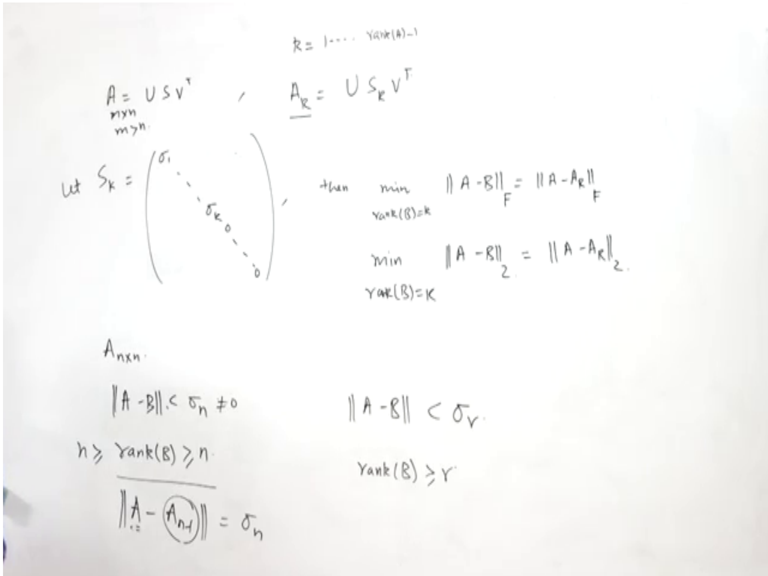
Numerical Linear Algebra
Dr. D. N. Pandey
Department of Mathematics
Indian Institute of Technology, Roorkee

Lecture - 48
Outer product expansion of a matrix

Welcome friends. In previous lecture we have seen that minimum of this norm A minus B, where B is a matrix of rank k is equal to norm of A minus A k.

(Refer Slide Time: 00:24)

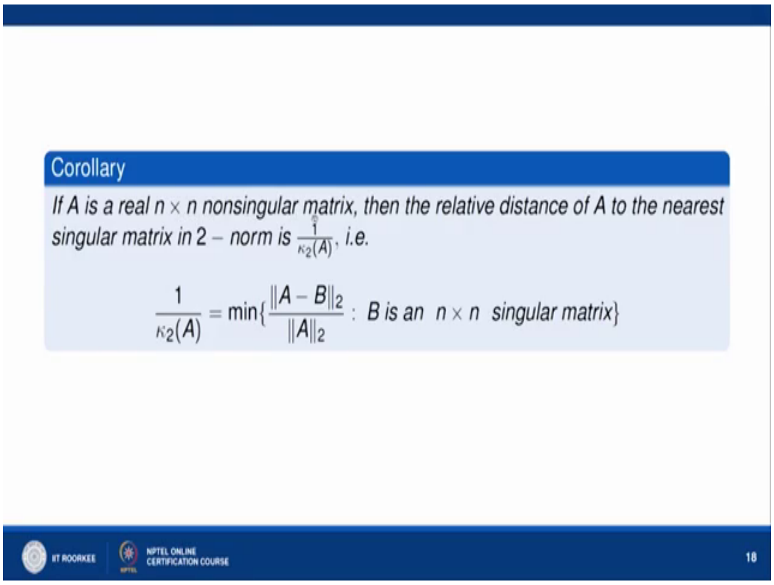


Now, if we apply a if we write Frobenius norm then this is the result given by Eckart and Young and if we replace this A minus B this Frobenius norm by 2 norm then we have also seen the same result that this same result also valid in 2 norm also.

So, it means that 2 norm of A minus B if we take the minimum over all the matrices of the rank k then it is achieved at A k then it means that that A k is the matrix at which we achieve the minimum of this quantity. It means that minimum of A minus B, where B is of a matrix of rank k whether it is 2 norm or f norm it is achieved at A k. So, A k approximate the matrix A and this is a closest approximation of the matrix A.

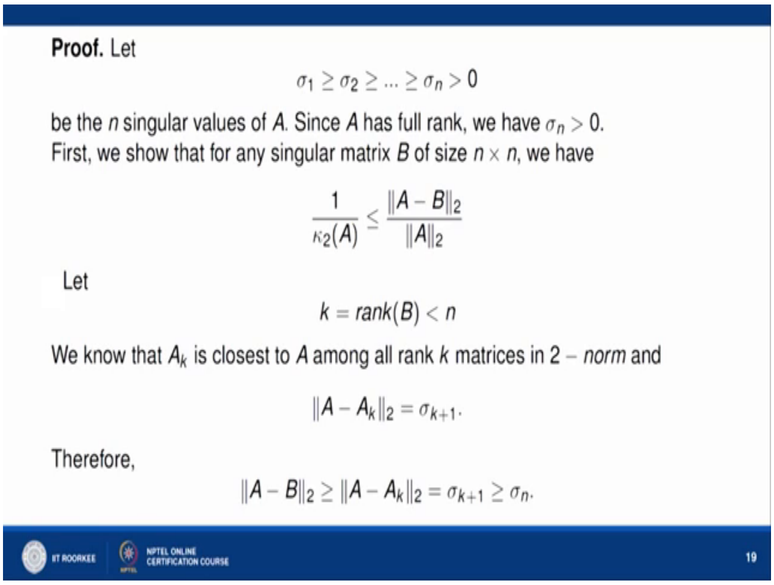
And we have also seen the corollary based on this that if we consider that rank of A minus B if we say that it is less than sigma r then your rank of B must be greater than or equal to r. So, this is the a last corollary we have proved and we also consider one more corollary based on this result and it is state like this that if A is a real n cross n nonsingular matrix then the relative distance of a to the nearest singular matrix in 2 norm A is 1 upon kappa A, where this kappa 2 A represent the condition number of the matrix A in 2 norm.

(Refer Slide Time: 02:14)



So, here we can say that 1 upon kappa 2 A is minimum of 2 norm of A minus B divided by 2 norm of A where B is an n n cross n singular matrix. So, we say that the distance minimum distance between this A minus B is 1 upon kappa 2 A. So, it means that A is nonsingular matrix and B is a singular matrix. So, minimum distance between is nonsingular matrix and singular matrix is given by 1 upon kappa 2 A.

(Refer Slide Time: 03:06)

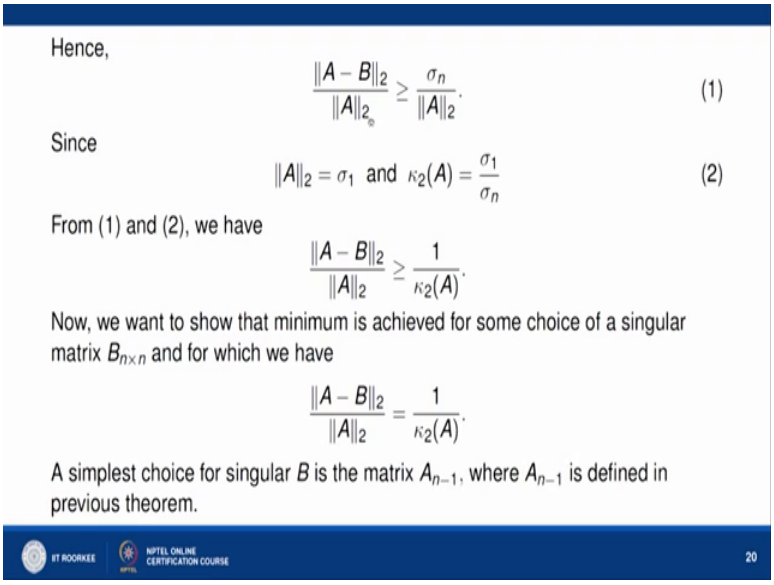


So, let us prove this important corollary and it says that say let sigma 1 to sigma and arranged in a decreasing in decreasing order and since A is nonsingular matrix so sigma n is going to be 0.

So, we say that since a has full rank we have sigma and greater than or equal to 0 and first we show that for any singular matrix B of size n cross n we have 1 upon kappa 2 A less than or equal to 2 norm of A minus B divided by 2 norm of A. And then we try to show that this minimum is achieved at some matrix B here.

So, let us say that rank of B is a less than n and let us say that it is k and we already know that A k is a closest to a among all rank k matrices in 2 norm or Frobenius norm and 2 norm of A minus A k is given by sigma k plus 1 and also we know that 2 norm of A minus B is bigger than or equal to 2 norm of A minus A k which is nothing, but sigma k plus 1 and since we have arranged in this order then sigma k plus 1 is going to be the greater than or equal to sigma and so A minus 2 norm of A minus B is greater than or equal to sigma n.

(Refer Slide Time: 04:29)

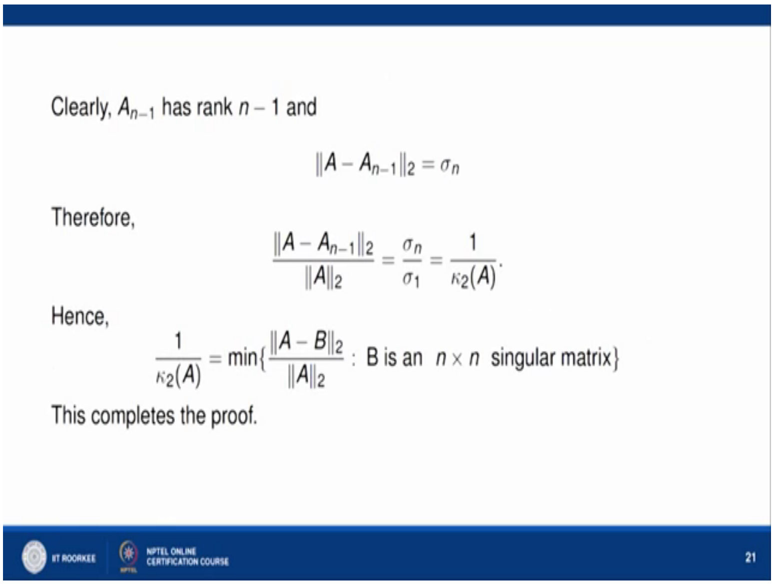


Now, if we divide by 2 norm of A then it is what 2 norm of A minus B is greater than equal to sigma n divided by 2 norm of A. So, we divide both side by 2 norm of A and we already know that 2 norm of A is positive. So, it will not disturb the in equality. So, now, if we recall the result that 2 norm of A is nothing, but the largest singular value of the matrix A then it is given by sigma 1. So, 2 norm of A is sigma 1 and we also know that the condition number of A in the 2 norm of 2 matrix norm is given by sigma 1 divided by sigma n.

So, using this information we can say that that norm a 2 norm of A minus B divided by 2 norm of A is equal to sigma n by sigma 1 and which is given by 1 upon kappa 2 A. So, we what we have shown here that the lower bound of 2 norm of A minus B divided by 2 norm of A is 1 upon kappa 2 A.

Now, we want to show that for some matrix B we can achieve this lower bound for and the simplest choice for this matrix B which is a singular matrix A is A min A n minus 1. So, it means that to show that minimum is achieved for some choice of a singular matrix B and for which we want that this should be equality. So, 2 norm of A minus B divided by 2 norm of A is equal to 1 upon kappa 2 A and we try to show that the simplest ways for singular matrix P is the matrix A n minus 1. So, A n minus 1 is a singular matrix and A n minus 1 is defined in previous theorem. So, here A n minus 1 is what? U sigma U S n minus 1 and V transpose.

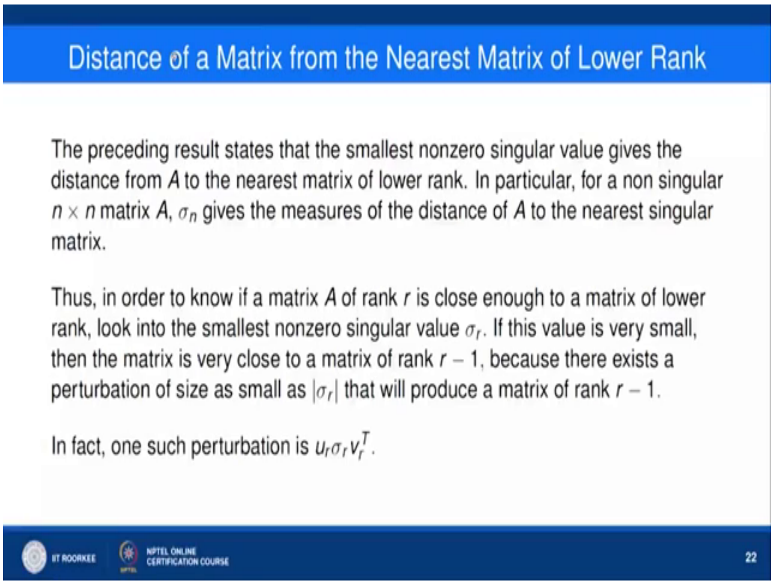
(Refer Slide Time: 06:29)



So, we know that clearly A n minus 1 has rank n minus 1 that is what we have proved earlier and the 2 norm of A A minus A n minus 1 is sigma n. So, it means that 2 norm of A minus A n minus 1 divided by 2 norm of A is nothing, but sigma n divided by sigma 1 and sigma n divided by sigma 1 is nothing, but 1 upon kappa 2 A. So, for the math for the choice B equal to A n minus 1 we have the equality.

So, it means we can say that 1 upon kappa 2 A is the minimum 2 norm minimum of 2 norm of A minus B divided by 2 norm of A and here B is any n cross n singular matrix and A is going to be a nonsingular matrix. So, it means that the nearest distance from and between is nonsingular matrix and a singular matrix is given by the 1 upon 1 upon the condition number of the matrix A. So, it means that if the condition number is very very large then your nonsingular matrix is very near to the singular matrix and for singular matrix A your condition number is infinite that we have already seen in previous lectures.

(Refer Slide Time: 07:57)



So, now, with the help of this we want to find out say distance of a matrix from the nearest matrix of lower rank. So, the proceeding results state that the smallest nonzero singular value gives the distance from A to the nearest matrix of lower rank. So, it means that if we have the nonzero singular the smallest nonzero singular value is very very small then it is very very close to the singular matrix.

So, in particular for a nonsingular matrix n cross n matrix A sigma n gives the measures of the distance of A to the nearest signal matrix. So, it means that it means that if the distance between A minus B, where B is a nearest singular matrix A, A must be given by 1 upon kappa 2 A that we have already seen.

So, here we want to show that in particular for A nonsingular n cross n matrix A sigma n gives the measure of the distance of A to the nearest singular matrix in particular. If you look at this corollary of this theorems that norm of A minus B if it is less than sigma r then rank of B is going to be greater than or equal to r. So, if I take this r as n and a as a nonsingular matrix then rank of A minus B if it is less than sigma n and sigma n is of course, nonzero in fact, it is positive value then such a matrix B must have rank greater than or equal to n and since we are considering the cases of square matrices then this rank of B greater than or equal to n means that B is a nonsingular matrix.

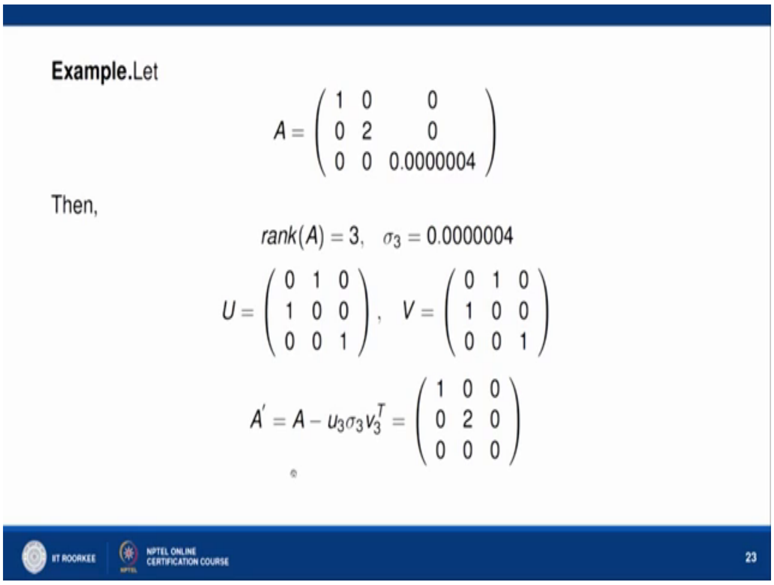
So, it means that the distance between 2 nonsingular matrixes is given by this sigma n. So, it means that if we are searching for a singular matrix then the distance between A minus B distance between A and B has to be bigger than sigma n. So, it means that if the distance is less than sigma n then we will get only nonsingular matrices. So, it means the minimum distance we are talking about a nonsingular matrix and a singular matrix is at least sigma n right. And we have seen that the choice is A minus A n minus 1. So, if you look at a is a nonsingular matrix and we have a matrix A n minus 1 which we have defined in previous theorem then this is a singular matrix and the norm of A minus A n minus 1 is coming out to be sigma n that we have already show seen.

So, it means that the distance between a nonsingular matrix and a singular matrix must be exactly equal to the smallest singular values of a matrix A. So, it means that if your sigma n is very very small then the distance between A and a nonsingular matrix and a singular matrix is very very small. So, or we can say that in case that the smallest singular values is very very small then your nonsingular matrix is very very near to a singular matrix.

So, that a clearly indicate that the smallest singular value gives the measures of the distance of A to the nearest singular matrix. Thus in order to know that if a matrix A of rank r is close enough to a matrix of lower rank look into the smallest nonzero singular value sigma r. If this value is very very small then the matrix is very very close to a matrix of rank r minus 1 because they exist a perturbation of size as small as sigma r that will produce a matrix of rank r minus 1.

Now, we can say that the exact perturbation we can say that it is nothing, but u r sigma r v r transpose where sigma r is the smallest singular value of this matrix A. So, if you perturb your matrix A by u r sigma r v r transpose then we can get a singular a matrix of rank r minus 1.

(Refer Slide Time: 12:29)

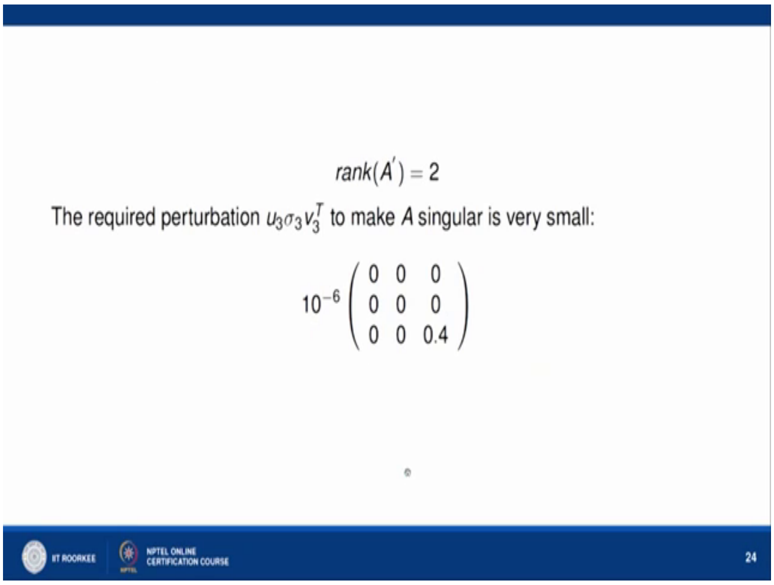


So, here you can say that let us consider this a as 1 0 0, 0 2 0, 0 0 and this quantity 0.0000004 it is 4 into 10 to power minus 6.

Then we see that it is rank of a is 3 and if you look at the lowest singular value is 4 into 10 to power minus 6. And here we can find out the singular value decomposition of a as U and V and we can calculate the A dash or you can say that it is nothing, but A 2 in if you look at the previous theorem notation then A 2 is given by A dash is given by A minus U 3 sigma 3, V 3 transpose and if you calculate it is coming out to be 1 0 0, 2 0 0, 0 0 0.

So, it means that A dash is a matrix of rank 2 and if you look at the initial matrix which we started with is a matrix of rank 3 or you can say that this is a nonsingular matrix, but if you perturb your matrix A by this is a simple perturbation and look we will see that what is a mod norm of this perturbation.

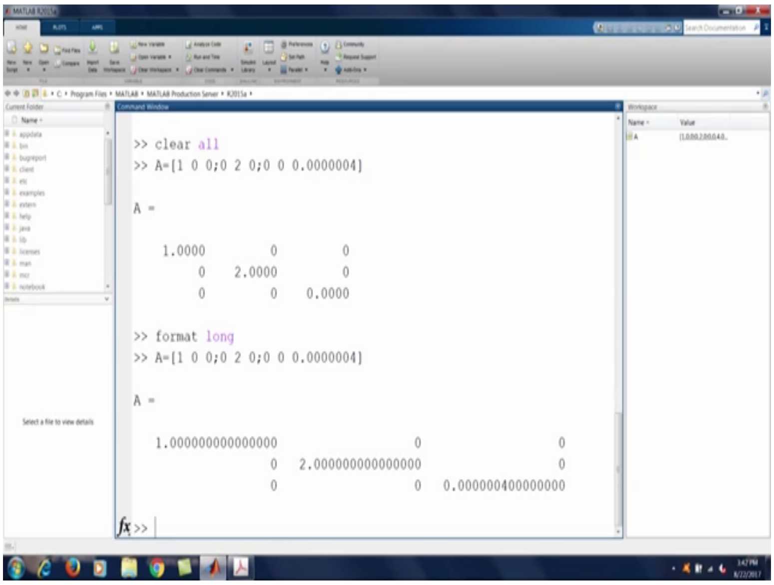
(Refer Slide Time: 14:02)



So, we get a matrix which is a singular matrix and rank is exactly less than 1 and here rank of A dash is 2 and the required perturbation U 3 sigma 3 V 3 transpose to make a singular is very small and it is nothing, but 10 to power minus 6 into this quantity. So, we can say that by knowing the smallest singular value we can always find out a small a small perturbation to the matrix A and the perturb matrix is a singular matrix and that perturbation you can write it like this A minus u n u r sigma u r transpose.

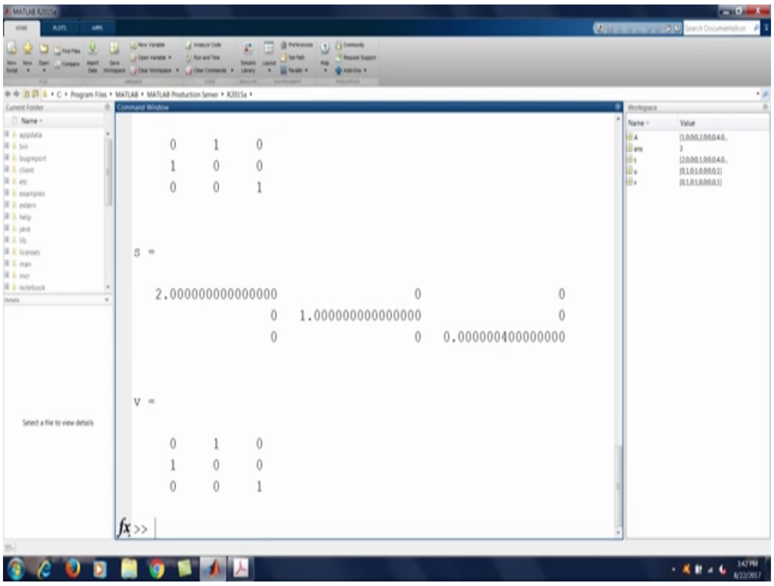
So, by this small, it means that that if you want to find out a matrix given a matrix A if you want to find out a another matrix having the rank r minus 1 we perturbed our matrix A by this matrix u r sigma r v r transpose and by writing A minus u r sigma r v r transpose we get a new matrix having rank exactly less than 1. So, that we have shown here we can easily verify in this using MATLAB.

(Refer Slide Time: 15:22)



So, if you look at you will look at a here. So, let us say that B since we have already used A or let us say that clear all. So, this will clear the matrix A. Now, define a as the matrix if you look at what is the matrix here 1 0 0. So, 1 0 0, 0 2 0, comma 0 0 0 and 0.0000004 and this is the matrix A which we are writing here, let me it is not showing here. So, let me write it format long and then write A as this. So, it is now, showing here.

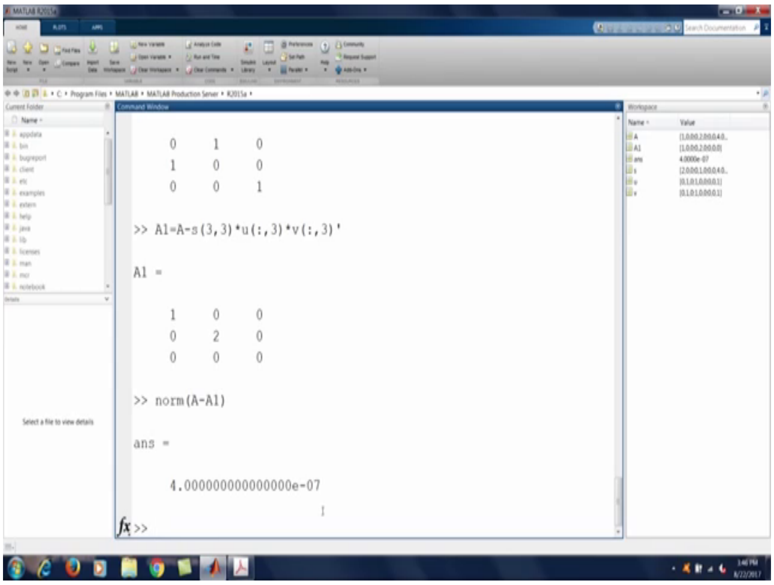
(Refer Slide Time: 16:15)



Now, we know that rank of A is what. Rank of A you can verify then it is given by 3 and we can always find out same SVD of this. So, USV is equal to SVD of this matrix A, here we have A. So, we can say that S is this V is this and this. So, that is what is given here U is given here V is given here.

Now, we can look at the smallest singular values that is given by this. And corresponding to this we can calculate U 3 and V 3 here.

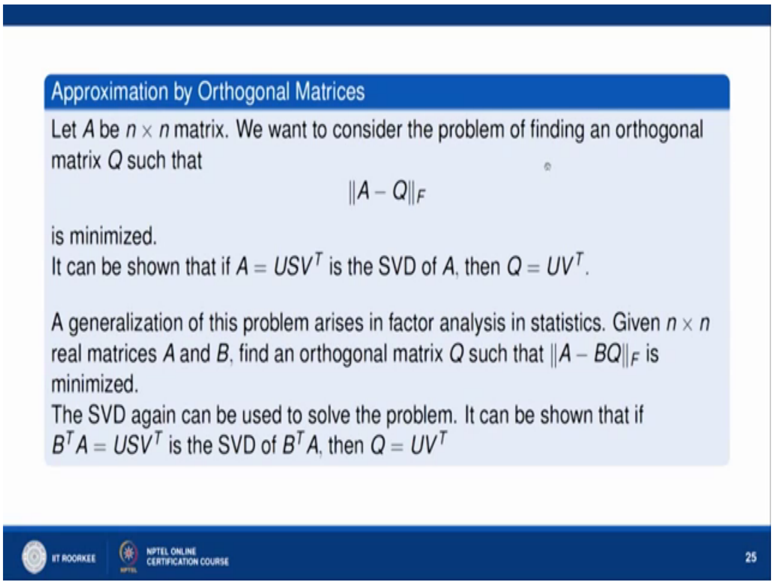
(Refer Slide Time: 17:00)



So, this is the singular value decomposition of A. Now, let us calculate a one which is A minus S 3 comma 3 which represent the this singular value of matrix A the lowest singular value of matrix A into third column of U and to a third column of V transpose. And if you calculate it is coming out to be point 0 0 0 2 0 0 0 0. if you look at this is what we have seen here that A dash which is given as A minus U 3 sigma 3 V 3 transpose and it is given by this.

And we can easily check that rank of A 1 is 2 and we can also check the norm of this quantity that is norm of here we can write it that A minus A 1 is what and it is coming out to be 4 into 10 to power minus 7 here. So, it is basically it is it written as point 4 into 10 to power minus 6. So, you can say that this perturbation is quite a small and we are able to find out a h a matrix which is singular matrix and it is very very near to your nonsingular matrix. So, that is what we have seen here.

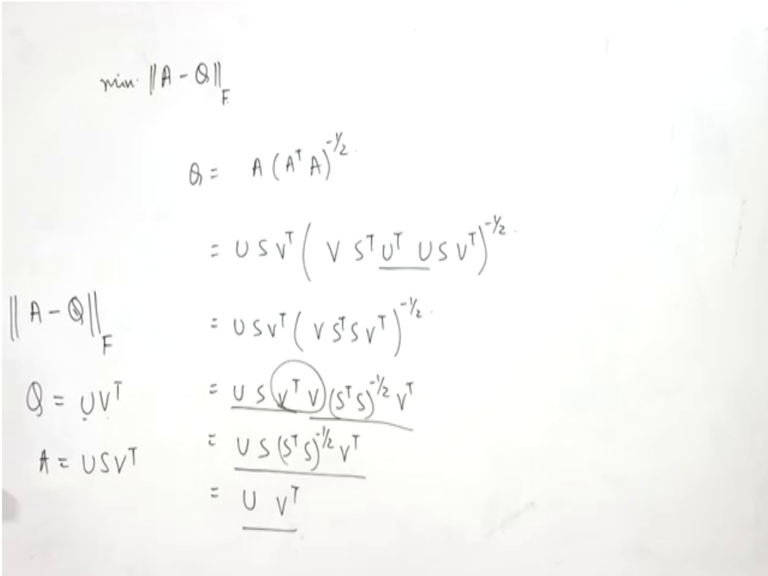
(Refer Slide Time: 18:24)



So, now, moving on further let us find out the approximation by an ortho orthogonal matrix. So, A is a any matrix of size n cross n and we want to consider the problem of finding an orthogonal matrix Q such that norm of a minus Q in the sense of Frobenius norm is minimized. And we can check that this can be given by this can be find out using a similar value decomposition of matrix A that if A is written as USV transpose is the SVD of A then you can take this Q as UV transpose.

In fact, the general form of this Q is given by A into A into A transpose A to power minus half and if you use the singular value decomposition of A matrix A then the representation of Q is given by UV transpose. I am just going to write that statement.

(Refer Slide Time: 19:32)



So, the problem is that we want to minimize the Frobenius norm of A minus Q where Q is an orthogonal matrix, then the here I am stating the form of Q without giving the proof of this result that the form Q is given by A, A transpose A to power minus half. Now, here if I use the singular value decomposition of a matrix A then if A it can be seen that us V transpose where UV are orthogonal matrices then A transpose is given as V S transpose U transpose and A as USV transpose to power minus half.

And if you simplify this since U is orthogonal matrix then U transpose U is going to be identic then this nothing, but V S transpose S V S transpose. And then if we write this as U S V transpose at it is as it is and this we can say that V S transpose S to minus half V transpose then we can say that we have this U S, S transpose S to power minus half here I am using the orthogonality of matrix A V then we can write it that it is nothing, but U S S transpose S to power minus half V transpose.

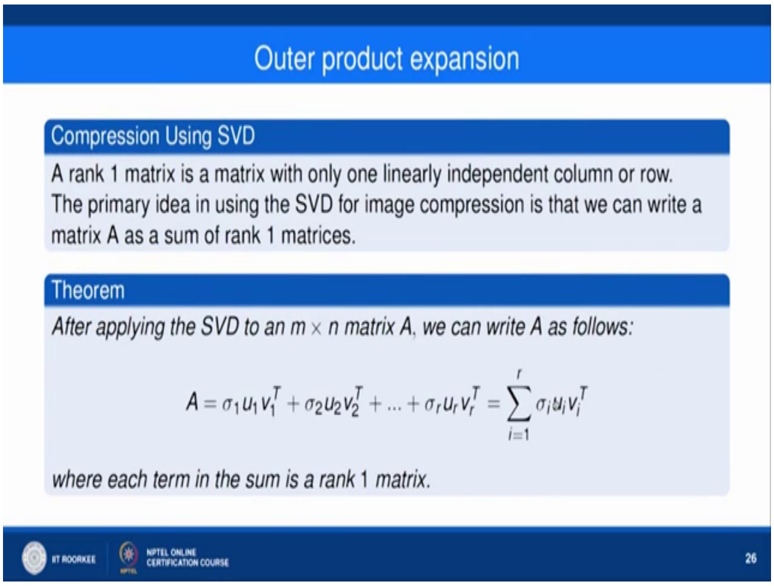
And if you simplify it is coming out to be UV transpose. So, it means that if you want to minimize or if you want to approximate this a by n orthogonal matrix then the in Frobenius norm then the representation of Q is given by UV transpose where U and V are orthogonal matrices such that a can be written as U S V transpose.

Now, if you look at a if we drop the condition that it is an orthogonal matrix we already know that the approximation of A can be approximated with the help of A k and where k is from 1 to n. So, that we have suite, but here we want to approximate with the help of orthogonal matrices and it is given by this. So, this result we are not proving and a generalization of this problem is used in many problem in statistics for example, in factor analysis.

So, we there we want to consider the problem that we want to minimize given a matrix A and B we want to find out a matrix Q which is orthogonal and norm of Frobenius norm of A minus B Q is minimized. And here again using the method given above we can say that the SVD again can be used to solve this problem.

So, here we find out say singular value decomposition of B transpose A rather than singular value decomposition of matrix A. So, here you consider S SVD of B transpose A it is given as USV transpose and we once we have this U and V with us then we can consider Q as UV transpose and we can show that Q which is given as UV transpose minimize this quantity in Frobenius norm.

(Refer Slide Time: 22:44)



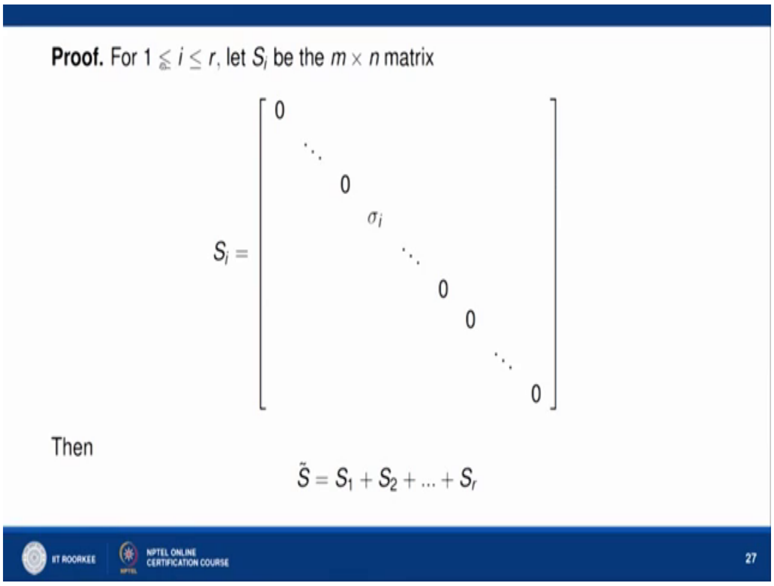
Now, now, let us consider one more very important application of singular value decomposition that is outer product expansion. So, this says that any matrix A of size m cross n can be written as sum of rank 1 matrices. Now, what is rank 1 matrices? So, a rank 1 matrix is a matrix with only one linearly independent column or row. And we utilize this content of this result content of this theorem in many application in particular we can use this idea in image compression.

So, given in any image how we can compress our image so that we are not losing more details. So, we can approximate a given image by a by another image using less data.

So, what is the theorem let us understand the theorem first and then we see that how we can utilize to use in image processing. So, the statement of this theorem is that after applying the singular value decomposition to an m cross n matrix A we can write a as a as sigma 1 u 1 v 1 transpose plus sigma 2 u 2 v 2 transpose to sigma r u r v r transpose. Here sigma 1 sigma 2 and sigma r are singular values of the matrix A and here we are using that rank of matrix A is r and this u i's are the columns of U and v i's are columns of V.

And this can be written as summation i equal to 1 to r sigma i u i v i transpose, where this each component sigma i u i v i transpose is a matrix of size m cross n and having rank 1 So, let us prove this theorem and we will see that how we can utilize this theorem and image processing.

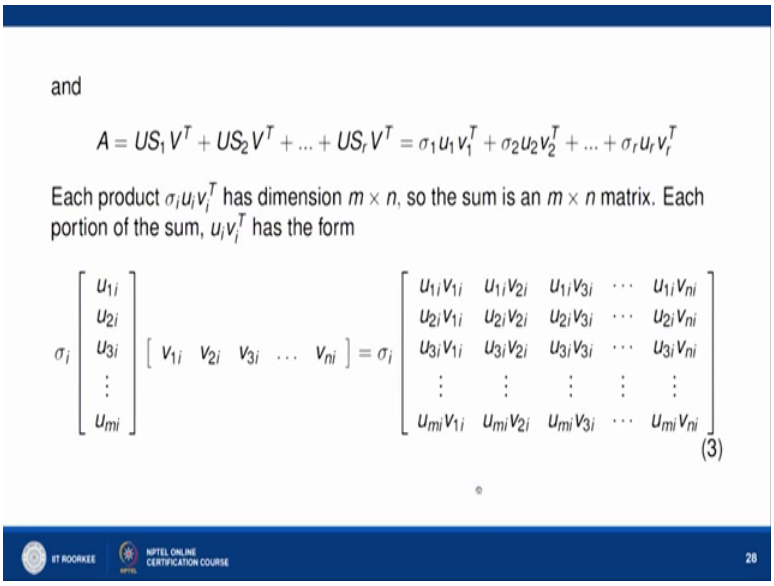
(Refer Slide Time: 24:59)



So, for proof let us say that i is from 1 to r where r is the rank of the matrix A and let S i be the m cross n matrix given by this the only thing is that the in the diagonal element ith i cross ith element is sigma I rest are all 0s. So, S i is given by this

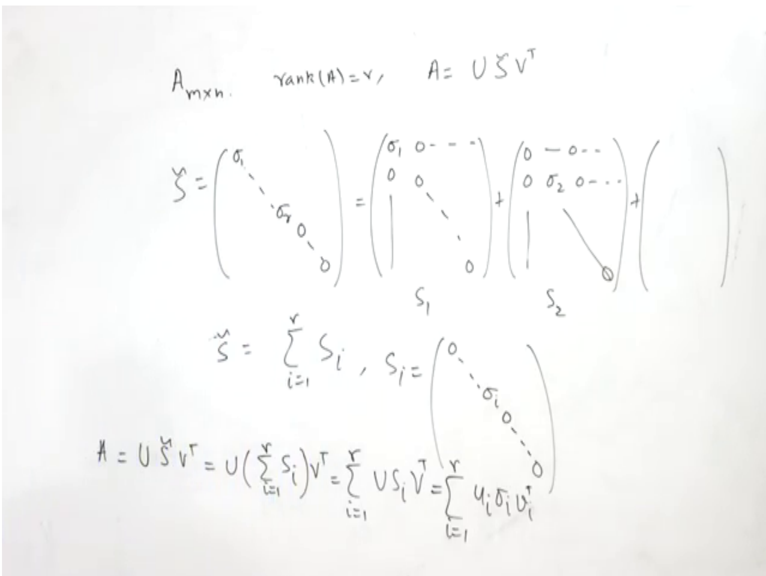
Now, with the help of this S i we can write S tilde as S 1 to S r, where S tilde is the singular valued we can obtain this tilde as singular value decomposition of A, that is A can be written as U S tilde V transpose and this S tilde can be written as S 1 to S r.

(Refer Slide Time: 25:43)



So, now, we can write A as U S tilde V transpose as U S 1 V transpose plus U S 2 V transpose up to us r V transpose and if you simplify we can get this. Just look at the proof here.

(Refer Slide Time: 26:03)



So, here we wanted to prove this theorem. So, A is a m cross n matrix, rank of A is equal to r and A is given by U S tilde V transpose. So, let us assume that the singular value decomposition of A is given as U S tilde V transpose, where S tilde is given by sigma 1 2 sigma r 0 0 all these are 0. And this can be decomposed in S i's, where S 1 is given by the first entries is sigma 1 rest all 0es here S 2 is the 2 element is sigma 2 and rest all 0's and similarly we can define. So, it means that S tilde can be written as i equal to 1 to r S i, where S i is given as that the i, ith diagonal entries sigma i rest of all 0.

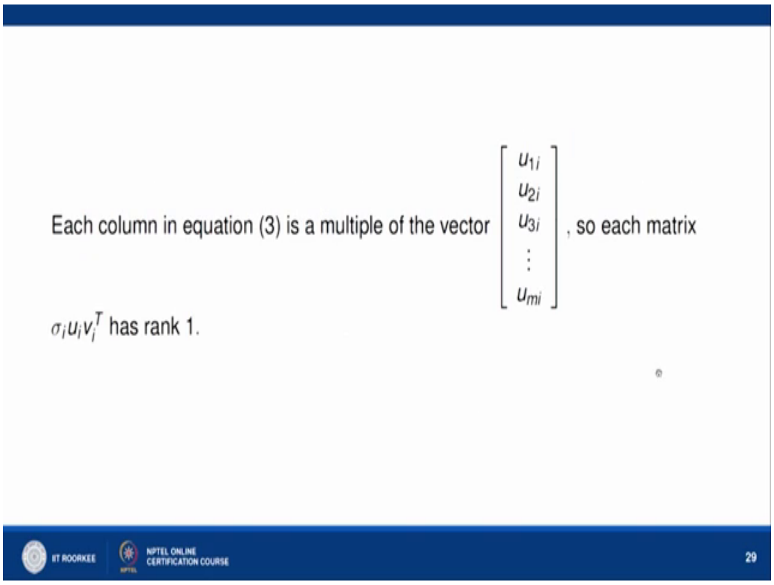
So, using this we can write U as U S tilde V transpose. Now, S tilde I am writing as i equal to 1 to r S i. So, it means given as U into I equal to one to r S i V V transpose. Now, this can be written as summation i equal to 1 to r US i V transpose. Now, calculate U S i V transpose then if you look at only this entry is nonzero. So, you can say that U S i V transpose is nothing, but summation i is equal to 1 to r it is u i sigma i v i transpose. So, we can write that a matrix can be written as summation i equal to 1 to r u i sigma i v i transpose here each u i sigma i v i transpose is a matrix of size m cross n, but having only rank 1.

So, let us see how it is. So, here we have seen that a can be written as us one v transpose plus us 2 v transpose till us r V transpose and if you simplify this is U S 1 V transpose is given by sigma 1 u 1 u 1 v 1 transpose and this is given as sigma 2 U 2 V transpose and so on. And each product sigma i u i v i transpose has dimension m cross n. So, the sum is an m cross n m cross n matrix which is what we wanted here.

Now, the good thing is that here each portion of the some u i v i transpose has this form. So, sigma i u i is this and v i transpose is this. So, u i is a can be written as u 1 u 1 i u 2 i u 3 i up to I. So, it is in a vector in rn rn cross 1 and it is the vector in r n cross 1.

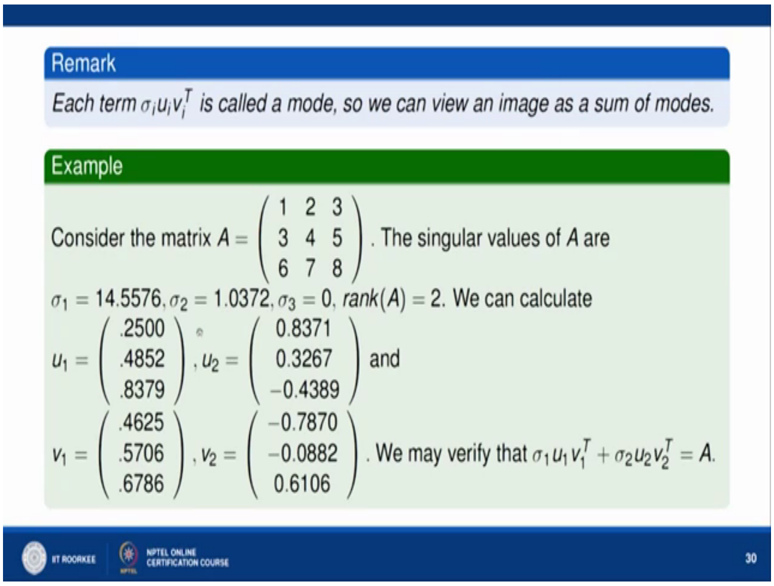
So, if you multiply this and you will get that it is nothing, but sigma I into this matrix. Now, if you look at this matrix carefully look at the first column the first column is what u 1 i, v 1 i, u 2 i, v 1 i, u 3 i, v 1 i and so on i v 1 i. So, I can say that the first column is a product of this by v 1 i. So, it means that first column is the product of the u u i by v 1 i, if you look at the second column then you can say that it is the product of the vector u i with v 2 i. So, you can say that precisely every column of this matrix is some scalar multiplication of u i matrix. So, it means that we have only one linearly independent column in this m cross n matrix.

(Refer Slide Time: 29:48)



So, we said that each column in equation 3 is a multiple of the vector this. So, we can say that each matrix sigma i u i transpose has a rank 1 and we can say that each term sigma i u i v i transpose is called a mode and we can say that the any matrix A can be written as sum of modes, where r represent the rank of this matrix A.

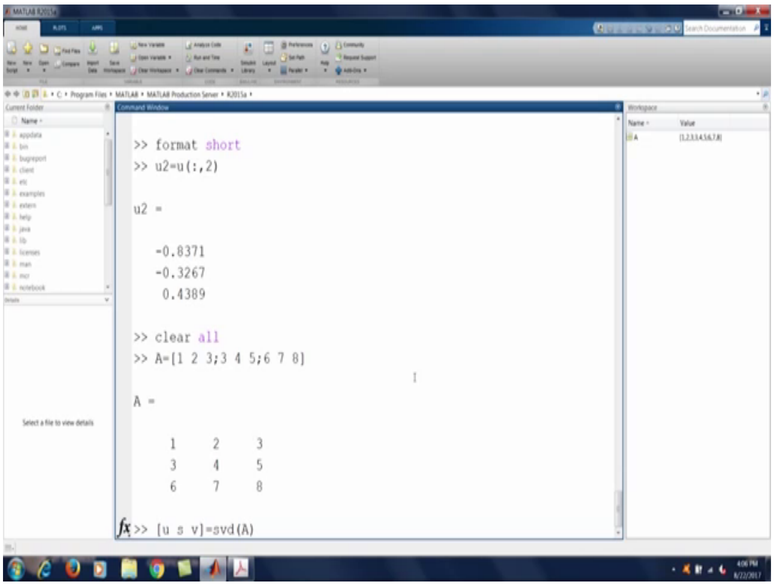
(Refer Slide Time: 30:00)



So, any matrix of rank r can be written as sum of r modes where each modes is given by sigma i u i v i transpose and having rank 1. So, let us verify the result. So, for that we consider a matrix A which is given as 1 2 3 3 4 5 6 7 8 and we can say that if we can calculate the singular values of matrix A as using MATLAB as sigma 1 as 14.5576 sigma 2 as this and we can say that sigma 3 is equal to 0. So, with this we can say that rank of a is 2.

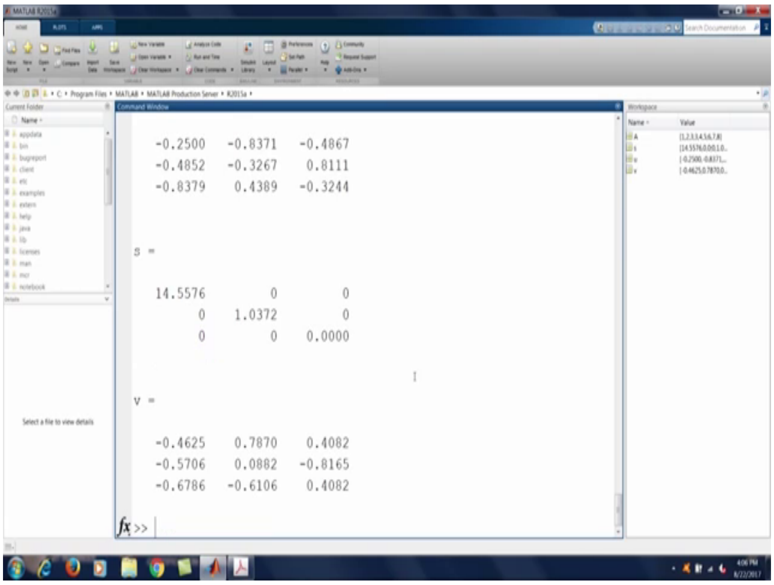
Now, we want to show that this matrix A can be written as this sum sigma 1 u 1 v 1 transpose plus sigma 2 u 2 v 2 transpose. So, we can calculate the first column of matrix U that is u 1 it is 0.2500, 0.4852, 0.8352; u 2 the second column of matrix U and v 1 first column of matrix V and v 2 as second column of matrix V and we can verify that sigma 1 u 1 v 1 transpose plus sigma 2 u 2 v 2 transpose is nothing but the matrix A here.

(Refer Slide Time: 31:44)



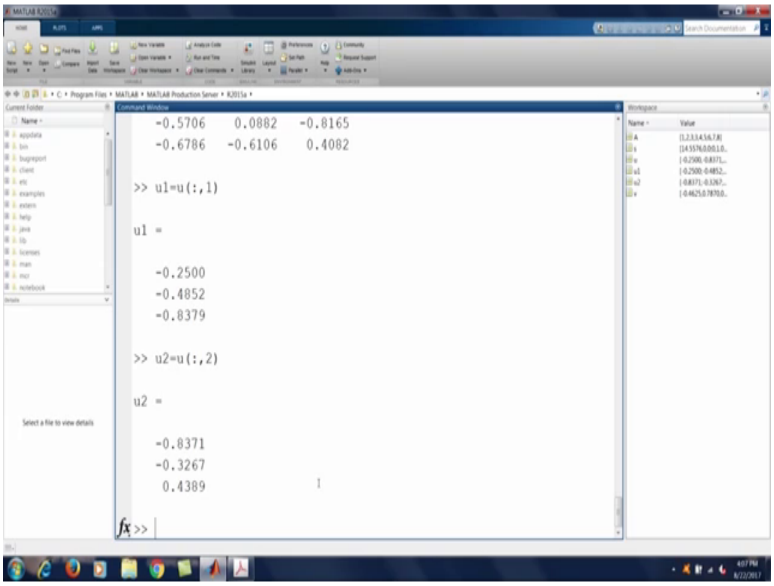
So, let me look at here. So, you define our matrix A as 1 2 and 3, 3 3 4 5 6 6 7 8. So, this is the matrix we are considering this is A. Now, let us consider its singular value decomposition as SVD of A and it is coming out to be this.

(Refer Slide Time: 32:13)



So, we can verify that a singular values are 14.5576, 1.0372 and the last one is 0. So, that is with clear with the agreement of this statement.

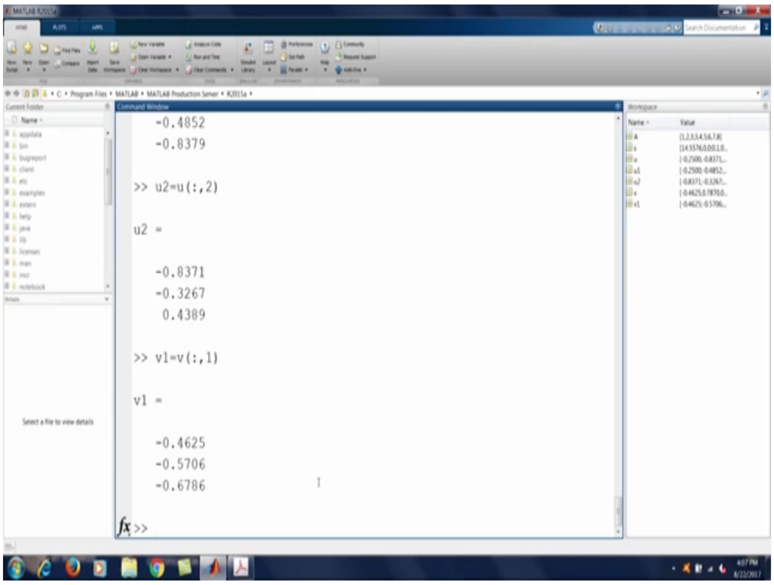
(Refer Slide Time: 31:30)



Now, we can write our u 1. So, you can say that u 1 is nothing, but the first column. So, to write the first column colon comma one this will give you the first column and this is minus 0.2500 this thing is given. So, it is given as the first column here.

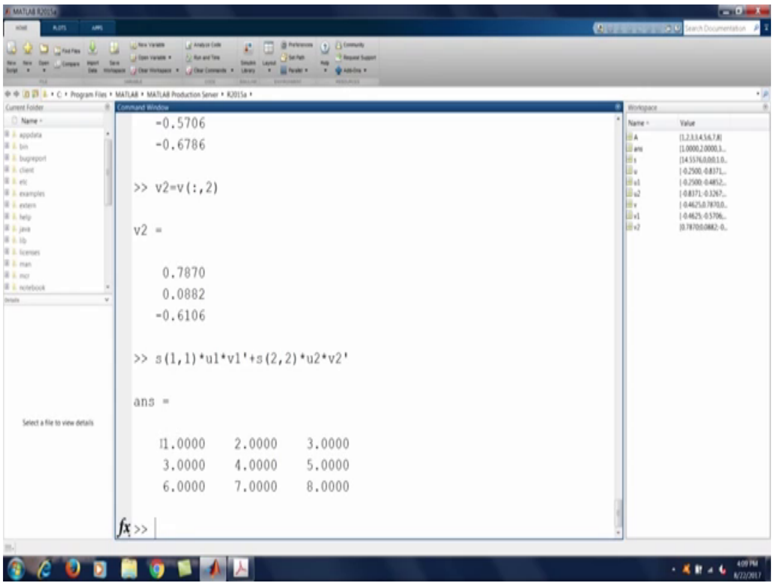
Now, similarly you can calculate u 2 which is u 2 is the second column of u, so that is given by u and this comma 2 and give you second column.

(Refer Slide Time: 33:02)



Similarly, we can calculate v 1 v 1 is the first column of v. So, it is v this comma 1 gives you first column. So, this is in agreement ok, the only difference is minus sign that is that if some is some vector is an eigen value then the minus 1 is also an eigen eigenvector.

(Refer Slide Time: 33:27)

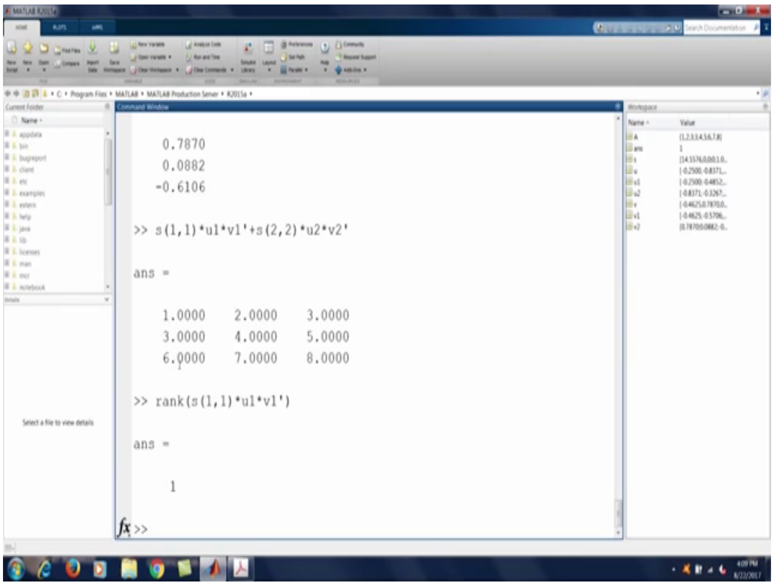


Someone, if a factor is eigenvector then constant multiple of that is also an eigenvector, so let us calculate v 2 which is the second column of V. So, it is coming out to be this.

So, now, we want to verify. So, we say that we calculate say S 1 1, S 1 1 is basically the first eigenvalue multiplied by u 1 say u 1 into say v 1 transpose plus S 2 2 2 comma 2 and then multiplied by u 2 and multiplied by v 2 transpose and see what you will get. And here you will see that it is the matrix A which you have entered that a as 1 2 3, 3 4 5, 6 7 8.

So, we you can say that your matrix A can be written as the sum of these 2 these 2 modes, where each mode is of the rank 1.

(Refer Slide Time: 34:47)



If you look at what is the rank of say this quantity. So, here we can write this as rank of S we can write it 1 comma 1, 1 comma 1 multiplied by u 1 multiplied by v 1 transpose and you can say that rank of S given by 1.

So, it means that any matrix can be written as sum of modes and the rank of each mode is equal to 1. So, it means that the important part of this is that if you want to save the each mode then if you look at the size of each mode is m cross n, but the rank is 1. So, if you want to save any m cross n matrix we required a many spaces, but if you if you if you save the rank m cross rank 1 matrix of size m cross n we need only m plus n spaces.

So, there is a huge difference between m plus n and m n, where this m n is very very large. So, this is quite efficiently used in an image processing. Their images are very very large or you can say that the matrix associated with images are very very large, there we can use this outer product of outer product expansion of this matrix A very efficiently.

So, I will stop here and in next lecture we will consider a example where we can utilize this outer product inspection of a matrix A very efficiently. And we can say that that we can remove the unnecessary details of a representation of a representation of an image as a matrix and we can still get a picture with the near nearer to your original image that we are going to see in a next lecture. So, here we stop our lecture. Thank you very much for listening.

Thank you.