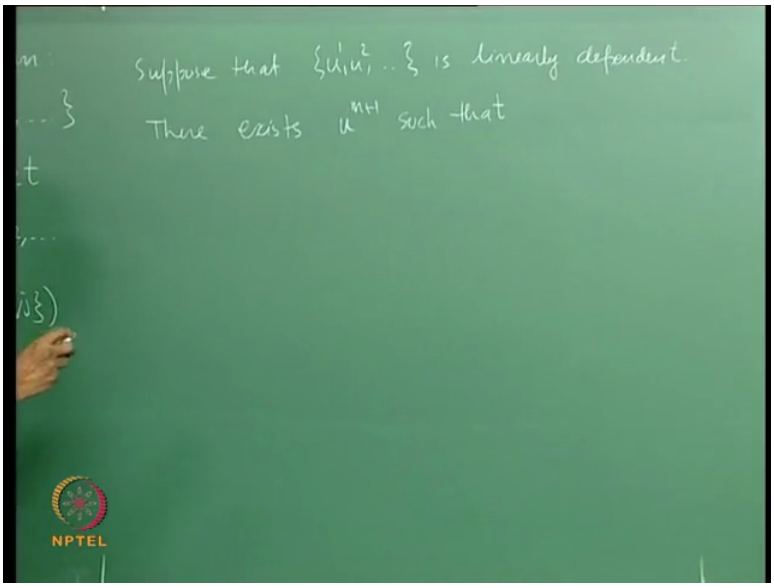
Linear Algebra
Professor K.C Shivakumar
Department of Mathematics
Indian Institute of Technology, Madras
Module 11-Inner Product Spaces
Lecture 41
The Gram-Schmidt Procedure 2

Ok so this is Gram-Schmidt Procedure the formula that I have written down. The idea is to construct an orthonormal set from a linearly independent set ok, this can also be used this procedure can also be used to determine if the set that we started with is linearly dependent ok this can also be used to determine if this U1 U2 etc is a linearly dependent set. Let me explain that.

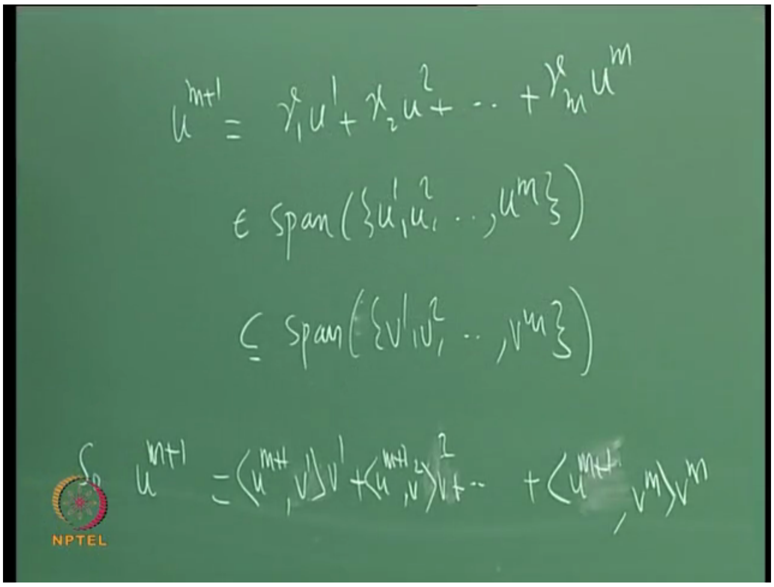
(Refer Slide Time: 00:49)



Suppose this set is linearly dependent ok, then we have learnt that there is atleast one vector which can be written as a linear combination of the preceding vectors ok.

So I will say that there exists offcourse I am assuming that zero vector does not belong to this, if zero vector is already there you don't have to check linear dependence ok. So I will assume that these are non-zero vectors and I want to see when it follows from the procedure that this set is linearly dependent. So there exists U let us say m plus 1 such that there is atleast one vector that can be written as a linear combination of the preceding vectors.

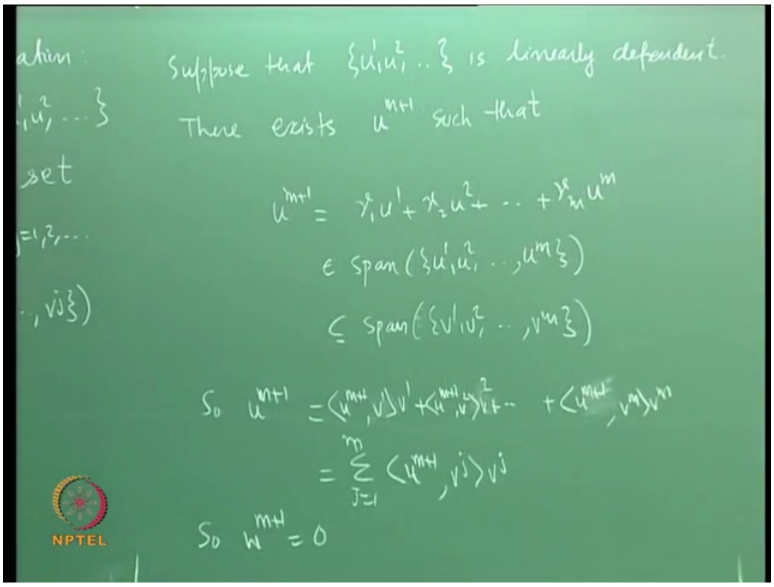
(Refer Slide Time: 01:56)



So Um plus 1 such that Um plus 1 can be written as let us say gamma 1 U1 plus gamma 2 U2 etc gamma m Um.

Now this belongs to span of U1 U2 etc offcourse which in turn is equal to contained in span of V1 V2 etc but what is extra about this V1 V2 etc, is that there are orthonormal vectors ok. So when I write Un plus 1 belongs to this span this vectors are orthonormal then we have this yesterday Un plus 1 can be written as scalar times V1 plus scalar times V2 etc, what are the scalars? Inner product of Un plus 1 with each of this, so that is Um plus 1 V1 this is Um plus 1 V2 plus etc U m plus 1 Vm into Vm this is the expression for Um plus 1 in terms of the orthonormal vectors V1 V2 etc Vm ok.

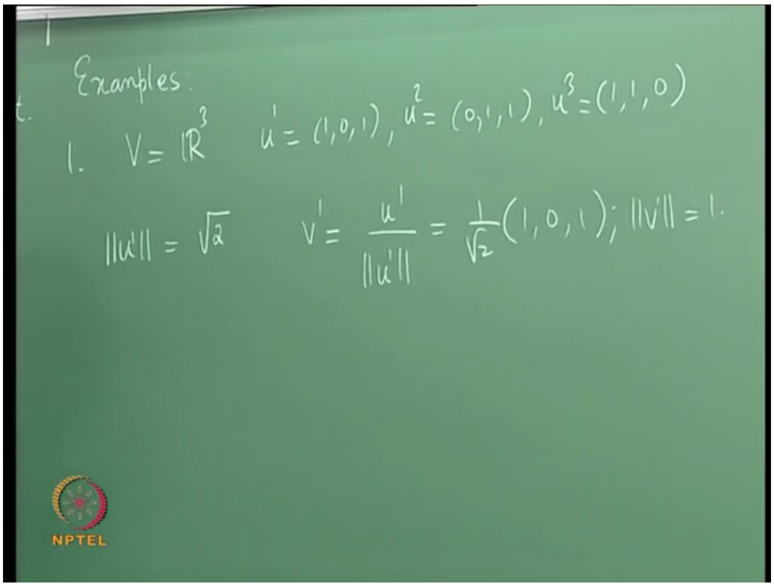
(Refer Slide Time: 03:51)



This is summation j equals 1 to m Um plus 1 is fixed we have Vj times Vj Um plus 1 is equal to this from this can you see that Wm plus 1 is 0 the expression for Wm plus 1 is here. So what does this convey? It means that any stage if you get the zero vector then you have linear dependence Wm plus 1 is the vector that we are determining at the m plus 1 stage, if that turns out to be the zero vector then the vectors that we started with much be linearly dependent.

Infact Um plus 1 is a linear combination of the preceding m vectors. See this whole thing can be traced back, Wm plus 1 is 0 if and only if Um plus 1 is equal to this which will tell you that Um plus 1 is a linear combination or the preceding m vectors ok. So one can check linear independence also by using the Gram-Schmidt Procedure, is that clear? Ok lets now look at two examples numerical examples and see how this procedure gives us orthonormal vectors.

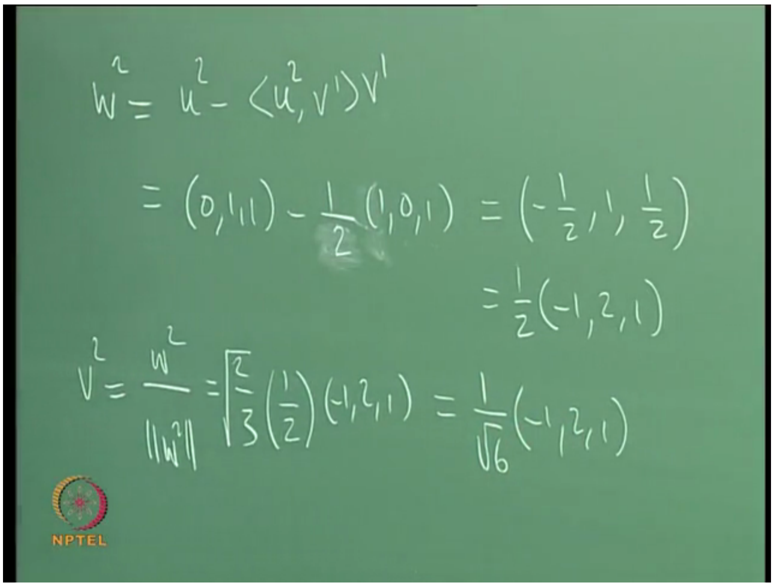
(Refer Slide Time: 05:22)



Lets look at the case of R3, V is R3 with the usual inner product I take the following vectors U1 is 1 0 1, U2 is 0 1 1 and U3 is 1 1 0 ok then you can verify that these vectors are linearly independent, so they form a basis for R3.

You would like to construct an orthonormal basis for R3 starting with this linearly independent set ok. So lets do the computation. The inner product is the usual inner product so the norm is the Euclidian norm. So if you look at norm of U1 that is root 2 and the formula for V1 is U1 by norm U1 which is 1 by root 2 into 1 0 1, this is V1 and it is clear that norm of V1 is 1, that is how it has been constructed.

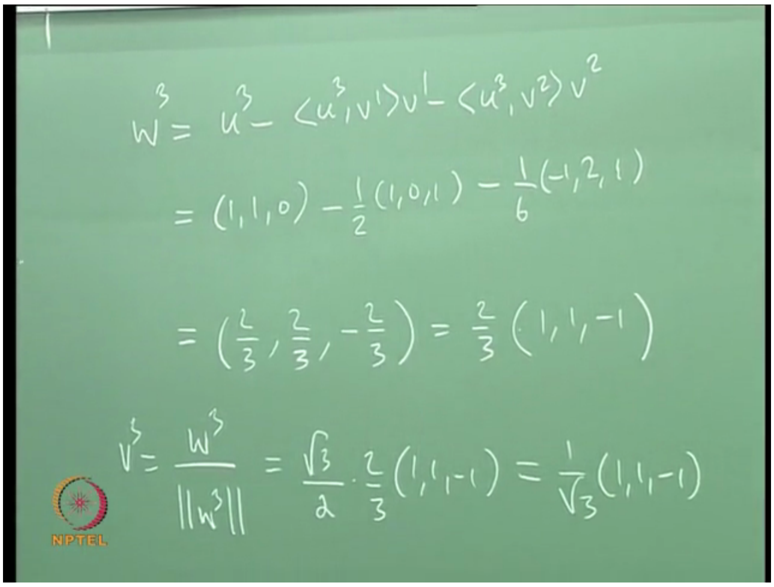
(Refer Slide Time: 06:40)



To construct a V2 we must construct W2, W2 is U2 minus the inner product of U2 with V1 into V1 from the formula so U2 is 0 1 1 minus inner product of U2 with V1 this two terms get cancelled you get 1, 1 by root 2 minus 1 by root 2 into V1, V1 again goes with a 1 by root 2 so I will write 1 by 2, V1 goes with a 1 by root 2, 1 0 1 so that is minus 1 by 2 and 1 minus 1 by 2 let me write this as 1 by 2 outside minus 1 2 1 that is W2 this does not have norm 1 so I must divide define V2 as W2 by norm W2, norm W2 is 1 by 4 into 1 plus 4 plus 1 that's 6 by 4 that's 3 by 2, 1 by that plus 2 by 3 into 1 by 2 I think root 2 by 3 into minus 1 2 1.

So let us compute this gets cancelled 1 by root 6 into minus 1 2 1 , this is my V2, for one thing norm V2 is 1, 1 plus 4 plus 1 by 6 the W2 must be orthogonal to V1 which is clear this is minus 1 by 2 this a plus 1 by 2 the dot product is zero I am just checking the calculations. So this is V2.

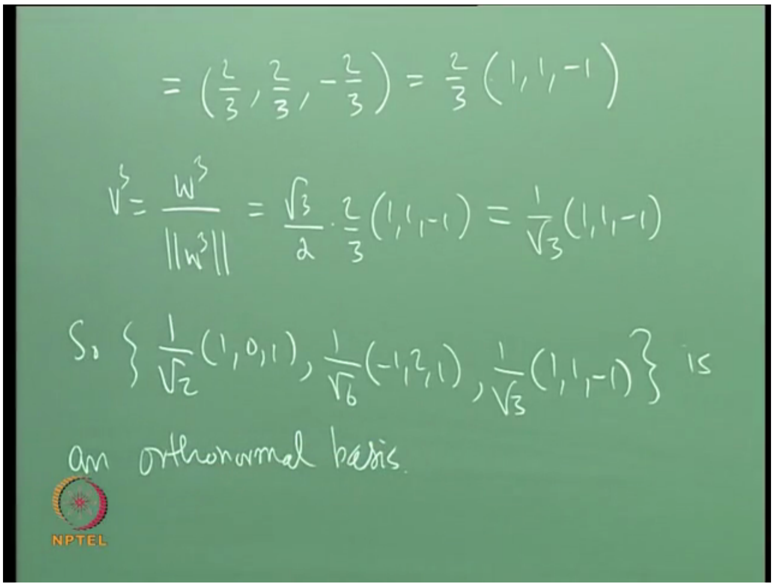
(Refer Slide Time: 08:47)



V3 is constructed using W3, W3 is U3 minus U3 V1 V1 minus U3 V2 V2, U3 is 1 1 0 minus U3 V1 V1 see this 1 by root 2 comes twice because V1 comes twice so you go with 1 by 2. So I will write 1 by 2 first what is U3 V1? U3 V1 first just 1, so it is just 1 multiply by V1. So V1 is 1 by 2 I have already written this 1 0 1 minus U3 V2 with U V2, V2 goes with a 1 by root 6 so I get a 1 by 6 outside U3 V2 that's just one again.

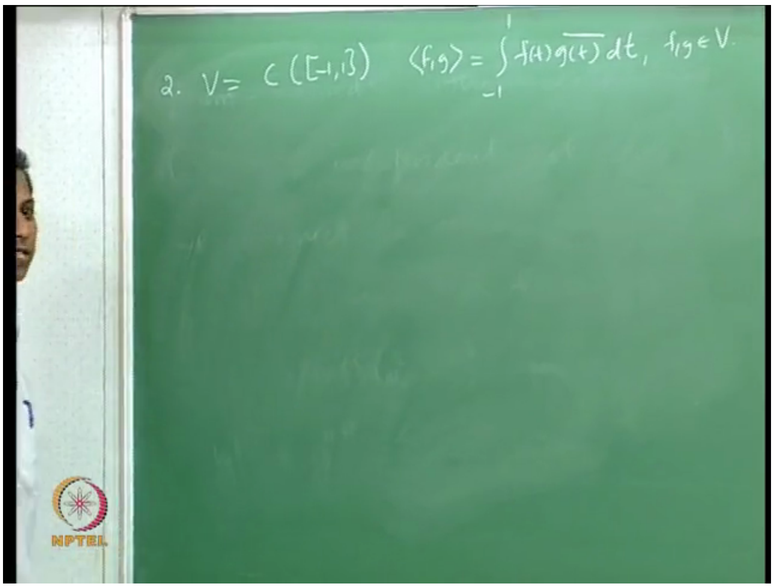
So 1 by 6 into V2 without the constant minus 1 2 1, that is 1 minus 1 by 2 plus 1 by 6, what is that? 2 by 3, 1 minus 1 by 3, minus 1 by 2 minus 1 by 6 minus 2 by 3. Please check this calculations so if you look at the dot product of this vector with V1 and V2 that is zero this will go with minus 1 minus 1 minus 2 plus 2 zero and W3 what is the norm of W3? V3 is W3 by norm W3, norm of W3 is 4 by 9 into 3 that is 4 by 3 so root 3 by 2 into 2 by 3 into 1,1, minus 1, so that is 1 by root 3 1,1, minus 1 that is V3. So you can now verify that norm V3 is 1 and we have already verify that this is orthogonal to the vectors V1 and V2 ok. So this gives an orthonormal basis.

(Refer Slide Time: 11:42)



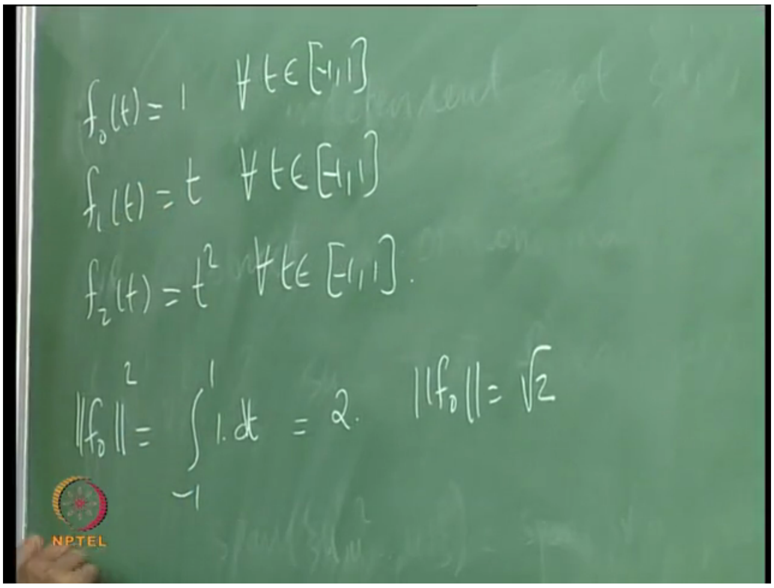
So let me just summarize V1 is here, 1 by root 2 1 0 1 V2 is here 1 by root 6 minus 1 2 1 and 1 by root 3 1 1 minus 1. So this is what we get after applying the GS process which has the required properties that span of V1 the span of U1 span of U1 U2 span of V1 V2 span of V1 V2 V3 span of U1 U2 U3 ok. So this is corresponding to the discrete case that is finite dimensional case. Lets look at a linearly independent set coming from C01 lets say C minus 1, 1.

(Refer Slide Time: 12:51)



So I wanted to do one example coming from C minus 1, 1 with the usual inner product ok.

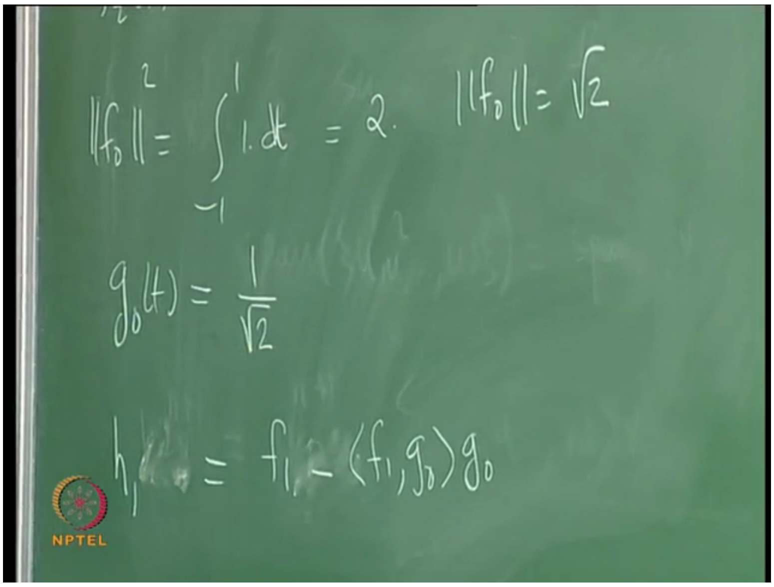
(Refer Slide Time: 13:22)



So let me start with the usual linearly independent set. So lets say I have F knot of t is 1 F1 of t is t F2 of t is t square, so I just take this see this vector space has an infinite has a basis constituting of infinitely many elements. So I am just taking three linearly independent functions I want to apply GS process for this set and obtain an orthonormal functions ok.

This will involve computing integrals ok. What does norm of F knot? See I will use F knot and G knot the intermediate W's I will use H ok, what is norm F knot? Integral minus 1 to 1 mod F t square dt that is just dt one times dt which is 2 norm F knot square is 2 and so norm F knot is root 2 I said I will define G knot G1 G2.

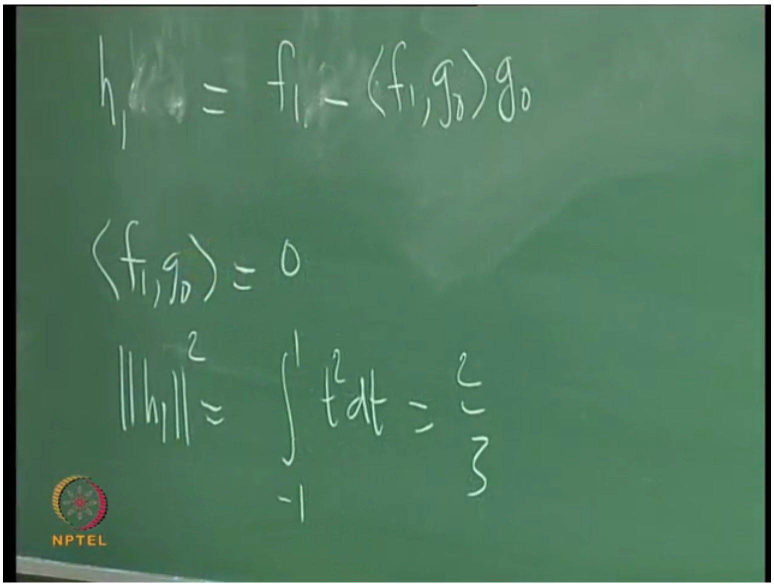
(Refer Slide Time: 14:52)



So G knot of t I will define that as 1 by root 2 then this has norm 1 ok, this is the first vector, this is V1. Construct the second function which is orthogonal to this function with respect to this inner product. The procedure I will use H from G knot we are trying to construct G1 so I will use H1, H1 of t is, is that F knot and G knot, see G knot is f knot by norm F knot ok, H1 of t is so what is the formula? I take the second vector, I must take the second vector so that is F1 that is what I start with, F1 of t minus ok let me write without t, F1 minus F1 G knot G knot is that ok?

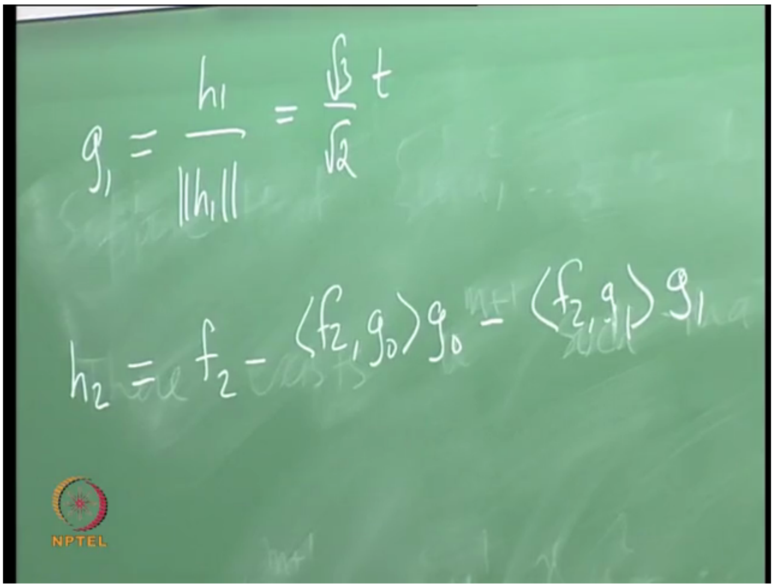
This are my U's this are my V's W's will be my H F knot is a first vector, F1 is a second vector.

(Refer Slide Time: 16:21)



So what is inner product F1 with G knot? Odd function that zero, so it is just F1 is that ok, F1 zero's G0 is 0 which means H1 is just F1 and I have to just divide by norm F1. So what is norm H1? That's norm F1 minus 1 to 1 t square, 1 by 3 and you apply its 2 by 3 or what, yes norm square is 2 by 3 ok.

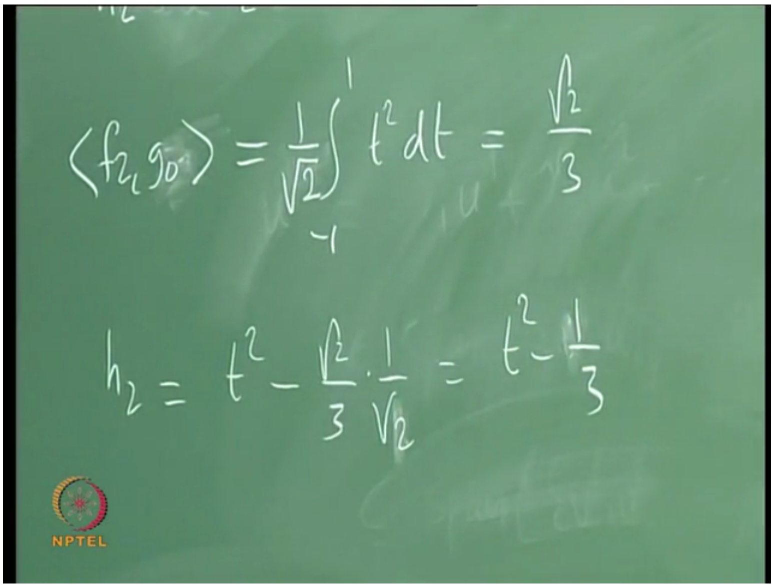
(Refer Slide Time: 17:17)



So what is my G1, G1 is H1 by norm H1 root 3 by 2, H1 is just F1, F1 is t ok this is H1 sorry this is G1.

Finally I need G2 so I need W2 which is my H2. H2 is F2 minus F2 G knot G knot minus F2 G1 G1, F2 G knot F2 is t square G knot so this is an odd (dot) even functions that is not zero F2 G1 F2 with G1 that becomes an odd function so that second term is third term is zero.

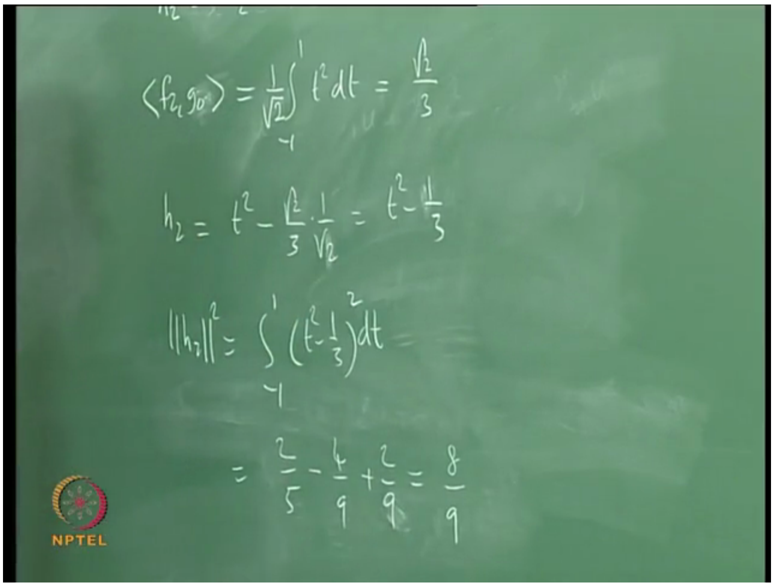
(Refer Slide Time: 18:24)



So this is F, ok what is F2 G knot? Minus 1 to 1 t square and G knot is constant, so 1 by root 2 I want F2, F2 is t square, so that is 2 by 3 root 2 by 3 just root 2 by 3 is that correct, 1 by 3 and 2 by 3 that gets cancelled root 2 by 3 that is F2 G knot that is this term so H2 is F2, F2 is t square minus F2 G knot is root 2 by 3 into G knot, G knot is just 1 by root 2. This is t square minus 1, 1 by 3.

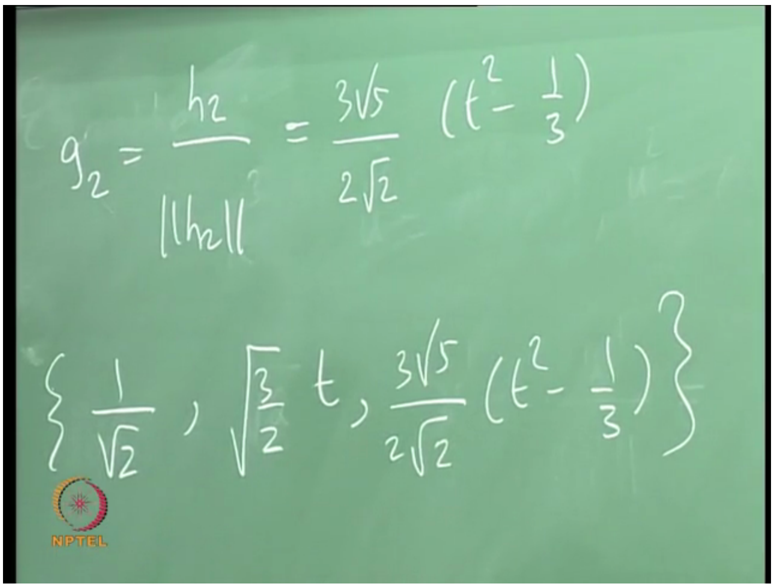
Finally we need to compute the norm of this and then divide H2 by the norm that's G2 ok.

(Refer Slide Time: 19:32)



So what is norm of H2? Minus 1 to 1 t square minus 1 by 3 dt that is 2 by 3 minus sorry ok so let do the calculation t square (())(20:05) whole power 2 that is t power 4 can we do mental calculation hopefully not make mistakes that's t power 4, t power 4 is 1 by 5, 2 by 5 minus 2 t square by 3, minus 2 by 3 t square, minus 4 by 9 and plus 1 by 3 into 2 by 3 is this correct, whole square. 2 by 5 minus 2 by 9 so what is the 8 by 9 that is norm square so 45 yes.

(Refer Slide Time: 21:00)



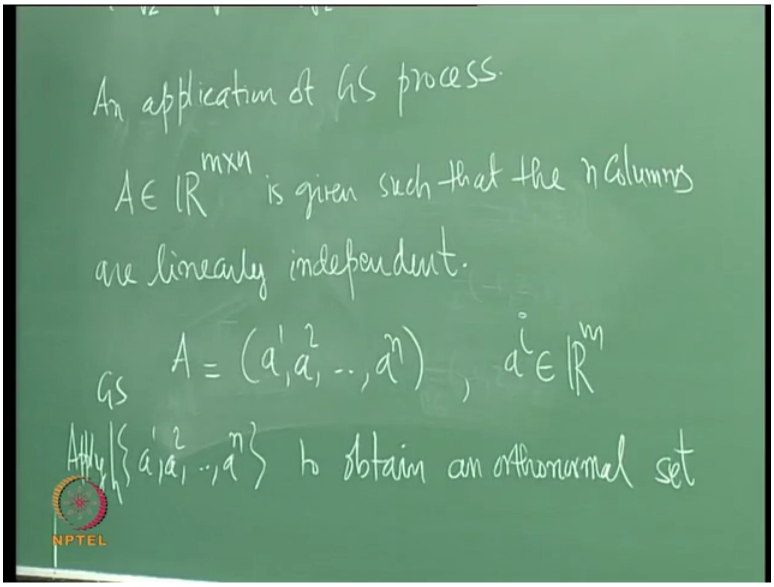
Which means my G2 is H2 by norm H2 that is 45, 9 into 5 that is 3 root 5 by root 8 that is 2 root 2 into H2, H2 is t square minus1 by 3 is that ok.

Root 45 by root 8, see the computations are (())(21:36) the only thing is this is a little more complicated but what is the moral of the story, let me write down this functions that we have got. G knot, G knot is the constant function 1 into 1 by root 2 G1 is the function that I have here root 3 by 2 times t and G2 another constant times t square minus 1 by 3 ok. See this functions appear have you seen this functions? Multiple of constant multiple of t, multiple of t square minus 1 by 3 etc.

That is the purpose of giving you this example the functions see suitable multiples of constant suitable multiples of t suitable multiples of t square minus 1 by 3 etc are Legendre polynomials. Legendre polynomials come when you model certain physical problems Legendre differential equations these are the Legendre polynomials ok.

They have some nice properties which you must have studied in your differential equations course, you apply what is called as the power series method to the differential equation and for some values of n you get polynomial solutions, this are those polynomials ok. So just to tell you that this are not different, they are interconnected differential equations linear algebra and many other subjects they are all interconnected only to illustrate that point I wanted to do this example ok. Ok that is GS process and examples I want to discuss one application of the GS process, lets look at the following it might appear to be completely different situation but you can apply the GS process to infer the following.

(Refer Slide Time: 23:50)

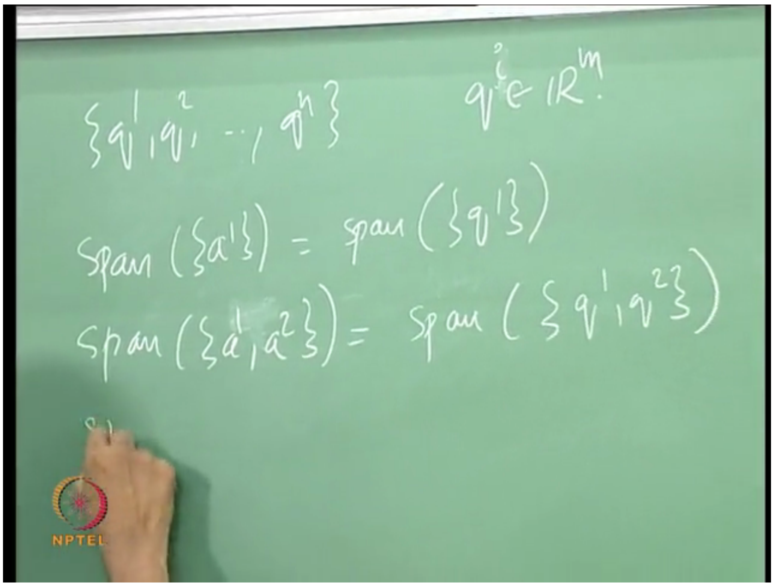


So I want to discuss an application, see what I would like to do let me tell you that beforehand is to derive certain decomposition for matrix with real or complex entries given that the columns are linearly independent ok it is called as a QR decomposition. So what exactly is a problem.

Lets say we are given a matrix A in terms (())(24:27) so we have a matrix with m rows and n columns, this is given such that the n columns are linearly independent let me denote the n columns by A1 A2 etc so I am writing A as a1 a2 etc a n the n columns so each ai is a column ai is the ith column of a and lets also remember that ai belongs to Rm ok, there are m rows so I am writing A just by using the columns of A. Now this columns are linearly dependent so I consider the set a1 a1 etc a n this is a linearly independent set of vectors I apply Gram-Schmidt process apply GS to this I get the following set I get an orthonormal set of vectors to obtain an orthonormal set.

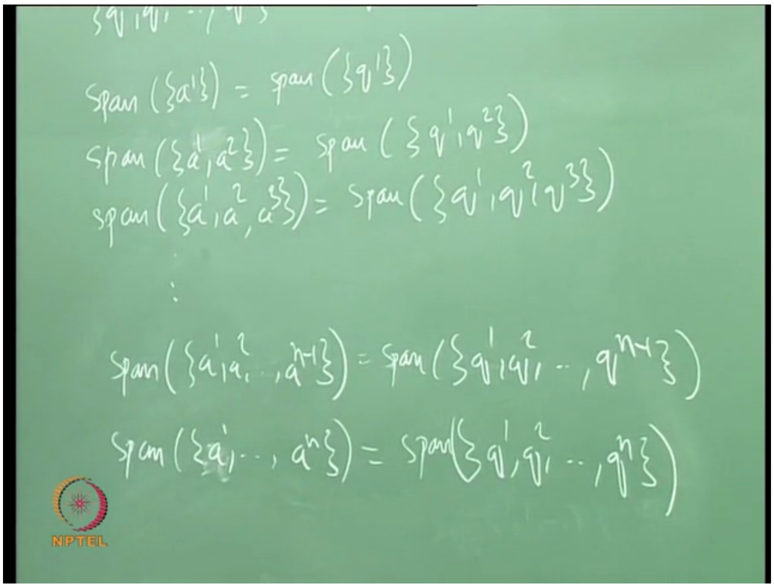
Given a matrix A with linearly independent columns the rectangular matrix with linearly independent columns consider the column vectors are linearly independent vectors apply Gram Schmidt process to get orthonormal vectors n orthonormal vectors in Rm.

(Refer Slide Time: 26:26)



I will call those vectors q1 q2 etc just to emphasize ok that each of these vectors belong to each belongs to Rm. There are some properties right that these two sets together satisfy let me write down. Span of a1 is span of q1 which means q1 is a multiple of a1 infact that is how we construct span of a2 sorry a1 a2 this is span of q1 q2.

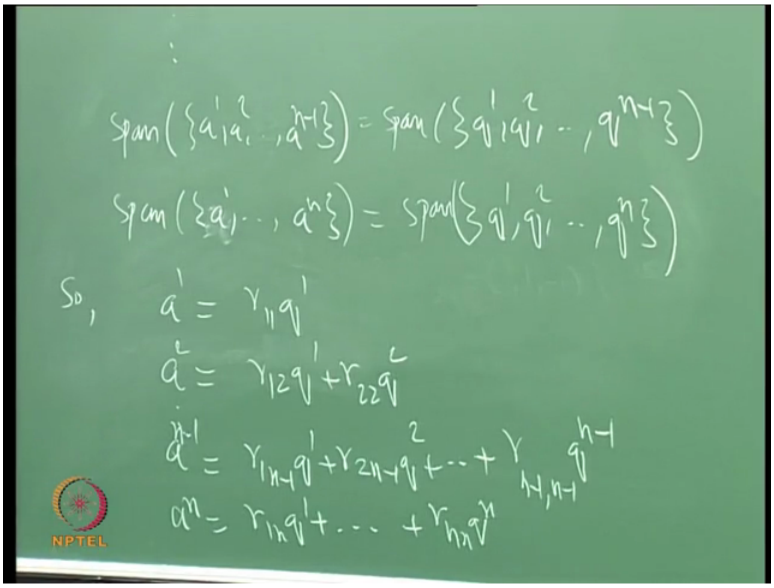
(Refer Slide Time: 27:17)



I will just write one more and write down the last relation a1 a2 a3 this is span of q1 q2 q3 etc we can proceed in this manner.

The last step is span of ok last but one span of a1 a2 etc an minus 1 that is equal to span of q1 q2 etc q n minus 1 and finally span of q1 sorry a1 etc a n is span of q1 q2 etc q n ok. So we have these equations between subspaces.

(Refer Slide Time: 28:24)

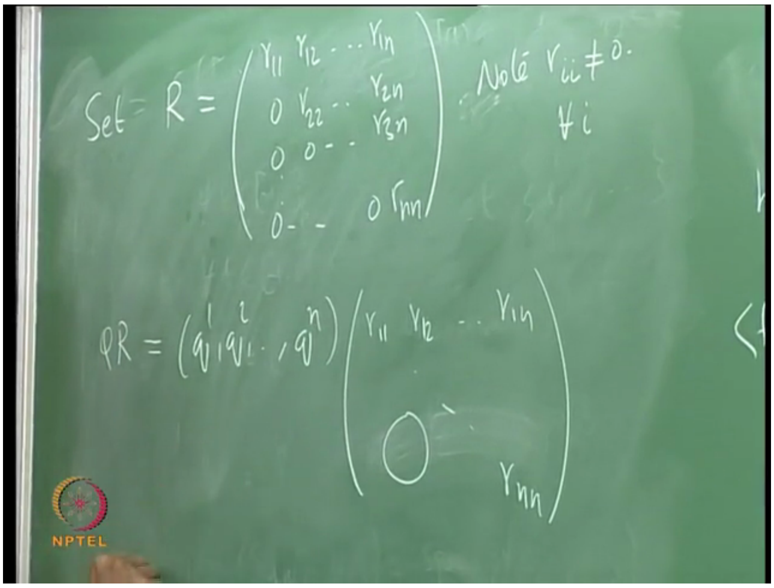


Write down the first equation from this a1 is a multiple of q1, q1 is a multiple of a1 really but that is an non-zero multiple you can divide. So a1 is a multiple of q1 I will call it r11 q1 ok. Look at the second look at a2, a2 belongs to span a1 a2 and so it belongs to span q1 q2 so a2 is a linear combination of q1 and q2.

I can write a2 as a linear combination of q1 q2 I will write it as r 12 q1 plus r 22 q2 the next equation is similar ok I will proceed like this the last one gives me a n minus 1 a n minus 1 is r1 n minus 1 q1, r1 to 22 the second the is fixed it is the first one that changes r2 n minus 1 q2 etc plus you go only upto the n minus 1th term r n minus 1n minus 1 q n minus 1 the last one will have all the n terms let me write the last one a n is r1 n q1 etc r n q n ok can you see it when I write it here? Ok can r1 1 be zero? Can be zero because I started with a linearly independent set so zero is not present if r11 is zero it means a1 is zero that is not possible, can r22 be zero?

Suppose r22 is zero then I am writing q1 in term a2 in terms of q1, q1 in term can be written in terms of a1 so I am writing a2 in terms of a1 not possible a1 a2 are independent etc r n minus 1 n minus 1 can that be zero for the same reason it cant be zero. If it is zero then I am able to write q1 etc q n minus 2 I am able to write a n minus 1 in terms of q1 etc q n minus 2 that is a1 etc a n minus 2 which means a1 etc a n minus 2, a n minus 1 is linearly dependent that is not possible. So all these last coefficients are non-zero ok. So I form a if I form a matrix with this as diagonal entries let us say a triangular matrix that matrix should be invertible.

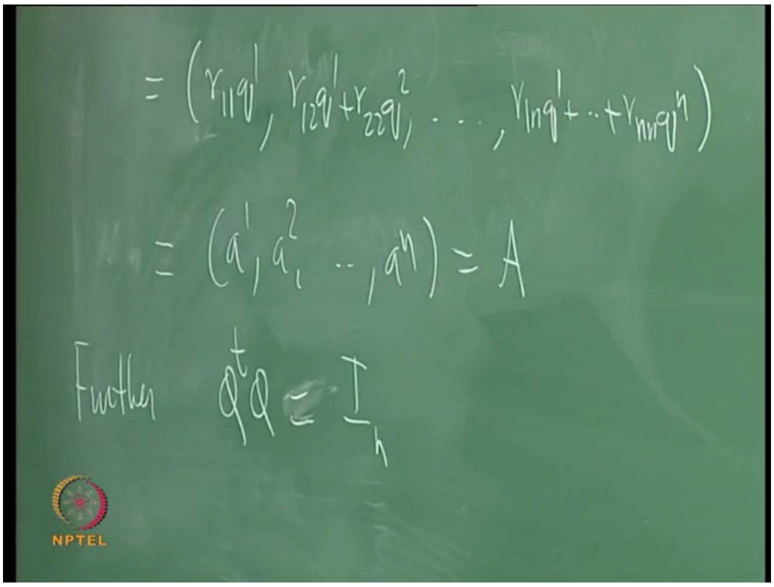
(Refer Slide Time: 31:25)



So lets say we call the matrix R as so I have written it in that manner r11 r1 to etc so it is an upper triangular matrix r11 r12 etc r1n this entry is zero r22 etc r2n 00 r3n etc all these entries are zero the last entries are n n, this is no matrix R, there is an upper triangular matrix the lower the entries beyond the below the principle diagonal are zero and the diagonal entries are not zero. Note rii is not zero for all i and so this is an invertible matrix ok this is an invertible matrix the diagonal entries are not zero this is invertible.

So if you look at QR for instance what is Q? Q is my matrix whose columns are q1 q2 etc qn into r11 r12.See if you look at this matrix QR then can you see that this is see you do this multiplication r11 into q1 plus this entry into q2 etc into qn which means I have only one term r11 q1.

(Refer Slide Time: 32:59)

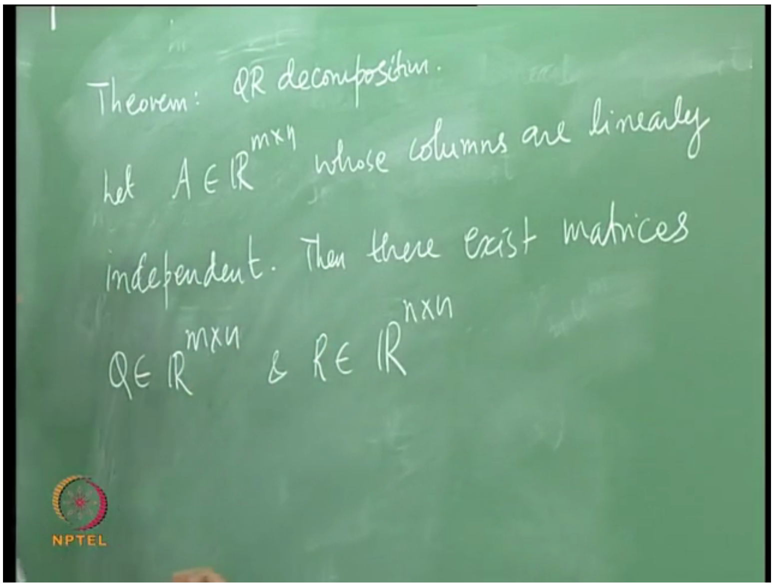


The next one is r12 sorry the next one is r12 q1 plus r22 q2 r12 q1 plus r22 q2 etc the last one has all this terms, r1n q1 etc rnn qn ok, but what this are is already there this equation is already there. R11 q1 is a1 r12 q1 plus r22 q2 that is a2 etc that is a n which is the matrix that we started with. So q times r is A this is called the Qr decomposition of a matrix given that the columns of the matrix are linearly independent.

We have not said anything about we have not made use of the fact that the Qi's are orthogonal orthonormal infact, what happens to this? See remember q is a rectangular matrix the order of Q is same as order of A so Q is m cross n ok. Q transpose q is n cross n ok, what can you say about q cross Q transpose q? What is Q transpose q? Or what do you know about the columns of, can you see its identity? See its identity of order n Q transpose Q the identity of order n but if you look at Q Q transpose you cant expect that to be this because in the first place Q Q transpose is of order m ok. Such a matrix is called an orthogonal matrix.

See this is a matrix with real entries such a matrix is called an orthogonal matrix if you have a matrix with complex entries that satisfies this Q star Q equal to identity it is called unitary matrix in the rectangular case, in the square case lets remember that in the square if Q transpose Q is identity then Q Q transpose also has to be identity ok. So this matrix Q has this extra property that comes primarily from the fact that the columns are orthonormal vectors ok.

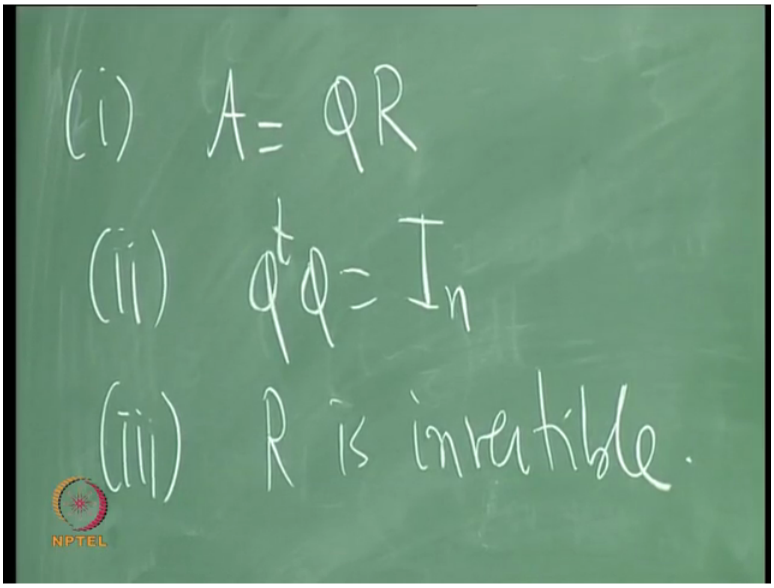
(Refer Slide Time: 35:44)



So what we have proved is the following theorem, Q R decomposition. Let A belong to R m cross n with whose columns ok, let me write like this whose columns are linearly independent then there exists matrixes Q and R Q has the same order as A.

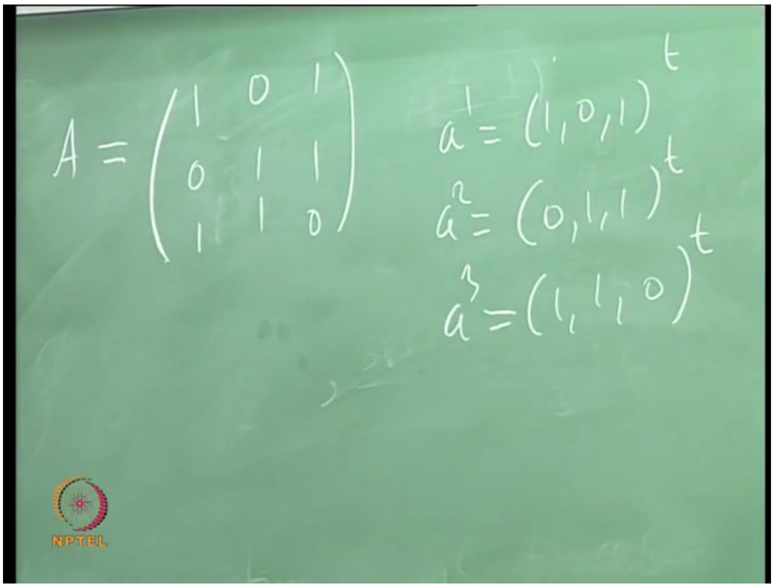
Their exists matrix is Q in R m cross n and R in R n cross n R is a square matrix such that the following condition holds.

(Refer Slide Time: 36:54)



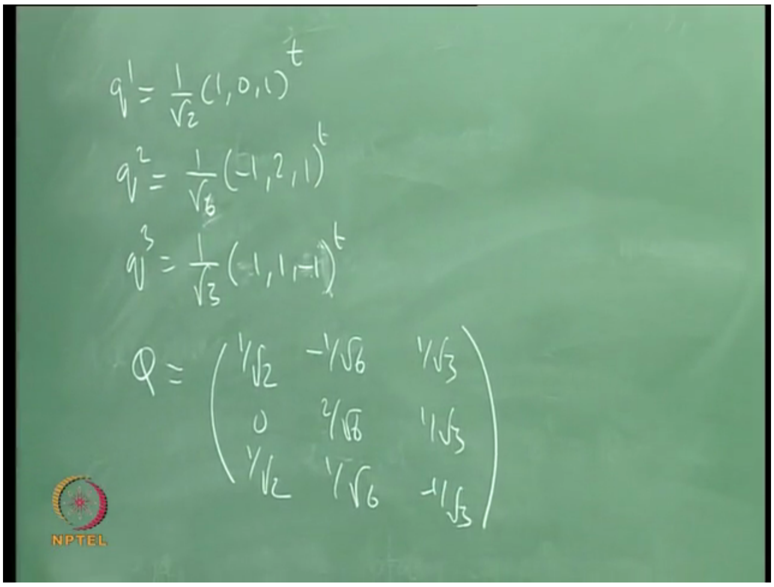
The first one is this decomposition A is Q times r second condition Q transpose Q is identity of order n and the third condition third property is that R is invertible ok. The idea behind studying decompositions of a matrix is that the question about problems involving the matrix A can be reduced to questions about problems involving Q R, it gives a kind of a reduction ok. I will explain that a little later but lets maybe work out an example. Ok so we had this theorem has been proved.

(Refer Slide Time: 38:01)



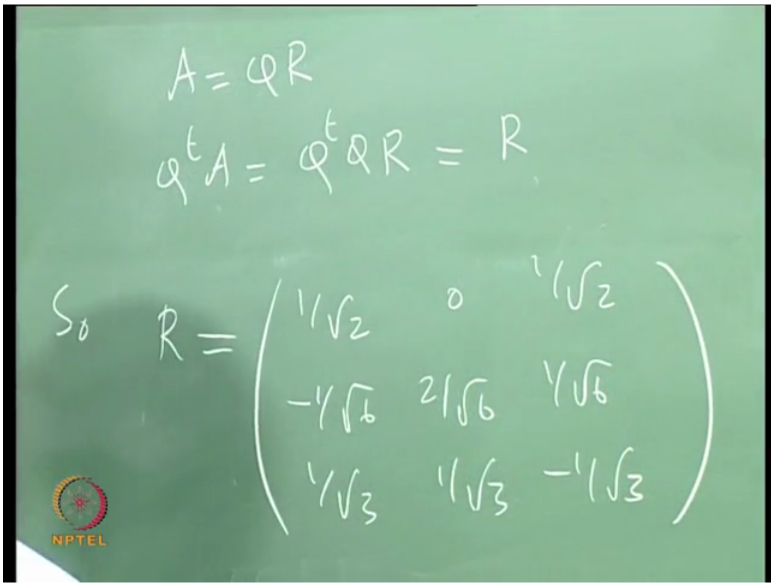
Lets look at an example that we have done before today this Gram Schmidt process for the first example so can you just tell me those vectors, lets say the matrix A the three vectors a1 a2 a3, what is the first one? 1 0 1, 1 0 1, 0 1 1, 1 1 0 this are linearly independent vectors. So I have a1 as 1 0 1, a2 is 0 1 1, a3 is 1 1 0 see by the way I am when I write vectors I interchangeably write them as rows and columns from the context it should be clear whether it is a row or a column, it is only for convenience that I write it as a row vector but remember that a1 a2 a3 are the column vectors of A ok. So strictly speaking I must write transpose ok but at times I might omit that but just make sure that the compatibility, matrix multiplication etc that is done correctly, that is appropriate.

(Refer Slide Time: 39:07)

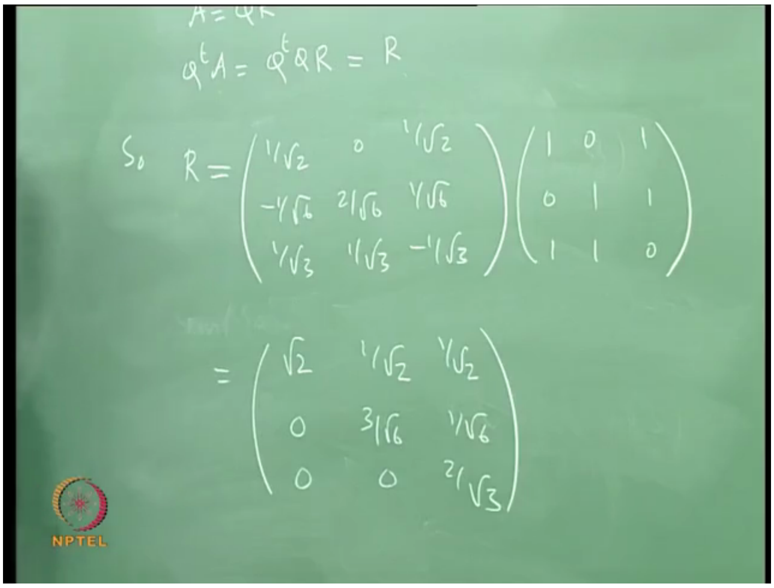


So these are the vectors we obtain q1 q2 q3 so q1 I can write down 1 by root 2 into 1 0 1, q2 you should tell me, minus 1 1 1 no, minus 1 1 1 q2? Is this correct? Ok and q3 I remember 1 by root 3 at must that much minus 1 1 ok. So the matrix Q in this problem is ok you have to just write down all this 1 by root 2 0, 1 by root 2, minus 1 by root 6, 2 by root 6, 1 by root 6, 1 by root 3, 1 by root 3, minus 1 by root 3 this is my Q. See to determine R I need to solve those equation but that is not necessary.

(Refer Slide Time: 40:28)



(Refer Slide Time: 42.05)



See A is Q R ok I will do that A is QR I can pre-multiply by Q transpose, Q transpose A is Q transpose Q R this where we use the second property, this is just R.

So R is Q transpose A, so just do the multiplication matrix multiplication to get R, R must ne an upper triangular matrix I am not simplifying just leaving the fractions as they are, ok this is R ok so computationally you don't have to solve those equations so the property that Q transpose Q equal to I saves effort of computing R ok just pre-multiply do one matrix multiplication which is much easier ok. Let me stop here and discuss in other application of

QR with regard to what are called as least square solutions of linear systems ok. So I will stop here.