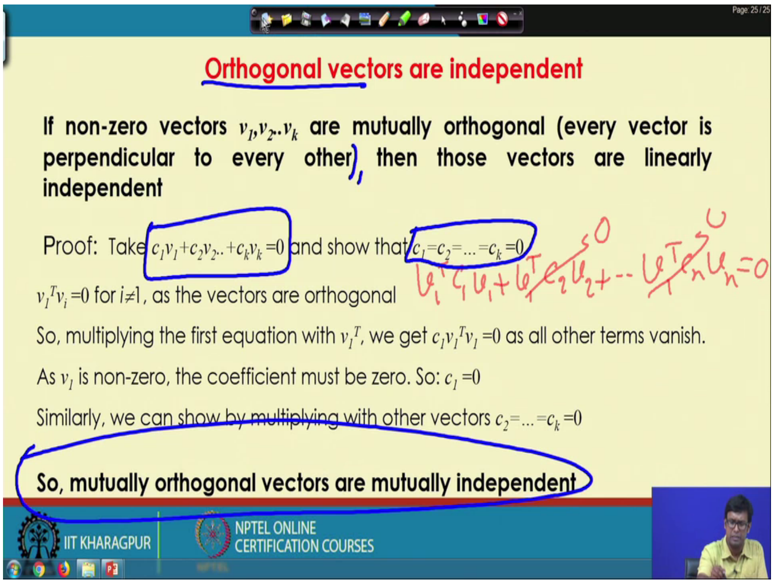
So, we can write if x and y are orthogonal they will give a , they will have a x transpose y is equal to 0 or y transpose x is equal to, when x and y they must be vector in the same dimensional vector space belongs to R n; is important because n can be anything, n can be more than 2 or 3. For 2 or 3 2 dimensional and 3 dimensional vector we know that perpendicular vectors we will give zero dot product. But they are higher dimensional vectors also if they are perpendicular they will give 0 dot product. We will go ahead with this definition of orthogonality.

(Refer Slide Time: 06:39)



If non--zero vectors are so, this idea is orthogonal vectors are independent. If non-zero vectors v 1 up to v k are mutually orthogonal, every vector is perpendicular to every other bracket close here. Then these vectors are linearly independent. So, you so, this is like if I have one vector 2 perpendicular vector they are linearly independent if I get another perpendicular vector in R 3 they will be linearly independent. In R 4 if we get some way another perpendicular vectors the all these vectors are perpendicular to each other. They will be linearly independent.

(Refer Slide Time: 07:35)

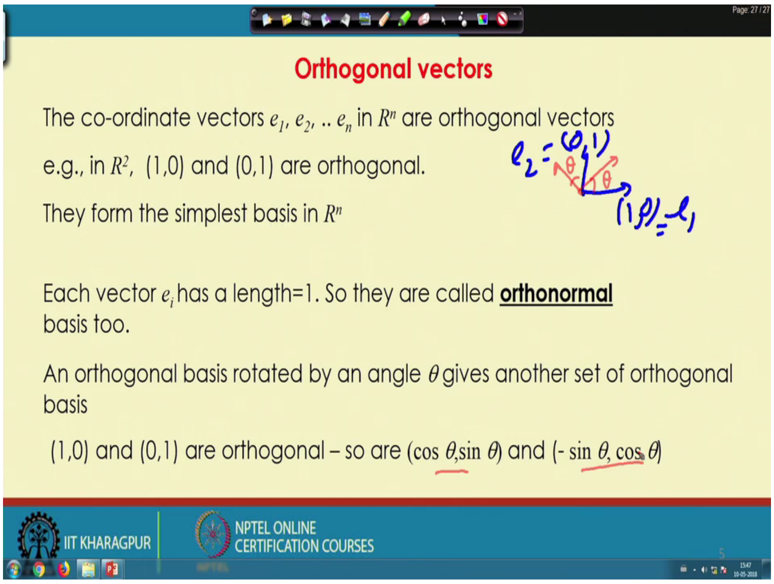


So, the proof is that considered the orthogonal vectors. These vectors are mutually orthogonal. So, each vector dot product with other any other of the vector is 0. So, and now for linear independence we have to combine them and show that the coefficients are 0. So, take c 1 v 1 plus c 2 v 2 up to c k, v k and show that c 1, c 2, c k is equal to 0. If they are independent, if they are orthogonal then v 1 transpose v i is equal to 0 for i naught is equal to 1 as the vectors are orthogonal.

So, what we will do? We will multiply this equation with v 1 transfer, which will be v 1, sorry, which will be v 1 transpose into c 1, v 1 plus v 1 transpose into c 2, v 2 plus v one transpose c n v n. And this should also be equal to 0. Because the left right hand side of the a 0. So now, v 1 transpose v 2 c 2 is the constant this come out v 1 c 2 is the scalar and this will count come out. So, we one transfer v 2 is equal to 0. So, this will be 0, similarly, this will be 0 and we will get v 1 transpose v 1 is equal to 0.

And now it has been said that the vectors are non 0. C 1 is 1 0. Therefore, c 1 is equal to 0. Similarly, we can multiply it with c 2, c 3 and c k and so that show that the sorry, you can multiply it with v. V 2 transpose v 3 transfers v k transpose, and show that c 2 up to c k is equal to 0. So, mutually so, this relationship will only hold when c 1, v 1, c 2, v 2, c k, v k will only hold when c 1, c 2 is equal to 0; is the vectors are orthogonal. And therefore, mutually orthogonal vectors are also mutually independent. This is trivial if we look into few examples. This looks very trivial.

(Refer Slide Time: 09:46)

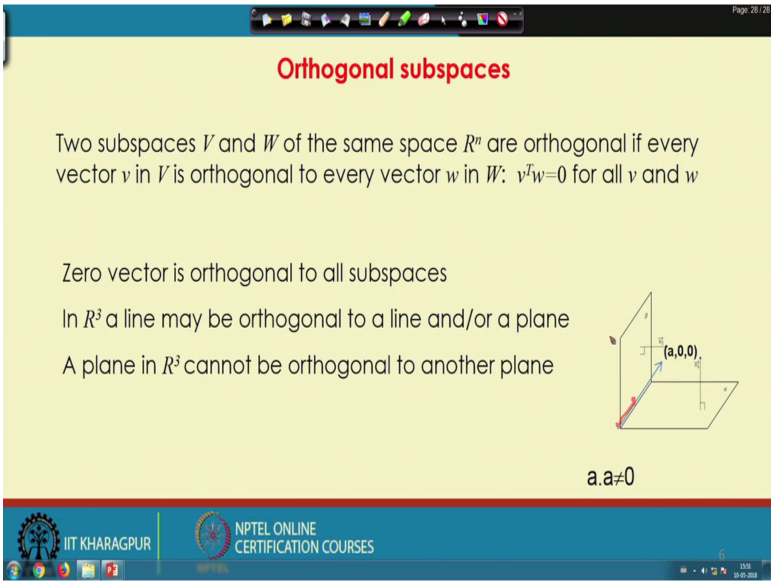


For example, the coordinate vectors e 1 e 2 e n in R n orthogonal vectors. Say we have R 2 and we have the coordinate 1 0 the e 1 this is e 1. And e 2 is equal to basically i e 2 is equal to 0 1. They are mutually orthogonal vectors and their also independent vectors. In R 2 1 0 and 1 0 1 are orthogonal. They form the simplest basis. The orthogonal vectors form the simplest basis in R n. Each vector e i has a length of 1, like i and j, j k they have a length of 1, and they are called orthonormal basis 2. A mutually orthogonal vector, if it is R n, n mutually orthogonal vector will give us n mutually independent vectors which will be a basis.

And if these, vector have done in magnitude of 1, then we call this as a orthonormal basis. Normal is associated with dimension magnitude 1 in this case that R. And orthogonal basis rotated by an angle theta gives another set of orthogonal basis. So, 1 0 0 1 and we if we rotate them by theta like this is 1 0 0 1, and if we rotate them by angle theta is they will be the intervect basis, basis will be rotated by angle theta, and they will be cos theta sin theta and minus sin theta cos theta. So, they will be again a orthogonal basis. So, one orthonormal basis rotate orthogonal basis one orthogonal basis or orthonormal basis rotated by another. Certain angle will remain a orthogonal basis.

So, there are several orthogonal vectors like that an orthogonal vectors are independent vectors. Therefore, if we have n orthogonal vectors in R n we should get a get an independent vector which will form a basis. And it is an important and useful if we can have orthogonal basis. Why because, we can very easily divide get the components of a vector along each basis decompose a vector along mutually orthogonal and mutually independent basis like we do for vectors, for vector calculus or vector algebra, when we have a i plus b j, we can say a is component along x axis, b is component along y axis, and thing becomes very easy to handle.

(Refer Slide Time: 12:24)



Then comes the idea of orthogonal subspace. Two subspaces V and W of same space R n are orthogonal; is every vector in v every vector small v in V is orthogonal to every vector w small w in W, or v transpose w is equal to 0 for all v ad w. So, any vector in one particular vector space is one subspace is orthogonal to a vector in another subspace. So, if we considered any vector along like the if we considered this vector and this plane and this slide, any vector along this plane is orthogonal to any vector along this line.

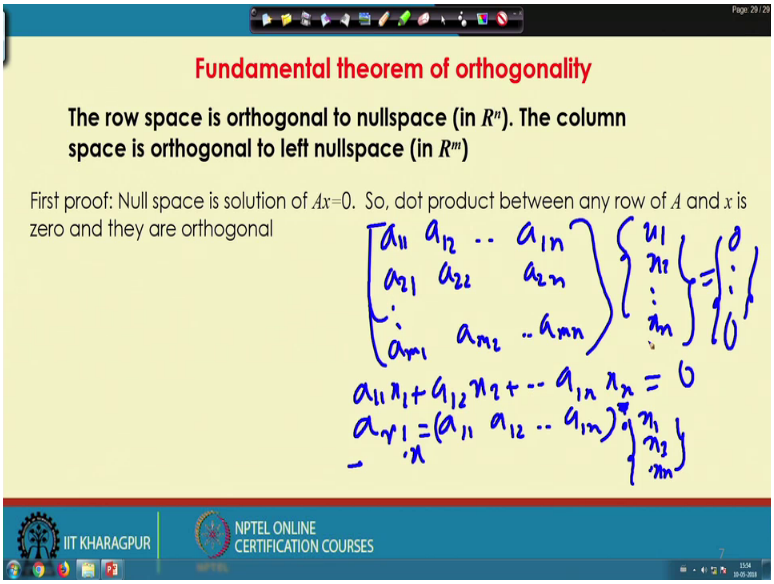
And now, we will tell that there orthogonal subspaces. This space is orthogonal to this subspace. So, one subspace is orthogonal to another subspace. Any vector in one subspace is perpendicular or orthogonal to any vector belong to the other subspace. And they are called orthogonal subspaces. Zero vector is orthogonal to a subspace. So, I take dot product of 0 vector with any vector belong to any particular subspace it will be always 0. So, 0 is always orthogonal to all subspace.

In a R 3 a line maybe orthogonal to a plane or a line. Like, like this line this is 3 dimensional space R 3 can be orthogonal to this particular line. So, this particular line this particular lines. So, a line can be orthogonal to a line or any multiple line. And a line can be orthogonal to a plane also. So, these are orthogonal subspaces; however, and this is again interesting a plane in R 3 cannot be orthogonal to another plane. Why? If we think of quickly we can think of these 2 planes, ok.

Are they orthogonal to each other? It will looks like any vector along this plane is perpendicular to any vector along this plane; however, they are not orthogonal to each other. And we can see that let us consider 2 planes in R 3, and there is at least one vector a, which is basically the common each of the planes. And this belongs to both the planes. If we take multiply dot product of this vector to itself, it is non 0.

So, plane is not orthogonal to plane. Line can be orthogonal to a plane in R 3. It is important that the ortho the dimension of the orthogonal subspaces when we add them up, like for plane that I mention is 2. So, 2 planes they have both have both of them have dimension 2. If we add them of we will get a dimension 4. It cannot be in orthogonal in R 3. In R 3 the dimension of orthogonal subspace as can be added and that should be less than equal to 3. This is important you can (Refer Time: 15:34) over it will more.

(Refer Slide Time: 15:44)

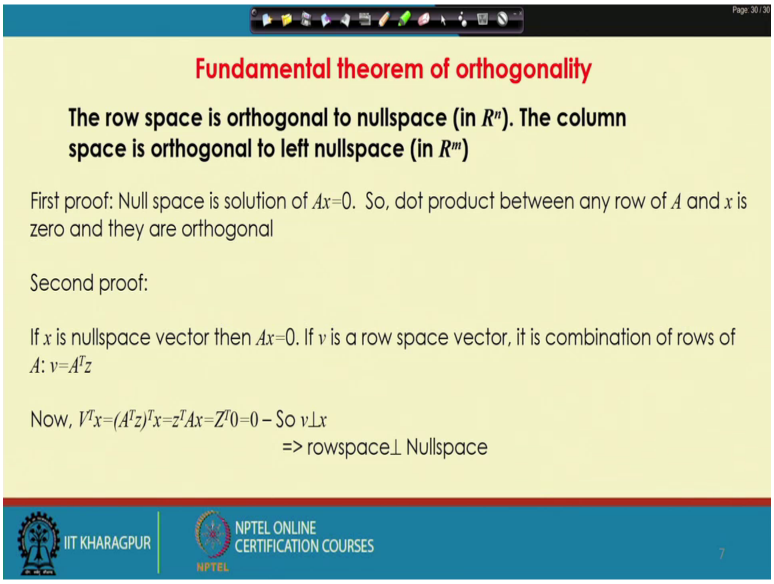


And it gives the fundamental theorem of orthogonality, which says the row spaces orthogonal to null space; the column space is orthogonal to left null space. Row space and null space are in R n and column space and left space null space are in R n. Remember the sighted in the last class, we found out that those space is coming orthogonal to null space, and column space is coming orthogonal to left null space; however, we will try to find out the proof of this, except looking into 1 or 2 theorem. So, it is important that by 1 or 2 examples, we can verify the statement; however, the proof should be more generalized and should not is not only varificable it should be provable.

So, we will look into the proof, there are 2 2 proofs. The first proof is that null space is a solution of Ax is equal to 0. So, what is Ax is equal to 0? If I try to write down the Ax is equal to 0, this is a 1 1, a 1 2, a 1 n, a 2 1, a 2 2, a 2 n n comes up to a m 1, a m 2, a m n. And x 1, x 2, x n is equal to 0. So, I can see a 1 1 x 1 plus a 1 2 x 2 plus a 1 n x n is equal to 0. Which is basically dot product of a r, first row of a. What is it? This is a a 1 1 a 1 2 a 1 n, dot x, dot n transpose rather transpose of this with x 1 x 2 x n. Now transpose of this dot x sorry transpose.

This will be 1 1 1 x 1 plus 1 2 x 2. So, each row multiplied with the like x vector gives us 0. Therefore, each row is orthogonal to the x vector. So, we can say that row space is orthogonal to the null space, when Ax is equal to 0; x is the solution of null space. So, null space is the solution Ax is equal to 0. So, dot product between any row of a an x is 0 and therefore, they are orthogonal.

(Refer Slide Time: 18:24)



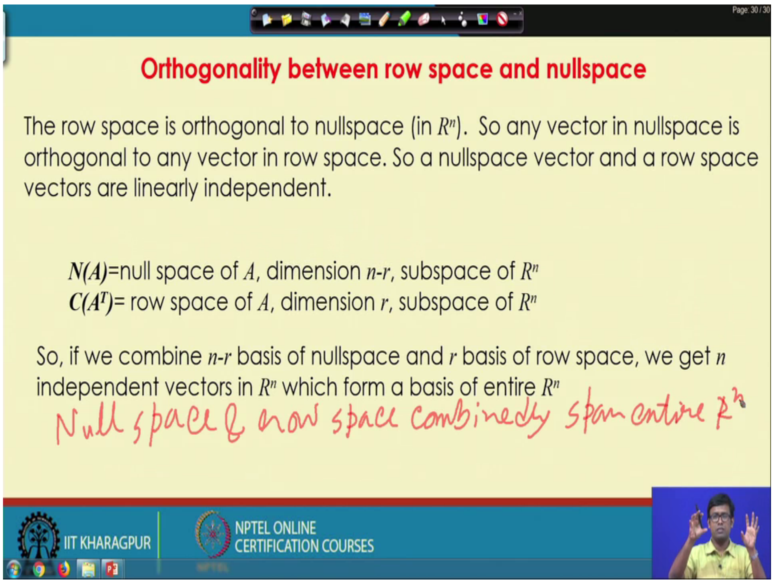
The second proof is that if x is null space vector, then Ax is equal to 0. Let us see v is a row space vector. So, v can be considered as combination of rows. And v transpose v multiply A transpose which the columns and now the rows of A, A in columns of A transpose with a vector x. So, is z sorry, so, z 1 first column of A transpose plus plus z 2 first column of a transpose; is basically z 1 into first row of A plus z 2 into second row of A and so on.

. So, v is equal to a transpose z is the row space vector, because this is linear combination of the rows. Now we if we write v transpose x, which is a transpose z v is equal to x transpose z whole transpose x. In this becomes a transpose Ax by the property of transpose which is z transpose 0. And as Ax is equal to 0 z trans z transpose Ax is equal to 0 so, z transpose 0 so, this is 0. So, v is perpendicular to x so, any row space vector becomes perpendicular to the null space vector. And row space is orthogonal to null space. So, we get row space is orthogonal to null space.

So, what we got? Row space and null space both belong to R n, dimension of row space is R dimension of null space is in minus R. So, if there are R vectors in row space, there are m n minus r vectors in null space. These are vectors any of this vector in R vector is orthogonal to n minus r null space vector. So, any of this vector is mutually independent with the null space vectors So, if we have n minus and null space basis whish are mutually independent, any of the row space which is mutually independent with that. Therefore, all R row space basis are mutually independent with the n minus n null space basis. So, this R and this n minus r gives me a total set of n mutually independent vectors, right. Rather than row space vectors are orthogonal to null space vector. Therefore, they are independent of a null space vectors.

So, each of the row space vector is mutually independent of with linearly independent with all of the null space vector. So, you get a of basis in null space, where we have n minus r independent vectors. And we get a basis in row space with they we have R independent vectors. So, there are R independent vectors, there n minus r independent vectors, and they are also mutually independent. So, you get a total set of n independent mutually independent vectors. This is in R n; R n are total set of n vectors we will form the basis. So, row space and null space combining we can span the entire R n, sorry.

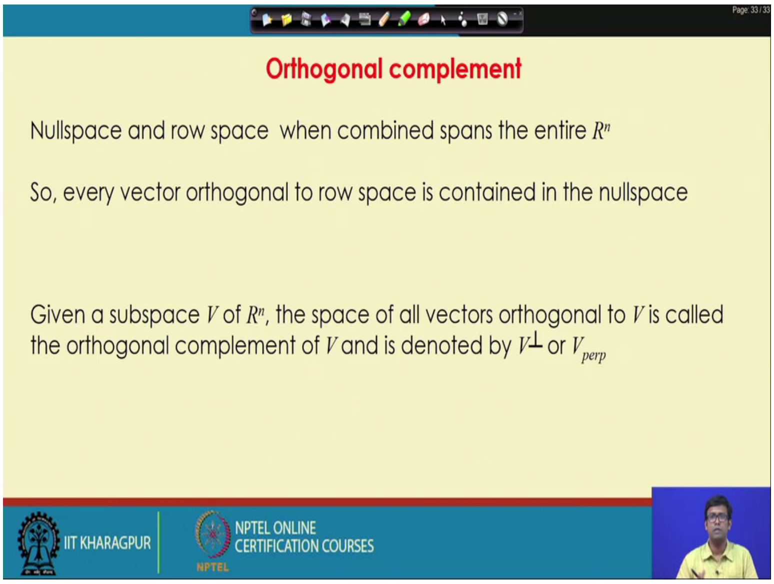
(Refer Slide Time: 21:44)



So, yeah the row space is orthogonal to null space in R n. So, any vector in null space is orthogonal to any vector in row space. So, A row space vector and the and null space vector are linearly independent. N A is null space of A has dimension n minus r subspace of n. C A transpose c 0 is row space of a dimension R subspace of R n. So, if we combine n minus r basis of null space and R basis of row space, we get n independent vectors of null space and row space. And they form basis of entire R n.

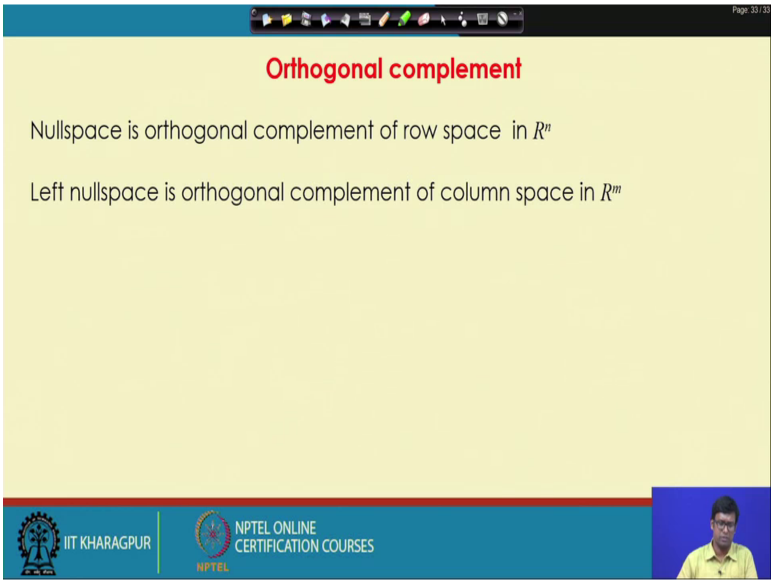
So, we can also write that null space and row space combinedly span entire R n. Because basis of row space and basis of null space they are farming basis of R n. So, this meaning entire R n. So, null space so, for example, this is null space and this is row space, this span the entire R n. So, in R n in the mean vector space if there is something which is not belonging to row space that n must be in the null space. Or something which is not belonging to null space, that must be in the row space, because combindely we get the entire space.

(Refer Slide Time: 23:26)



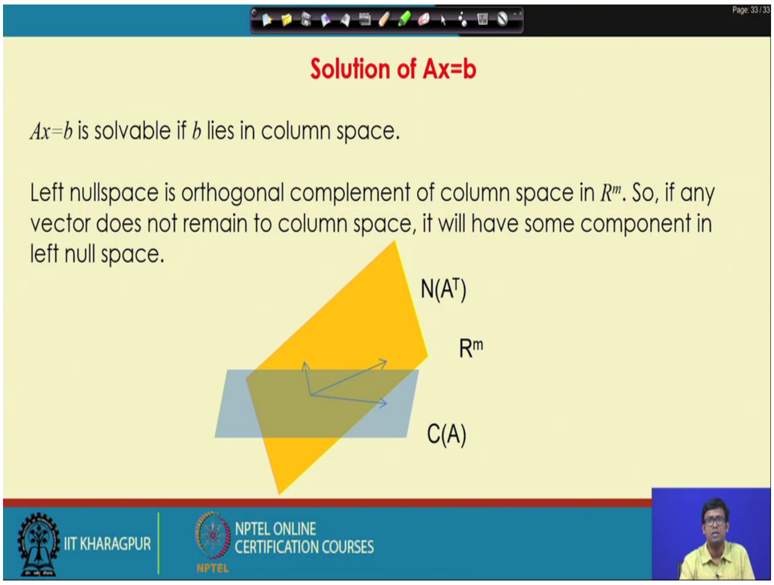
Null space and row space when combine spans the entire R n. So, every vector, every vector orthogonal row space is contain in null space. Every vector which is not in row space is a member of null space. And given a subspace V of R n the space of all vectors of orthogonal to is call the orthogonal complement of V, and divided by V perpendicular, perpendicular sign or perpendicular. Therefore, null space becomes an orthogonal complement of row space. Whatever is orthogonal to row space, whatever does not lie in row space must be null space.

(Refer Slide Time: 24:10)



And if we can write null space is orthogonal complement of row space. And similarly we can show left null space n is orthogonal complement of column space.

(Refer Slide Time: 24:22)

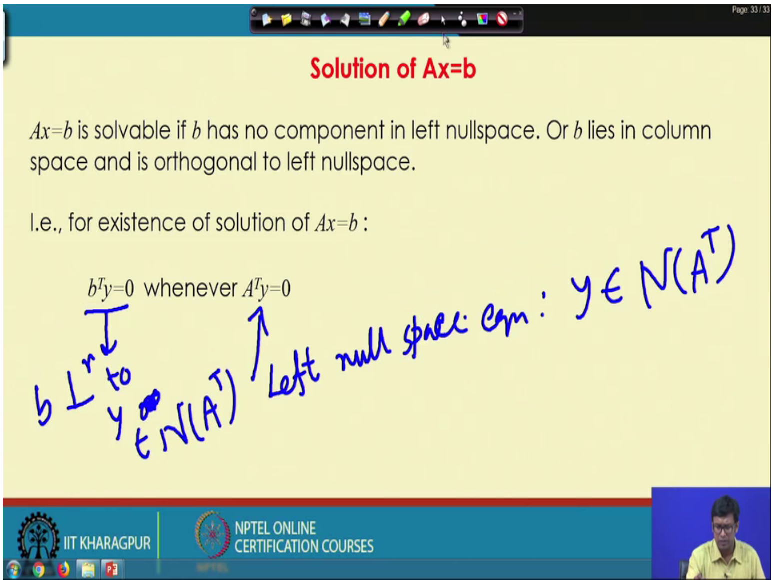


So now, if we start looking into Ax is equal to b. We may last class we took an example and observe that they are the particular and null space row space column space left null space. So, started from last class row space looked into column space left null space. So, that they are perpendicular. Now in from more general sense we are coming that they are orthogonal complement. Left null space and column space is also orthogonal compliment. I am interested specifically in column space at this stage, because something which does not belong to column space now we know that they belong to left null space.

If the b vector does not belong does not lie and column space, then Ax is equal to b is not solvable. That is why I am specifically interested in column space and null left null space. So, if I can say that b has a component along left null space, b is away from columns presented b is does not is not a column space vector and Ax is equal to b is not solvable. So, we will look into solution of x is equal to b from the ideas of orthogonality, orthogonal complements etcetera; Ax is equal to b solvable if b lies in column space. Left null space is orthogonal complement of column space in R n. So, if any vector does not remain to columns space, it will have some component in left null space. It may be entirely in left null space or maybe combination of some use a some basis in left null space and some basis in columns space may be a combination of that.

Or we can say that say this is the in R n this is the null left space and this is column space. Now, we get a vector b which is away from which is not on the column space, which comes along the column space. And it must have 2 component, one is along left column space and one is along left null space. So, if a vector b has some component along left null space, it must not be in column space. Therefore, Ax is equal to b will not be solvable; Ax is equal to b solvable when it is entirely in column space, and b does not have any component with left null space.

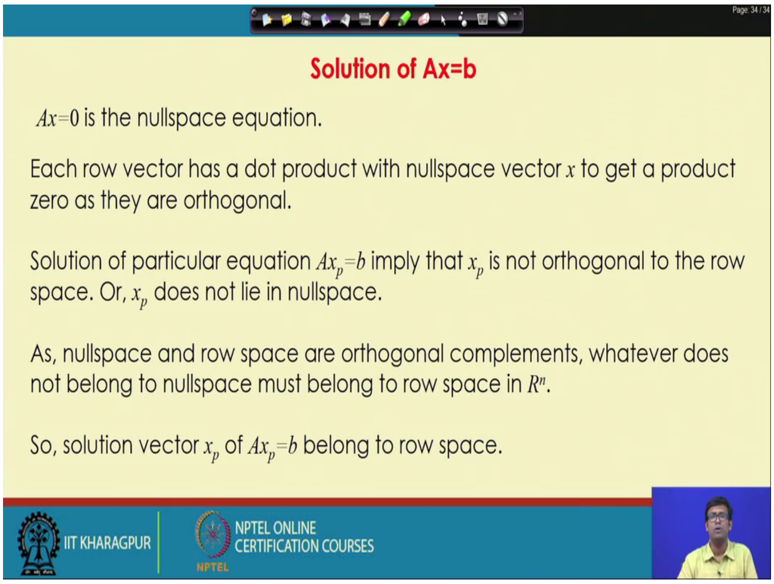
(Refer Slide Time: 26:38)



Ax is equal to b solvable when b has component in left null space, or b lies in column space and is orthogonal to left null space. That is for existence of solution Ax is equal to b, b transpose y is equal to 0, b whenever a transpose y is equal to 0.

A transpose y is equal to 0 is left null space equation. And this implies that y belongs to left null space. And then b transpose y is equal to 0 will imply that b is perpendicular to y or y which belongs to left null space. So, if orthogonal to left null space or b transpose y is equal 0 when a transpose y is equal to 0, then b is orthogonal to left null space means b lies in column space. And Ax is equal to b will be solvable.

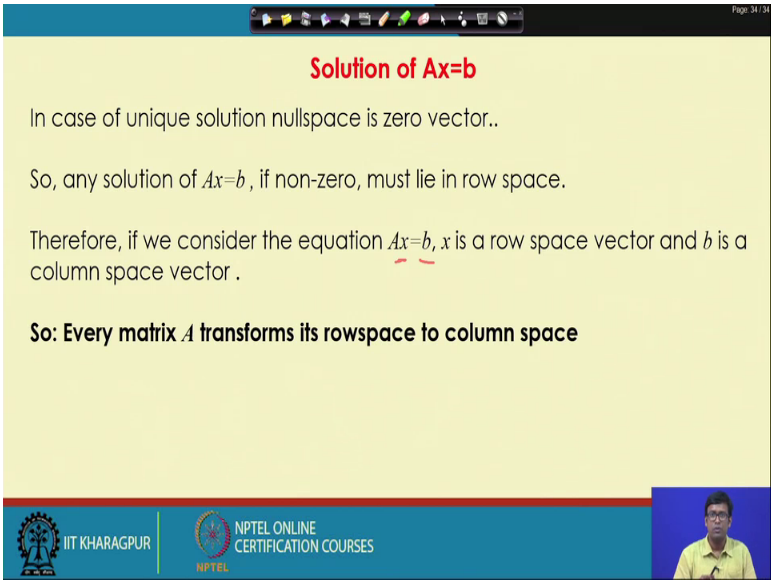
(Refer Slide Time: 27:52)



Ax is equal to b, x is equal to 0. So, we look into solution of Ax is equal to b further. Ax is equal to 0 is the null space solution. Each row victor has a dot product with null space vector x to get a product zero as they are orthogonal. Solution of particular equation Ax p is equal to b imply that xp is not orthogonal to the row space, or x p does not lie in null space. If Ax is equal to 0, then x has a 0 dot product with a, and x is a null space vector. Remember for particular equation is first found out particular solution, we first found out the null space equation and then did not consider the null space equation, only solve for particular solution when xp is equal to b in case we have multiple solutions.

So, this equation x p is equal to b says that x p and not orthogonal. They have a non-0 dot product. So, x p does not lie in null space. As null space and row space are orthogonal complements, if it does not lie in null space is must lying row space. So, the solution x p of Ax p is equal to b belong to row space. Therefore, the solution vector always lies in the row space. Solution vector of a non-null non homogeneous equation; x p must lie in the row space. Whatever does not trying null space in R n must lie in the row space. Therefore, if I am when I am solving Ax is equal to b, the x I am finding that is the row space vector. That lies in the row space.

(Refer Slide Time: 29:40)

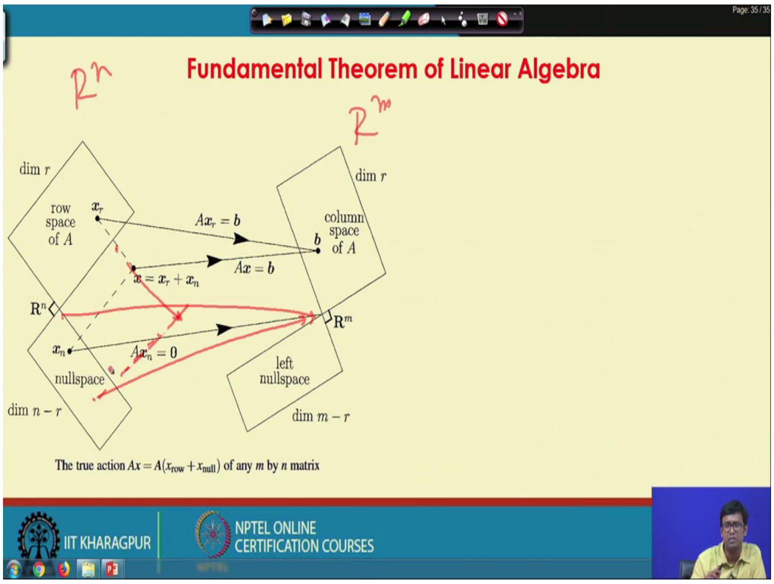


In case of unique solution null space is zero vector so, everything will lie in row space. So, any solution of Ax is equal to b if non-zero, 0 means it is both 0 is also a part of null space. But for non-homogeneous equation Ax is equal to b must give a non-zero solution. And if in case non-zero it must lie in the row space. Therefore, if we consider the equation Ax is equal to b, x, x is a row space vector, right? And b is a column space vector. So, if I multiply a row space vector by the make it is A, it will transform it to the column space vector, the product is column space vector.

And we can say every matrix a transforms it is row space to a column space vector. It is in the sense we can try that, we can take row and multiply with A and the product will be nothing but a column space vector, a combination of columns. Now, this is a this 2 or 2 few like last few discussions give us a few important one very important theorem of linear algebra, which is we call fundamental theorem of linear algebra. That is, Ax is equal to b solution x lie in the column space; lie in the row space, sorry.

The row space vector x is any row space vector let us consider multiplied with a will give me some b which is in the column space vector. Whatever does not lie in b in R n must lie in the left null space. Whatever does not lie in row space in R n must lie in the left null in the null space. And any null space vector when multiplied with the main matrix A will give me the 0 vector. Ax is equal to 0 the null space equation. And this is pictorial representation of the fundamental theorem of linear algebra.

(Refer Slide Time: 31:40)

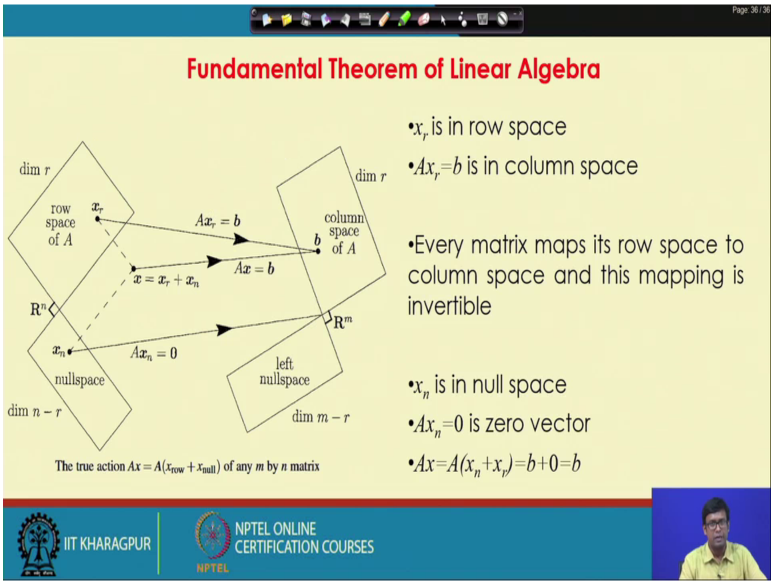


And I have a row space which was dimension R, which is perpendicular to a null space which is dimension n minus r. Now and there is a column space which has n dimension R, left null space as dimension n minus r. Now, I take any vector in so, the left hand side is R n and the right hand side is R m. Now I take any vector in R n which is an x vector, it has 2 component. One is along the row space; another is along the null space. The row space component when multiplied with the matrix A, we will go to the column space Ax is equal to bx.

The null space component when multiplied with A we will take us to the 0 solution. And now if we get Ax R is equal to b which is in the column space. Now if I invert it, this is the mapping the row space is being map to the column space. If I invert it, column space will always have an unique mapping in the row space because, this solution is unique. The null space solution is non-unique, anything multiplied with null space can take me. So, if I have another null space vector that would have been taken me here. If I had another null space vector that would have been taken me here.

So, and based on this null space, we could have different values of x, we could have different values of x. So, outside null space if x lies outside row space, Ax is equal to b can have infinite solution. If x is in row space the x is equal b has exactly one solution. Ax is equal to b solvable if b is in the column space. If b is away from the column space Ax is equal to b is not solvable.

(Refer Slide Time: 33:41)



And x r is in row space, Ax r is equal to b is in the column space. Every matrix maps it is row space to column space, and this mapping is invertible. So, we can get exactly same x, if I invert b, if I put multiplied b with A, A inverse we will get back to x r. However, the null space mapping is not invertible, it can take us to anywhere in the null space 0 inverse 0 can take us anywhere to the null space, x n is in null space. Ax n is equal to 0 is a 0 vector. And the final solution Ax is Ax n plus x r is equal to b plus 0 is equal to b plus 0 is equal to b.

And this is pictorial representation of the fundamental theorem of linear algebra. Now what remains in question is that incase b is not in the column space. Of course, it has no solution. But if we think from statistical point of view from a point of view of optimization, there can be something which approximately satisfies, the equation b is not in the column space, but b has a component along column space can we solve for that part. We and will look into subsequent classes, what should we do in case Ax is equal to b has no solution, b is not on column space, but b has some component in column space, and some in left null space, what is the solution for that. So, we will we will do it in the subsequent classes.

Thank you.