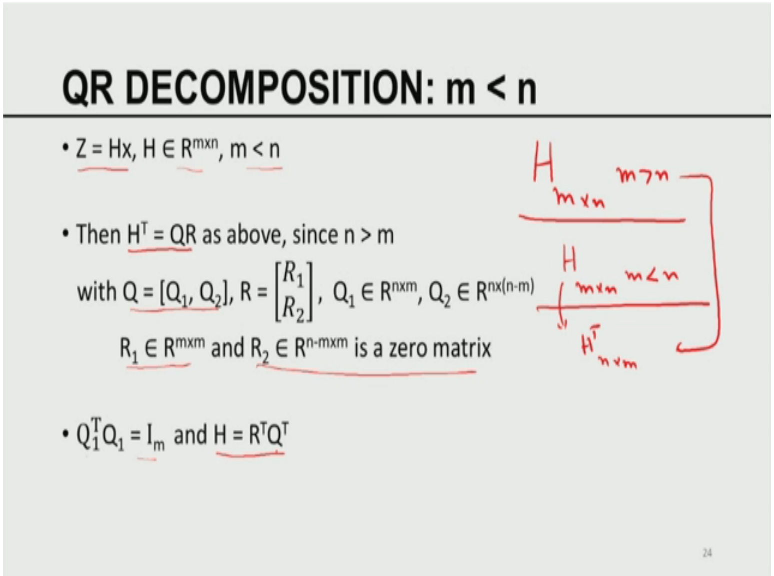
Dynamic Data Assimilation
Prof. S Lakshmivarahan
School of Computer Science
Indian Institute of Technology, Madras

Lecture - 17
Matrix Decomposition Algorithms Continued

(Refer Slide Time: 00:13)

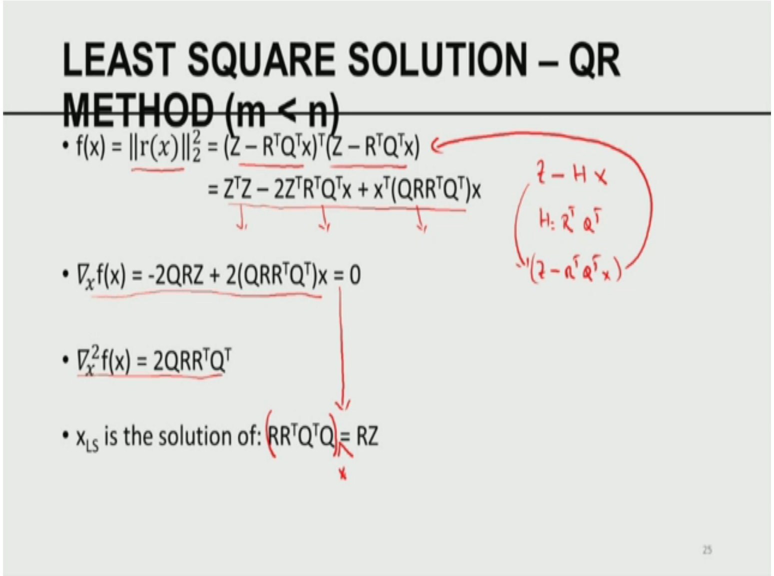


In the previous talk we talked about QR decomposition for the over determined case. QR decomposition can also be applied for the under determine case. I am going to quickly point out some of the key steps. Let again Z is equal to H of x, H is m by m, m is less than n, m is the number of observation, n is the number of unknowns. I have more unknowns than the number of observations. In this case, please remember I have done the hue QR decomposition of H m by n, when m is greater than n. We have already done with that. In this case I have H m by n, here m is less than n.

If I took the transpose of this H; that is n by m n is greater than m. So, these two cases are the same if you interchange m and R. Therefore, QR decomposition of H for the over determined system, is the same as the QR decomposition of H transpose for the underdetermined system. That is a quick mathematical enterprise that you can utilize. If you utilize this you can see the QR decomposition for the lower triangular, for hundred determine case on the over determine case, are not too different from each other.

So, I can compute H transpose QR as above. Since in this case, n is greater than m by interchanging the role of m and n, I can apply everything that we saw thus far, Q again is Q 1 and Q 2 R is again R 1 and R 1 R 2. R 1 is an upper triangular matrix. R 2 is a 0 matrix, Q 1 transpose Q 1 is I m, and H is equal to R transpose Q transpose. So, essentially the whole thing flies without any much trouble at all.

(Refer Slide Time: 02:22)



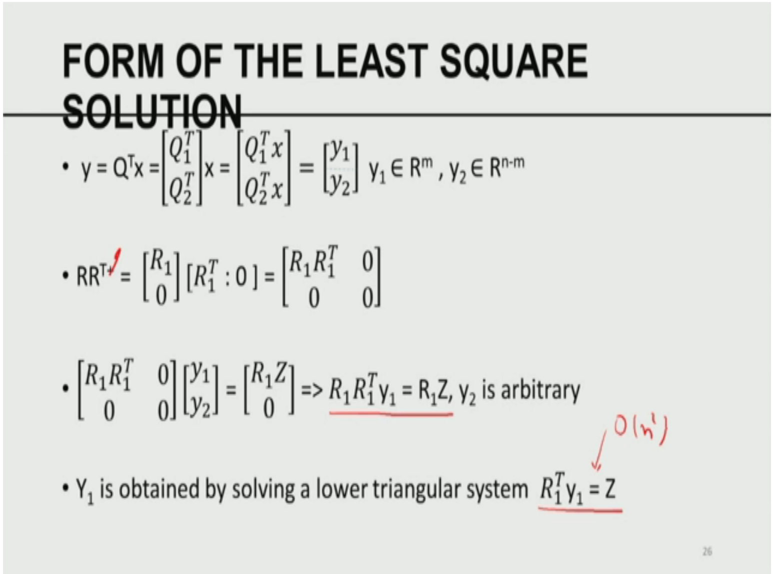
So, in this case again, my f of x is equal to square of the normal residual. So, Z is equal to. So, let us go back to the previous case. The problem reduces to Z is equal to. I am sorry H is equal to R transpose Q transpose, in view of that H Z minus H of x. I am sorry, in view of that Z minus H of x. I am sorry once again; let me correct myself Z minus H of x.

Z minus H of x, H is equal to R transpose Q transpose. Therefore, Z minus H of x now becomes R transpose Q transpose x, and this is what is used in here. So, f of x is the square of the sum of the errors represented this way. So, this is an alternate expression for the residual, which can be, when multiplied can be given by this. You can compute the gradient of this. This is independent. The first term is the independent of x. this is linear in x.

This is quadratic in x. From the module on multivariate calculus, we know how to compute the gradient of a linear function; we know how to compute the gradient of a quadratic function. By applying those rules, it can be verified that the gradient is given by this expression, at the minimum gradient must be 0. I can also simultaneously compute the hessian. So, the solution for the least square problem is obtained by solving the gradient to be equal to 0, and that gives raise to this equation. So, R R transpose Q transpose x that must be. I am sorry.

X must be equal to R Z therefore, if I change if I can multiply both sides by R inverse.

(Refer Slide Time: 04:39)

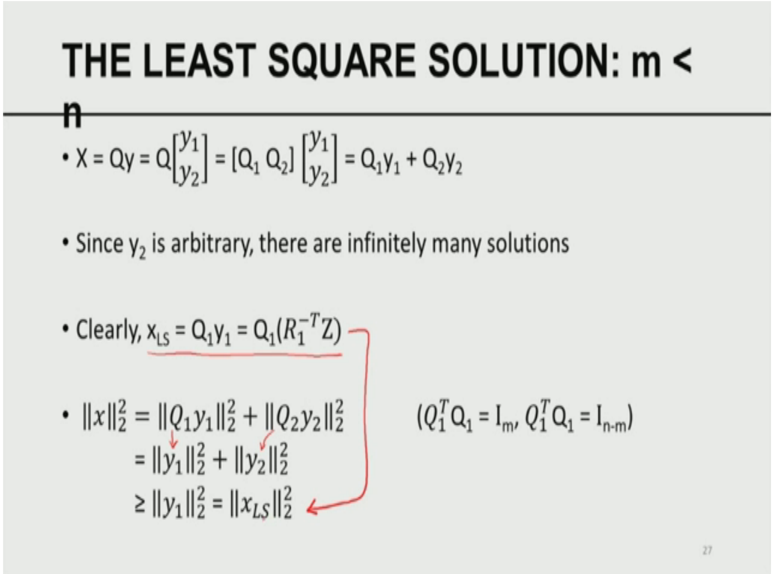


If I multiply R inverse, I get the overall solution. So, to that extent, I am now going to express various operations, and here y is equal to Q transpose x, which is given by this partition form. R plus R transpose is essentially given by. So, there is no plus here. So, sorry there is no plus here. This is given by this matrix. There is R 1 R 1 transpose 0 0. Therefore, if I assemble all these things I get this equation, which is the equation I need to solve R 1 R 1 is a non singular matrix, if I multiply both sides by R 1 inverse, I simply obtain the least square solution by solving R 1 transpose, y 1 is equal to Z. So, this is again R 1 is upper triangular system R 1 transpose a lower triangular system. This can be solved in o of n square.

So, the whole theme of the exercises, whether it is over determined undetermined by invoking the QR decomposition of the rectangular matrix H in either case, one can reduce the solution least square problem, to one of solving a lower triangular system are an upper triangular system which is much easier to solve. This is the motivation for using the QR decomposition, because the pathway leads to a very simple problem in the end.

And we already know solution of upper triangular system lower triangular systems are very simple.

(Refer Slide Time: 06:20)



This is another look at the least square solution. The least square solution is now given by x is equal to Q of y in the case of underdetermined system. y is y 1 y 2. Q is Q 1 Q 2. Therefore, x is equal to Q 1 y 1 plus Q 2 y 2.

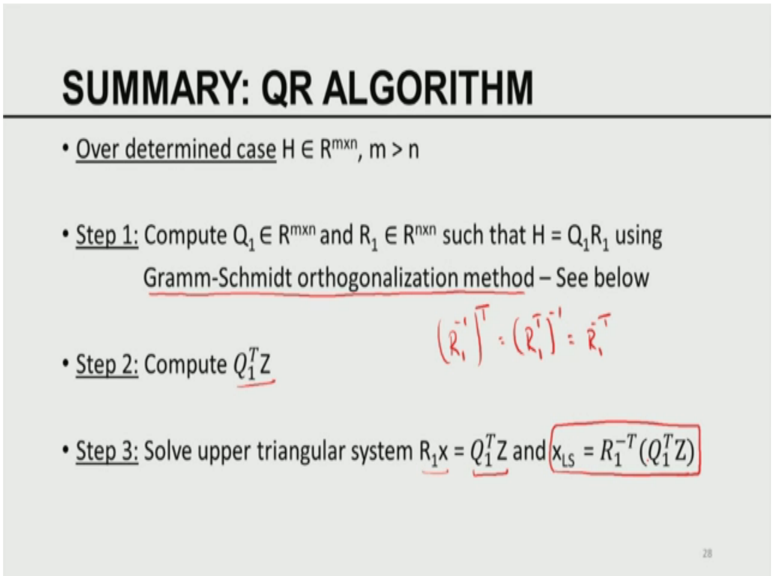
Now, please remember y 2 is arbitrary. So, there are infinitely many solutions, and that is consistent with the undetermined system. In the case of there are infinitely many systems. Infinitely many solutions for the equations whether, because there are more unknowns than the known's from the previous analysis, we get the linear solution of the linear system, arising from the least square problem, is given by this; x therefore, the norm of X L S.

So, this must be the norm of X L S, the same X L S in here that is equal to the norm of Q 1 y 1 square plus Q 2 y 2. Q 1 transpose Q 1 is I m Q 1 transpose Q 1 is I n minus m, I m is the mth order unit matrix, I n minus m is the n minus mth order unit matrix. Q 1 is an orthogonal transformation, ortho are the length is invariant orthogonal transformation. Therefore, this term reduces to this term, because of orthogonality. Again this term reduces to this by orthogonality.

So, y 2 is arbitrary; therefore, the least square solution is larger than, or equal to y ones at the. The norm of the square of y 1, we are going to any solution x is greater than the least square solution; therefore, the least square solution is given by the formula that we had already given which is here.

So, this corresponds to this therefore, any arbitrary there are infinitely many solution, any one of the infinite many solution has a norm, that norm is greater than least square solution. Therefore, the least square solution is the unique solution of minimum norm. So, that is the end gain the end result, and this is result of applying the QR decomposition to the underdetermined system as well.

(Refer Slide Time: 09:13)



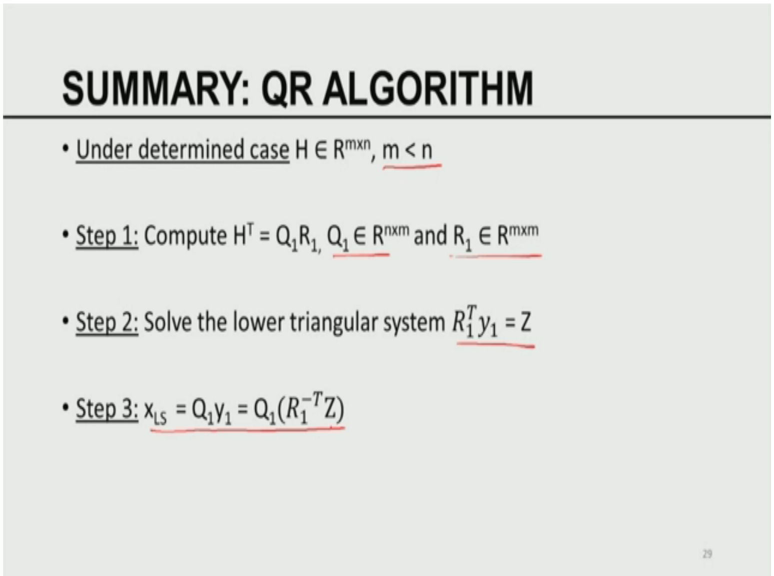
So, summary of the QR algorithm over determine system H is, m is greater than n compute Q 1 compute R 1; such that H is equal to Q 1 R 1 using a method called Gramm Schmidt Orthogonalization. So, that is where we are now leading to, that is the reason for the summary. I only said that such a thing can happen, but we do not know how to make it happen. So, if I can express H as Q times R Qs orthogonal R is upper triangular, we saw all the beauties in the analysis, that essentially assumed I can do now it is time for us to be able to tell how to make the decomposition happen, that decomposition H is equal to Q 1 R 1; that is the reduced QR decomposition is often done by a very famous procedure called Gramm Schmidt Orthogonalization method. We will talk about it shortly.

So, once you have Q 1 R 1, I can compute Q 1 transpose Z. Once you have Q 1 transpose Z that defines the right hand side. I need to solve the lower triangular system R 1 x is equal to Q 1 transpose Z, that gives us to the linear least square solutions R 1 minus T. Minus T means what. This is R 1 inverse transpose. It is denoted by, is also equal to R 1 transpose inverse.

So, these two operations commute; because they commute for simplicity, in mathematics we simply call it minus t. So, if I say minus T, is a combination of 2 operations the order in which you apply these operations is a material these two operators commute. So, the least square solution is given by your solution of a lower triangular system, which is given elegantly by this.

So, the whole method is beautiful and very nice. So, we need to concentrate on being able to express H as Q 1 times R 1, the reduced decomposition, and that is done by Gramm Schmidt Orthogonalization. So, that is our next term, given H How do we convert H into product of Q and R. So, we in our method, what is that we assume, let us pretend we can do the decomposition H is a QR Q and R, and how if I can do that, the solution algae, the solution, the structural solution becomes. After having achieved that it behooves us to ask a question, how do I make it happen, how do I make H x process the product of Q and R.

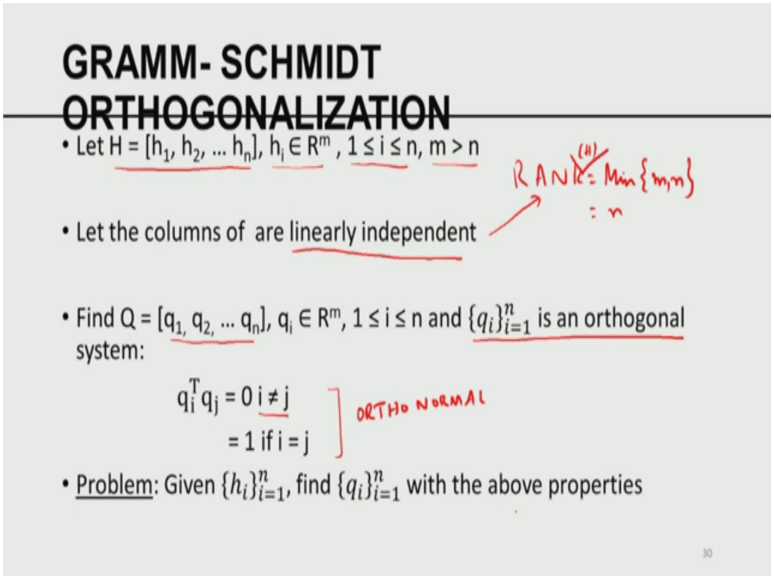
(Refer Slide Time: 11:56)



So, this is the next part of the exercise. I also want to touch upon the under determine case, when H is, when m is less than n, it is an undetermined case. Again it is a complete summary of all the analysis we have done H transpose. So, whatever we did for H, now I am going to do for H transpose. Operation and H and operation of H transpose are not too different from each other, they are mathematically equivalent. Again H transpose can be expressed with the product of Q 1 and R 1.

Q 1 R 1 are given by these respective sizes. Solve the lower triangular system, and least square solution is given by Q 1 times R 1 inverse transpose time Z. So, with this we have completed a discussion of the QR transpose algorithm, QR decomposition algorithm modulo the method for decomposition; QR decomposition.

(Refer Slide Time: 12:56)



And that is given by the method called Gramm Schmidt Orthogonalization. So, now, I am going to conjure up case and illustrate how given a matrix H, and I can deliver the Q factor and the R factor. To that n, let H be a matrix, is the m by n matrix, each of the H is are a vector in R m. So, there are n vectors each of size n, each of size m. I runs from 1 to n, m is greater than n. So, there are n columns, each column have m rows.

Let the also. I am going to assume the columns are H are linearly independent. So, what does imply, H is a full rank. So, rank of H is equal to minimum of m n. In this case m is greater than n. So, minimum of m and n is n, if the rank of. I am sorry rank of H that is what I should have said. Rank of H is the equal to minimum of m and n is equal to m.

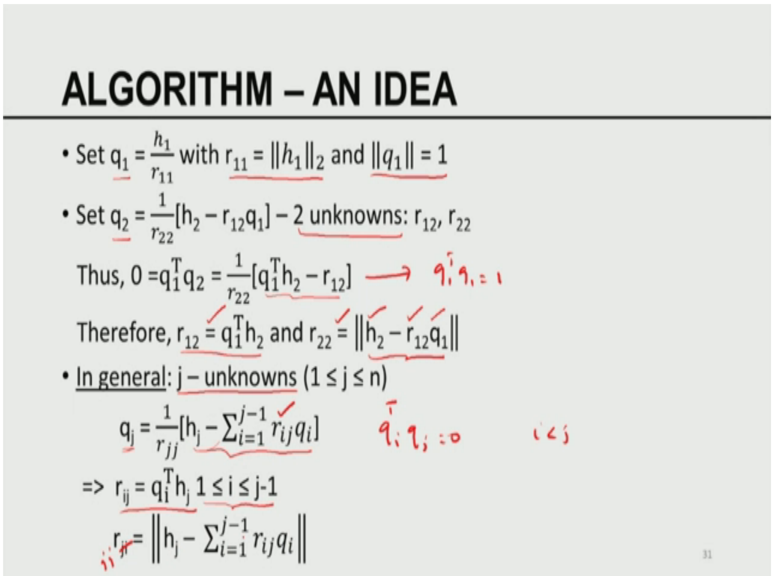
So, it is a full rank matrix, and that is guaranteed by the columns of H being linearly independent. Now you can see the basic concepts from vector space theory comes into play, right through. So, what is our aim? Our aim is to be able to find a Q matrix with n columns, which are q 1 q 2 q n.

I would like the columns of Q q1 and q 2 to be ortho orthogonal system. In fact, it has been ortho normal system, that is given by the following if I take any 2 vectors with distinct entities i and j not equal q i transpose q j is 0. If i is equal to j q i transpose q i is 1. So, this is orthonormal. So, what is the problem? Given the columns of H, I need to find the columns of Q; such that the columns of H are linearly independent, but the columns of QR orthonormal.

This is simply a procedure for converting a set of linearly independent columns, to a set of orthonormal columns, is a very simple procedure, and this procedure is very fundamental, is used repeatedly in several different applications in linear algebra. We have already talked about the applications, in solving linear least square problems. Now we are simply going to concentrate on how to do the decomposition itself.

So, let me go back and say this. So, given h 1 h 2 h n find q 1 q 2 q n, given that H are linearly independent. I need to find a set of vectors Q that are orthonormal .

(Refer Slide Time: 16:20)



So, what is my first vector q 1? First vector q 1 is arbitrary I can start at any place. So, what I am going to do. I am going to start with q 1 is equal to my h 1 divided by the norm of h, the divided by the norm of h. Norm of h refers to the length of the vector h 1. So, r 1 1 is the first element of R matrix, it refers to the norm of the vector h 1, is the two norm of that. So, if I divide the vector by its norm I get a unit vector. So, q 1 is a unit vector. So, first vector q 1 is very simple, it is simply normalized version of h 1. Now I need to compute q 2. I am going to compute q 2 as a linear combination of h 1 and h 2.

q 1 has already information about h 1. So, I am going to now express q 2 as 1 over r 2 2 times h 2 minus r 1 2 times q 1. Please remember q 1 is known, h 2 is known, r 1 r 1 2 is not known r 2 2 is not known. So, there are 2 unknowns, r 1 2 and r 2 2. How do I find these two unknowns? These two unknowns are found by imposing two condition. What is the first condition, q 1 q 2 must be orthogonal to q 1 that is give rise to value of one of the unknowns. Then the second one is obtained by forcing q 2 2 be of unit length. So, orthogonality condition, and the normality condition two conditions are enforced by fixing two parameters.

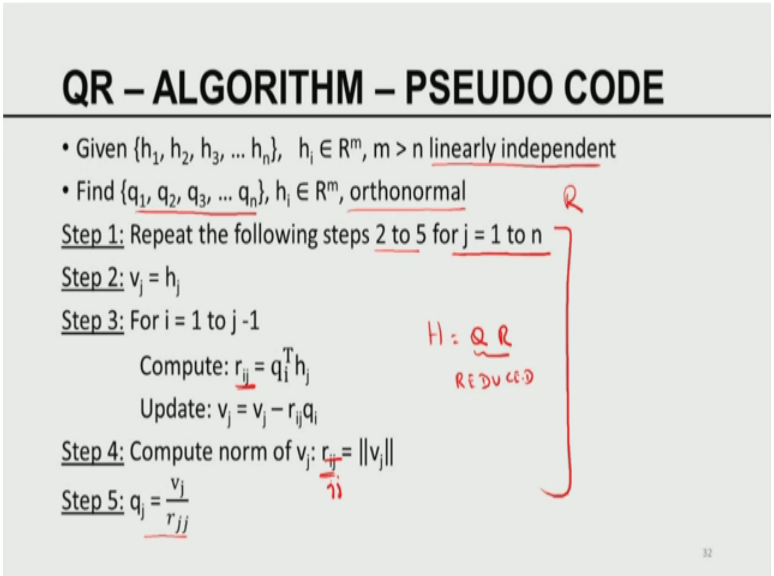
So, if I were to require q 1 to be orthogonal to give to, it requires 0 is equal to q 1 transpose q 2 the assumed to form of q 2 is above I substitute this in here. Therefore, why this is r 1 2, I would have q 1 transpose q 1 is 1. Please remember that I have already utilized in here q 1 transpose q 1 is equal to 1. Therefore, when I equate this to 0 r 1 2 is simply q 1 transpose h 2. So, I got 1 1 1 constant. What is r 2 2. Now I know r 1 2 I know q 1 I know h 2 that is the vector, I can compute the norm of the vector, r 2 2 essentially divides this vector by its length. So, r 2 2 must be the length of that vector which is h 2 minus r 1 to q 1.

So, this gives raise to a formula for computing one of the core unknowns. This gives you a formula for computing the other unknown. So, from the known I am computing the unknown. So, here look at this now I have recovered q 1, I have recovered q 2 I have recovered r 1 1, I have recovered r 1 2 I have recovered r 2 2. So, I am recovering both the system simultaneously q and r. So, let us go back to a general, go to a general case when there are j unknowns. In other words I can now express q j. So, let us pretend I have found out q 1 q 2 q 3 all the way up to q j minus 1; I am going to compute q j. q j is going to be expressed as a linear combination of all the previous vectors.

h j q i has information about h i; therefore, I can express q j by the linear combination that is given here divided by a constant r j from the previous calculations you can readily see r j j is going to be the normalizing constant, r i j are the constants use in which I am going to force the orthogonality. So, q j has to be orthogonal to q j 1 j 2, I am a q 1 q 2 q 3 q j minus 1. So, there r j minus 1 constant that are required to fold the orthogonality one conscience is required to force the normality, and that is the that is a clear story. So, by multiplying q j, by multiplying q i transpose q j for i less than j, and equating to 0, I compute all the values of r i j to be q i transpose h j.

So, I have computed j minus 1 such constant, then r j j. So, this must be r j j, is computed by the norm of the entire vector q j j, is therefore, r j j is equal to h j minus, thus the norm of the vector.

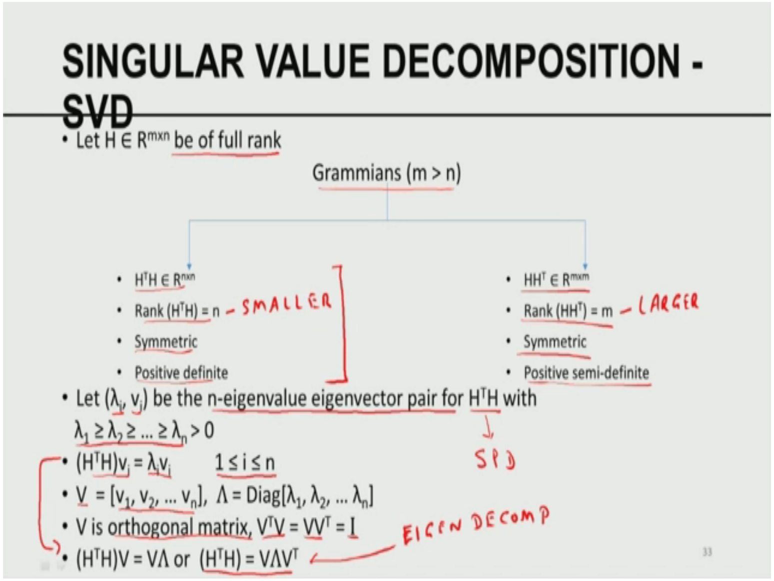
(Refer Slide Time: 21:15)



So, I can now embed this in a in a pseudo code. Given h that are linearly independent, that are linearly independent. I want to find this are orthonormal, repeat the following steps 2 to 5 for j running from 1 to n. Step 2 v j is h j, for i is equal to 1. Now this must be v 1 must be equal to ah h 1 that is correct, sorry 1 second, what sorry once again, yeah that is correct.

v j is right, because j is running from 1 to n. So, the overall do loop is on j that is right, that is correct the overall do loop is on j. So, step 2 v j is equal to h j; that means, j is equal to 1. So, initially j is equal to 1 v 1 is equal to h 1. Step 3 i for i is equal to 1 to j minus 1, I am going to make v j perpendicular to the previous vectors. So, that is how I compute this, and now compute the norm of v j, and then q j is given by this. So, this procedure is repeated, this procedure delivers q. It also delivers the matrix r as r 1 1 r j i r j j, compute the norm this must be I am sorry r j j, or this is r j j. Therefore, in every case for every j i am trying to find all the i running from 1 to j minus 1 in here r j j here.

(Refer Slide Time: 23:10)



So, together I compute all the required quantities for every j, and that completes the procedure for computing the QR decomposition, with this algorithm, with this pseudo code, with this pseudo code sorry.

With this pseudo code, with this pseudo code we are able to express any H to be Q and R R is upper triangular Q isorthonormal, and this is the reduced to QR decomposition, the reduced the QR decomposition. So, here we saw how to multiply, how to decompose H into a multiplicative decomposition QR. I have already utilized if H is equal to QR, how to utilize it in my solution process for the linear least square problems. These two together help you to solve the linear least square problem by the QR decomposition. So, that is method 2.

Now, I am going to go for the last of the three matrix decomposition methods, which is called the singular value decomposition. Singular value decomposition is economist S V D. This is the third alternative method for solving linear systems, I am going to assume my matrix H is again a full rank matrix. Given a rectangular matrix there are two grammians H transpose H H H transpose. Both the grammians are of rank n are rank m there are full rank, both of them are symmetric, both of them are positive definite. In this case it is positive semi definite; if it is positive definite means all the Eigen values are positive. If it is positive semi definite means the Eigen values are 0 or positive that is the only difference between the two. And when m is greater than n when m is greater than n this is the smaller grammian.

Because n is smaller. This is the larger grammian. When the rank of H transpose H is n, it has n non 0 Eigen values, when the rank of H transpose H, the rank of H itself is n. Therefore, the rank H H transpose is deficient, if this rank deficient some of the Eigen values in this case are allowed to be 0; that is why it is symmetric and positive semi definite. That essentially comes from the rank conditions for product of matrices and the given rank of and the given rank of H, at the given rank of h.

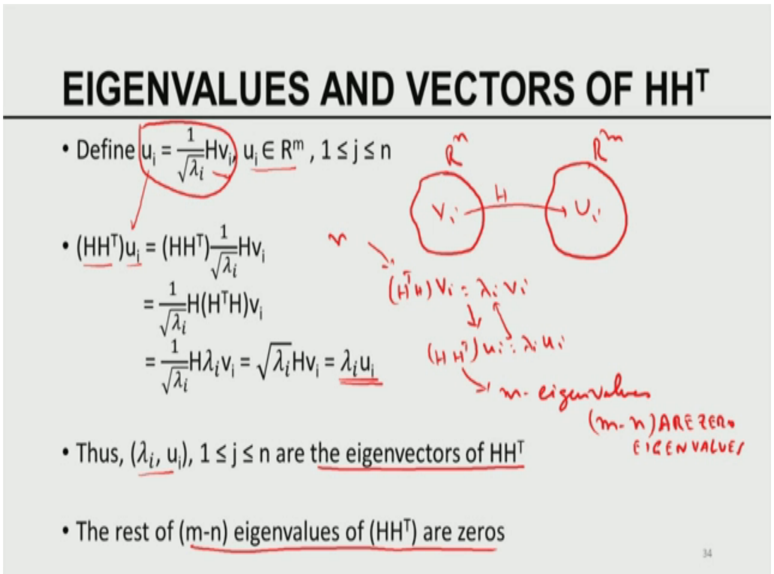
Because we are considering the case m is greater than n; that is fixed, but these two are two different grammians, one is smaller another is larger, because it is cheaper to work with smaller matrices. Now I am going to consider first the analysis, Eigen analysis of H transpose H. So, consider H transpose H, which is there an n by n matrix. H transpose H by this definition is S P D; therefore, it has the n Eigen values which are the largest of them is lambda 1 the smallest of them is lambda n, even the smallest of them is strictly positive.

If lambda I is an Eigen value, let V i be the corresponding Eigenvector corresponding Eigenvalue lambda I for each I Eigenvalue pair. So, if these are the n different Eigenvalue Eigenvector pair for H transpose, it; that means, H transpose H V i is equal to lambda I V i; that is the fundamental relation that comes from the definition of Eigenvalue. There are m such relations; I am going to by invoking to matrix relation, succinctly denote the den relation is our own single relation. To that end I am going to concoct a matrix V. The matrix V is simply n by n matrix, whose n columns are the n Eigenvectors of H transpose H lambda is a diagonal matrix whose diagonal elements are lambda 1 to lambda n.

So, V is a n by n matrix diagonal matrix lambda is also n by n matrix. We have already alluded to the fact, that the Eigenvectors of a real symmetric positive threat matrix are orthogonal to each other. Therefore, V is again an orthogonal matrix. Now you can see orthogonal matrices are occurring again and again, 1 is you are decomposition. Now in Eigen decomposition of symmetric positive definite matrices, there are very many different uses of orthogonal matrices in different applications. If V is orthogonal v transpose v and v v transpose is for the identity. Therefore, this relation can be expressed succinctly as H transpose H V is equal to v lambda. If a post multiply both sides by v transpose, both this relation comes to be; that means, H transpose H is equal to v lambda v transpose, this is called the Eigen decomposition of H transpose H. This is called the Eigen decomposition.

So, so far so good, what does it tell you? Given any matrix H of size m by n, H is a full rank. I can consider I can compute two grammian smaller larger, compute the Eigen structure of the smaller grammian. I can express the grammian as the product of v lambda v transpose; V is the matrix of Eigen vectors of the grammian lambda is the diagonal matrix of Eigenvalues of the grammian. So, this is the very basic fundamental results, that comes from the symmetric possible, theory of symmetric positive definite matrices.

(Refer Slide Time: 29:35)



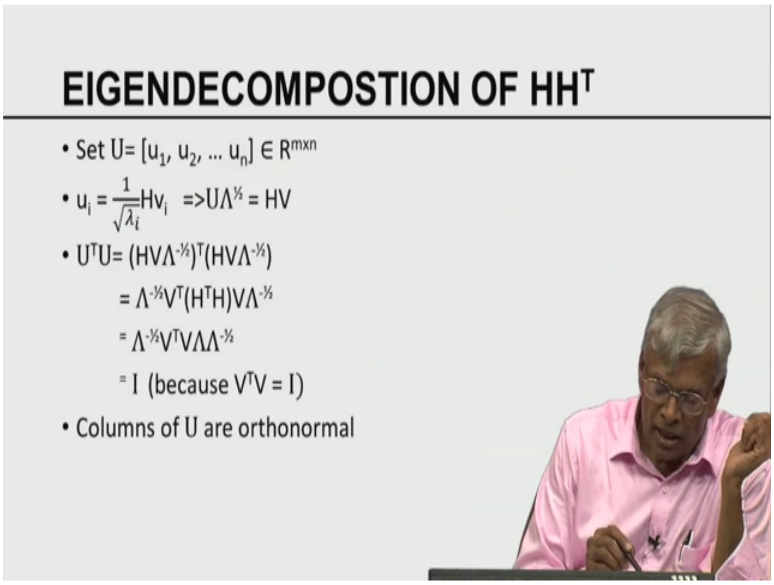
Now, I know V, I know lambda, I know H. So, I am going to now consider a linear transformation define a vector u i is equal to 1 over square root of lambda I lambda is are positive I can take the square root H is I m by n matrix, V i is a m n by 1 vector. So, H V i is a m vector, this is the constant. So, u i are m vectors I would like to remind you. So, this converts your vector from this transformation converts the vector from R n to R m. So, this is R n, this is R m. If there is V i, this is u i this transformation H, helps to convert an n vector into a m vector; that is a linear transformation u i. Now I am going to multiply H H transpose. Please remember H transpose H is a smaller the grammian, this is the larger of the grammian. I am going to take the larger the grammian that is m by m matrix; I am going to multiply that by m vector.

I already know the value of u, I substitute this value which is here. I now simplify, if I simplified it and I am, I get H H transpose u i is equal to lambda i u i, I am going to let you follow through the simplification procedure, its a very simple sequence of arguments in matrix algebra. So, what does this tell you? This essentially tells you the following, if V i is such that H transpose H is equal to lambda i V i, that immediately tells you H H transpose u i is equal to lambda i u i; that is a fundamental result. So, that goes to tell you if V i is are the Eigenvectors of H transpose H, u i are the Eigen vectors of H H H transpose, but they share the same Eigenvalue. So, they share the same non 0 Eigenvalues. This matrix must have m Eigenvalues.

This matrix must have n Eigen values, m is larger than n. So, in this case they share the same set of non zero Eigenvalues the m minus n Eigenvalues are 0, are 0 Eigenvalues in here, 0 Eigenvalues; that is where the semi definiteness comes into play. therefore, lambda I u i are the Eigenvectors of Eigenvector value Eigen vectors of H H transpose, the rest of the m n Eigenvalues of H transpose are 0. So, we talked about the semi definiteness of this, positive to semi definiteness of this matrix H H transpose.

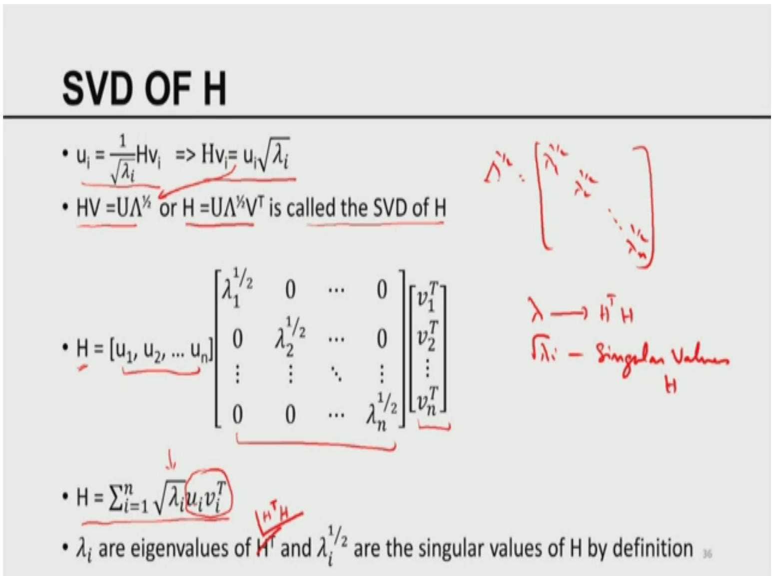
Now, these are the general results. I would like to concentrate on the definition of this, and that is what is going to help us to be able to look at the linear transformation leading to S V D.

(Refer Slide Time: 32:56)



So, let u 1 to u n be a matrix U, u i is equal to please remember, this is equal to this is equal to H v a 1 over square root of lambda I H V i H V i. So, U U transpose you can readily see is I therefore, U U transpose has also the orthogonal property that 1 would require, I would like you to follow through this simple relation. So, that essentially tells you the columns of u are also orthonormal. So, we have proved the properties of v, we have proved the properties of u, u and v are related to the metric H transpose H, and H H transpose. So, we have seen all the Eigen structure, Eigenvectors of the two grammians H transpose H and H H transpose.

(Refer Slide Time: 33:58)



Now, we are going to concentrate on what is the singular value decomposition of H, that comes from basically the definition of u i under the relation to V i, please remember this that is the way we define u i. So, this by cross multiplying both sides by square root of lambda I this can be rewritten as H V i is equal to u i lambda I this is for each i. So, this relation if I concoct for all I, can be written in a matrix form as H V is equal to u lambda to the power half, what is lambda to the power half, lambda to the power half is simply the diagonal matrix, which is lambda 1 to the power half, lambda 2 to the power half, and lambda n to the power half, the square root of the diagonal matrix lambda.

V is orthogonal. So, I can multiply both sides by v transpose, if I multiply both sides by v transpose because v v transposes I, I get this relation look at this. Now H I have now expresses the product of u lambda to the power half and v transpose, and this decomposition is a new decomposition, it is not L u, it is not Cholesky, it is not QR, it is a new animal, is a new form of decomposition, it is a multiplicative decomposition.

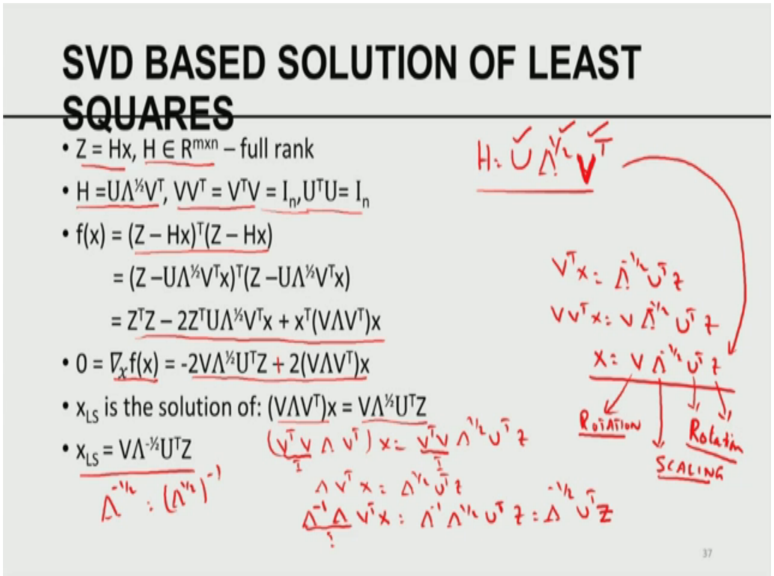
This decomposition is called singular value decomposition of H, why this is called singular value; lambdas are the Eigenvalues of H transpose h. So, square root of lambda are called the singular values H. We have already alluded to these things in our module on matrices, because square root of lambda I or the singular values of H this is called singular value decomposition. therefore, H can be expressed has a product of the matrix u, the product of lambda to the power half, and the product of v transpose like this, if I do the multiplication you can readily see H can be expressed as a product of square root of lambda i. So, that is a scaling factor u i, is a column vector, V i transposes a row vector. So, this is an outer product matrix.

Each matrix is multiplied by square root of lambda i. So, this is the weighted sum of outer product matrices, where lambda is are the Eigenvalues of H transpose H, are the Eigenvalues of H transpose H. I am sorry I have Eigenvalues of a, let me correct that once second Eigenvalues of. This must be H transpose H, lambda i are Eigenvalues a H transpose H and lambda I to the power half or the singular values of H by definition. therefore, this expression, this expression, this expression. They are all called singular value decomposition of H, is a very powerful device. In fact, it is 1 of the most fundamental algorithms in numerical linear algebra, 1 of the important application of this in numerical linear algebra is the following.

Suppose somebody gives you a rectangular matrix H, and asks you to find what is the rank of H. A computational process for determining the rank of H is to compute the grammian H transpose H, and do an Eigenvalue decomposition, and organize the Eigenvalue in the decreasing order. The number of nonzero are the number of positive Eigenvalues is equal to the rank of the matrix. Therefore, this algorithm for computing the rank of the matrix, using S V D is a very stable procedure for computing the rank. Not only it is used in computing the rank, but also it can be used in solving the least square problems.

So, the Gramm Schmidt procedure, the S V D the things that we have developed, even though we have developed within the context for specific aspect of inverse problems, these algorithms are so fundamental to numerical linear algebra in computation, each 1 of these find multitudes of application. That is why these are considered to be some of the nuts and bolts of computational linear algebra.

(Refer Slide Time: 38:46)



Now, what I have to come back and close the loop. I have expressed H in the form of S V D. If I am able to express H in the form of S V D, how does it help me to solve the linear least square problem that is what I would like to come back to, because the whole aim in these lectures is to solve inverse problem.

So, H Z is equal to H of x I want to solve, H is a full rank. I can express H as S V D u lambda to the power half v transpose. Here v v transpose v transpose V is I n, u transpose u is I n, the orthogonality of u and v are already established. So, f of x is equal to Z minus H of x transpose Z minus H of x, but I can express H as u lambda to the power half v transpose, u lambda to the power v transpose. I multiply both sides, I utilize all these properties, it can be shown f of x is essentially this term.

This is the constant term; this is the linear term that is a quadratic term. So, if I compute the gradient and equated to 0, the gradient of this expression is given by. The linear term is given by the first term, the quadratic term is given by this this must be a plus instead of a minus. Therefore, if I solve this system I simply have v lambda v transpose, x is equal to v lambda to the power half u transpose Z. I can multiply both sides on the left by v transpose, I can multiply. So, if I multiply both sides by v transpose, v transpose v lambda v transpose x must be equal to v transpose v, lambda to the power half u transpose z, but this is I, this is I though that reduces to v u transpose. I am sorry 1 second.

That leads to, this is I, that is way also equal to I. We are left with lambda v transpose x is equal to lambda to the power half u transpose Z. now I can multiply both sides by lambda to the power inverse, lambda inverse. So, if I multiply lambda inverse lambda v transpose x is equal to lambda minus 1 lambda to the power half u transpose Z. This is equal to I, this is equal to lambda to the power minus half. You can readily see that is equal to lambda to the power minus half, I hope you can see, that may be, maybe I will write that clearly.

This. So, this is equal to, this is equal to lambda to the power minus half u transpose u transpose Z. therefore, v transpose x is equal to lambda to the power minus half u transpose Z. I again multiply both sides by v, v v transpose; x is equal to v lambda to the power minus half, lambda to the power minus half u transpose z. So, this gives it x is equal to v lambda to the power minus half u transpose Z, say that is the least square solution. Therefore, if I am able to express my H, to be equal to u lambda to the power half v transpose, v transpose. I have found u, I know lambda to the power half, I know v transpose I i use v, I use lambda to the power half here and u. So, all the factors are used here. So, I can readily use the factors that I find in the S V D, and express my least square solution to be this.

Now, this are the very beautiful interpretation. So, Z is the observation vector V is given to me, u transpose is an orthogonal transformation, orthogonal transformation essentially rotates they do not elongate or shrink. So, u transpose Z is simply a rotated vector Z, then I multiplied by lambda to the power minus half diagonal elements. So, that essentially scaling. So, this is rotation.

So, I am going to give a geometric feeling. This is rotation, multiplying this is scaling, and then V is again orthogonal that is again rotation. Therefore, the least square solution, is obtained by first rotating the given observation, and scaling the given observation by the diagonal matrix, which is the inverse square root of lambda, and then again rotation by v. So, you get the least square solution; that is a beautiful way to be able to express the least square solution, using the method of singular value decomposition.

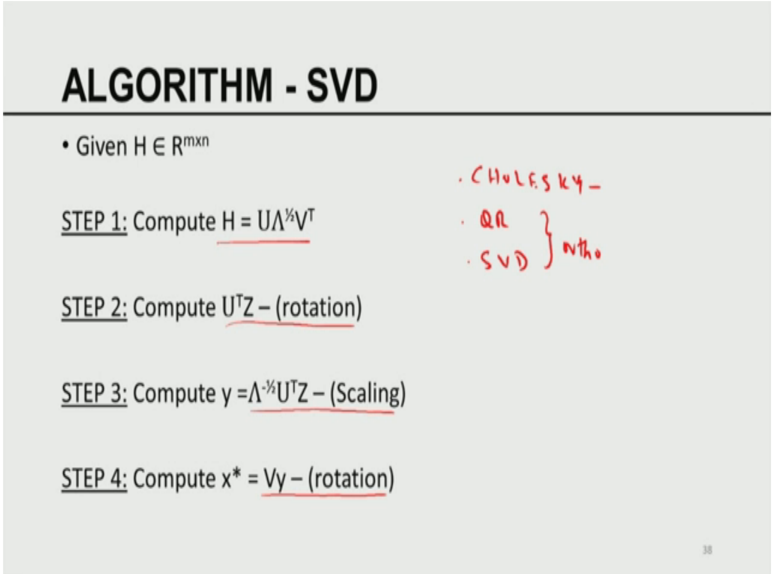
So, it has a beautiful geometric view as well, in addition to be able to analytically express the solution in some perform. Now look at this now, in here I do not have to solve any system. In the previous case I have to solve, either a cholesky decomposition, I had to solve an upper triangular system lower triangular system.

In the case of QR decomposition I had to solve an upper triangular system. Here I do not have to solve any system. Why, the inverse occurs only for lambda to the power minus half. What is lambda to the power of minus half? Lambda to the power minus half is equal to lambda to the power half inverse. Inverse of a diagonal matrix is they are essentially inverse of the diagonal elements, inverse of the diagonal elements, there are n elements I can compute the reciprocals of each, stick them along rely of this. So, that is the only inverse I need to compute.

So, there is no cost associated with computing inverse except for n reciprocal computation. So, once u lambda to the power half, and v are available, it is simply. You do not have to do any operation except rotation simple scaling, and another rotation, but I want to quickly add, to compute u lambda and v, we have to solve an Eigenvalue problem. Eigenvalue problems are intrinsically more expensive. Therefore, this method could overall be very expensive, but this is a very good stable method and very often recommended as one of the alternate methods.

So, among the three methods we have seen; the cholesky, the QR and the S V D. QR and S V D relate to orthogonal transformations. Any algorithm in matrix theory that involves orthogonal transformations, are generally more stable numerically. therefore, these methods, these two lateral methods are much preferred. The method of normal equation based on cholesky method is very good method, but its subjected to lot more numerical round of errors. So, the three methods have different strengths and weaknesses.

(Refer Slide Time: 46:43)

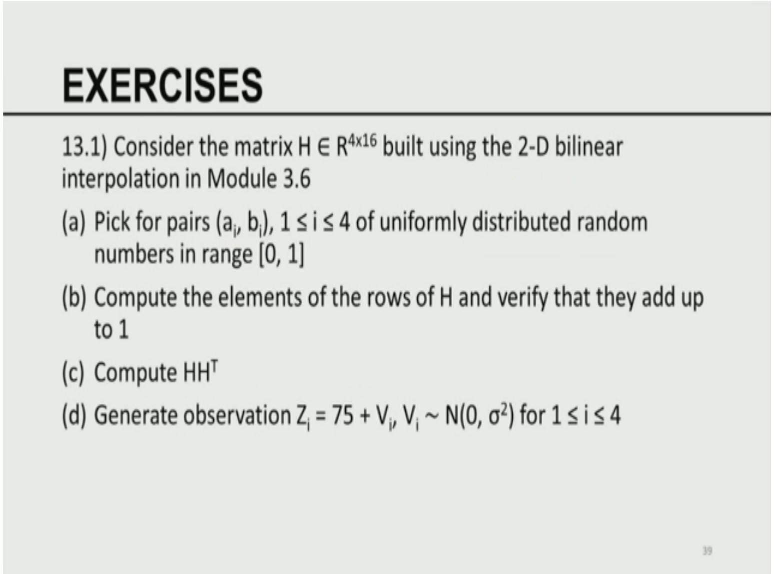


But before I go to compare I would like to summarize the S V D quickly here, H is U lambda to the power half v transpose.

That is the S V D decomposition, compute u transpose Z that is rotation I am sorry U transpose, see that is simply a rotation, I have already talked about that in the previous slide. This essentially a scaling and this is essentially another rotation. Therefore, we can very easily see the extra, is the least square solution. So, with this we have covered three methods cholesky sorry. With this we have covered three method; cholesky, QR, and S V D, these two involve orthogonal methods, orthogonal matrices. This method essentially is a variant of L u decomposition. Cholesky method is computationally not stable in the sense of numerical round off, as QR and S V D.

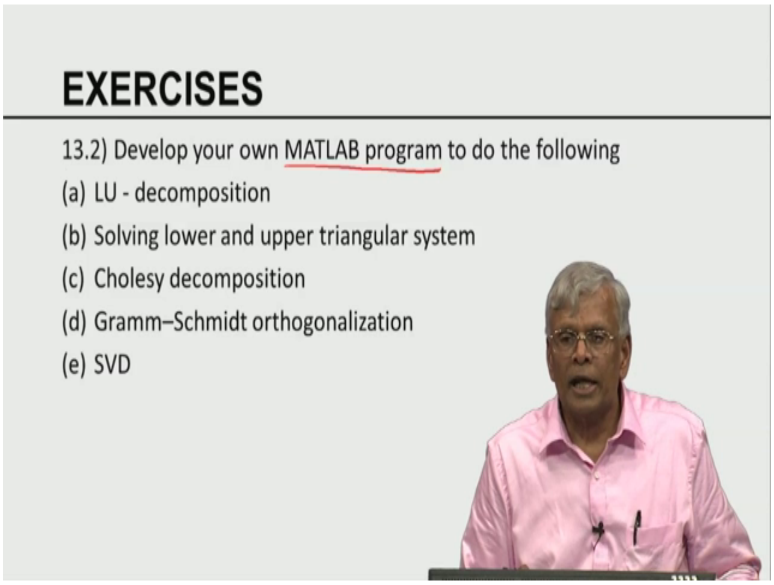
So, then you are trying to learn these methods, it is better to take 1 particular problem and solve it by three methods. When you when I say solve it the three methods I have given you algorithm for each of these. So, you can develop your own subroutine to do cholesky, your own subroutine for QR, your own subroutine for S V D. So, this way you can develop your own mathematical packages ground from ground up. So, that you do not have to depend on, any built in library; such as MATLAB library are C library or Fortran library, this way you can develop mathematical software to solve data assimilation problems from ground up rather independently from, by using most of the basic algorithms.

(Refer Slide Time: 48:57)



There are lots of exercises we have we have given. I would like to encourage the reader to work or through all these examples.

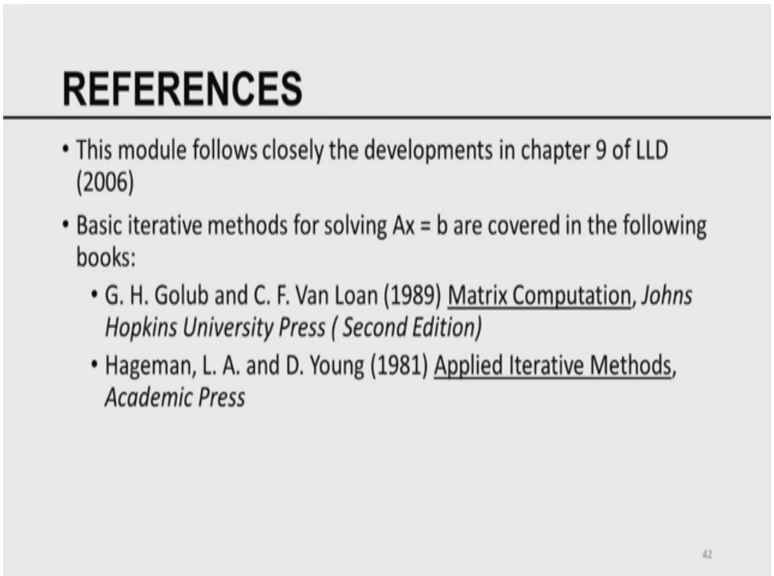
(Refer Slide Time: 49:09)



In fact, working through these examples is a fundamental aspect of thorough understanding. I also would like you to develop your own MATLAB program that is the important thing; you can develop your own MATLAB program. You may say hey MATLAB as already has this program why do I do this, but that is an exercise. Suppose you want to go to work forever national agencies such as the meteorological forecasting agency. They may try to develop all these systems ground up independent of any existing libraries, because they would like to have a better control over everything.

So, in order to be employed, in order to be able to develop such systems, you must do an exercise in developing these programs, L u decomposition, solving lower upper triangular systems, cholesky decomposition, Gramm Schmidt method, S V D. I would like you to be able to program these methods, and compare them, and that is a very good part of the exercise.

(Refer Slide Time: 50:01)



These are other examples these modules follow closely developments in chapter 9 of our book. We are not considering iterative methods. Iterative methods for solving linear systems are covered in several excellent textbook, Golub and Van Loan, as well as Hageman and Young. So, in view of time, we will not be able to cover the iterative methods. Dark methods we covered iterative methods are pursued in here. With this I believe, I have provided you a very good summary of matrix method, especially a direct methods to solve linear least square problems.

Thank you.