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

Lecture - 25
Creating Orthogonal Basis Vectors

Welcome, in the last few classes starting from the fundamental subspaces we observed that the fundamental subspaces are mutually orthogonal to each other. And from there, what we to the case that b vector does not lie in column space. So, b vector has a component in left null space also; as left null space and column space are mutually orthogonal to each other. We projected b vector in 2 column space and the remaining part we will say that; this is in a left null space. Using the projected part of the b vector into the column space, we form the normal equation and we solve normal equation to get best estimate.

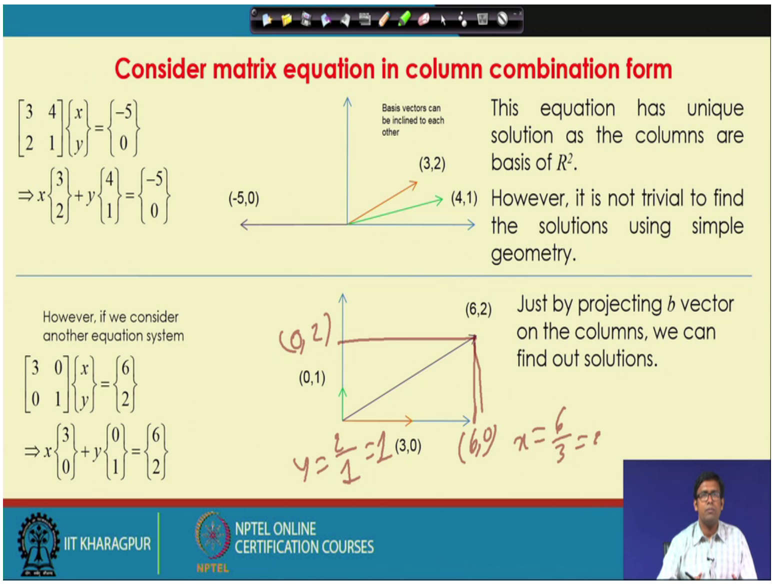
The idea from this exercise we got is that a vector can be projected into a space and the left out part will be the orthogonal part to the projection. So, if we have one particular vector we can project this vector into another space and get the projection and subtract the projection from the main vector and get another component which is perpendicular to the projection.

So, we can constitute mutually orthogonal components of a vector or we can decompose a vector into mutually orthogonal components. In this area, now we will explore more that we will have a set of vectors which are mutually independent, but not orthogonal to each other. And now, we will start projecting one vector to other and subtracting the projection from it and then that, way we will create a set of mutually orthogonal vectors.

Or if essential we will create a set of orthogonal basis vectors for any vector subspace, and this process we will start from a set of mutually independent vectors and which are not necessarily orthogonal to each other. And end up with the set of mutually orthogonal basis vectors and also these basis vectors; each of the basis vector will have a length unity and this process is called a Gram Schmidt orthogonalization process. We will look into the Gram Schmidt orthogonalization process, we will see; what are the advantages and how this can be further applied for developing better matrix solvers, shift little bit from matrix solvers. We will see how to orthogonalize the vectors given any set of basis vectors and then we will again come back to matrix solvers; all this can be used in order to solve ax is equal to b.

As well as in order to solve the normal equation a transpose a x tilde is equal to a transpose b, when a x is equal to b does not have any solution.

(Refer Slide Time: 03:04)



So, we go to the first slide what is given here sorry that we have matrix equation. We considered it is in a column combination form and this is in the vectors are in R 2. It is real coordinate space of dimension to see it is easy to visualize it. So, these 2 columns are 3 2 and 4 1 and they are combined to get a resultant vector which is minus 5 0. So, if you look into the column space the basis vectors are 3 2 and 4 1 and this basis vectors are inclined to each other.

(Refer Time: 03:43) Linear independent vector, but they are not than any combination of this vectors can give me r (Refer Time: 03:48) vector in the column space. However, these vectors are not perpendicular to each other. So, this is some point certain cases we make this mistake that basis are always perpendicular to each other because, we deal with basis like xyz or e theta e z e r something like that in Cartesian or cylindrical coordinate. However, basis vectors can be inclined to each other also we consider matrix equation in R 2. That means, in real coordinate space R 2 or in 2-dimension space, the vectors are member of the real coordinate space R 2 and we consider the column combination form of this equation. So, this can be written as some constant multiplied with 3 2 plus some constant multiplied with 4 1 is equal to minus 5 0. And we have to find out what are these constant coefficients through solution of the matrix equation.

So, if I look into the vector space of the column space here, the column space is constituted by the victors 3 2 and 4 one they are basis of the column space. Because, in R 2 we need only 2 vectors which are linearly independent to form the basis and this is interesting to observe that this basis vectors are not mutually orthogonal, they inclined to each other. Sometimes, when we think of basis of any real coordinate space we think of x y or x y z or r theta. If you think of anything in polar coordinates which are perpendicular basis. Vectors that is very easy to use perpendicular very helpful to use perpendicular basis vectors in geometry, but and will see why is it is important in just in a moment. However, basis vectors may also be inclined to each other. In general basis, vectors are inclined to each other. There they are not at 90-degree angle in between them, but we our exercise will be to form orthogonal basis vectors that the next thing.

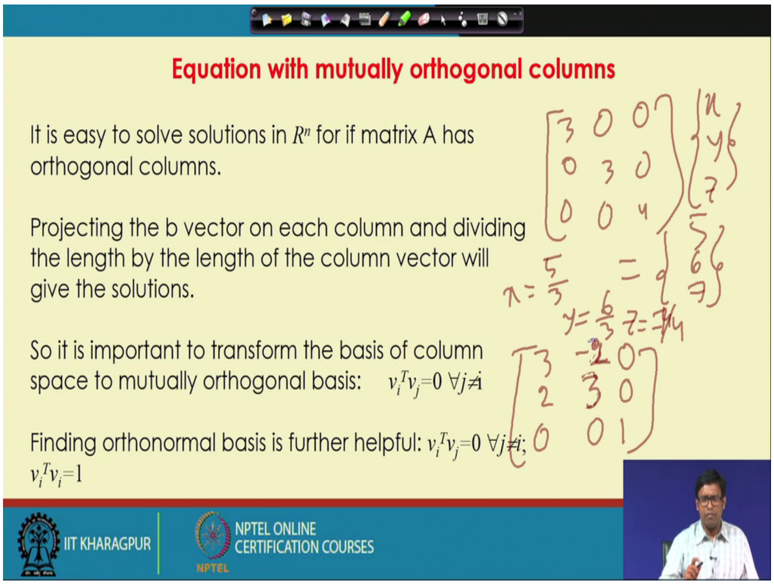
So however, in this particular case when the basis vectors are inclined each other of the column space if I try to solve this equation geometrically; that means, I take the first column vector multiply it with the coefficient. And then add it with the second column vector multiplying into another coefficient and see what are the coefficients. So, geometrically this is not trivial for 2 d event we can think of solving it like this will extend this line and will extend this line. If we can extend in this line and then start growing parallel of this and then we will see; we will do multiple parallel of this and. Then, we will see that this parallel of this particular vector meets at this point and this is minus 2 times into this particular column and time into this particular column. So, minus 2 and one is the solution.

For 2 d they can be visualized and it is it is difficult exercise if we think of doing it for 3 even for a 3 d case. You have to draw line and several parallelistic parallels in the plane it is not trivial. So, what we can say that when the equation has unique solution as the columns are basis of R 2; however, it is not trivial to find the solutions using simple geometry using simple geometry construction. Here, for 2 d we can find it if it is 3 d it is difficult to find. However, if we have an equation where the columns are not inclined to each other. The columns are perpendicular to each other, if we consider this equation system 3 0 and 0 1 these columns are mutually orthogonal. The solution can be very easily found out simply by project the b vector into 1 column the columns are perpendicular to each other. Project the b vector into one column so this is 6 0 x will be 6 by 3 is equal to 2.

And the remaining part project it into another column this will be 0 2. So, y is equal to 2 by one is equal to 1 in this case. So, solution becomes extremely straight forward just by projecting b vector on the columns. We can find out the solutions and dividing it getting the length of the projection dividing it by the length of the column vector, you should be able to find out the solution. So, if we have a matrix where the columns are orthogonal to each other solving the matrix equation also becomes very simple just only. We have to take each column project in and we have seen how to project a column into a vector onto another vector to project the b vector onto that column. Find out the length divide by the length of the column is the solution associated of the is a coefficient, associated with this particular column and go on.

So, this is one advantage where will can for which we should speak for mutually orthogonal columns.

(Refer Slide Time: 09:08)

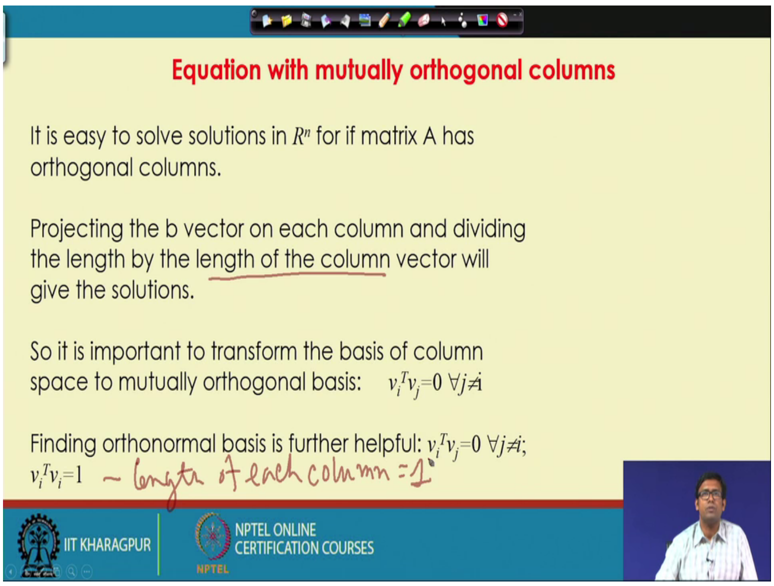


Now, if you look into equations with mutually orthogonal columns. It is easy to solve equation in Rn if a has mutually orthogonal columns. And we have seen that projecting b vector into each column and dividing by the length of the length of the column the projected length. By the length of the column vector, will give us the solution it is also important. So, it will be important if you can transform the basis of the column space to mutually orthogonal basis. So, what will happen when we have some vector like that see 3 0 0 0 3 0 0 0 4 x y z is equal to 5 6 7.

So, I can see that x is equal to if I project this into 3 0 0 this will be 5 0 0. So, x is equal to 5 by 3, y is equal to 6 by 3, z is equal to 4 by 7 by 4, something like that we can find out instead, if we have a columns in a rotated form. So, these are not 3 0 0, but for example, if I have column like 3 2 0 and then minus 3 2 0 minus 2 minus 2 3 0 minus 2 3 0 and 0 0 1. There this is not 3 0 0 they are not in diagonal form, but also here the solution can be very well found in that way. You just multiply the project the vector onto the b vector on to each of the column and find out the length and divide by.

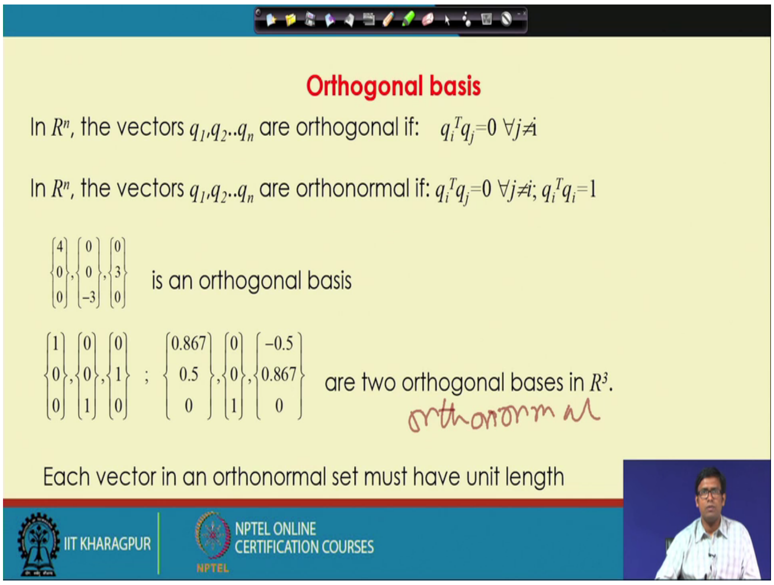
There can be the thing and the idea of having mutually orthogonal column is that one column vector transpose dot product with another column vector transpose of one column vector multiplied with another column vector v will be a 0 if the indices are not equal. This will be there will be a further advantage if we can have orthonormal basis. That means, the columns are perpendicular to each other and the length of each column vector is one. So, when we are doing this division divide this. So, when we dividing the divide the projected length by the length of the column if the columns are divide the projected length by the length of the column.

(Refer Slide Time: 12:13)



If the columns are orthonormal then this length; length of each column is equal to 1. So, if we only find out the projected length of b vector along each column that is the solution that is a coefficient multiplied with each column. So, getting orthonormal column vectors will be of further importance further use.

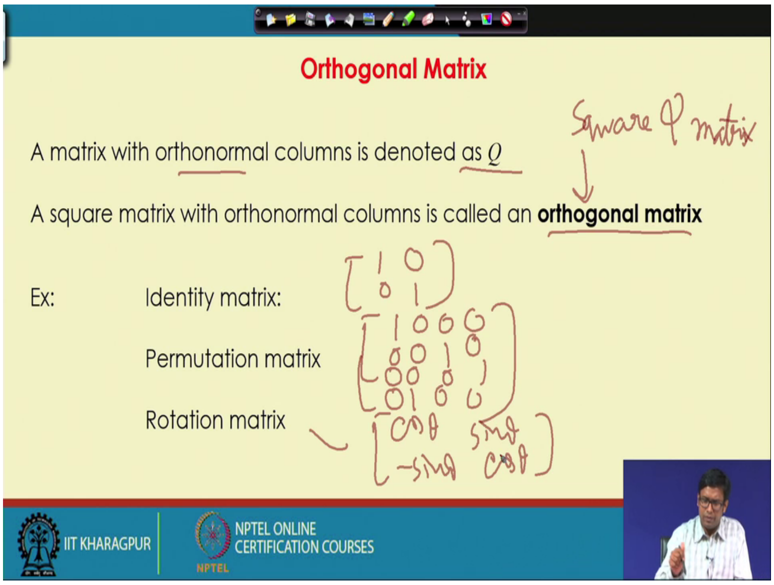
(Refer Slide Time: 12:46)



So, in R n vectors these vectors will be called orthogonal q 1 q 2 Q n. If Q I transpose q, j is equal to 0 if j is not equal to I the vector q 1 q 2 Q n are called orthonormal; if their orthogonal along with Q I transpose Q I; that means, length of each vector is equal to 1. So, we can see some of the exercise examples like 4 0 0 0 minus 3 0 3 0 is an orthogonal basis. However, length of the first vector is 4 second 2 or 3. So, they are not orthonormal in one 0 0 0 1 0 0 0 1 0 this is an orthonormal basis this is ones 0.8 6 7 0 0 1 minus point minus 0.5 0.8 6 7 0 7. They are orthogonal basis as well as they are orthonormal even orthonormal basis each vector must have unit length, because each of this vectors are unit length.

So, they form a orthonormal basis or orthonormal basis is orthogonal basis, but it is normal; that means, each vector has unit length and this is important in terms of solving equations. We saw that that only the length of the projection we give a solution here.

(Refer Slide Time: 14:24)

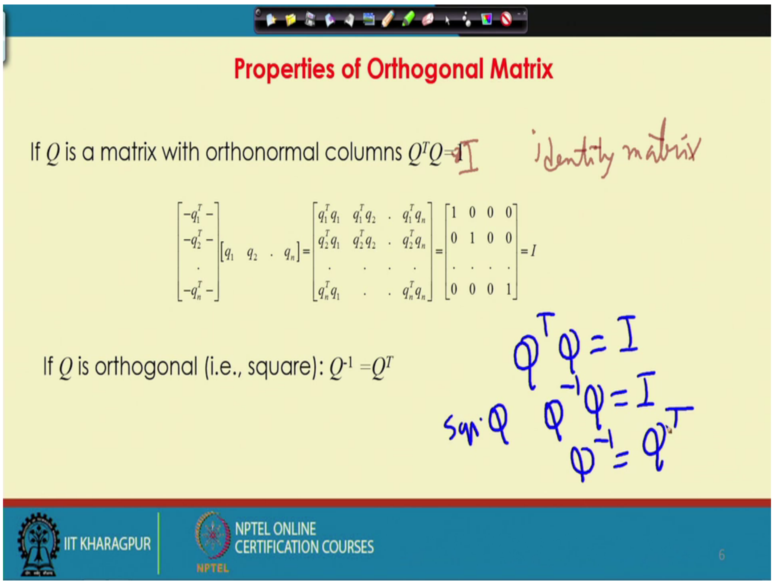


A matrix with orthonormal columns is denoted as a Q matrix; Q is general form of a matrix which is orthonormal columns. Now, there can be any number of columns it can be a rectangular matrix; however, it is a square matrix a square matrix with orthonormal column or Q matrix which is squared is called an orthogonal matrix remember, these are terminologies it is not called orthonormal matrix it is called orthogonal matrix, because it is conventionally it is being set.

So however, for an orthogonal matrix the columns are orthonormal. There can be some example of orthogonal matrix in all these cases the columns are orthonormal. So, this is basically, we can write a square Q matrix; Q matrix is a matrix with orthonormal columns. If Q matrix is squared it is matrix it is called orthogonal matrix. An identity matrix for example, 1 0 0 1 is an orthogonal matrix a permutation matrix like 1 0 0 0 0 0 1 0 0 0 0 1 0 1 0 0. This is also an orthogonal matrix of matrix a rotation matrix for example, cos theta sin theta minus sin theta cos theta ok.

So, if I take that product among these 2 things it is cos theta sin theta minus sin theta cos theta 0 and each column has cos square theta plus sin square theta length which is, 1 they are orthogonal matrices.

(Refer Slide Time: 16:18)



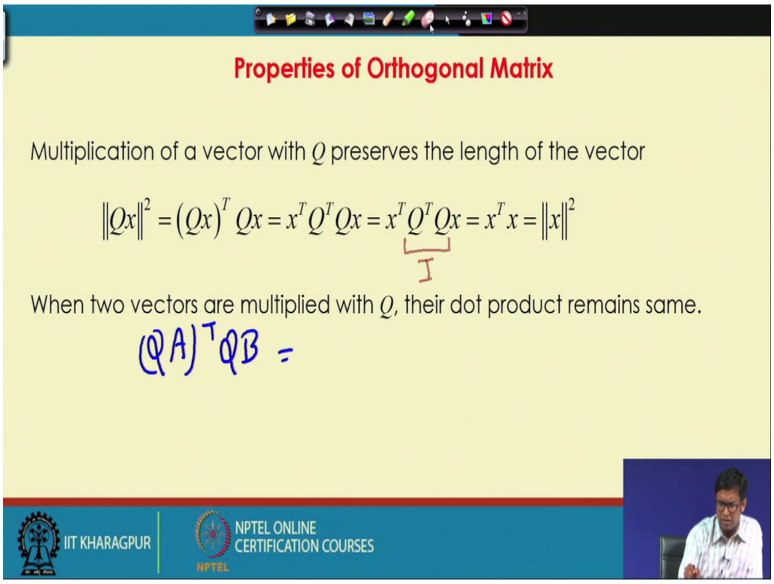
Looking to some important properties of orthogonal matrix if Q is an orthogonal matrix Q transpose Q is equal to 1. So, Q transpose Q is not 1 Q transpose Q sorry is equal to identity matrix identity matrix.

So, these are each rows of the Q transpose which is orthogonal row of the orthonormal matrix column of orthonormal matrix, multiplied with the columns of orthonormal matrix. So, this becomes first column of orthonormal matrix and it is transpose with the first column of the orthonormal orthogonal matrix and this is one and for any other value of q 1 and 2 q 1 transpose Q two this will be 0. So, it becomes 1 0 0 0 1 0 0 0 1 so it becomes an Q transpose Q becomes an identity matrix and if Q is orthogonal that is square.

So, this is for any Q matrix not necessarily has to be a square matrix; however, if Q is square we can say Q inverse is equal to Q transpose why because, for Q. As we have seen Q transpose Q is equal to I for square matrix square Q Q inverse Q is also identity matrix. So, it gives us that Q inverse is equal to Q. So, for an orthogonal matrix it is transpose is it is inverse so if we have an orthogonal matrix multiplied with a vector. And we have to find out the solution Q x is equal to b it is x is equal to Q inverse. B is x is equal to Q transpose b or it is trivial to find out the solution.

In this case, in we just make transpose of Q matrix take it into right hand side multiply with the b vector and that will be your solution. So, this is a very utility that inverse is equal to transpose of the Q matrix Q inverse is equal to Q transpose.

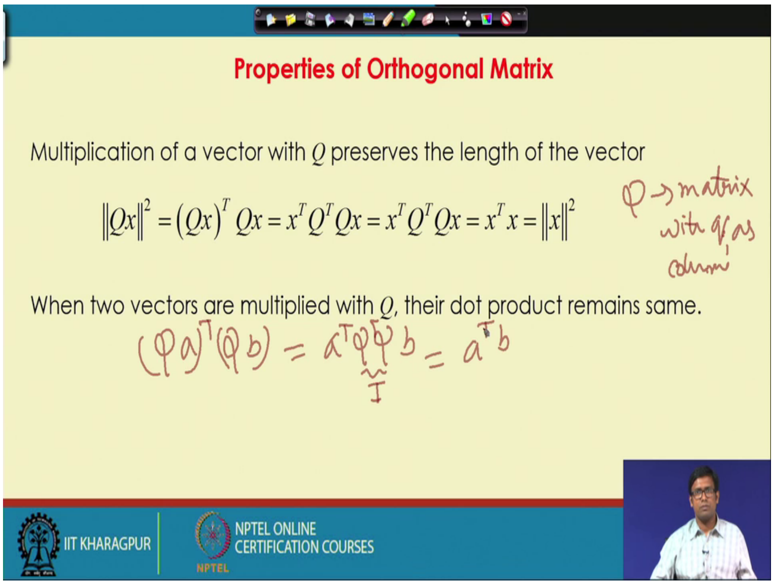
(Refer Slide Time: 18:50)



Multiplication of a vector with a Q matrix a matrix which has orthogonal column preserves length of the vector and.

We can see that if I have a vector Q multiply with Q vector x multiply with Q Q x and take it is l 2 norm Q x transpose. Q x is equal to x transpose Q transpose Q and which is the term Q transpose Q is nothing but an identity matrix. So, it is x transpose x is equal to x square. So, if a vector is multiplied with Q the length will be essentially same. When 2 vectors are multiplied with Q their dot product remains same and the proof is very similar that, I have Q a transpose Q b is the dot product of a and b and I think they are vectors.

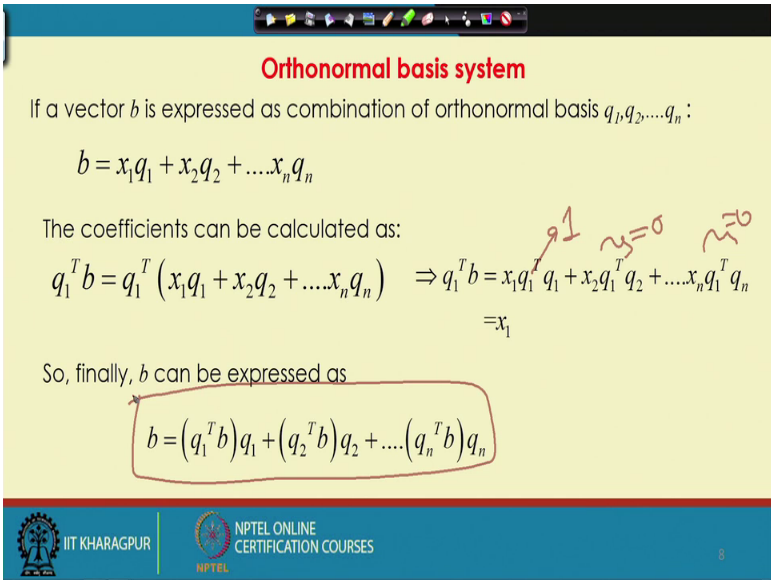
(Refer Slide Time: 19:59)



So, I think I should not write this the proof is like Q a transpose Q b. So, I deleted the capital A and writing small a because for vector. We usually use small letters and A is usually, used for matrices which will not column or row vectors which is m into n order or 2 dimensional element for 1 dimensional element, we use small letters anyway.

So, this is a transpose Q transpose Q b similarly like Q is a matrix with Q as Q I as columns this Q small Q is a column vector and when we have number of column vectors. We get the Q matrix anyway and this Q transpose Q is identity. So, this is a transpose b so dot product between 2 vector remains same when, we multiply the both the vectors with Q and the angle between the vectors should also then remains same when we multiply them. Because, dot product is we have seen that is it was inequality, dot product is related between the angle between the vectors.

(Refer Slide Time: 21:17)



If a vector b is expressed as a combination of orthonormal basis q 1 to 2 or we write b is equal to x1 q 1 plus x 2 q 2 up to x n Q n. So, there is an orthonormal basis for R n and b is vector in Rn. So, we can express b as the combination of this orthonormal basis some constant coefficient into each of the column vectors basis vectors and the combination is vector b. The coefficients can be calculated as take q 1 transpose multiply it with b.

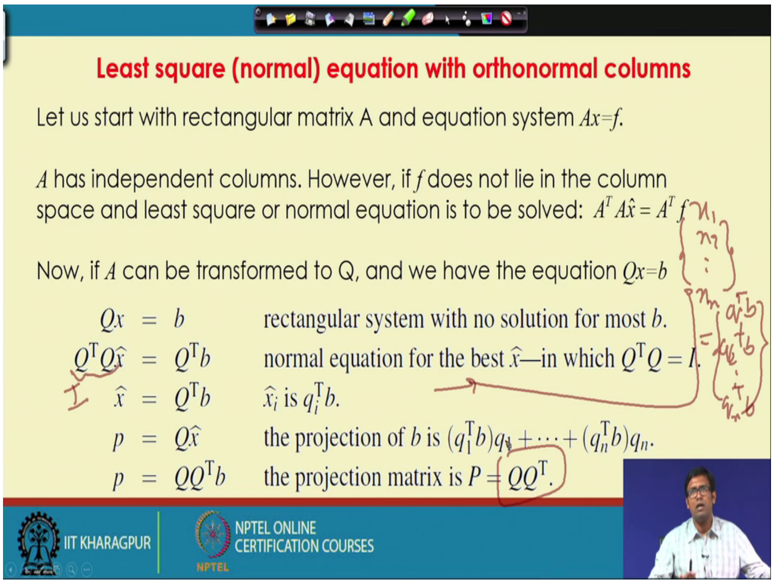
So, q 1 transpose x1 q 1 q 1 transpose q 1 x is a x is a scalar it will come outside q 1 transpose q 1 is equal to 1 q 1 transpose. X 2 q 2 is x 2 into q 1 transpose q 2 and as q is an orthonormal q 1 transpose q 2 is equal to 0. So, there will be q 1 transpose b is x1 q 1 transpose q which is this is one this is 0 this is 0. So, this will be x 1. So, finally, b can be expressed as b is equal to. So, x1 is q 1 transpose b. Similarly, x 2 will be q 2 transpose b x 3 will be q3 transpose b and so on b is equal to q 1 transpose b q 1 plus q 2 transpose b q plus q n transpose b q n.

So, coefficients associated with each of the basis vectors is equal to projection of the vector of the main vector. On each basis this is a very important extremely important way of expressing a function vector this can be utilized for vectors like when we write a vector is equal to 3 I plus 4. That means, it is projection on I along I vector is 3 projection along j vector is 4 also in certain cases we need to Bessel function of a Fourier functions Fourier series expressions.

When there is an expression with mutually of there is expression of continuous function continuous functions of vector spaces in r infinity we have seen that are here. So, there is something r infinity there will infinite orthonormal basis and we can express any function as combination of this infinite orthonormal basis. if you can remember Fourier series the coefficient for each term. In Fourier series, it is sin or cos term in Fourier series is found out by taking the main function then multiplying it with that cosine function or sin function or cos omega x or sin omega is and then they are doing that integration and normalizing it.

So, this is basically projecting the main function along each of the basis vectors and finding it is Fourier series it is important it is Fourier and Bassel series Fourier series within the polynomial. All this series functions are composed of orthogonal basis vectors in r infinity and we can express any function like that so never less Fourier series. Is the as of now out of purview of this particular syllabus we move on.

(Refer Slide Time: 24:35)



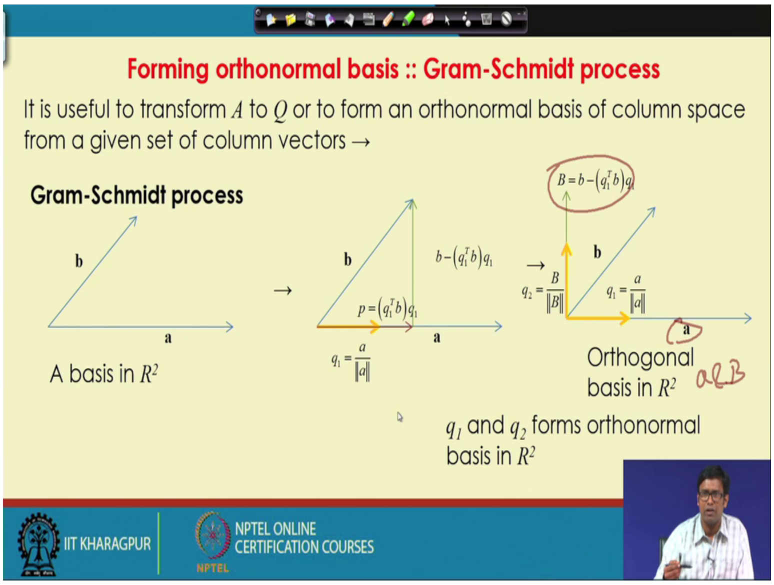
So, we will look into normal equation with orthogonal columns let us start with rectangular matrix a with equation system ax is equal to f A has independent columns. Therefore, there has terming at most one solution. However, f does not lie on the column space. And so, a least square solution or a normal equation is to be solved now a somehow I will be transformed a to Q you will look into this formation and now the equation is Q x is equal to b.

So, we did some matrix transformation matrix operations on a and got a form a is equal to Q, Q is a set of Q has orthogonal columns. So, you get the equation Q x is equal to b Q, x is equal to b is a rectangular system with no solution; for most b if, b does not lie on the column space of a. There is a Q, there is no solution the normal equation is Q transpose Q x tilde is equal to Q transpose b.

So, normal equation for base test tilde in which we can now say that Q response Q is equal to I because, you has orthogonal columns. So, the normal equation reduces to x tilde is equal to q i q transpose b. So, we have basically we have the solution that it here x 1 x 2 up to x n is equal to q 1 transpose b q 2 transpose b. So, on Q n transpose b so the solution is in one line if we can express a as Q. And then we can see that, P is equal to Q x tilde which is the projection of b into the normal into the column spaces.

Basically, x Q multiplied by x so Q Q tilde of b. So, this is q 1 transpose b q 1 plus this physically b expressed in terms of the Q vector q 1 transpose b q 1 up to Q n transpose Q n b. And the projection matrix is Q Q transpose. So, in this these are interesting to look that how is Q Q transpose all this things. However, what we can see that the least square or normal equation becomes much very simple to solve if we can express a as Q and that will give us the next drive that.

(Refer Slide Time: 27:22)



If we can transform A to Q or we can form orthonormal basis of column space form a given set of column vectors. We have a set of column vectors basis of the column space which are not orthogonal.

So, you cannot write Q x is equal to B and solve it very easily. You have a x tilde is equal to f and then we have to solve it like the normal equation a transpose a x tilde is equal to a transpose b. We will come into it later that normally solving normal equations especially when will use literary methods has certain implications will come into it later, but Q transpose what you got Q x is equal to Q transpose b. You do not need to solve it is already solved if you can expressed a is equal to a transform a is as Q that in the normal equation or we observed in the last slide. So, if you if instead of solving a transpose a x is equal to a transpose, f you can solve the normal equation. In this form, which is very simple it is already solved you do not have to do anything.

So, that that is our focus in next slide that it is useful to transform A to Q to form or to form an orthonormal basis of column space, from a given set of column vectors and this is called gram Schmidt process. So, you have a basis which are not perpendicular to each other in R 2 which are not orthogonal basis vectors a and b. What we do we transform make it and unit vector out of it, because, our target is to be orthonormal basis. So, you divide a by it is length and get an unit vector q 1 which is an unit vector q 1 it project b onto a and this projection as we have seen q 1 is the unit vectors.

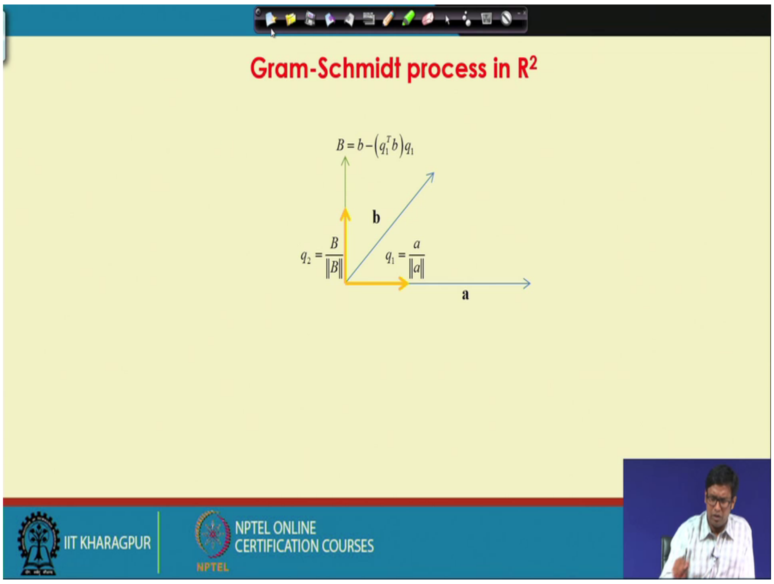
So, this projection; so let s project it into the unit vector. So, the length it is accept results (Refer Slide Time: 29:34) So, becomes q 1 transpose b q 1 because q 1 is along unit vector along a projection along b. Projection along a and projection along q 1 will be same the projection will be same and the left out part is b minus q 1 transpose b q 1. So, this left out part should be perpendicular to the projected part p or perpendicular to q

So now, we can make another basis using the left out part. So, what we will do we have found this term we have found q 1 basis another q one is. Let us say q 1 is one basis this is also this should be independent b capital B which is b minus q 1 transpose b q 1 is independent from q 1 (Refer Time: 30:04). They should also form a basis and will normalize it divided by it is length. So, you get an orthonormal basis before that sorry; here we got an orthogonal basis in R 2 a and the b vector b minus q 1 transpose b q 1 and then these 2 the vector a and this vector a and b they form an orthogonal basis in R 2.

And now, we divide both of the vectors by their length we get q 1 and q 2 form an orthonormal basis in R2. So, this is the idea of gram Schmidt process you take 2 vectors project one vector onto other and the subtract part from the main vector. And the projected vector will be perpendicular to the projection use this 2 and form an orthogonal basis

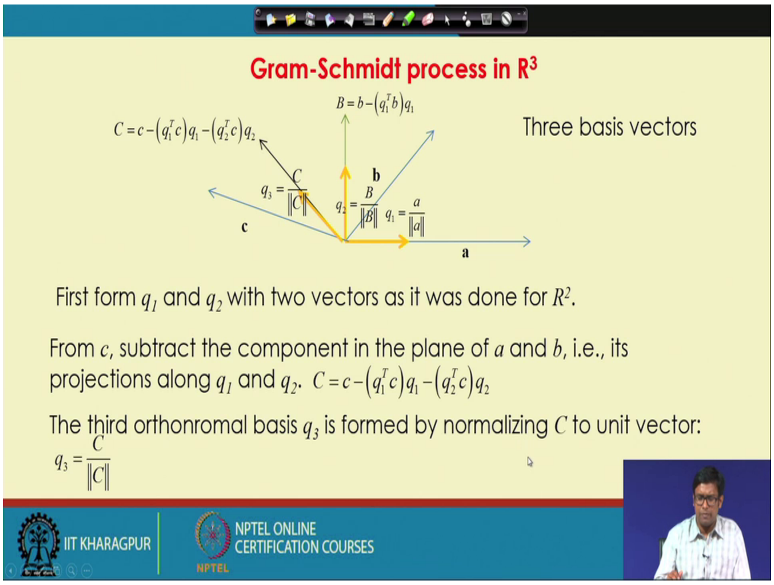
So, this is very straightforward for 2 vectors finding taking a considering a vector finding it is component along one particular vector and perpendicular to that particular vector that when.

(Refer Slide Time: 31:50)



So, this is what we get gram Schmidt process in R 2 we have the vector B project B to a. Sorry, we have the vector a and then find out first orthogonal basis by dividing a by it is length which is q 1. Then project vector b 2 vector, and get the subtract subtraction which is b minus q 1 transpose b q 1 divide it by it is length. So, you get q 1 and q 2 which are orthonormal basis. So, these 2 q 1 q 1 and q 2 they are orthonormal vectors and they give a basis in R2.

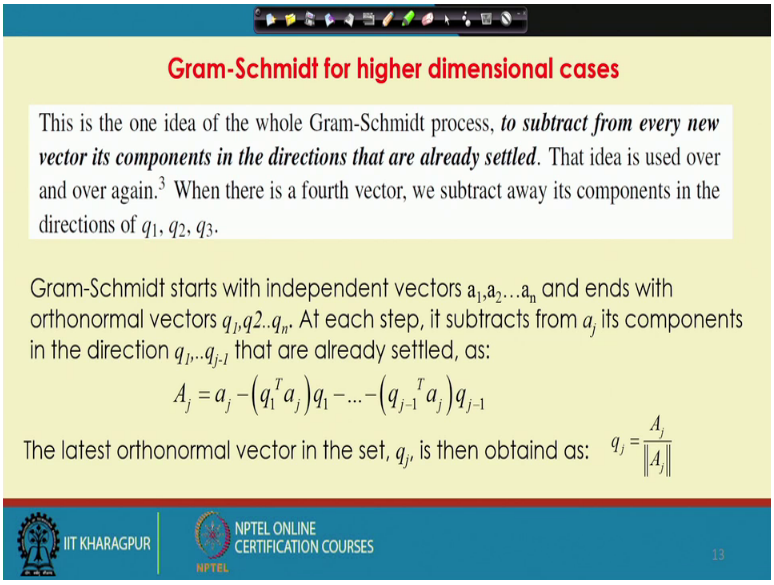
(Refer Slide Time: 32:34)



Now, if we go to R3 here 3 basis vectors. We do the exact same part for first 2 basis vectors the third one is away from the plane. For 2 vectors, we will form a plane on that plane we found; similarly, q 1 and q 2 2 basis vectors for that particular plane for the third one first from q 1 and q 2 with 2 vectors as it was done for R 2. Now for the third one form c subtract the component which is in the plane in form of a and b.

Now, this plane now has an orthonormal basis q 1 and q 2. So, project c alone q 1 project c along q 2 subtract it form the c and you get capital C c minus q 1 transpose c minus into q 1 minus q 2 transpose c Q 2. So, you get capital c and then third orthonormal basis Q 3 is formed by normalizing c to unit vector c by modulus of c and then we can these 3 becomes a orthonormal basis vectors.

(Refer Slide Time: 33:40)

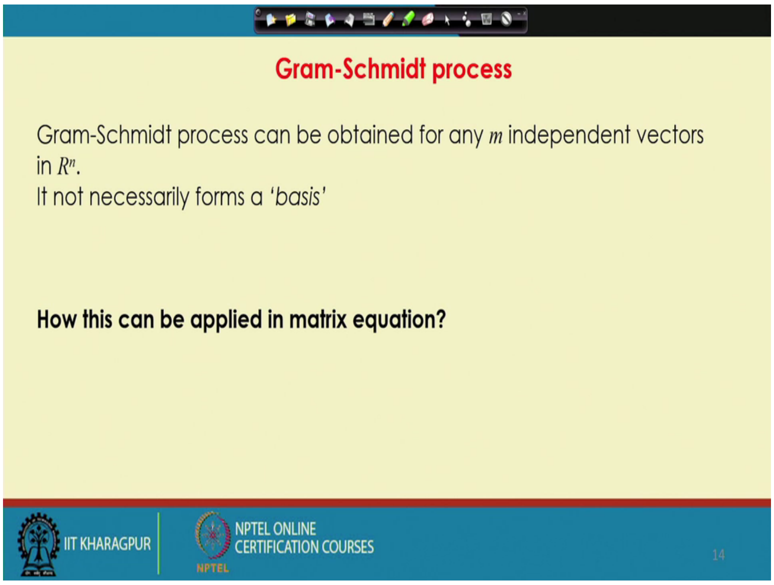


So, then we can continue it for higher dimension this is the one idea of whole gram Schmidt process subtract from every vector every new vector, it is component in the directions that are already settled. So, you get the first you get 2 vectors settle this like get to orthonormal basis vector settle. These 2 vectors then you have the 3rd vector you subtract the component of the 3rd vector, which are when you project the 3rd vector onto the 2 already settled vectors and then subtract the component from there.

So, you get another vector which is perpendicular to the first 2 orthonormal basis. So, this and normalize it get the third orthonormal basis. So, you subtract every new vector from the components that are already settled and then normalize it divided by it is length. The idea is used over and over again when there is a 4th facto we subtract away it is components in the direction of q 1 q 2 q3. So, gram Schmidt starts with independent vectors a one a 2 up to a n and ends with orthonormal vectors. q 1 q 2 up to q n at each step it subtracts from a vector a j which was part of the earlier basis it is components in direction q 1 up to q j minus 1, which are already been found out. And there that are already settled as a j is equal to this.

And then it divides a j by it is length and get the latest orthonormal vector in the set Q j and so on. It can go for a number of vectors and find out the; it can take number of basis vectors and find out the same number of orthogonal orthonormal basis for any set.

(Refer Slide Time: 35:31)



So, gram Schmidt process can be obtained for m independent vectors in Rn. So, it is not it not necessarily that, it has to form a basis you just have to have few independent vectors and then create mutually orthogonal independent vectors out of that. So, take independent vectors which are inclined to it and then create a set of mutually orthogonal vectors. So now, the question is how we can apply it? For the matrix equations if the number of vectors independent vectors in the initial set or number of columns of the matrix is. All these are independent all these are an independent columns.

So, the number of columns in this space is equal to the dimension of the real coordinate space. Then, we have a square matrix and we have solid perfect solution. If it is not, we may have a normal equation we have to find out the best estimate. However, how can we start with the matrix a normalize it is columns orthonormalize it is columns get a q form and solve a x is equal to b. That will be the question which will address in the next few classes.

Thank you.