Matrix Solvers

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

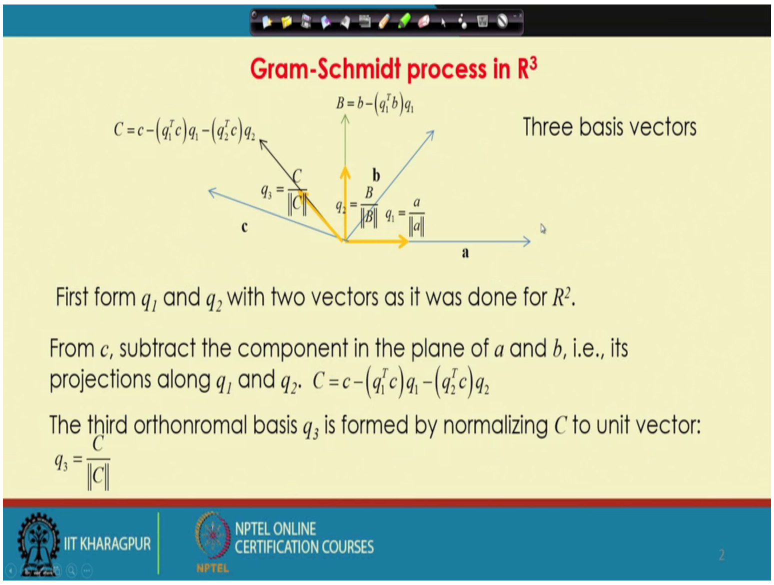
Lecture- 26
Gram-Schmidt and Modified Gram-Schmidt Algorithms

Welcome. In last class we looked into Gram Schmidt process this process essentially starts with a set of independent vectors and then generate a set of orthogonal vectors from those independent vectors it is not only orthogonal vectors the vectors are also orthonormal vectors which we obtain through or Gram Schmidt process. So, in that process if we have a matrix with independent columns the matrix turns into a q matrix or the transform matrix turns into a q matrix or a matrix with normal columns in case of a square matrix we get a matrix, which is called an orthogonal matrix or matrix which has all orthonormal columns. So, in today's class we will start look into more detail on Gram Schmidt process.

And we have earlier discuss that getting orthonormal columns is important in terms of matrix solutions for example, if we have a matrix ax is equal to b and we can convert it to a form q x is equal to c where q is the orthonormal transformation of the matrix a and then we can take the c vector and project it against all of the basis orthonormal basis of q or all the columns of q and the projection will be solution x 1 x 2 x n. So, the matrix solution becomes essentially match simple if we can transform a matrix from the matrix a to an orthonormal matrix q.

So, today's class we will look into more detail of the Gram Schmidt process do a quick recapitulation of what we discussed in last class and then will try to do some examples of Gram Schmidt process we will see how the Gram Schmidt algorithm should be for practical implementation purpose and the will go to qr decomposition of the matrix and see how it is useful in matrix solvers.

(Refer Slide Time: 02:17)



So, Gram Schmidt process in real coordinates space R 3. We have discussed this exactly, this slide in the last class; I have 3 vectors a b and c. This 3 are independent vectors, but they are not mutually orthogonal to each other.

And the goal of Gram Schmidt process is to get a set of mutually orthonormal basis vectors. Orthonormal means, the vector should be orthogonal to each other and each of the vector should have a link equal to unity. So, we take the first vector a and divide it by it is lengths. So, we get an unit vector along a which is first vector in my Gram Schmidt set the next vector q 2 which will be orthogonal vector a. As well as it should be it should contains some component of b q and it should have a unit length. So, what we do? We project b on q 1 q 1 is an unit vector.

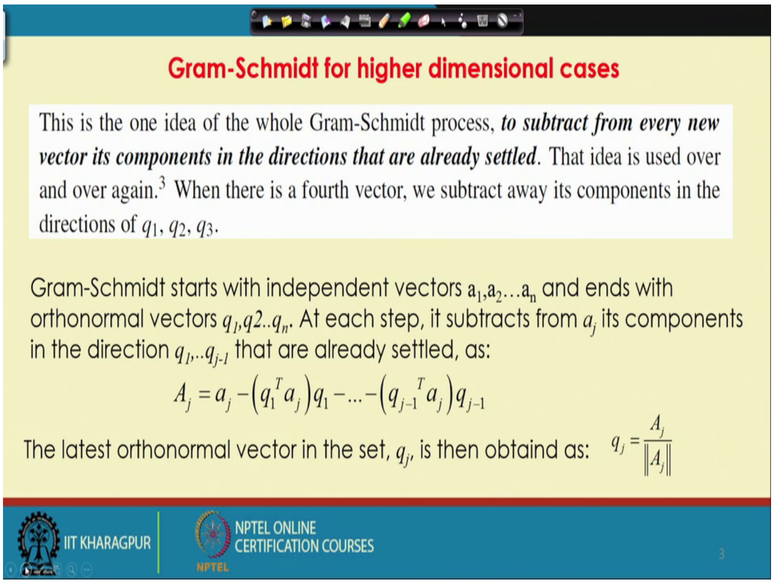
So, the project if I take b dot q 1, the dot product is the length of the projection of b on q 1 and then we subtract q 1 transpose b or b dot q along the direction q 1 from b. So, we get a vector B which is orthogonal to q 1 or 2 a because, p is decompose can be decomposed into 2 parts; one is along q 1 another is perpendicular to q 1.

So, when we take projection of b along q 1 and we subtract it from the main matrix B, the capital the main vector B, the new vector B because, orthogonal to a and then we take or orthogonal to q 1 q 1 is along a and then we find out it is length, divide the vector by it is lengths. We get another unit vector q 2. Now, we have the third vector C.

So, we keep on doing this can keep on doing this for any number of vectors. We project c both alone q 1 and along q 2 and subtract the projection from the sorry subtract the projections from the vector C and get a vector C this C is now orthogonal to both q 1 and q 2 and we divide this C by it is length and get a unit vector which is orthogonal to each other q 1 and q 2.

So, in this process we get 3 vectors q 1 q 2 q 3 this 3 vectors are orthogonal to each other and each of them have unit link. So, we call we tell that we have arrived into a orthonormal set of vectors.

(Refer Slide Time: 04:50)

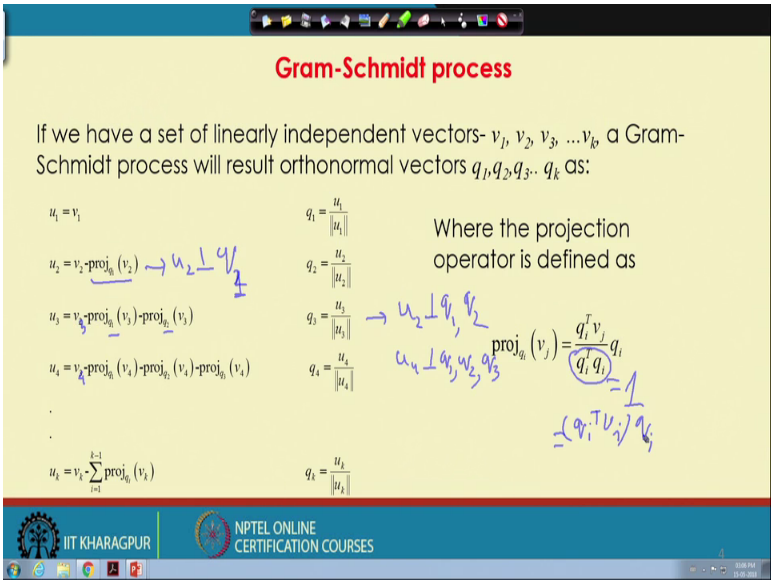


And now, for any number of vectors, this any number of independent vectors. This idea can be repeated. So, we subtract from one particular vector which is independent too few other vectors and these other vectors. We got an orthogonal set orthonormal set of vectors. So, we project these particular vectors to the previous set of orthonormal vectors and subtract the projection from these vectors. So, the remaining part is perpendicular to the already obtained set of orthonormal vectors and then we divide it by it is length and get a new vector which is orthonormal to the previous set of vectors.

So, Gram Schmidt starts with independent vectors a 1 a 2 a 3 and ends up with orthonormal vectors q 1 q 2 q n. In each steps, it subtracts from the vector a j. It is components along the direction q 1 to q j minus 1 which are the orthonormal vectors which have already been settled as like this and then from capital it divides A by it is length to get a unit vector around this direction. So, it essentially creates number of unit vectors which are mutually perpendicular to each other and what is the number of mutually perpendicular vectors, that will be exactly the number of mutually independent vectors with which we have started.

Because, perpendicular vectors are also independent vectors. So, we take a particular vector subspace which is spanned by the independent vectors whose set was given to us initially and then, we get a new basis for that set which are mutually orthogonal vector and each has length 1 which are q 1 to q 1.

(Refer Slide Time: 06:35)



Now, we can write down the algorithm. We will try to write down the algorithm like that. We have set we have set of linearly independent vectors v 1 to v n and then, we make that the first vectors we get another vector, not another vector u 1 which is equal v 1 for the first vector and unit vector along the direction u 1 is q 1 is equal u 1 by mod u. For the second vector, we subtract from v 2 the projection of v 2 along q 1. So, this becomes a vector perpendicular; that means what we get? U 2 is perpendicular to q 2 and we divide u 2 by it is modulus and get sorry u 2 is perpendicular to q 1 which is already settled.

We divide u 2 by it is modulus and get a unit vector q 2. Similarly, what we do? We subtract from v 2, it is projection along q 1 and it is projection along q 2. So, what we get? A new vector u 2 which is perpendicular to both q 1 and q 2 and we divide it by the modulus of the vector itself. So, we get an unit vector along this direction for the forth vector. We do the same thing from u 2, sorry this is v 3. This will be v 3 from v 4, this will be v 4. From v 4, we subtract the projection of v 4 along q 1 the projection of v 4 along q 2 and the projection of b 4 along q 3. So, we get a new vector u 4 which is perpendicular to q 1 q 2 and q 3.

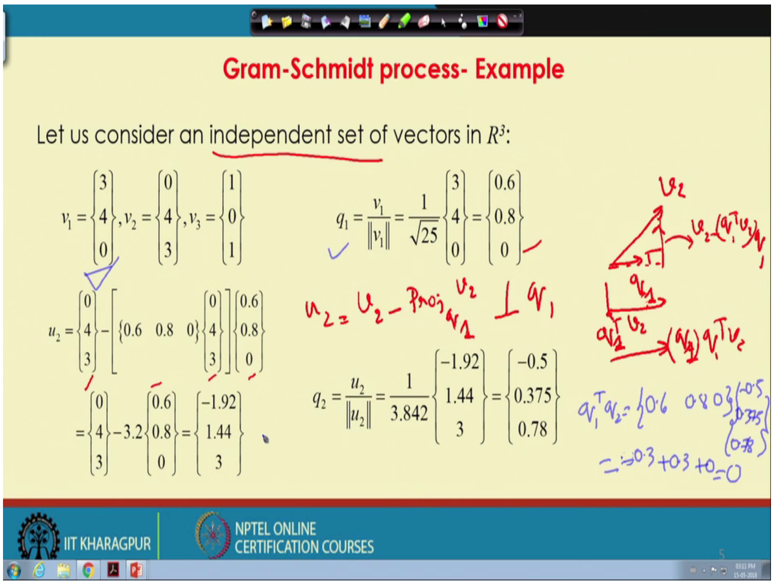
And we find a unit vector like that. So, that way, we can go up to k th vectors. Any number of independent vectors in a particular subspace, only the number of independent vectors cannot be more than the dimension of the subspace or if you think of real coordinates R n. We cannot have more than n number of vectors there and each vector is from each vector. We subtract it is projection from the previously settled orthogonal vectors.

So, what remains is that, the component of the victor which is perpendicular to the set of orthogonal vectors. We have obtained already like for step 4. What remains as u 4 when we subtract from v 4? It is projection along q 1 q 2 q 3 what remains in u 4 is a vector which is perpendicular to both q 1 q 2 and q 3 q 1 q 2 and q 3.

So, in that way, we have to take a vector, we have to project it along few orthogonal vectors were already settled in and we have to subtract the projections from that particular vector. What will remain with us is a vector which is perpendicular to the previous or set of orthogonal vectors and will find the unit vector along that. So, when the projector the projector operator is defined as q i transpose v j by q i transpose q i as q is an unit vector, this sorry as q is an unit vector this length. This is equal to 1 as q is an unit vector.

So, projection operator here will basically give as q i transpose v j into q i. This is a dot product. So, this is a scalar quantity. It is a scalar into q i. So, some length along the direction q i. Q i is the unit length. So, this is the amount the magnitude of the projection along the direction of q i, ok. So, this we have discussed in detail in last class.

(Refer Slide Time: 11:00)



So, we now take an example we have vectors 3 4 0 0 4 3 and 1 0 1. They are independent vectors.

We can quickly check and now we have to. So, this is an independent set of vectors. We can only work with an independent set of vector. If we are working on Gram Schmidt, now we have to get a set of mutually orthogonal vectors. So, the first step is very easy. You take the first vector v 1. The first vector v 1 and get q 1 out of it, just divide v 1 by it is length. So, 3 4 0 the length is 3 plus 4 square and root over of that which is 9 plus 16s root which is 5 and the unit vector becomes 0.6 0.80 and we can quickly check the length is 0.6 square plus 0.8 square is 36 plus 0.36 plus 0.36 plus 0.64 and if I take a square root of that, this is 1, fine.

Now, the second vector is obtained as the first vector; the second vector for second orthonormal vector, what we do? We take the second vector, also take a dot product of the second vector with q 1 and multiply with a q 1 the dot product along q 1. So, this will give as the projected vector of v 2 along q 1 and subtract it from v 2.

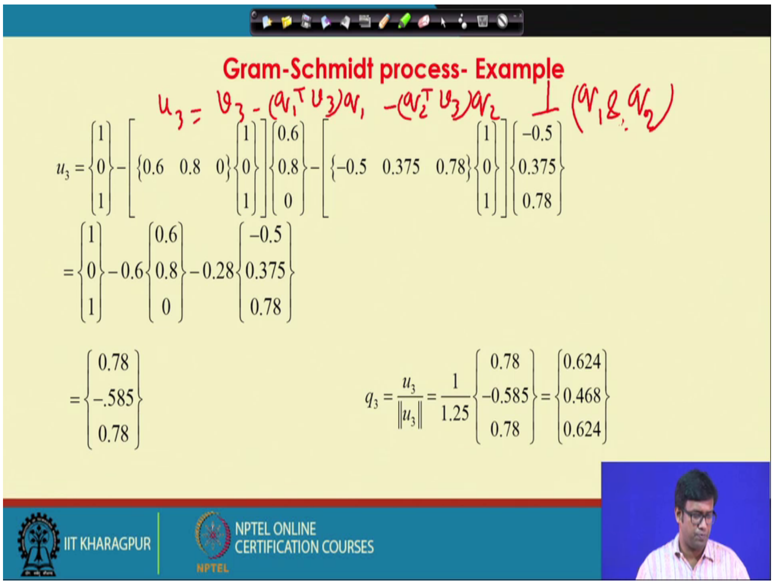
So, this will be what we can write u 2 is equal to v 2 minus projection of v 2 along q 2 and what is projection of v 2 along q 2 if this is v 2 and this is q 2? So, this length, this length is v 2 q 2 dot v 2 and the particular victor is q, sorry this is q 1, this is not q 2. I am sorry, 1 second.

This is q 1, this is q 1. So, u 2 is obtained as v 2 minus projection of v 2 along q 1. So, this is v 2, this is q 1 this is projection of v 2 along q 1. It is length is q 1 transpose v 2. The vector is q 1 transpose q 1 along q 1 is an unit vector. So, along this direction, we have q 1 transpose v 2 and the sum the difference v 2 minus this is v 2 minus q 1 transpose v 2 along q 1 this vector is perpendicular to q 1. So, from v 2, we subtract from v 2 we subtract the first part q 1 transpose v 2 q 1 and we get v 2 minus q 1 transpose v 2 q 1 which is perpendicular to v 1 or q 1. So, this, we get a vector which is minus 1.92 1.443.

And this vector is perpendicular to. So, I can say that this is perpendicular to q 1. So, my q 2 q 2 must be perpendicular to q 1. So, q 2 will be this vector y 2 divided by it is modulus which gives us minus 0.5 0.375 0.78 and we can try to do q 1 dot q 2. Here, a sorry we can try to find verify it q 1 transpose q 2 is equal to 0.6 0 0.80 into minus 0.5 0.375 0.78 which is 0.5 minus 0.5 into 0.6 is 0 sorry is minus 0.3 minus 0.3.

0.8 into 0.375 is plus 0.3 and this is 0 into 0.780. So, this is 0. So, q 1 and q 2 are orthogonal. That is also verified. So now, we have to find out the third vector for the Gram Schmidt process because, there are 3 vectors. So, we will also get 3 mutually orthogonal vectors; 3 orthonormal vectors when is do the Gram Schmidt process.

(Refer Slide Time: 16:06)

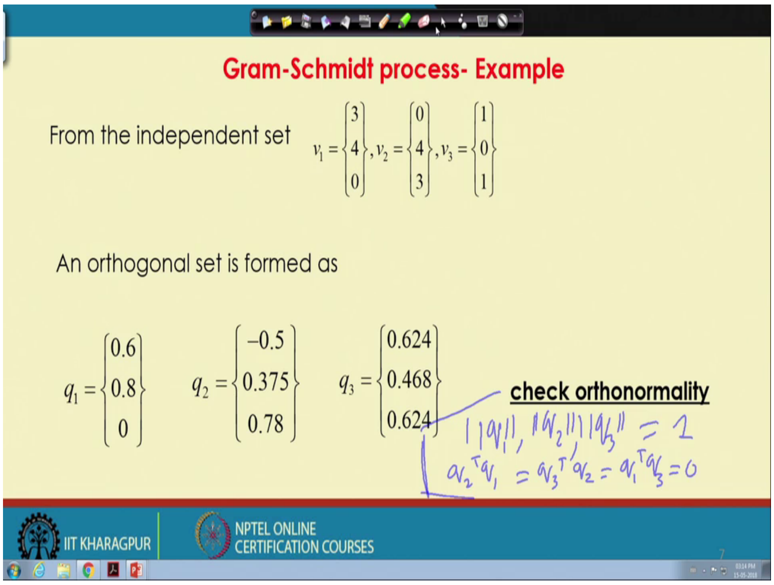


So, if we look into the third vector, that is u from u 3, we from first vector, we subtract it is projection along q 1 and it is projection along q 2. So, which is now basically u 3 is equal to v 3 minus q 1 transpose v 3, q 1 minus q 2 transpose v 3. So, this is v 3 minus 0.6 0.8 0 q 1 transpose 1 0 1 along q 1 minus 0.0 0.5 0.375 0.78 1 0 1 with minus 0.5 0.375 0.78.

And this gives us 1 0 1 minus this length is the projection magnitude is 0.6 the projection magnitude is 0.28 here. So, minus 0.6 into q 1 minus 0.25 into q 2 and we get minus 0.78 0.585 0.78. So, this u 3 is now perpendicular to both q 1 and q 2 this is perpendicular to both the vectors q 1 and q 2. So, we get a third vector u 3 which is perpendicular to both q 1 and q 2.

However, we do not know whether q 3 is also with is unit vector or not. But, we can very easily verify that, we will find it is length and this the length is obviously, not 1. We divide u 3 by the length and we will get q 3. So, q 3 0 3 by the length of u 3 which is 0.624 0.468 0.624.

(Refer Slide Time: 18:12)

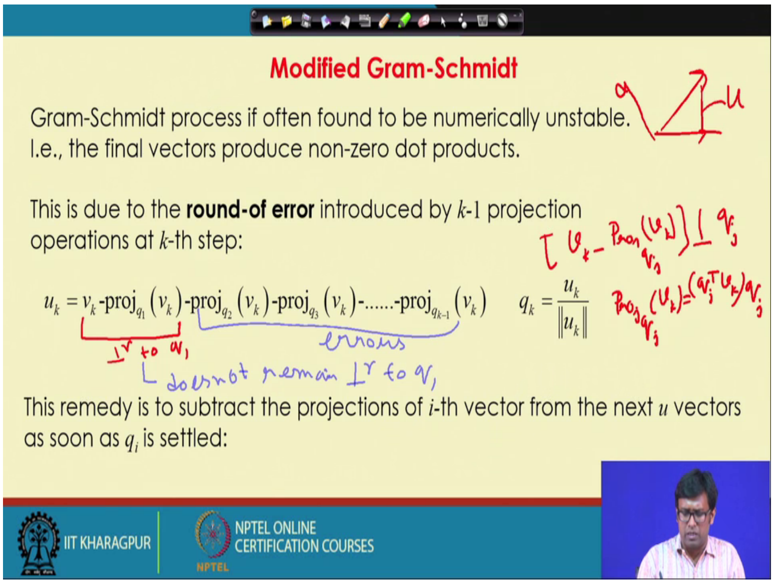


So, from the independent set of vectors v 1 v 2 v 3, we got an orthonormal set which is q 1 q 2 q 3 and we can check for the orthonormality in a sense we should check that what this the l 2 norm for both this case q 1 q 2 q 3. This norm is equal to 0.36 0.8 0.641 root over 1 is 1.25 0.78 square plus 0.37 square root. This is also 1 and this is all these are 1 and q 2 transpose q 1 is equal to q 3 transpose q 2 is equal to q 1 transpose q 3, all these are 0.

So, this is what finally, can be checked and this satisfies that the set of vectors we obtained are orthonormal set. So, this process is essentially very simple. You only have to follow certain steps. The step is that first start with take the first vector, divide it by it is length, get the first vector in the orthonormal set.

Now, for any subsequent vector in that orthonormal set, you take the linearly independent vector which is remaining with you. Now, project it with the already found out orthonormal vectors and from the vector subtract the projections. So, what we will get is the new vector which is perpendicular to the already found orthonormal set of vectors. And divide this vector by it is length and you will find an unit vector along this. However, this particular process when we try to implement it in a computer program gives some issues.

(Refer Slide Time: 20:12)



Gram Schmidt process is often found to be numerically unstable, that is when we try to check for orthonormality, the final vectors when we check that dot products, the dot products are often non 0. However, the formulation is sounds. So, we should not get a on 0 dot product. When we get q 2, any q q i q q i plus 1 etcetera because every q we are finding from a vector which is perpendicular to previously settled vectors. However, what you see that there is a round off error introduced by k minus 1 projection operation at k th step.

So, what is the k th step? The vector perpendicular to already settled orthonormal vectors q 1 q k minus 1 is obtained as v k minus it is projection from v k minus 2 q 1 minus it is projection on q 2 minus 1 it is projection on q 3 etcetera. So, it is supposed that v k. Vk minus projection of v k along q j projection of v k along q j this vector should be.

So, from v k, we are subtracting the component of v k which is along q j. So, what will remain is the component which is perpendicular to q j from vector v k. From this vector, we are subtracting it is component which is along the slide. So, what will remain is a vector is this vector which is perpendicular to this line on which I am projecting it.

So, this should be perpendicular to q j. So, if I write v k minus projection of v k q q along q 1, this should be perpendicular to this particular part should be perpendicular to q 1. Now, from that perpendicular part, we are subtracting few other components which are perpendicular to q 2 q 3 q 4 etcetera. However, in each subtraction or each projection what is the projection operation? Typical projection operation looks like projection of q j on v k variation of v k on q j is equal to q j transpose v k along q j. So, this is dot product between 2 vectors.

So, it needs several multiplications and then also, when we are finding out q 3 q q 2 q 3 q k plus what we have dividing it by the modulus of this vector. So, all these calculations are introducing a round of errors. Because, all these are dealing with real numbers and if I have some number which is 3 by 7, that has to be truncated after 8 decimal or 12 decimal or 16 decimal place in a computer program. So, we cannot write a division up to infinite digit.

Similarly, will get irrational numbers also like here. Also, because we are finding out root over of certain values and finding out length and this root of this, irrational number will also be truncated after certain values. So, as the round of error set being introduced and each step. We are introducing some of the errors. So, due to this error, this term does not remain. If there is no error from a perpendicular vector, we are subtracting something. It will still from a vector which is perpendicular to this, we are subtracting something it root of still remain perpendicular to this.

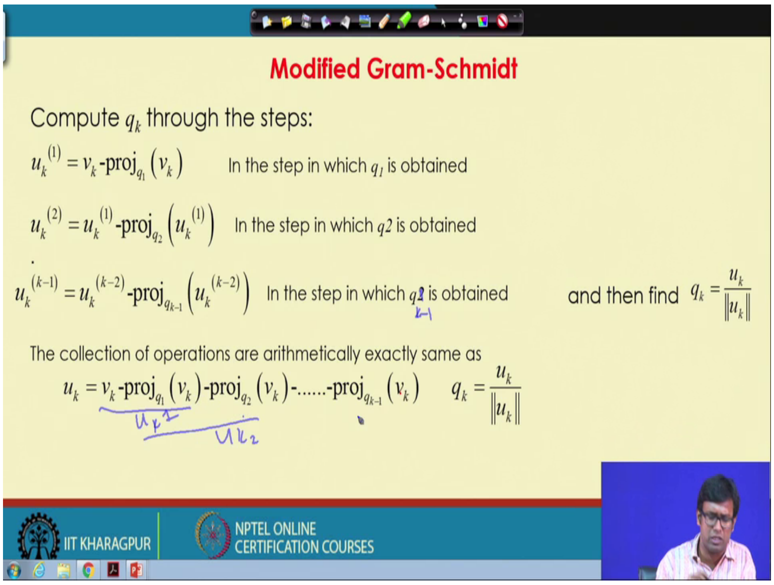
But, what we have subtracting now is introducing some error. So, the this vector is being reduced along the this reduced random. So, it can reduce at any direction.

So, it might know from perpendicular it might not remain perpendicular to q 1. So, and it happens that this does not remain perpendicular to q 1. So, so what we get is a set of ortho is a set of vectors which are not perpendicular to each other or as as as the product, as a final result, we get a sectors ; set of vectors, we cannot producing 0 dot product.

So, what will be the remedy? The remedy is to subtract the projections of i th vector from the next u vector as soon as q i is settled and will check it quickly. So, once we find out this particular like this particular vector, now instead of projecting v k on q 2, sorry instead of projecting v k on q 2, will project this subtraction on q 2 and subtract it.

So, in a sense, if I have a vector like this and I project it on this q and got that this is my u. Now, instead of again projecting v on to a third vector which we are doing, here will project u on the third vector. Because, this is already an ortho perpendicular vector only projective and there that will reduce the issues due to round off error. So, the solution will be subtract the projection of i th vector from the next u vector as soon as q i is settled.

(Refer Slide Time: 26:50)



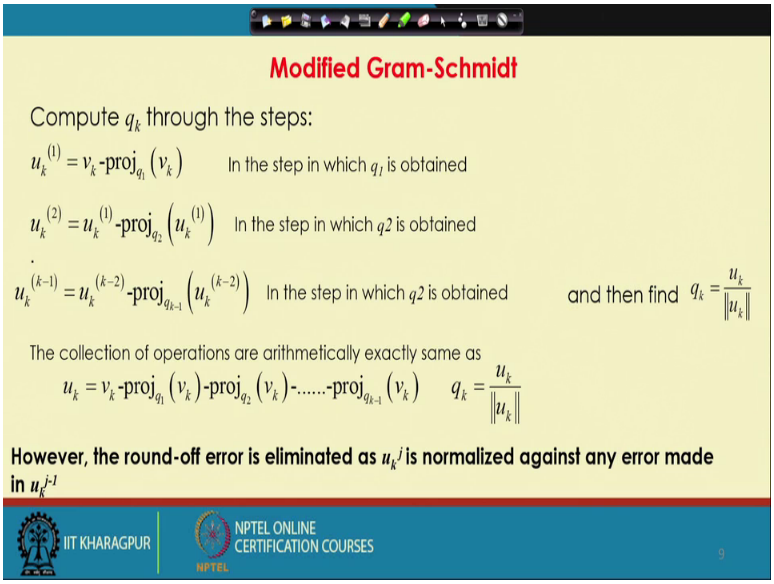
And the sorry the steps will be first find u k 1 v k minus q 1 v k. Once we have found q 1, subtract it is projections. Project all the vectors, all the remaining vectors and q 1 and subtract it from them.

And for k th vector, you got u k 1 which is v k, the original k th vector from which the it is projection on q 1 is subtracted. Now, when you found q 2, you subtract the projection of u k 1 on q 2 from u k 1 and get u k 2 and in the step when in which q 1 is obtained, you do it for k minus 1th step.

And then finally, in this step where q sorry k minus 1 is obtained and then finally, find q k is equal to u k by mod u k. The collection of operations are arithmetically same as what we are doing in Gram Schmidt method only. Instead of projecting, we say for example, for finding out u k, you project v k on q 1 q 2 up took k minus 1 instead you project v k on q 1 and get this as u k 1.

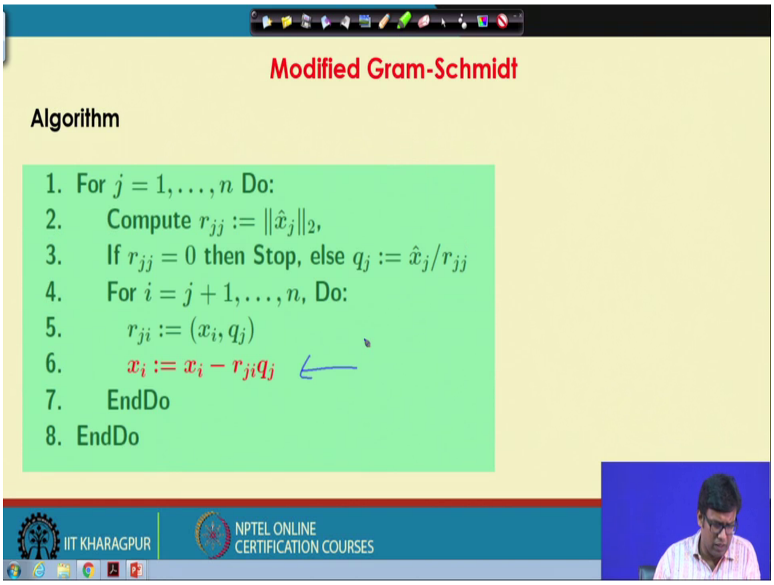
Now, project u k 1 on q 2 and subtract this, then get u k 2 and then project u k 2 on q 3 and subtract it and get u k 3 and so on.

(Refer Slide Time: 28:42)



And this essentially gives you a set of vectors without round of error because, round of error always eliminated when you normalize, when you orthogonalize, normalize when you orthogonalize u k with the with another vector. So, you have a vector which has some round off error. You do a projection with another vector and subtract that part.

(Refer Slide Time: 29:10)



So, around off error is in a way also subtracted. It has been shown that this is the modified Gram Schmidt algorithm where basically you subtract.

So, you start with start with i is equal to j is equal to 1, find the length of the vector and find the unit vector. Then, you take that unit vector and take dot product of the unit vector with all other vector in that set and subtract it from all the vectors. So, in this state and then you go for 2 and do this thing. So, subtract the projection of on j th vector and in the next step, you project the remind remaining part on j plus 1 th vector.

So, essentially, this is a much stable algorithm and in next class, will see that if we compare gauss Gram Schmidt with modified Gram Schmidt in certain cases where there is some component, where round off error can be vital or there is some calculation in which round off error can be vital. For example, I am subtracting from 1 1.001. If there is no round off, it will be 0.001.

But, if there is a round off error after second decimal place, this will be 0 and division with them will give me a infinitely large number. So, in case is where round off error is important, we will see that ground Schmidt modified ground Schmidt gives much better result than ground Schmidt. And actually, ground Schmidt algorithm fails you fails to provide you a orthogonal set of vectors. So, in next class, we will see an example with modified Gram Schmidt and ground Schmidt algorithm and then we will see a q, what is called a qth decomposition or how a matrix can be decomposed into q matrix and some other matrix which is an R matrix and that can be used for matrix solution of matrix equations.

Thank you.