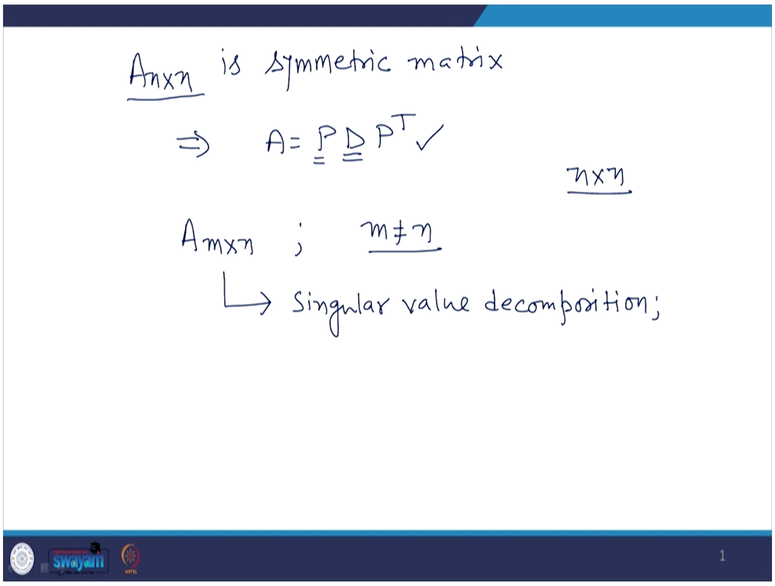
Essential Mathematics for Machine Learning
Prof. Sanjeev Kumar
Department of Mathematics
Indian Institute of Technology, Roorkee

Lecture – 11
Singular Value Decomposition

Hello, friends. So, welcome to the lecture number 11 of this course. In the last lecture we have learned about least square approximation and minimum norm solution. So, by now you are quite comfortable for solving rectangular linear system; whether we are talking about over determined system or under determined system. Today, again we are going to learn a very useful concept related to the rectangular matrices.

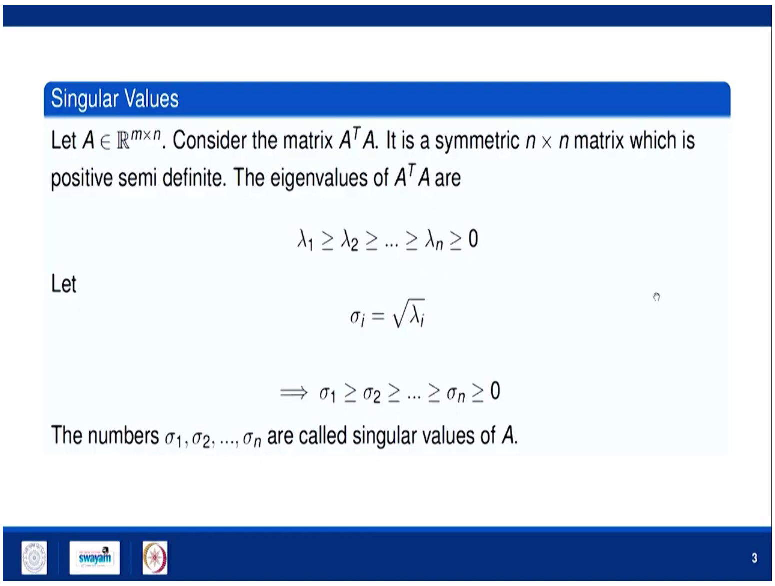
(Refer Slide Time: 00:59)



So, if you remember if A is a n by n symmetric matrix so, in 1 of the previous lecture we have learned that I can write A equals to P D P transpose where D is a diagonal matrix having the eigenvalue values of A. P is a matrix called model matrix which is having columns from the ortho normal eigenvectors of A and if it is a symmetric it will be having real eigenvalues and the same time it will be having n linearly independent even an orthonormal eigenvectors; because it will be having an orthogonal basis for r n vector space.

So, here A is a square matrix that is it is a n by n matrix and we can have this kind of decomposition of A. If we A is m by n matrix where m is not equals to n in this case such kind of diagonalization is not possible.

(Refer Slide Time: 02:49)

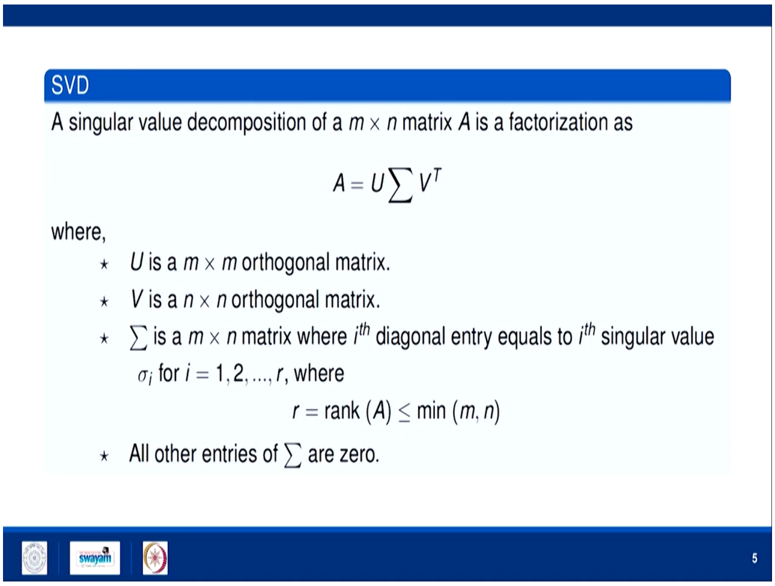


So, what to do? So, in this lecture we will learn singular value decomposition. So, let us learn first what we mean by singular values. So, let A be a m by n matrix having real entries. Consider the matrix A transpose A. So, A transpose A will be a n by n metric which is symmetric and positive semi definite. So, what is the meaning of positive semi definite here that all the eigenvalues of A transpose A are non-negative means greater than equals to 0.

So, now the eigenvalues of A T A, let us suppose these are lambda 1 which is the biggest 1 lambda 2 lambda n and these are in decreasing order, all are non negative. Now, define sigma h is square root of lambda i, all are non negative. So, a square root is well defined and more over if sigma 1 is square root of lambda 1 sigma 2 is a square root of lambda 2 and so on.

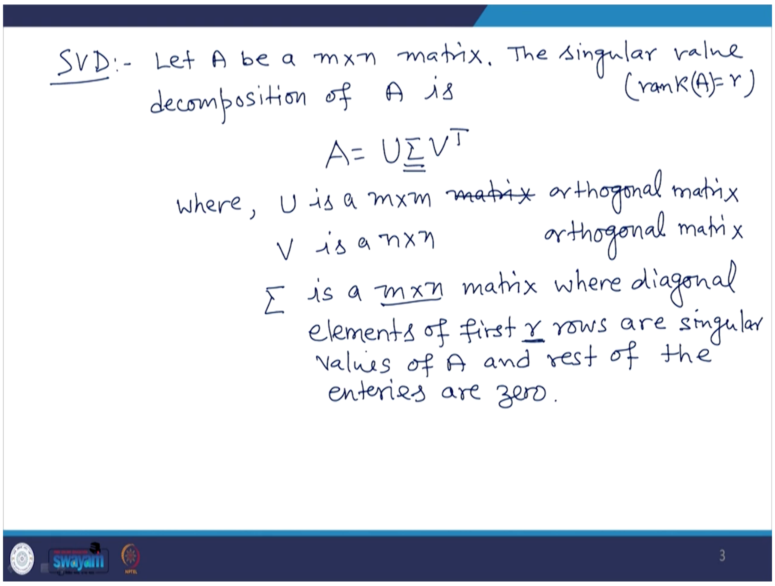
So, sigma 1 will be greater than equals to sigma 2 greater than equals to sigma 3 and so on and all those sigma i's will be non-negative. The number sigma 1, sigma 2, sigma n are called singular values of A.

(Refer Slide Time: 04:16)



Now, come to the Singular Value Decomposition in short I will write it as SVD.

(Refer Slide Time: 04:20)

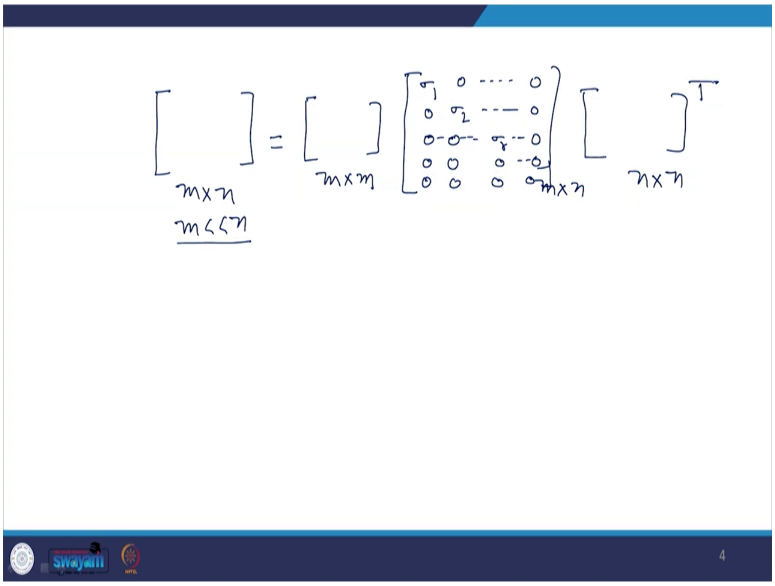


So, let A be a m by n matrix. Here in machine learning usually we play with real data. So, here we are assuming that entries are real numbers. The singular value decomposition of A is A equals to U sigma V transpose; where, U is a m by m matrix or better to write m by m orthogonal matrix such that U transpose equals to U inverse or columns of U are pair wise orthonormal. V is a n by n orthogonal matrix.

Now, what about this sigma? Sigma is a m by n matrix, where diagonal elements of first r rows are singular values of A and rest of the entries are 0. Why r we are taking? Let us assume that rank of A is r. So, here we are assuming that rank of A equals to r.

Now, it is a m by n matrix so, how we are talking about these diagonal elements and so on. So, let us see this concept in next slide.

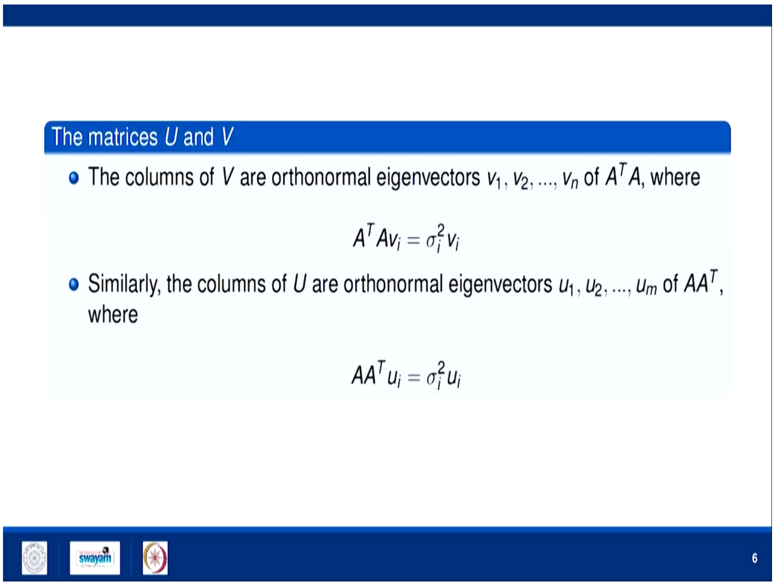
(Refer Slide Time: 07:31)



So, suppose A is m by n matrix. So, U will be m by m orthogonal matrix and sigma will be let us say m by n and here we are assuming that it is a (Refer Time: 7:59) matrix that is more number of columns than row. Then 0 sigma 2 0 0 0 sigma r 0 and rest of the entries will be 0 and then we will be n by n orthogonal matrix and singular value decomposition is U sigma V transpose.

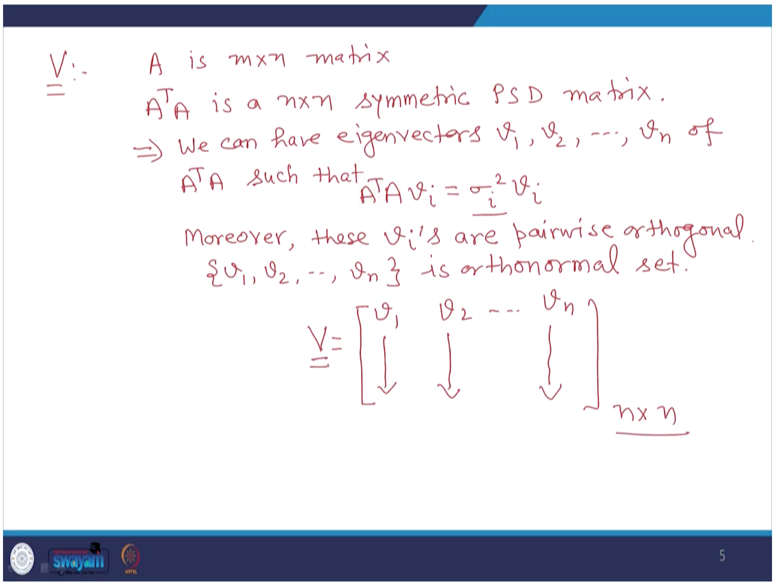
So, as I told you sigma is a m by n matrix where ith diagonal entries equal to i th singular value sigma i for i equals to 1 to r where rank of A equals to r all other entries of sigma are 0.

(Refer Slide Time: 08:46)



Earlier we have seen that how to find out singular values, but how to find out the matrices U and V that is again very important. So, the columns of V are orthonormal eigenvectors v 1, v 2, v n of A transpose A.

(Refer Slide Time: 09:10)

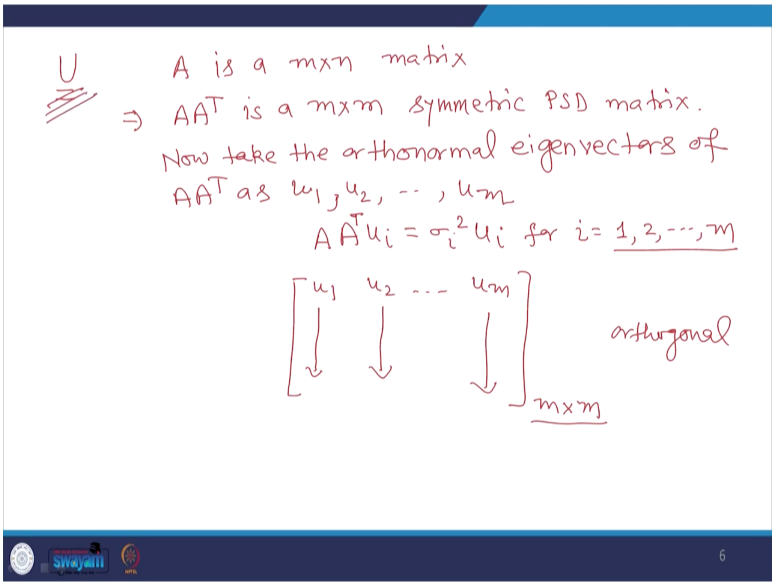


So, so, let us talk about matrix V. So, A is m by n matrix, then A T A is a n by n symmetric positive semi definite matrix.

Now, what we are having we can have eigenvectors v 1, v 2, v n of A transpose A such that A transpose A v i equals to sigma i square v i. Why I am writing this sigma i square because it is basically lambda i which is the eigenvalue of A transpose A moreover these v i's are pairwise orthogonal as well as the set v 1, v 2, v n is orthonrormal set. Because you if these are orthogonal you can make them orthonormal just by dividing the length of v 1, v 2, v n.

So, this is normal set. Hence what you do? You make the matrix V as so, v 1 is the eigenvector corresponding to sigma 1 square where sigma 1 is the largest eigenvalue. So, you write this v 1 column first. So, this is my v 1 column; similarly, second column will come from v 2 in that way the last column will be v n and in this way I will be having this n by n matrix and this will be the matrix v in singular value decomposition of A.

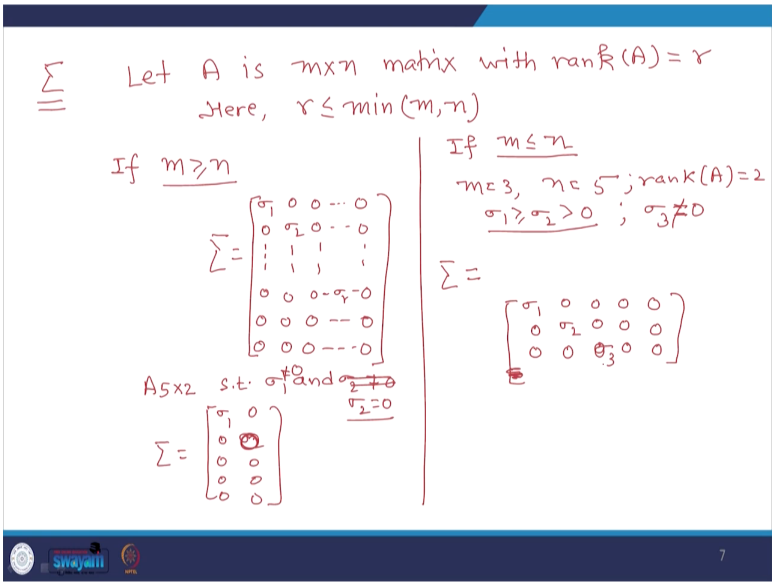
(Refer Slide Time: 12:01)



Now, let us talk about matrix U. So, again A is a m by n matrix then talk about AA transpose is a m by m symmetric positive semi definite matrix. Now, take the orthonormal eigenvectors of A into A transpose. As let us say u 1, u 2, u m. So, u 1, u 2, u m are the orthonormal eigenvectors of A and we are having A u i equals to sigma i square sorry, not A, A into A transpose.

So, now, A into A transpose u i equals to sigma i square u i for i equals to 1, 2, m more over u 1, u 2, u m are orthonormal's. So, what you do? You make a matrix where first column is u 1, second column is u 2 and last column is let us say u m. So, it will be a m by m matrix which is orthogonal and this matrix will be matrix U in the singular value decomposition of A.

(Refer Slide Time: 14:21)



Now, how to write sigma? So, let A is m by n matrix with rank of A equals to let us say r. So, here r will be less than equals to minimum of m, n.

So if m is bigger than n, so, my sigma matrix will be a skew matrix which is having more number of rows. So, I will be having sigma 1 0 0 0. So, these are m rows n columns 0 sigma 2 0 0 0. So, in that way I will be having 0 0 0 sigma r 0. So, these are r rows and these are n columns and then 0 0 0 0 0 0 0 0.

So, for example, let us take A as 5 by 2 matrix such that both will sigma 1 and sigma 2 are non 0 means rank of A is 2. So, what will be sigma in this case? Sigma 1 0 0 sigma 2 0 0 0 0 0 0. So, this will be my matrix sigma. If sigma 1 is nonzero and sigma 2 equals to 0, in this case this also will become 0. So, in that way I can write my matrix sigma.

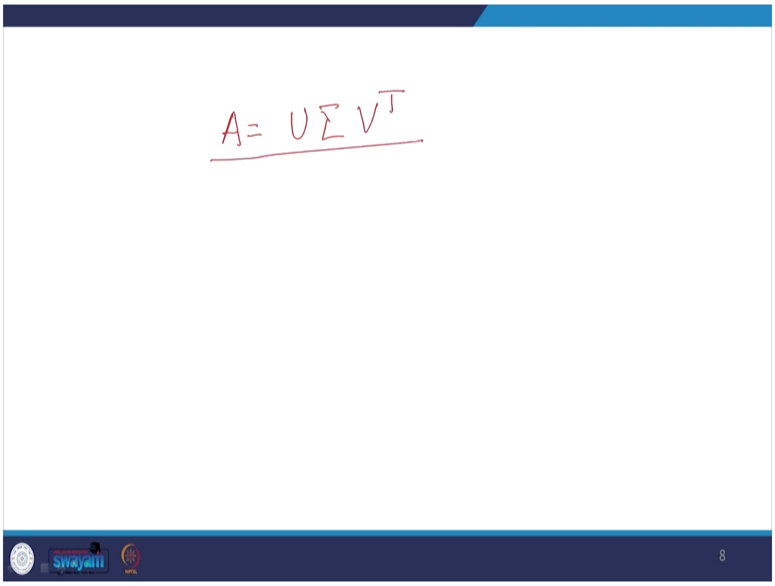
Similarly, if m is less than n means you are having more number of columns in a when compared to the number of rows. In this case, we will be having a (Refer Time: 16:45) matrix sigma means more number of columns when compared to the number of rows. So, for example, if you take m equals to 3 and n equals to 5. So, number of rows are 3 and number of columns are 5 and rank of A is let us say 2.

So, it means sigma 1 greater than sigma 2 and they will be strictly greater than 0 because rank is 2. So, 2 of the singular values will be nonzero and sigma 3 equals to 0. So, in this case sigma can be given like this. So, sigma 1 0 0 0 0, 5 columns. So, dimension of sigma is same as the dimension of matrix A; then 0 sigma 2 0 0 0 0 0.

Here if sigma 3 is nonzero it will come sigma 3, since sigma 3 we are taking 0 so, 0 0 0. This will become my matrix sigma if sigma 3 also nonzero, in this case I will be having here sigma 3. So, in that way by knowing the singular values of A which is coming from the eigenvalues of A A transpose or A transpose A I can write my matrix sigma.

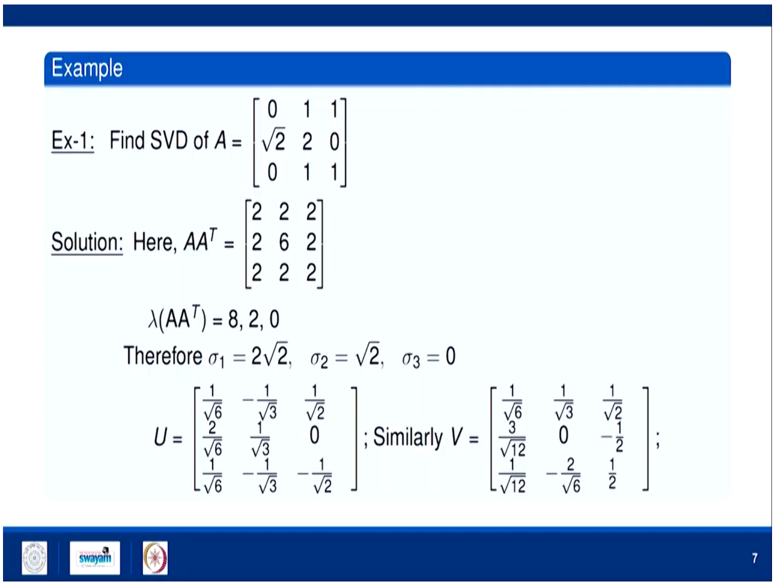
So, V is coming from the orthogonal eigenvectors of A transpose A, the matrix U is coming from the orthogonal eigenvectors of A into A transpose and sigma that is the singular matrix we are writing from the eigenvalues of A into A transpose or A transpose into A.

(Refer Slide Time: 18:54)



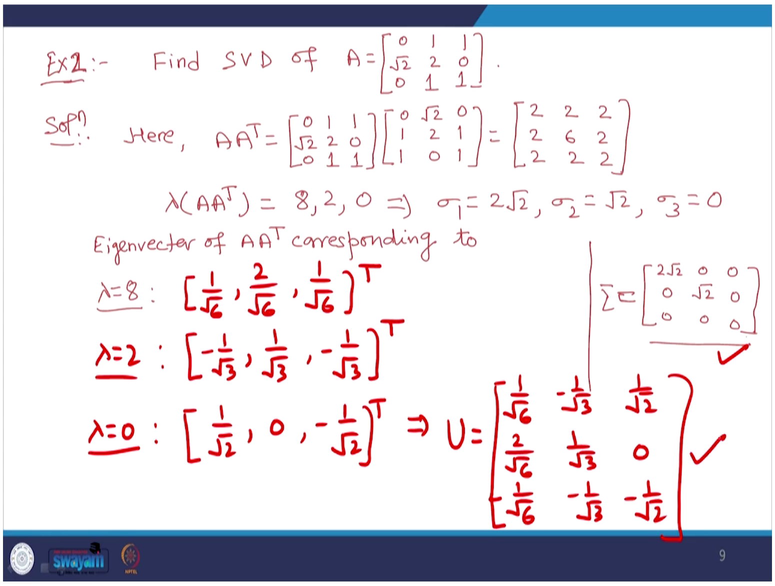
And, in that way the singular value decomposition of A is U sigma V transpose.

(Refer Slide Time: 19:00)



So, let us take couple of example for finding the singular values. So, first example I will take for a square matrix and the second example I will take for a rectangular matrix.

(Refer Slide Time: 19:14)



So, example 2 of this lecture. So, find singular value decomposition of A equals to 0 1 1 root 2 2 0 and 0 1 1. So, here first I calculate A into A transpose. So, A into A transpose comes out to be 0 1 1 root 2 2 0, 0 1 1 into A transpose. So, A transpose will become 0 1 1 root 2 2 0, 0 1 1.

So this will be 2 and then it will become root 2 again 2 2, then this will become 2 6 2, 2 2 2. Now, if we calculate the eigenvalues of A into A transpose so, I am not going to compute it I am writing directly. So, eigenvalues comes out be 8, 2, 0. So, from here what we can write that sigma 1 equals to a square root 8 that is 2 root 2 sigma 2 equals to root 2 and sigma 3 equals to 0.

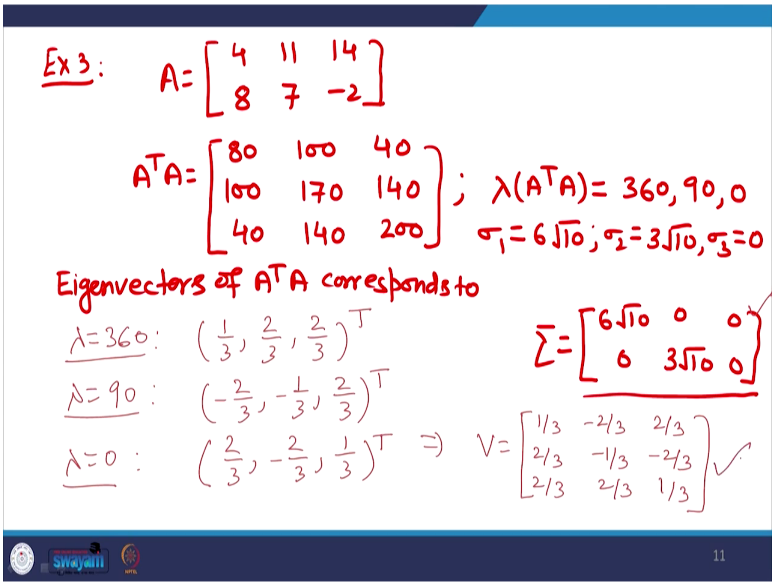
So, these three are the singular values of A and here rank of A also I can directly tell that 2 of the singular values are nonzero. So, rank of A is 2. Now, so, here I can write briefly what is my matrix sigma. So, sigma equals to 2 root 2 0 0 0 root 2 0 0 0 0. So, this is the matrix sigma. Now, I will calculate the eigenvectors of A into A transpose.

So, eigenvector of A into A transpose corresponding to lambda equals to 8 comes out to be 1 by root 6, 2 by root 6 and 1 by root 6 transpose. So, corresponding to lambda equals to 2 it is minus 1 upon root 3, 1 upon root 3 and minus 1 upon root 3. Similarly, corresponding to lambda equals to 0 it is 1 by root 2, 0, minus 1 by root 2.

So, you can easily verify that all these three eigenvectors corresponding to lambda equals to 8, lambda equals to 2 and lambda equals to 0 are mutually orthogonal and they are orthonormal also. So, from here I can write my matrix U. So, the first column of U will be 1 upon root 6 that is this eigenvector 2 upon root 6 and 1 upon root 6. The second column will be minus 1 upon root 3, 1 upon root 3 and minus 1 upon root 3; the third column will be 1 upon root 2, 0, minus 1 upon root 2.

So, I am having sigma, I am having U, now I need to calculate the matrix V.

(Refer Slide Time: 24:31)

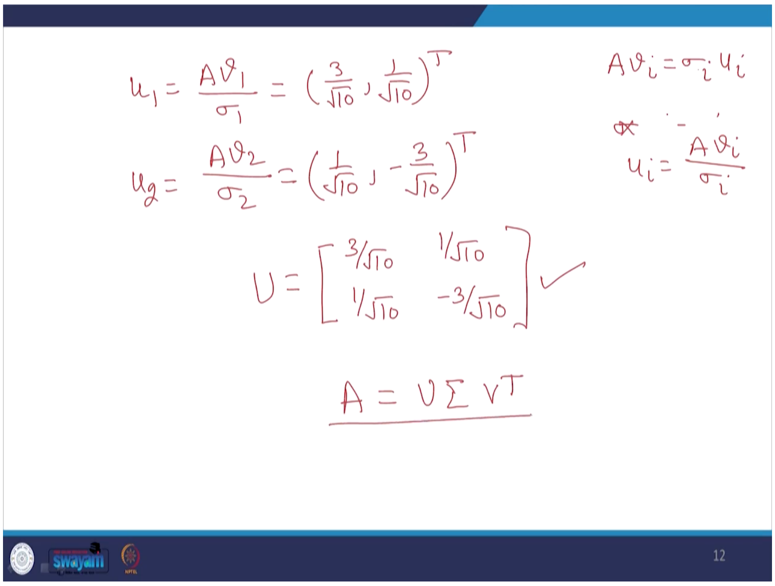


So, example 3 A is a matrix 4 8 11 7 and 14 minus 2. So, A transpose A will be 80 100 40 100 170 140 40 140 and 200. Now, eigenvalues of A transpose A comes out to be 360, 90 and 0 which gives the singular value of a sigma 1 is root 360 that is 6 root 10; sigma 2 is 3 root 10 that is your 90 and sigma 3 equals to 0.

So, hence my matrix sigma is 6 root 10 0 0 3 root 10 and then 0 0, this is sigma. Now, I will calculate the eigenvectors of A transpose A corresponds to lambda equals to 360. This comes out to be 1 by 3, 2 by 3, 2 by 3 transpose. Then lambda equals to 90 for this eigenvector is minus 2 by 3, minus 1 by 3 and 2 by 3 transpose and for lambda equals to 0 we are having eigenvector as 2 by 3, minus 2 by 3, 1 by 3 transpose. So, all these three are orthonormal vectors and pair wise orthogonal.

So, from these I can write my matrix V as 1 by 3, 2 by 3, 2 by 3, minus 2 by 3, minus 1 by 3, 2 by 3, 2 by 3, minus 2 by 3 and 1 by 3. Now how to calculate U?

(Refer Slide Time: 27:45)



So, here u 1 will become A into v 1 upon sigma 1 which comes out to be. So, instead of calculating the usual way I am doing this one because from the singular value decomposition you are having A v i equals to sigma i u i.

So, from here I can write u i equals to A v i upon sigma i. So, u 1 will be A v 1 upon sigma 1 which comes out to be 3 by root 10 and 1 upon root 10 transpose. Similarly, I am having u 2 which becomes A v 2 upon sigma 2 which is 1 upon root 10 minus 3 upon root 10 transpose.

So, from here my matrix U becomes 3 by root 10 1 upon root 10 1 upon root 10 minus 3 upon root 10 and the singular value decomposition of A is U sigma V transpose, where U is given by this sigma is here and V is here.

So, in that way we have seen two examples where in the first example we have done singular value decomposition of a square matrix, while in the second example we have seen the singular value decomposition of a rectangular matrix. In the next lecture, we will see certain applications and properties of singular values and how to use singular value decomposition for computing different or solving different problems. So, I hope you have enjoyed this lecture.

Thank you very much.