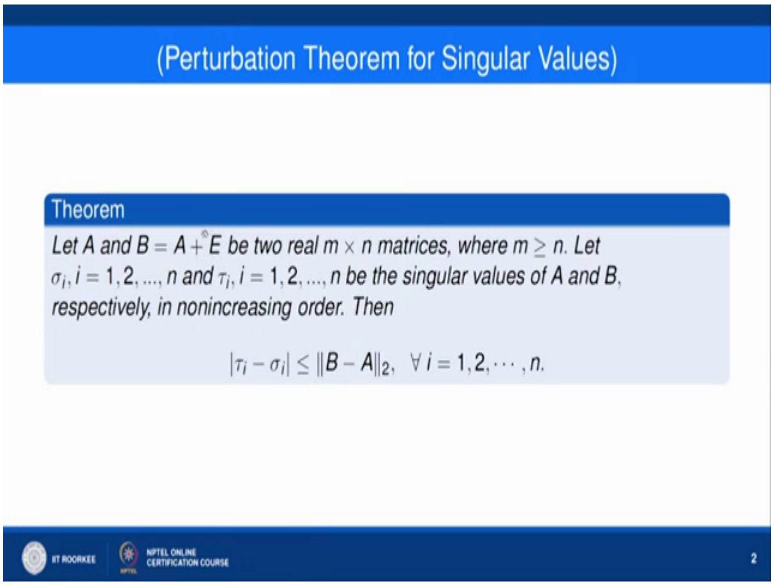
Numerical Linear Algebra
Dr. D. N. Pandey
Department of Mathematics
Indian Institute of Technology, Roorkee

Lecture – 47
Perturbation theorem for singular values

Hello friends, welcome to this lecture. Here in this lecture we will continue our study of singular value decomposition. If you recall in previous class in previous lecture of single value decomposition we have seen that the singular values of a any matrix A is well posed.

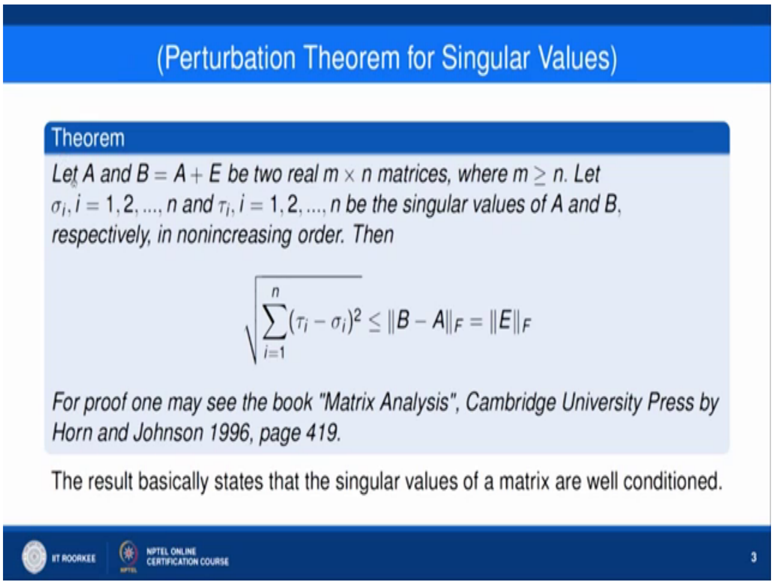
(Refer Slide Time: 00:48)



So, here we have discussed this theorem that let A and B where B is can be written as A plus E, E is considered to be a perturb perturbation matrix to this matrix A, be two real m cross n matrix, where m is greater than equal to n. And let sigma i, i from 1 to n and tau i i from 1 to n with a single values of A and B respectively. And it is arranged in non increasing order then modulus of tau i minus sigma i is less than or equal to 2 norm of B minus A and this is true for every i from 1 to n.

So, this is the result which we have discussed in previous class and we say that the that if we perturbed our matrix A by this matrix E then the eigen values of be singular values of B and a is differed by 2 norm of this perturb matrix E. So, in a the similar kind of result we can also state in terms of Frobenius norm.

(Refer Slide Time: 01:47)

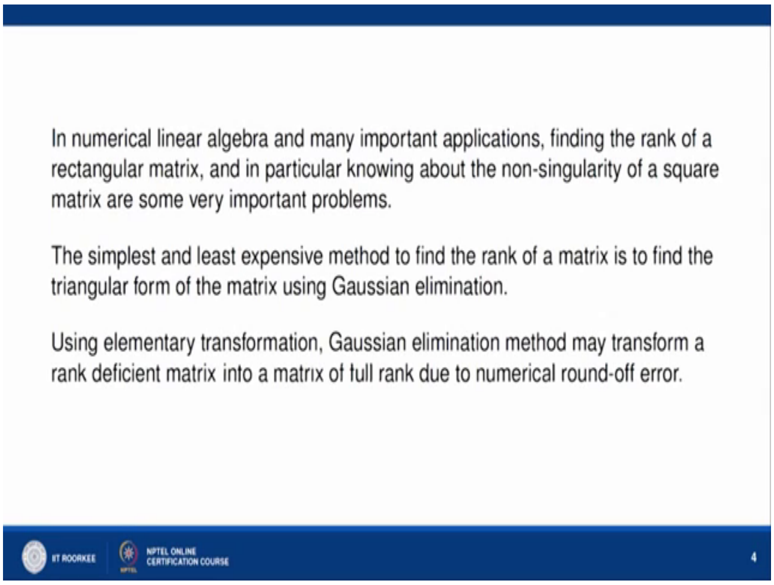


So, the similar kind of result which we a state a state in this theorem that let A and B, where B is A plus E be two real m cross n matrices, where m is greater than equal to n. Let sigma i i from 1 to n and tau i from i from 1 to n be the singular values of A and B respectively and it is arranged in non increasing order then under root of summation i equal to 1 to n tau i minus sigma square is bounded above by norma Frobenius norm of B minus A. And here Frobenius norm of B minus A is nothing but the Frobenius norm of the pert perturbation matrix that is this E. So, it means that into Frobenius norm this is bounded by the Frobenius norm of the perturbation matrix.

Here we are not providing any proof and if you want to see the proof of this theorem then we refer the book matrix analysis published by Cambridge University Press, and written by Horn and Johnson in the year 1996 and at page number 419 you can see the proof of this theorem.

And since we have already seen the proof here and we say that the similar kind of proof we can produce for this theorem. And the purpose of studying these two theorems is the following that the result basically state that the singular values of a matrix are well conditioned. So, it means that if you perturb your matrix your singular values will not change much.

(Refer Slide Time: 03:37)

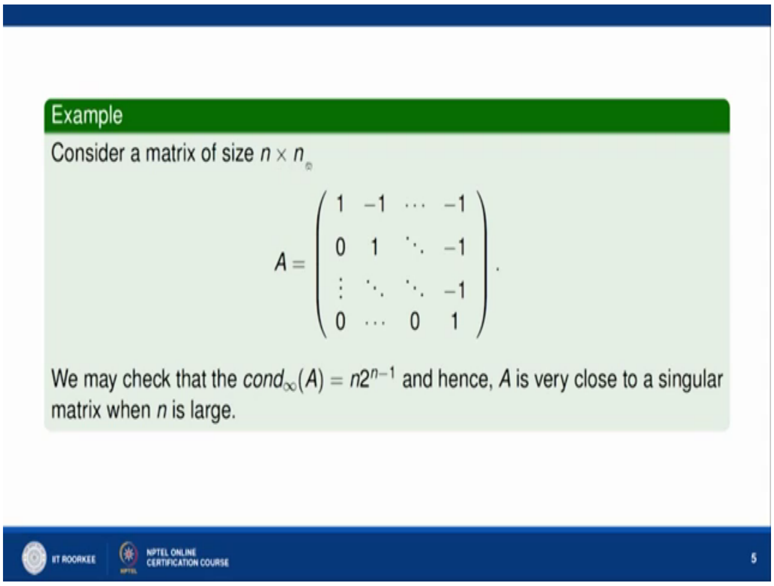


Now, with the help of the last result let us discuss some more important results. So, here we start with this that in numerical linear algebra and many important applications finding the rank of a rectangular matrix or in particular find finding the rank of a square matrix means, whether a square matrix is a singular or nonsingular is these problems are very very important problems. And we take will these problems of finding the rank of a given matrix by using the Gaussian elimination method and in Gaussian elimination method we use elementary root trans root transformation to convert a given matrix into a triangular matrix.

And in, but theoretically it is a very convenient method to find out the rank of a matrix. But if we consider this the same situation in practical life then it is a very it may create a lot of problem the problem is because of this you will elementary transformation and this using elementary transformation Gaussian elimination method may transform a rank deficient matrix matrix into a matrix of full rank due to numerical round off error. So, this is a fact which is very very important and we already know that the matrix which we are considering a may is already perturb or may be polluted by the equipment error or many other error

So, it means that there we are we are never sure that the matrix we are handling is the actual matrix are a perturb matrix. So, we cannot ignore these kind of thing that this elementary row transform may convert a rank deficient matrix into a rank full rank matrix. So, to avoid these kind of thing we take the help of singular value decomposition to find out the rank of a given matrix A.

(Refer Slide Time: 05:47)



And for example, if you look at this matrix here, where A is a matrix of size n cross n and it is a upper triangular matrix. So, the diagonal elements are one and the all the other element in same in the triangular part is minus 1. So, upper above diagonal entries are minus 1 and the lower diagonal entries are simply 0. So, it is basically a upper triangular matrix whose diagonal entries are 1 1 1 1.

So, it looks that this matrix is a nonsingular matrix and it is nowhere apparent that this matrix is very very near to a singular matrix how we can check that if you look at the condition number of this matrix and the infinity norm and it is you can check that it is coming out to be n times 2 to power n minus 1. And hence you can say that for large size n this condition number is very very large number and hence we can say that by the property of condition number we can say that A is very close to a singular matrix when n is large.

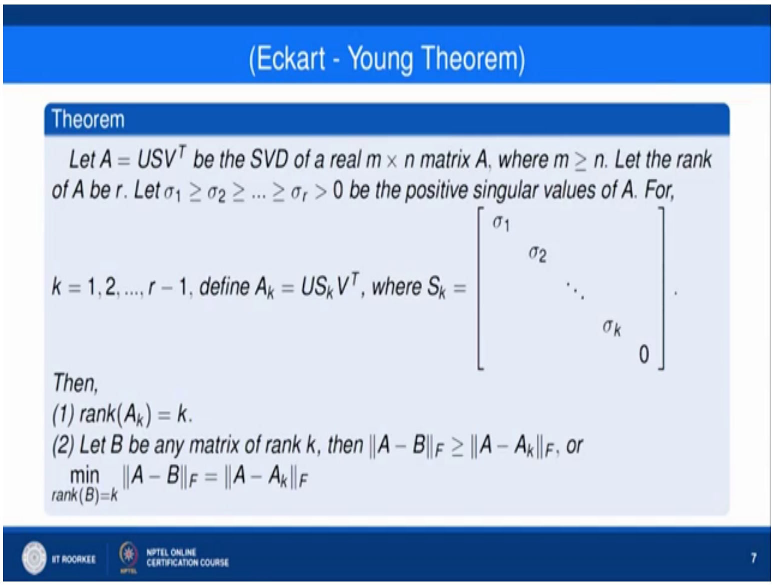
So, it means that handling since you can say that this matrix you can obtain from a any matrix using not any matrix for a given matrix using Gaussian elimination method and here looking at this a we are not able to find out whether this matrix A is very near to a singular matrix or not. So, looking at all these drawbacks that a the matrix given in hand A is a original matrix or a polluted matrix or due to numerical round off error this

Gaussian elimination may convert a rank deficient matrix to a full rank matrix, and even in a triangular form we really do not know whether a given matrix is near to a rank deficient matrix or not we prefer to use SVD, single value decomposition method.

Now, since the number of nonzero singular value determine the rank of a matrix we may say that a matrix A is arbitrary close to a matrix of full rank by changing all these 0 singular values by a very small number say epsilon which is very very small and it is positive, any nonzero number we can say. Then you have a same matrix and it may have some 0 singular values, you change your singular values by an arbitrary small number say epsilon and then this matrix is A changed matrix is a full a matrix of full rank, and you can say that it is arbitrary close to the given matrix.

So, in practical it is more important to know if a matrix is near to matrix of a certain rank rather than knowing the rank of the given matrix. And the singular value decomposition may exactly answer this query. The most useful way to determine the rank and nearness to rank deficient matrix is to use the SVD. So, we try to verify the these statement by same methods.

(Refer Slide Time: 09:14)

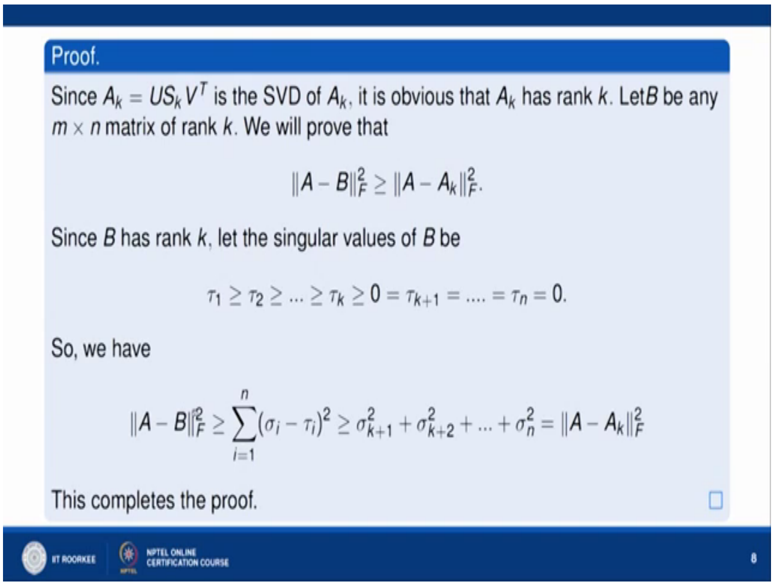


So, first result in this direction is the following theorem which is Eckart young theorem and given by Eckart and young in 1936. The statement of theorem says that, let A is equal to USV transpose be the singular value decomposition of a real m cross n matrix A, where m is greater than equal to n and we let us assume that the rank of this matrix A is r and sigma 1 greater than equal to sigma 2 greater than equal to sigma which is positive be the positive singular values of A.

Now, for k ranging between 1 2, r minus 1 we can define A k as US k V transpose where this matrix S k is given by this diagonal matrix where the first these diagonal elements are sigma 1, sigma 2 to sigma k and rest are all 0s. Wait, then we say that we claim that rank of A k is k and if we have any matrix B of rank k then Frobenius norm of A minus B is greater than or equal to Frobenius norm of A minus A k.

So, it means that if you calculate the Frobenius norm of A minus B, where B is any matrix of rank k then the Frobenius norm of this is always bigger than or equal to Frobenius norm of A minus A k. And you can say that the equality you can achieve that it means that and minimum of rank B is equal to k norm Frobenius norm of A minus B is a Frobenius norm of A minus A k. So, it means that the equality achieved when B is equal to A k.

(Refer Slide Time: 11:09)



So, let us consider the proof here. So, first thing is that the rank of A k is k and this is quite easy because if you look at the rank of A k is rank of US k we transpose. Now, U and V are orthogonal matrices and hence of full rank. So, rank of A k is nothing, but rank of S k and rank of S k is nothing, but lead nonzero values in the diagonal in the diagonal of S k which is k. So, it means the rank of A k is nothing, but the k. So, it is obvious that A k has rank k.

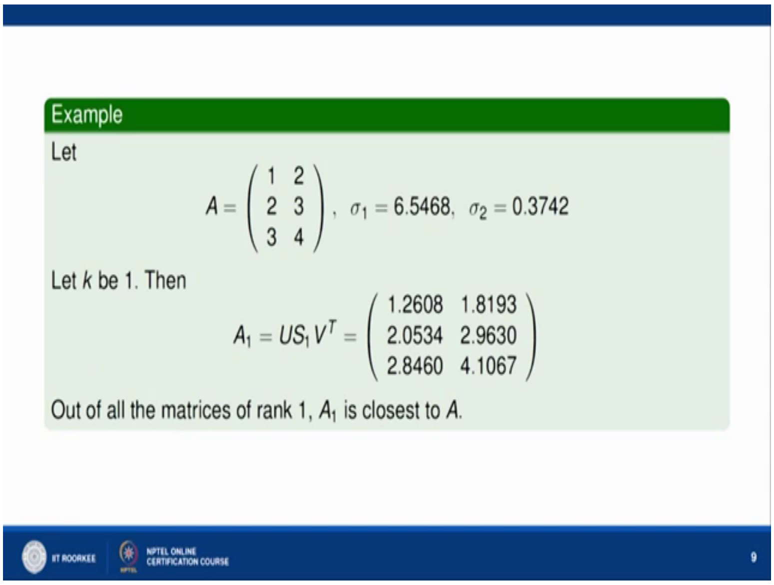
Now, we want to show that if we take any B which is a m cross n matrix of rank k then we want to show that the Frobenius norm of A minus B square is greater than or equal to Frobenius norm of A minus A k

So, let us take B and any m cross n matrix of rank k then the singular values of B, B let us say tau 1 tau 2 and tau k and we assume that rest are all tau k plus 1 to tau n are all 0 singular values. So, up to tau 1 to tau k we have positive singular values and rest are 0 singular values.

Then we already know that the result we have shown in the first as this result if you look at this result that, let A and B be two real m cross n matrix then we know that the Frobenius norm of B minus a is bigger than this. So, using this result we can say that that Frobenius norm of A minus B whole square is bigger than i equal to 1 to n sigma i minus tau i whole square, and if you if you look at this is what this is greater than or equal to sigma i square k plus 1 plus sigma i square k plus 2 and so on sigma n square.

Now, if you look at this is nothing, but the Frobenius norm of A minus A k square. So, it means that that if it calculate the Frobenius norm of A minus B square then it is greater than or equal to Frobenius norm of A minus A k square and the equality follows if you replace B as A k. So, this complete the proof of the theorem that Eckart young theorem that if you define A k like us k V transpose then minimum of Frobenius norm of A minus B, where B rank of B is a matrix of rank A, then minimum of rank B equal to k Frobenius norm of A minus B is equal to Frobenius norm of A minus A k.

(Refer Slide Time: 14:01)

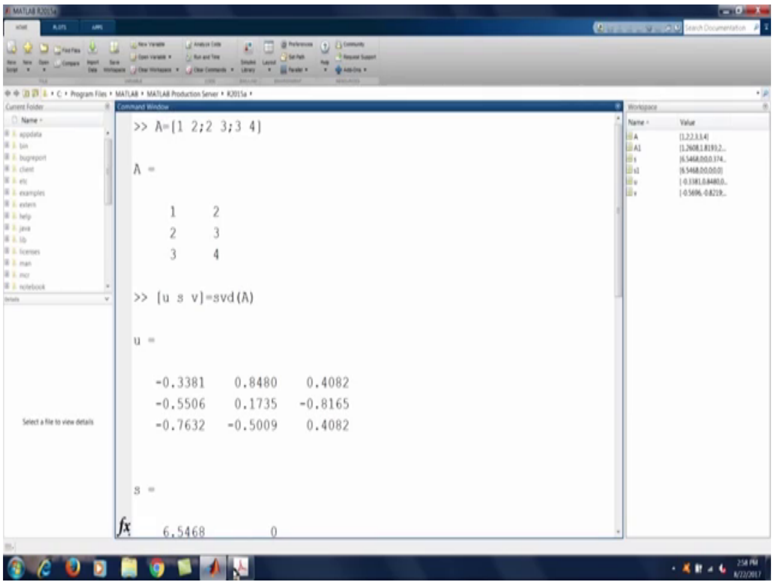


Now, consider the similar proof in 2 norm, but before that let us consider the following example here we have matrix A given as a 1 2, 2 3, 3 4 and we can easily calculate the singular values of this matrix A as sigma 1 is equal to 6.5468 and sigma 2 as 0.3742. And we want to find out a matrix B which is of rank 1 and it is closest to the matrix A.

So, for that you define A 1 as US 1 V transpose, where S 1 is a diagonal matrix where the diagonal entry is sigma 1 and rest all 0. Then we can calculate that A 1 as US 1 V transpose and it is given by this. So, by previous theorem we can show that this A 1 is the matrix of rank 1 and it is closest to the matrix A and this you can verify with the help of your MATLAB. So, you can see whether it is true or not.

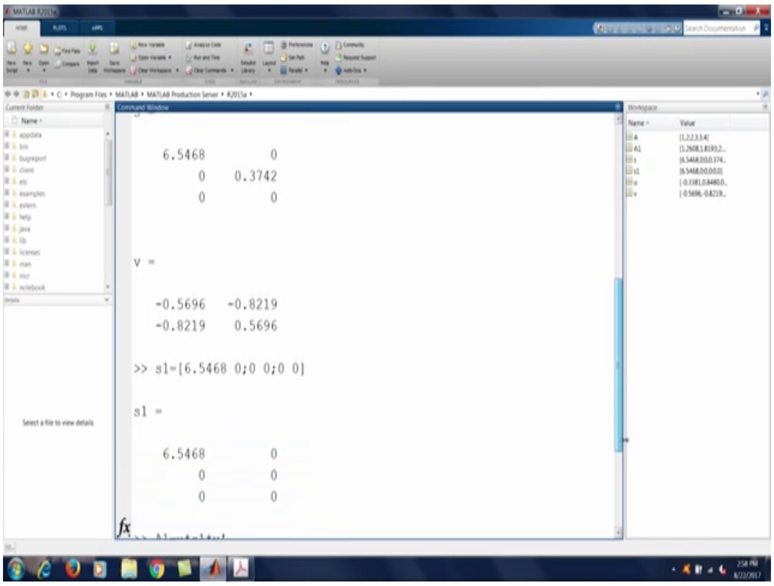
So, we want to show the verification of this example using MATLAB.

(Refer Slide Time: 15:20)



So, if you look at we define our matrix A as 1 2, 2 3, 3 4. If you display it and then A is given as 1 2, 2 3, 3 4 and then we try to find out say singular value decomposition of this matrix. So, it means that U is an orthogonal matrix given as this and S is this and V as this and this is given here.

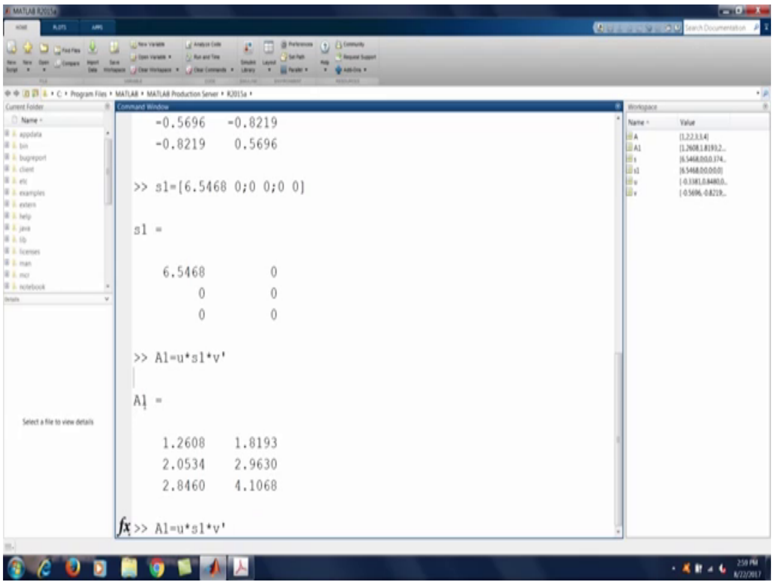
(Refer Slide Time: 15:40)



This here we have simply noted down sigma 1 and sigma 2 as 6.5468 and 0.3742 and which is given here say 6.5468 and 0.3742.

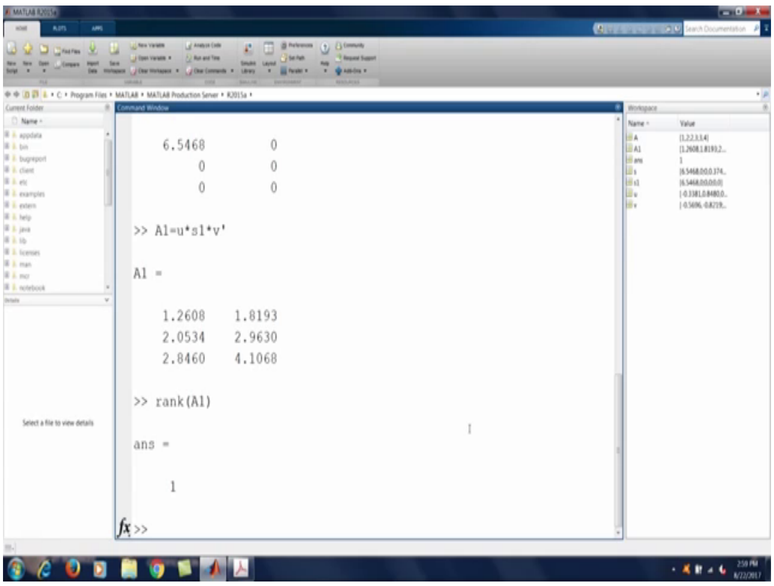
So, here if we define a matrix S 1 where S 1 will consist only first singular values the bigger singular value 6.5468 and rest entries are 0. So, it means that S 1 is your matrix like this 6.5468 0 0 0 0 and then we calculate A 1 as US 1 V transpose then your matrix A is given as 1.2608, 1.8193, 2.0534 and so on.

(Refer Slide Time: 16:15)



And we can say that this matrix A 1 which is listed here is a matrix of a rank 1 and A 1 is closest to A, that you can check that rank of A 1 is what.

(Refer Slide Time: 16:43)

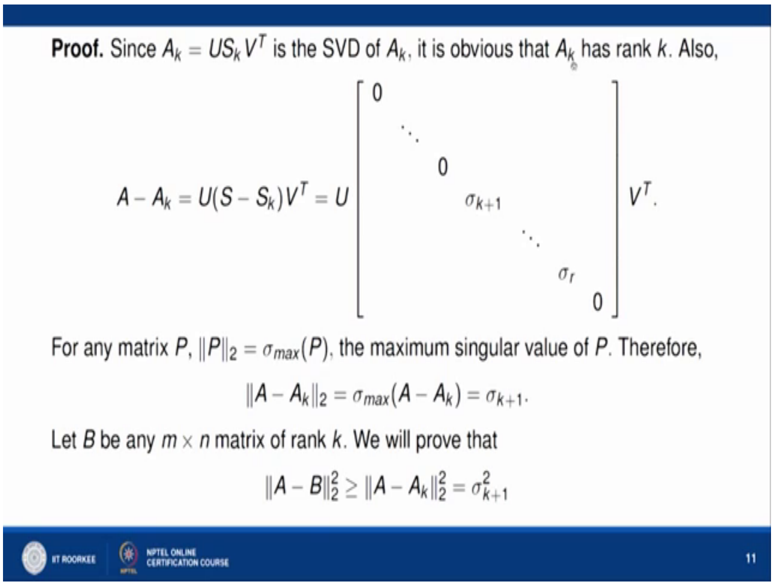


So, here rank of A 1 you can say that it is 1. So, it is coming out to be here. So, it is a of rank 1 matrix which is closest to matrix A with the in Frobenius norm.

Now, we try to see the similar result in terms of 2 norm and we have the following theorem, The statement of this theorem is that let A is a USV transpose be the singular value decomposition of real m cross n matrix A where m is greater than equal to n and let the rank of a matrix A is r and we arrange our singular values as this order. And be in the sigma 1, 2 sigma r be the positive singular values of a and for k from 1 to r minus 1 we can define A k as US k V transpose where S k is as defined as before. Then we claim that rank of A k is k and minimum of 2 norm of A minus B, where B is any matrix of rank k is your 2 norm of A minus A k which is nothing, but sigma k plus 1.

So, it is a similar kind of result as we have seen in Eckart Young theorem and let us try to prove this the proof is slightly different from the proof for Eckart Young theorem.

(Refer Slide Time: 18:15)



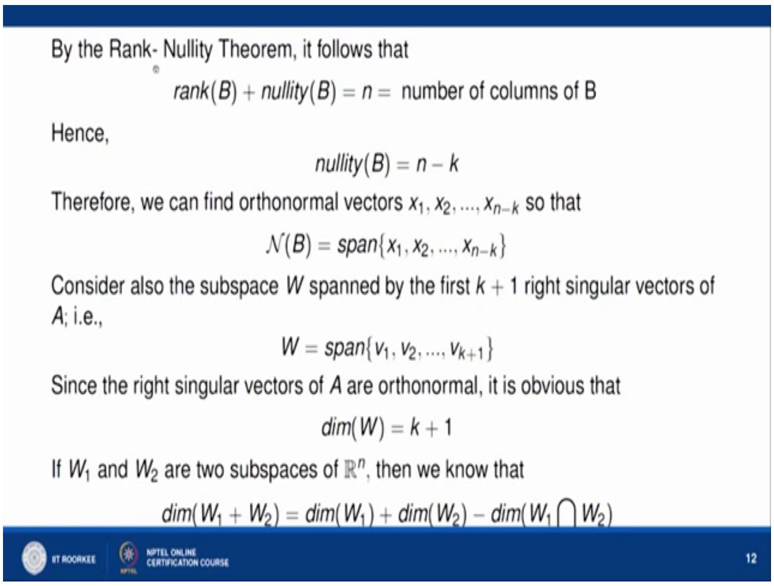
So, let us say that here since A k as US k we transpose is the SVD of A k it is obvious that A k as rank k or otherwise you can simply say the rank of A k is a rank of S k. And rank of S k is easily we can see that it is k because the nonzero entries in S k is only k.

So, now, you calculate A minus A k which is nothing but U S minus S k V transpose. So, if you look at it is given by this. So, the first diagonal k-th diagonal entries are 0 and then we have sigma k plus 1, 2 sigma r and rest all 0. So, the remaining r minus k entries are nonzero and rest are all 0.

Now, we already know that the 2 norm of a matrix P is sigma max of P singular here sigma represent the singular maximum singular values of matrix P, then 2 norm of P is the maximum singular value of P then we want to find out say 2 norm of this matrix A minus k A minus A k. So, 2 norm of A minus A k is the sigma max of A minus A k, and if you look at sigma sigma max of A minus A k is nothing but sigma k plus 1, as we have already arranged our single values in this fashion that sigma 1 is greater than equal to sigma 2 and so on. So, sigma k 1, so among these singular values your sigma k plus 1 is the maximum singular value of this matrix A minus A k. So, what we have proved here is that the 2 norm of A minus A k is sigma k plus 1.

Now, what remains is to prove that two norms of 2 norm of A minus B is bigger than this quantity. So, let B be any m cross n matrix of rank k, we will prove that square of 2 norm of A minus B is greater than or equal to 2 norm of A minus A k whole square and that is nothing, but sigma i square k plus 1.

(Refer Slide Time: 20:19)

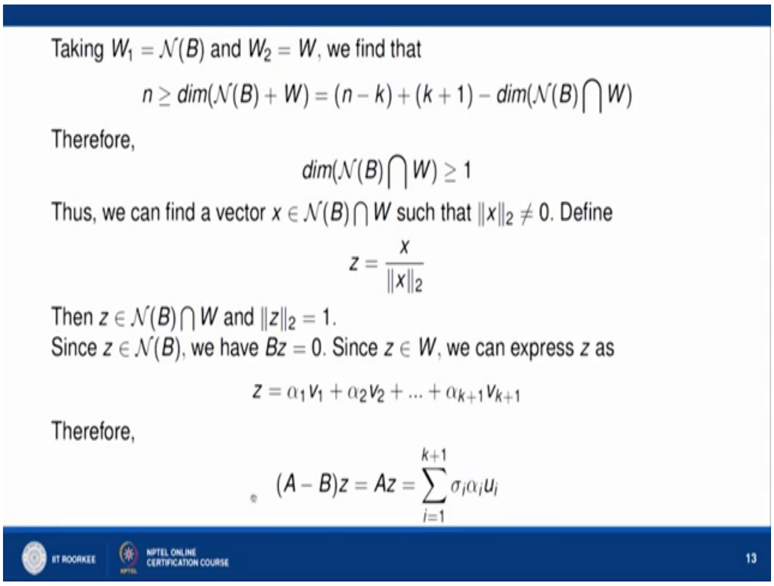


For that let us assume that for this rank of B which is already known as r k. Then by the rank nullity theorem we can say that nullity of B a is n minus k, here rank of B plus nullity of B is equal to n, where n is the number of columns of B and since rank of B is your k then nullity of B can be written as n minus k. So, if nullity of B is n minus k then we can always find out orthonormal vectors x 1 to x n minus k so that a nullity null space of B can be spanned by these n minus k orthonormal vectors.

Now, also consider a subspace W which is spanned by the first k plus one right singular vectors of A. It means at the first k plus 1 columns of matrix V. So, let us say that W is span of v 1 to v k plus 1, and we already know that by construction the right singular vectors of a are orthonormal. So, we can say that this rank dimension of this W is going to be k plus 1.

Now, look at the dimension of W plus N B. So, we already know that if W 1 and W 2 are two subspaces of R n then dimension of W 1 plus W 2 you can find out using this formula, that is dimension of W 1 plus dimension of W 2 minus dimension of W 1 intersection W 2. So, this we have already seen.

(Refer Slide Time: 22:01)



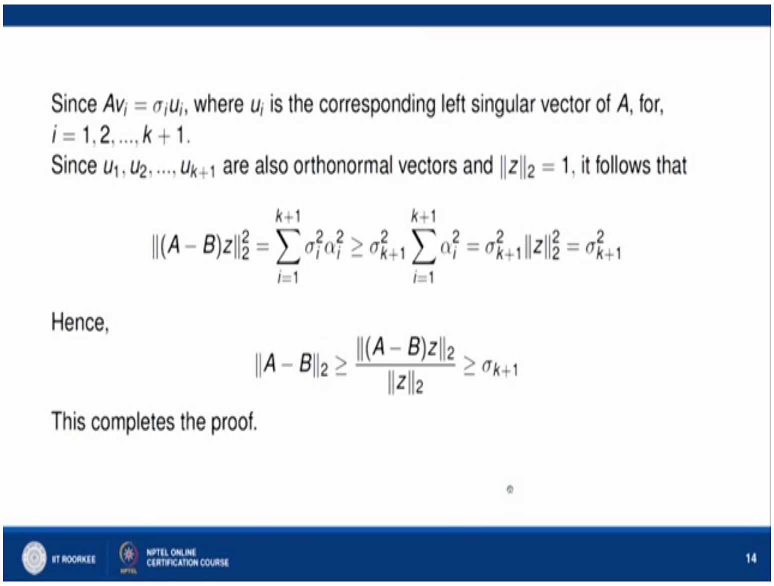
And we, so if we take W 1 as a null space of B and W 2 as W then we can say that your dimension of N B plus W it cannot exceed n. So, n greater than or equal to dimension of N B plus W is equal to n minus k which is the dimension of null space of B plus a dimension of W minus dimension of null space of B intersection W.

So, if you look at this is going to be n plus 1 and it can since this n plus 1 minus dimension of N null space of B intersection W cannot exceed this n. So, this implies that the dimension of null space of B intersection W has to be greater than or equal to 1. It means that they exist a non0 vector in this at least one nonzero vector in this in this intersection. So, it means that we can always find out a vector x belonging to null space of B intersection W such that the it is nonzero or you can say in particular to norm of this vector x is nonzero.

Now, with the help of this x we define a vector z as x by 2 norm of x so that the 2 norm of z is equal to 1. So, it means that z belongs to null space of B intersection W and 2 norm of z is 1. Now, with this z we already know that z belongs to null space of B it means that Bz is going to be 0, and since z also belongs to W it means that z can be expressed as linear combination of v 1 to v k plus 1. So, we can say that z is equal to alpha 1 v 1 plus alpha 2 v 2 and so on alpha k plus 1, v k plus 1 therefore, if we calculate a minus B z then it is basically Az only because Bz is going to be 0. Now, Az is basically what? a operating on this.

Now, we know that alpha 1 a v 1. Now, a v 1 is basically sigma 1 u 1 and a v 2 is sigma 2 u 2 and so on then we can write that A minus B is z is nothing but i equal to 1 to k plus 1 sigma i alpha i u i.

(Refer Slide Time: 24:25)

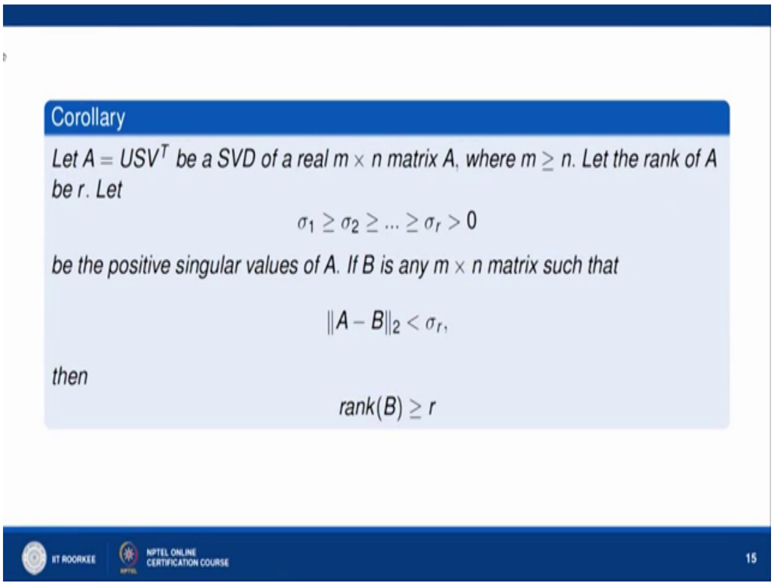


And we this is follows because Av i equal to sigma i u i, where u i is the corresponding left singular vectors of a for i equal to 1 to k plus 1.

Now, we already know that this u 1 to u k plus 1 which are left singular vectors of A or orthonormal vectors and we also that 2 norm of z is equal to 1 and hence we can say that sigma that 2 norm of A minus B z whole square is greater than equal to sigma k plus 1 square. Now, if you divide by 2 norm of z whole square then you can say that 2 norm of A minus B is greater than equal to 2 norm of A minus B z divided by 2 norm of z and it is greater than or equal to sigma k plus 1.

So, if we take any matrix B of rank k then 2 norm of A minus B is greater than equal to sigma k plus 1. Also we have seen that that if we take B as A k then 2 norm of A minus B is exactly sigma k plus 1. So, if we combine these two facts then we can say that the minimum of this 2 norm of A minus B, where we are taking minimum over all the matrix B of rank k then it is nothing, but 2 norm of A minus A k and it is equal to sigma k plus 1 and this complete the proof of this theorem.

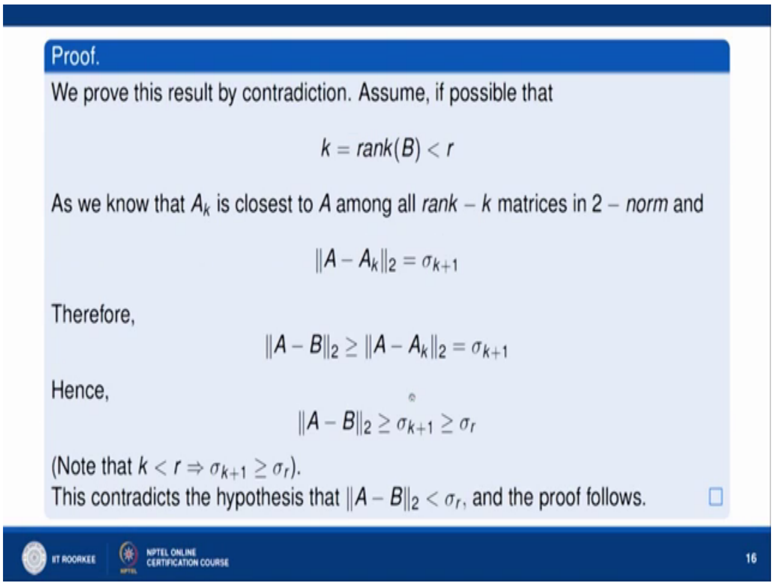
(Refer Slide Time: 26:04)



So, now, let us consider a important corollary based on the previous theorem and it says that let A is given as USV transpose B a singular value decomposition of a real m cross n matrix A, where m is greater than equal to n and let the rank of A be r and we arrange the singular values as this. And these are all positive singular values of A, and if B is any m cross n matrix such that 2 norm of A minus B is less than sigma r then the rank of B is going to be greater than or equal to r.

So, let us see what it how we can prove this corollary and then we try to see what the geometric interpretation of this corollary.

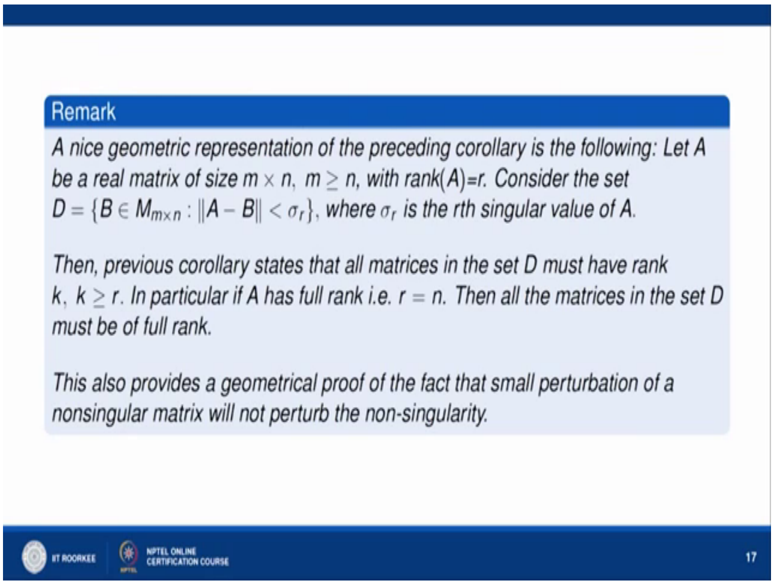
(Refer Slide Time: 26:48)



So, proof is quite simple. So, we prove this result by contradiction. So, assume it possible that rank of B is less than r. What we want to wanted to prove is that rank of B is greater than equal to r. So, let us assume that rank of B is less than r and let us say that it is k and we already know that this A k is the closest to a among all rank k matrices in 2 norm. So, that we have just proved in previous theorem and we say that 2 norm of A minus A k is sigma k plus 1. So, we, so it means that 2 norm of A minus B is greater than equal to 2 norm of A minus A k which is nothing, but sigma k plus 1 that is precisely the theorem we have just proved. And hence 2 norm of A minus B is greater than equal to sigma k plus 1.

And since your k is less than r then sigma k plus 1 is bigger than sigma r, but this contradict the hypothesis that 2 norm of A minus B is less than sigma r. So, this complete the proof of this simple corollary. So, what we have shown here that if we assume the rank of B is less than r then 2 norm of A minus B is greater than equal to sigma, but we have already assumed that 2 norm of A minus B is less than sigma r. So, we get the contribution.

(Refer Slide Time: 28:16)



So, we can understand the by the following remark that what is the meaning of the previous corollary. So, we can say that a nice geometric representation of the preceding corollary is the following that let A be a real matrix of size m cross n where m is greater than or equal to n with rank of A is r. Then we can consider this set S it is a set of all matrices of size m cross n such that 2 norm of A minus B is less than sigma r, where sigma r is the rth singular values of A.

Then the previous corollary states that all matrices in this set D must have rank k where k is greater than equal to r or you can say that any matrix in this set must have rank greater than or equal to r, or in particular if we assume that rank of a has full it means that r is equal to n then we can say that any matrix in this set D must have full rank. And it also provide a geometric proof of the fact that small perturbation of a nonsingular matrix will not perturb the non singularity. So, it means that if we perturb a nonsingular matrix by a little then it will not perturb the non singularity status of the matrix A.

So, we will this we stop our lecture. So, in this lecture what we have seen here is that singular value decomposition not only give the the best way we can say to find out the matrix rank of the matrix A, but also it also gives the approximation of the matrix A by A k, where A k approximate our matrix A in 2 norm and F norm and it is known as the best approximation of the rank k.

So, in next lecture we continue our study and we will see some more important result obtained through singular value decomposition.

Thank you.