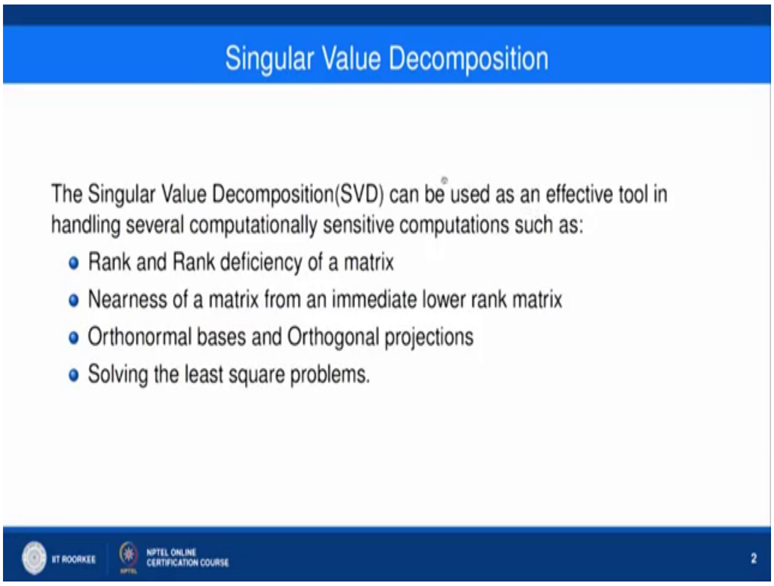
Numerical Linear Algebra
Dr. D. N. Pandey
Department of Mathematics
Indian Institute of Technology, Roorkee

Lecture - 42
Singular value decomposition of a matrix- I

Hello friends, welcome to this lecture, in this lecture we will discuss the very important tool in numerical linear algebra that is singular value decomposition and this here, after we call this singular value decomposition as SVD. And this SVD has a lot of very important application in various field of engineering and sciences. For example, we need it and control theory, image processing and several other field. So, let us start, what do you mean by singular value decomposition?

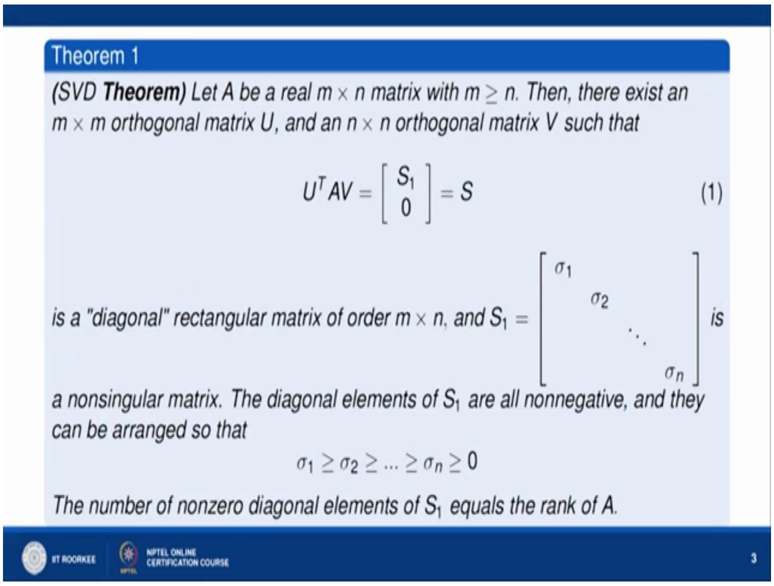
(Refer Slide Time: 00:58)



So, this singular value decomposition can be used as an effective tool in handling several computationally sensitive computation such as, rank and rank deficiency of a matrix. Sometimes it may happen that, your matrix which we are handling is having very is ill conditioned or we can say that, the lowest eigen value of a given matrix is very, very small. So, it may create problem in handling many of the calculation using this matrix and second is that, nearness of a matrix from an immediate lower rank matrix or you can say that, approximation of a matrix with an immediate lower rank matrix.

So, there this SVD is very, very useful and then, finding orthonormal bases for row space, column space and they are orthogonal complement subspaces and respective orthogonal projection. So, calculating this orthonormal bases and orthogonal projection we use SVD and also we use SVD in solving the least square problems, that all these all these problems we will discuss in coming lectures. So, let us first start what do you mean by SVD?

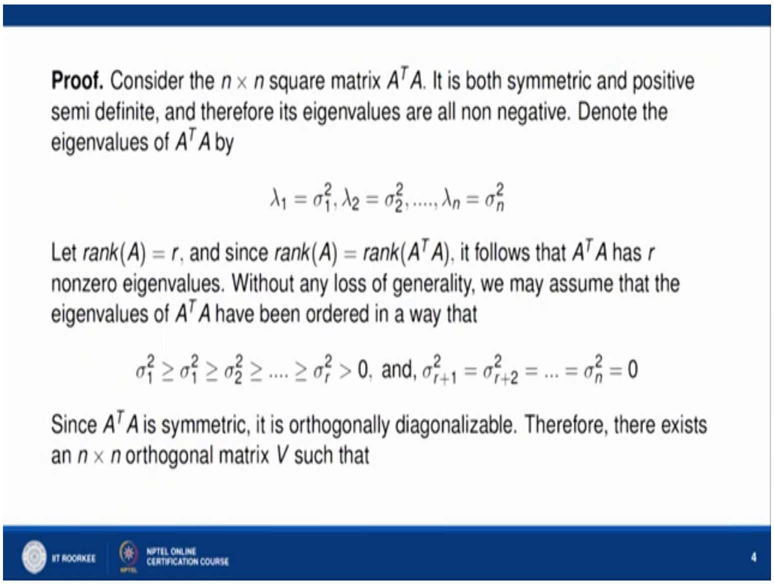
(Refer Slide Time: 02:18)



So, so here, let A be a real m cross n matrix with m greater than or equal to n, then they exist an m cross m orthogonal matrix U and n cross n orthogonal matrix V such that, U transverse AV equal to S1 0 equal to S here, 0 is a matrix of appropriate size and we call this S1 0 as S where, we say that, it is a diagonal rectangular matrix of order m cross n and S1 we are denoting as this diagonal matrix whose, diagonal elements are sigma 1 to sigma n is a nonsingular matrix.

So, the diagonal element of S1 are all nonnegative and they can be arranged in a way such that, sigma 1 is greater than sigma 2 and it is greater than equal to sigma n greater than equal to 0 see the number of nonzero diagonal elements of S1 equals the rank of A. So, here we want to get this kind of theorem first of all and if you look at look at this representation U transpose AV equal to S if you remember that, if A is a square matrix and if A is a diagonalizable matrix, then we can have this kind of matrix P transverse AP equal to some D matrix, but what happen if this matrix A is not diagonalizable then, we can utilize this SVD theorem to your nearly diagonalizable near nearly you can say that, U transverse AV is a nearly diagonal matrix or we can say that, we have a diagonal rectangular matrix and we will see that, why this form is very, very important? And how to find out this capital U and capital V?

(Refer Slide Time: 04:14)

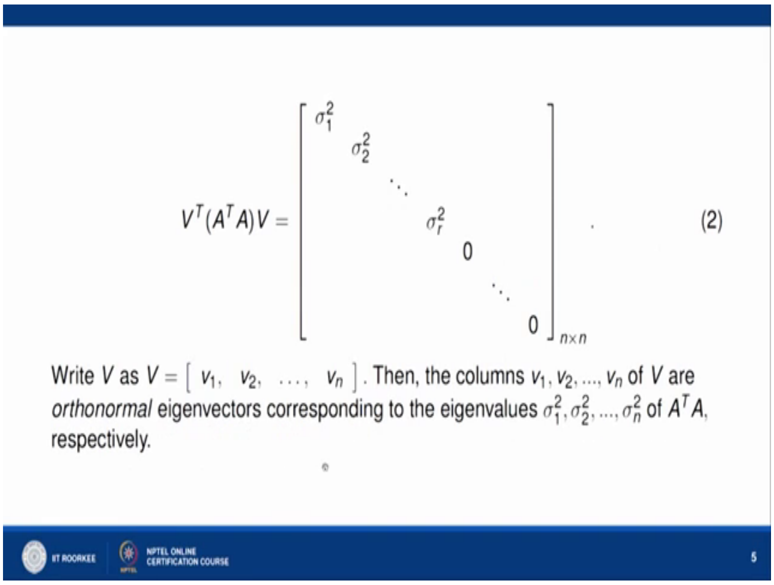


So, let us prove this theorem1. So, here consider the n cross n square matrix A transverse A now, since A is a n cross n matrix. So, A transverse A is we can define a n this is coming out to be symmetric and positive semi definite. So, it means that, it is eigenvalues are all non-negative. So, let us calculate all the eigenvalues of A transverse A call these eigenvalues as sigma 1 square and sigma n square and we are denoting as lambda 1 as sigma 1 square, lambda 2 as sigma 2 square, lambda n as sigma n square and here, let us assume that, rank of A is some r where, R may be less than or equal to n and we already know that, rank of A is same as rank of A transverse A.

So, it means that a transverse A has a nonzero eigenvalues. So, it means that, out of these n eigenvalues only r eigenvalues are non 0 rest of all 0 0 eigen value. So, without any loss of generality we may assume that, the eigenvalues of A transverse A have been ordered in a way. So, what we try to do we try to order them order the eigenvalues of A transverse A and in a way such that, sigma 1 square is the largest eigen value and sigma r r square is the smallest eigen value and rest are all is n minus r eigenvalues are 0 eigen values.

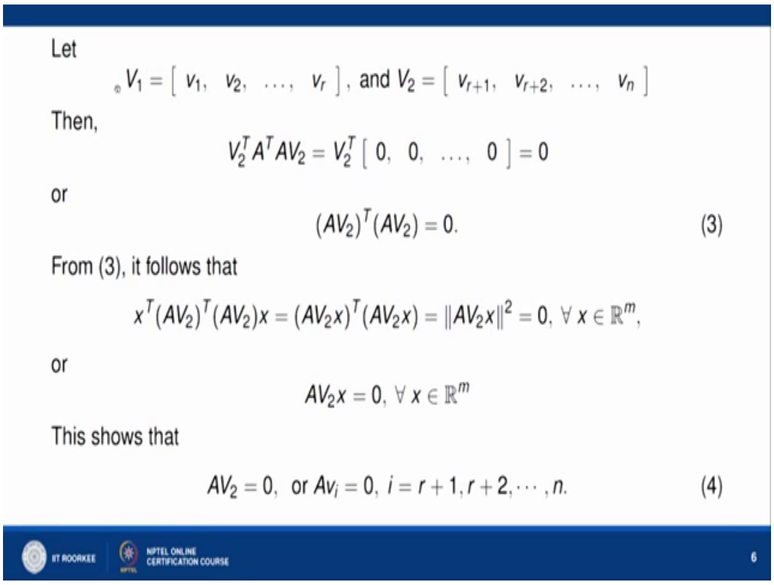
Now, here we also know that, since A transverse A is a symmetric matrix. So, it means that, A transverse A is orthogonally diagonalizable matrix. So, it means that, we they exist a n cross n orthogonal matrix V such that, V transverse A transverse AV is a diagonal matrix.

(Refer Slide Time: 06:00)



So, here diagonal matrix consist the eigenvalues of A transverse A here. So, here now, we can say that this V we can write as V1 to Vn where, V1 to Vn are eigen vectors of A transverse A. Now, then the columns V1 to Vn of Vr orthogonal eigenvectors corresponding to the eigenvalue sigma 1 square, sigma 2 square and sigma n square of A transverse A. So, this information we have got because, A transverse A is a symmetric real matrix. So, we can orthogonally ortho normally diagnose this A transverse A with the help of the matrix V here, V contain the orthonormal eigenvector corresponding to A transverse A.

(Refer Slide Time: 06:55)



So now, let us write this V as V1 and V2 where, V1 consist of first r eigenvectors. So, V1 is a matrix consisting first r eigen first r columns of V and V2 consist of the remaining n minus r columns of V.

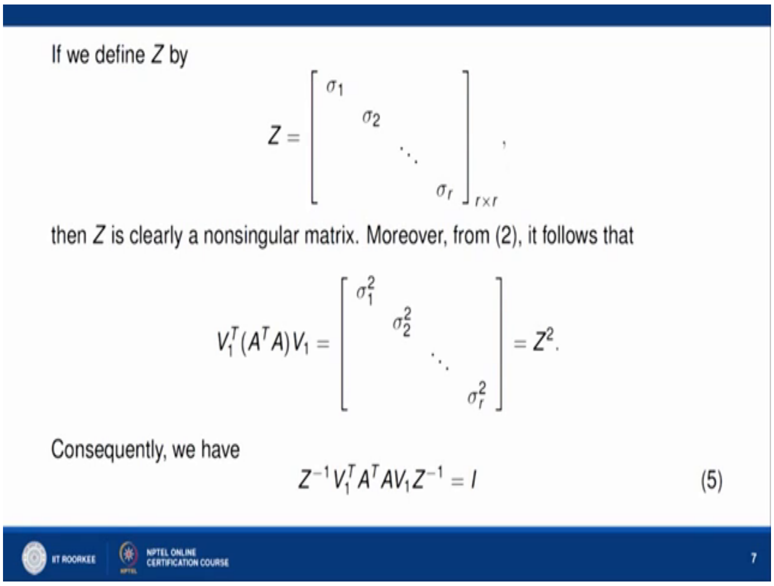
So, we can calculate that, V2 transverse A transverse A V2, which we can say that, it is nothing but A transverse A operating on V2. So, if you look at A transverse A operating on V2 means, A transverse A operating on Vr plus 1 to Vn. Now, this Vr plus 1 to Vn are the eigenvector corresponding to sigma r square sigma r plus 1 square, sigma r plus 2 whole square and sigma n square. Now, these are the eigenvector corresponding to 0 eigenvalue. So, it means that, A transverse A V2 is going to be 0.

So, it means that, A transverse A V2 is a 0 matrix 0 comma 0 comma 0 n minus r 0 are here. So, if you apply V2 transverse on this 0 matrix you will get a 0. So, it means V2 transverse A transverse A V2 is coming out to be 0, which we can written as A V2 transverse A V2 equal to 0 and here, we can say that from this we can verify that x transverse A V2 transverse A V2 x is equal to A V2 x transverse A V2 x and this is coming out to be norm of A V2 x whole square and it is coming out to be 0 because, of this equation number 3.

So, this is true for every x in Rm. So, in particular we can say that, AV2 x is equal to 0 because, if norm of any vector is 0. So, it means that that vector is nothing but a 0 vector. So, it means that AV2 x is equal to 0 for every x in Rm.

So, in particular we can take x as standard basis of Rm and we can show that, A V2 is a 0 matrix. So, or you can say that, A V2 is equal to 0 means AV i is equal to 0 where, i is running from r plus 1 to n. So, it means that AV i is equal to 0 for i equal to r plus 1 to n. So now, what we have done here, we are almost able to find out our V now, we try to find out the matrix U for that, we define Z by sigma 1 to A diagonal matrix whose diagonal elements are sigma 1 to sigma r.

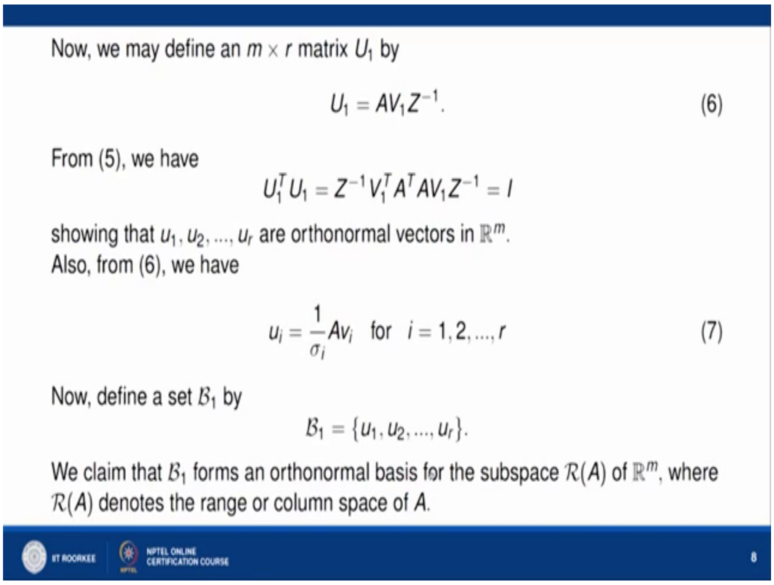
(Refer Slide Time: 09:23)



So, these are all nonzero eigen square root of eigenvalues of A transverse A. So, we can say that, Z is clearly a nonsingular matrix. So, from we can say that, let us calculate V1 transverse A transverse A V1. So, if you look at what is V1 here, V1 is is a matrix of consisting the first r column of V and these are eigenvectors corresponding to sigma 1 square, sigma 2 square and sigma r square. So, using this information we can say that, A transverse A V1 is basically sigma 1 square, V1 sigma 2 square, V2 and sigma r square V2 and when we apply V1 transverse it is coming out to be A diagonal matrix and whose, diagonal elements are sigma 1 square, sigma 2 square up to sigma r square.

And if you want to write this in terms of Z then, we can say that it is nothing but Z square. So, using this representation we can further simplify this and we can say that, applying Z inverse from the left-hand side and the right-hand side we can say that, this can be re written as Z inverse V1 transverse A transverse A V1Z inverse equal to i. So, let us if you look at this is a kind of this can be utilized as finding the matrix U. So, here we call this matrix A V1Z inverse as U. So, if you call this a V1Z inverse as U 1.

(Refer Slide Time: 11:09)



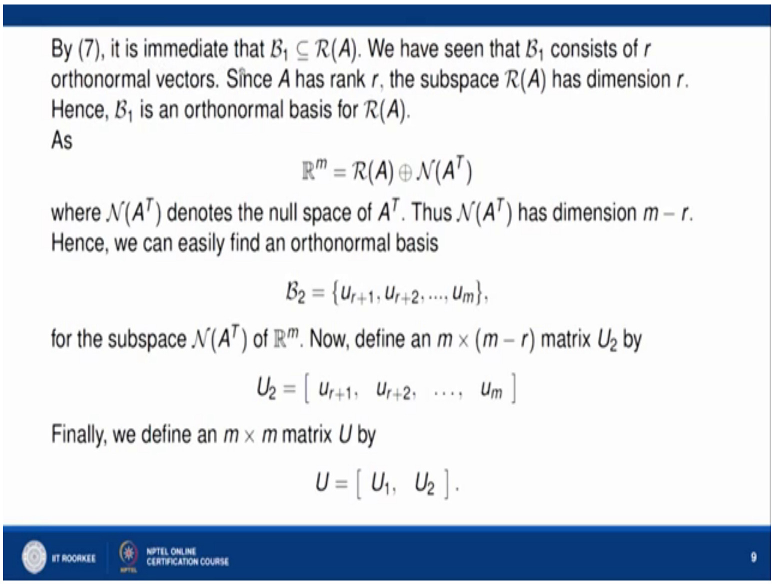
So, we define we define n m cross r matrix U 1 by U1 equal to AV1Z inverse. So, if you look at if we take this as U1 then, this is nothing but U1 transverse into U1 equal to i we simply say that that, U1 is an is an orthogonal matrix or we can say that U1 transverse U1is basically i or we can say that U1, U 2, U r are orthonormal vectors in Rm. So, columns of U1 forms an orthonormal say orthonormal set in R f Rm

So, what we have found here, is the matrix of size m cross r U1 such that, U1 transverse U1 is equal to i we can say that, columns of U1 matrix which is which are U1 to U r are orthonormal vectors in Rm or if we look at we can from this U1 if you look at this from here, we can say that U1, U i is nothing but 1 upon sigma i A of V i. So, here if you look at what is V1, V1is basically consisting of V1 to Vr eigenvectors corresponding to A transverse A. So, here we can simplify this equation number 6 and we can write it U i equal to 1 upon sigma i A of V i for i equal to 1 to r.

So, here we have started we have defined this capital U1 with the help of matrix Z, but if you directly if you can you can directly define your U1 from this from the matrix V1 and we can say that U1 is nothing but 1 upon sigma i A of V i where, i is running from 1 to r. So, we can start defining our matrix U1 from this point also, but anyway this is quite convenient if you define your U1 in terms of Z. So, here U i is 1 upon sigma i A of V i for i equal to 1 to r. Now, if you look at if you form the set using these U i. So, we have a set B 1 consisting r vectors defined as equation number 7 and we say that, this B1 forms an orthonormal basis for the subspace R of A of Rm.

Now, for that please observe that these U1to U r are all orthonormal vectors in Rm. So, it means that, if we can prove that range space of A is having a dimension r then we can say that, B1 is an orthonormal basis for this subspace R of A where, R of A denotes the range or column space of A.

(Refer Slide Time: 14:22)



So, here we can say that, B1 is subset of R of A, why B 1 is subset of R of A? Because, if you look at the equation number 7, 7 says that U i is nothing but 1 upon sigma i A of V i. So, here this implies that, this U i belongs to the range space of matrix A. So, it means that, all this U1to U r belongs to the range space of A. So, we have seen that B1 consist of r orthonormal vectors. So, since A has rank r it means subspace R A has dimension of r, why? because, range space of A is basically nothing but column space of A. Now, the column space of A has dimension n.

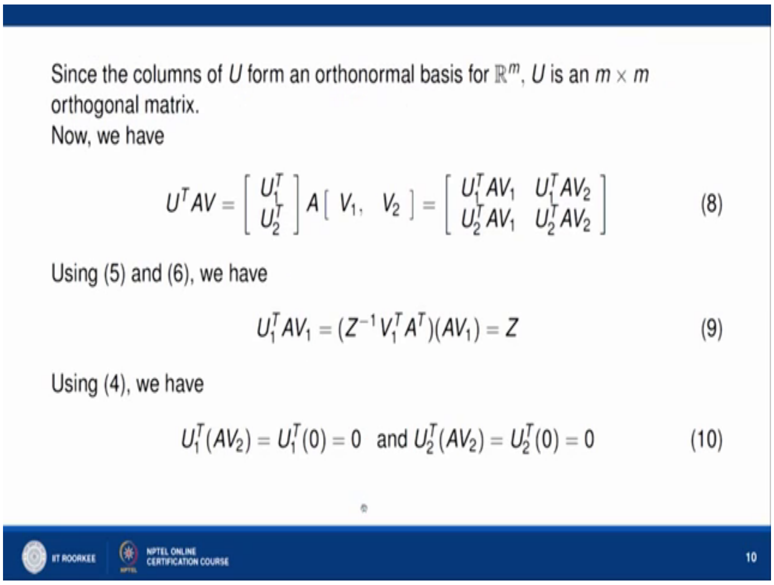
So, it means that range space of A has dimension r. Now, since A has rank r the subspace RA has dimension r and we can say that, B1 is basically a linearly independent because, it is orthonormal set of vector. So, we can say that, B1 is a linearly independent set of vectors whose, dimension is basically r.

So, we can say that, this B1 generate A subspace of dimension r and range space of A has dimension r. So, it means that, B1 generate the range space of A or we can say that, B1 is an orthonormal bases for this range space of A. Now, we already know that, Rm can be written has range space of A null space null space of A transverse and. So, here we can say that, what is null space of A transverse? Null space of A transverse denote the null space of A transverse.

Now, we can say that by (Refer Time: 16:12) theorem here, the null space of A transverse (Refer Time: 16:15) has dimension m minus r. So now, we can do here, we have the basis for this range space of A now, and we want to find out say basis for Rm. So, we can extend the basis of range space of A to the basis of Rm and we can find out the vectors Ur plus 1 to Um and we if we ortho normalize it using gram (Refer Time: 16:42) organization process. So, we can say that, we can easily find an orthonormal basis B2 for the subspace of null space of A transverse of Rm.

So here, with the help of this orthonormal basis, which is the orthonormal basis of null space of A transverse we can define the matrix U2consisting of these n m minus r column vectors, and we write U2as U r plus 1, U r plus 2 and U m. And with U1 and U2we can define an m cross m matrix U by U is equal to U1 comma U2. So, it means that the first r columns are coming from U1 and the remaining m minus r columns are coming from U 2. So, we claim that, this U is the desired matrix m.

(Refer Slide Time: 17:35)



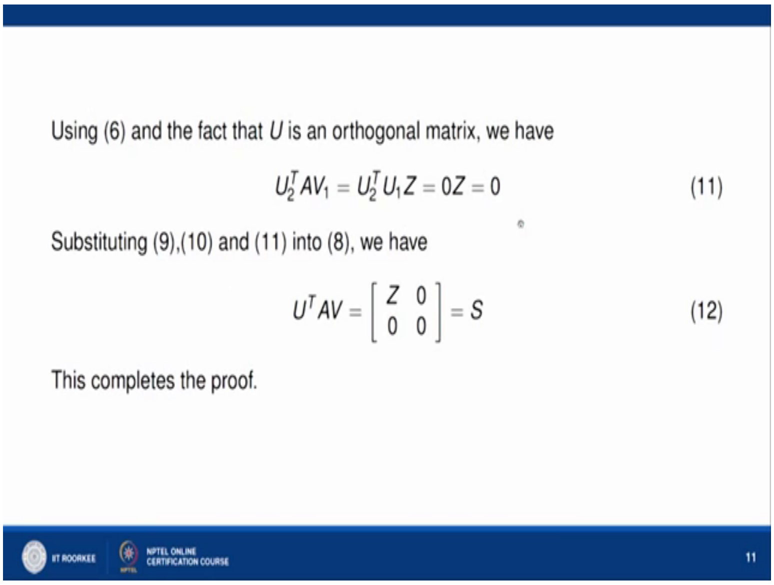
So, here with this U we simply say that, since the columns of U forms an orthonormal basis of Rm. So, we can say that, these columns of U1 and columns of U2 are basically orthogonal to each other and we say that, this U forms an orthonormal matrix of size m cross m. So, we want to calculate this U transverse AV. So, writing the expression for U transverse it is nothing but, U1 transverse, U2 transverse, A and V can be written as V1 and V2. So, when you simplify it is coming out to be this block matrix and whose blocks are given as U1 transverse AV1U1 transverse, A V2 U2transverse, A V1and U2ah transverse A V2.

So, to calculate these blocks let us start with this U1 transverse AV1. So, here U1 transverse a V1 is what using the expression for U1, U1is defined as Z inverse V1 transverse A transverse A V1and if you look at this V1 transverse A transverse A V1 is nothing but, your Z square we have already defined this as here. So, here we have defined V1 transverse A transverse A V1 as Z square. So, using this expression we can simply say that, Z inverse and this is your Z square.

So, we can say that it is Z. So, the first block is coming out to be Z here. Now, now let us we have already calculated that, this quantity A V2 is a 0-matrix for that, if you look at here we here we have proved that, A of V2 is equal to 0. So, using this A V2 is a 0 matrix we can say that, U1 transverse AV2 is basically U1 transverse applying on 0 matrix, which is coming out to be 0. So, this implies that, this block matrix is nothing but, a 0-block matrix similarly we can calculate U2 transverse A V2 because here, also we can say that since AV2 is a 0 matrix. So, operating U2 transverse on A V2 it is also coming out to be 0 matrix.

So, we here we can say that, this is Z this is 0 this is 0. Now, let us calculate this quantity U2 transverse A of V1.

(Refer Slide Time: 19:53)

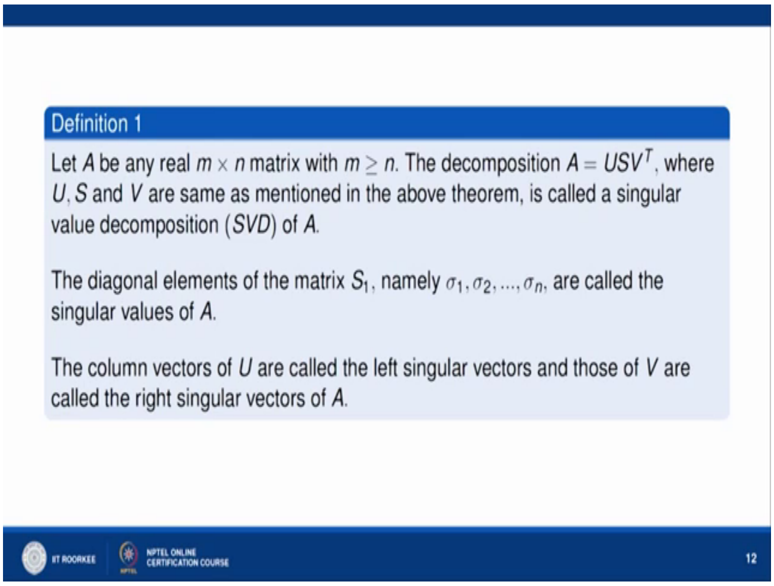


So, you look at U2 transverse AV1. So, here we are using the here, what is the expression for U1? So, here A of V1 can be written as U1of Z, how we can write it? If you look at the equation number 4 here. So, here we can say that here, equation number 6 basically U1 is equal to A V1Z inverse. So, we can apply Z here. So, you can say that U1Z is equal to AV1. So, using the expression 6 we can write the last here that, U2 transverse AV1 and in place of a V1 we are writing U1Z and you can say that, this is nothing but U2 transverse U1 now, U2 transverse U1 is basically 0, why?

Because, the columns of U and columns of U1 are orthogonal to each other. So, here because, of ortho gonalty of matrix U this is coming out to be 0 operating on Z and it is coming out to be 0. So, U2 transverse A V1is also coming out to be 0. So, it means that here, in equation number 8 this is coming out to be 0 because, of AV2 and this is coming out 0 because, of a V2 and here we have utilized the equation number 6 that, A V1 is equal to U1Z.

So, again we can say that, U transverse AV is coming out to be this matrix Z 0 0 where, Z is basically diagonal matrix consisting the diagonal entries as sigma 1 to sigma r here, and we call this block matrix as S and we say that, U transverse AV is S and we say that, this is the required thing we wanted to prove if you look at what is the theorem we wanted to prove? We wanted to have U transverse AV equal to S here, S is having this S1 comma 0 here, we have denoted this S1 as Z. So, Z is basically sigma 1 to sigma n here, in this case our n is coming out to be r. So, here we have proved the that, A can be decomposed as U transverse AV equal to S. Now, with the help of this decomposition we can find out we can define, what is singular value decomposition?

(Refer Slide Time: 22:14)

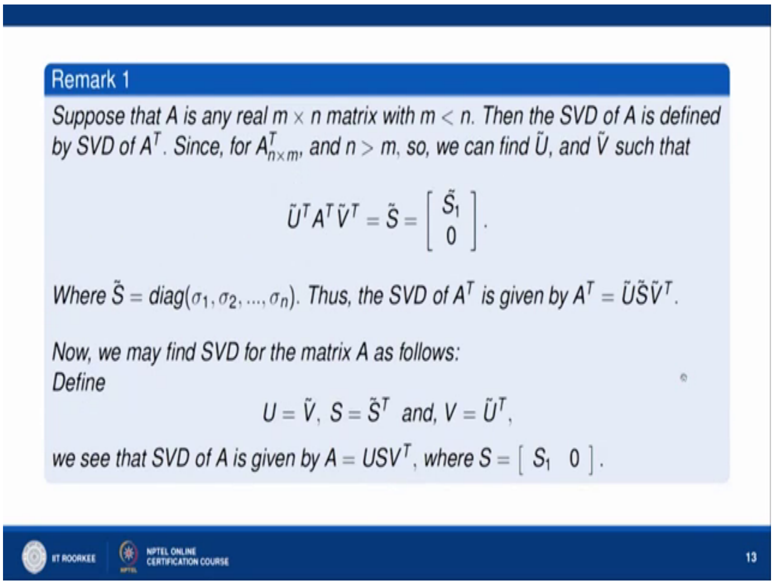


So, let A be any real m cross n matrix with m is greater than or equal to n the decomposition A, which is given as USV transverse. So, we have U transverse AV equal to S and since, U and V are orthogonal matrix we can write this as a as USV transverse which is we are writing here.

So, it means that where, U, S and V are same as mentioned in the above theorem is called a singular value decomposition of A. So, it means that, A written as USV transverse is known as singular value decomposition of A here, and the diagonal element of the matrix S1 namely sigma 1 to sigma n are called the singular values of A. The column vectors of U are called the left singular vectors and those of V are called the right singular vectors of A.

So, here we also define what is left singular vectors and right singular vectors of A. So, you can say column vectors of U are left singular vectors and columns of V are known as right singular vector vectors of A. We will try to see why we call this as left singular vectors or right singular vectors and because, of this thing that sigma 1 to sigma n are called singular values of A we say that, this decomposition of A is known as singular value decomposition of A.

(Refer Slide Time: 23:43)



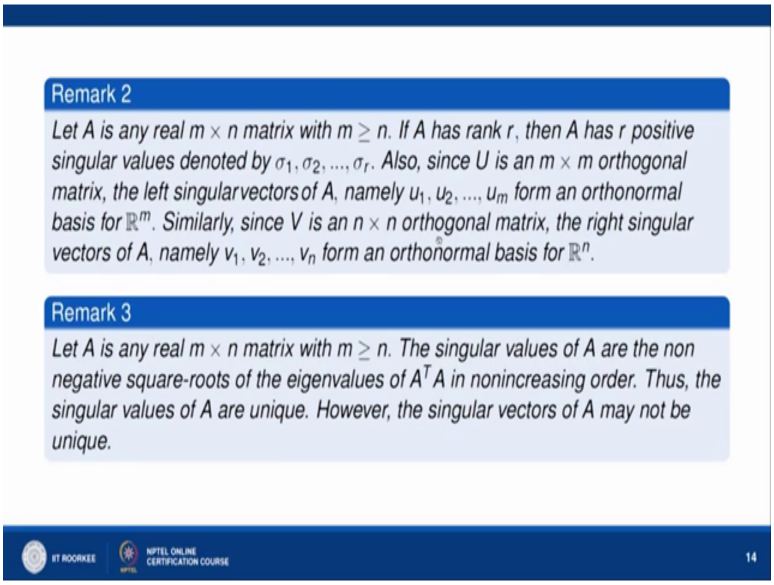
Now, remark 1. So, here we say that we have given the theorem for any real matrix m cross n where, m is greater than or equal to n, but what happen that if we have A matrix A of the size m cross n and m is strictly less than n. So, in this case, how we can say that how we can obtain the SVD.

So, if you look at in this case when A is any real m cross n matrix with m is less than n we can say that, SVD of A is defined by SVD of A transverse. Now, what is a transverse? So, if you define A transverse A transverse is of size n cross m and n is greater than equal to m. So, for this our theorem will be applied and we can find out U tilde and V tilde orthogonal matrixes of appropriate size such that, U tilde transverse A tilde AT A transverse V1 V tilde T can be written as S tilde where, S tilde is S1 tilde 0 where, S S1 tilde is the diagonal matrix consisting the singular values sigma 1 to sigma n thus, the SVD of A transverse is given by A transverse equal to U tilde S tilde V tilde transverse.

Now here, we can apply the transverse of this whole expression and we can say that, if you take the transverse of the whole expression what will have we define U as V tilde S as S tilde transverse and V as U tilde transverse and from this we can say that, taking the transverse of the whole expression we can say that, A can be written as USV transverse where, U is V tilde, S as S tilde transverse and VT V transverse is nothing but U tilde transverse and here, S is what S is given by this matrix S1 comma S1 0.

So, it means that whether it is a any m cross n matrix whether, m is greater than or equal to n or m is less than n we can always find out the singular value decomposition of a matrix A. So, without loss of generality we can always find out SVD by assuming that, your number of rows are bigger than number of columns here.

(Refer Slide Time: 26:07)

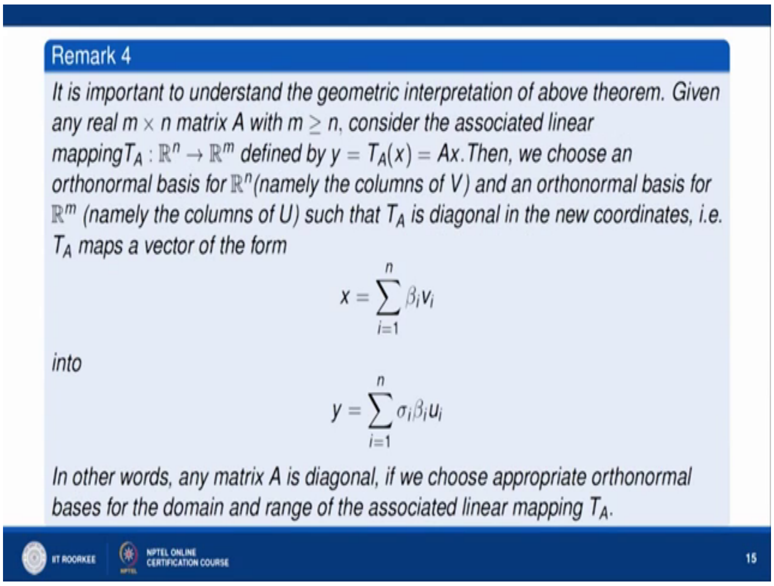


So now, let us consider this remark 2, remark 2 says that let a is any real m cross n matrix with m greater than equal to n and if a has rank r then A has r positive singular values denoted by sigma 1 to sigma r and here, if you remember we have we are able to find out this matrix U consisting m orthonormal vectors and we can say that here, U is an m cross m orthogonal matrix, which we are defined as left singular values of lift left singular left eigenvectors of A namely U1 to Um. And this will form an orthonormal basis for Rm similarly, V is an n cross n orthogonal matrix and we call this as right singular vectors of A and which are namely V1 to Vn and this will form an orthonormal basis of Rn.

So, what we have is that, if we have a singular value decomposition of A as U as V transverse then, columns of U will form a orthonormal basis for Rm and columns of V will form an orthonormal basis for Rn. So, with the help of SVD we can find out say orthonormal basis for domain space and codomain space now, coming out to coming to remark 3. So, here let A is any real m cross n matrix with m greater than equal to n, the singular values of A r the nonnegative square roots of the eigenvalues of A transverse A in non-increasing order. So, we can say that singular values are going to be unique, but the singular vectors of a may not be unique the we can say like this, that if you look at A transverse A and since it is symmetric non semi definite matrix we can say that, if in the case of repeated eigenvalues then, the corresponding eigenvectors can be taken as any linearly independent vectors which can spend the eigenspace.

So, here basis since basis of any vector space is not unique. So, we can say that basis spanning the eigenspace is also not unique and we can say that your V matrix is not unique and we can say that, the singular vectors of a may not be unique because, U is very much related to capital U is very much related to V. So, here we can say that this singular vectors of a may not be unique. So, it means that singular values are unique, but not the singular vectors.

(Refer Slide Time: 29:06)



So now, let us coming out to be remark 4, which discuss the geometric interpretation of the S singular value decomposition theorem and it says that, given any real m cross n matrix A with m greater than equal to n it is by default we are assuming all the time consider the associated linear mapping TA, which is map which map from Rn to Rm and is defined as y equal to Tax, which is given as A of x where, A represent the matrix representation of this linear mapping T of A and we can choose an orthonormal basis for R n and it is what it is namely the columns of V and an orthonormal basis for Rm again it is the columns of U we can use we can choose the columns of U such that, T of A diagonal in the new coordinates, that implies what that if you take any vector in R n ah, which can be written as i equal to 1 to n beta i V i.

Then our claim is that, it is image is going to be y, which is which has representation i equal to 1 to n sigma i beta i U i and if you look at how we are here that, if you operate your A on this x then, A of x is going to be what i equal to 1 to n beta i A of V i now, what is A of V i? A of V i is given as sigma i U i and this is how we define our U i. So, we say that in other words any matrix A is diagonal if we choose appropriate orthonormal basis for the domain and range of the associated linear mapping T of A. So, earlier this representation is we know in terms of diagonalizable square matrix, but with the help of this singular value decomposition theorem, we can extend the similar kind of notation for any rectangular matrix of size m cross n. So, here I will stop this discussion.

So, in today's lecture what we have done we have done we have found the singular value decomposition of any real m cross n matrix where, m is greater than equal to n and we have seen how to find out the component of the singular value decomposition, that is how to find out capital U and how to find out capital V and how to find out the singular values of S, and will end here this discussion.

In next lecture we will try to take some example and try to find out how we can find out this U S and V. So, thank you for listening us we will meet in a next lecture.

Thank you.