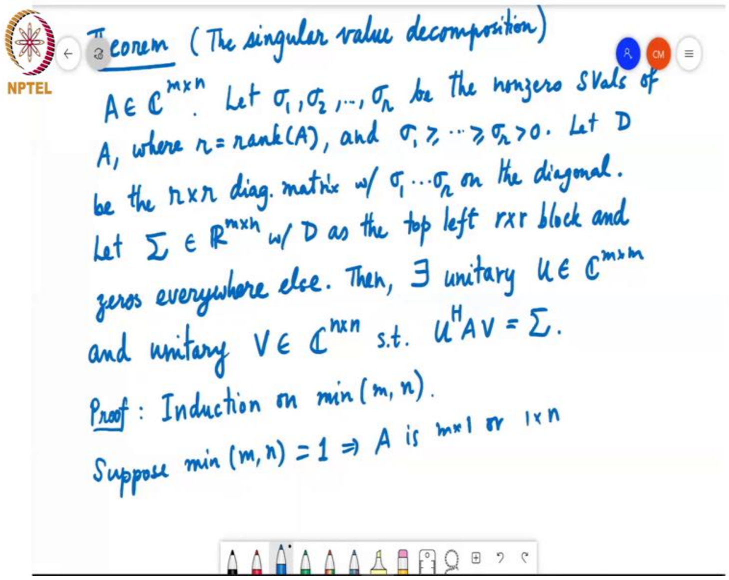
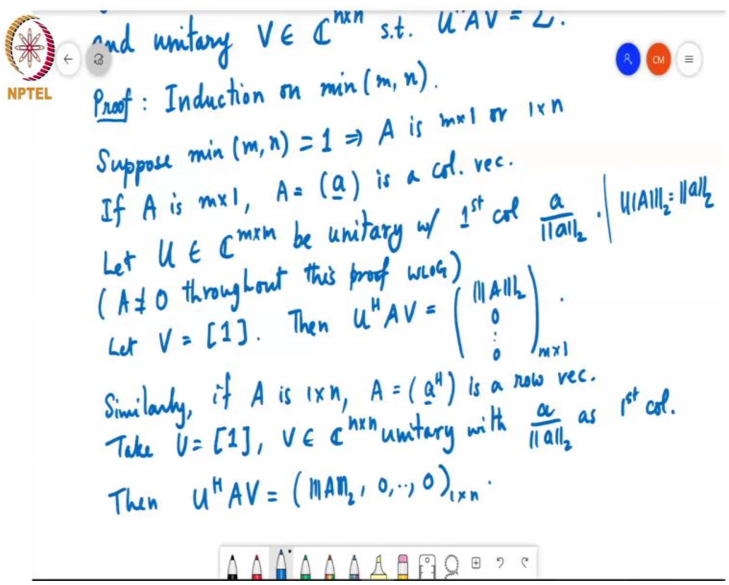
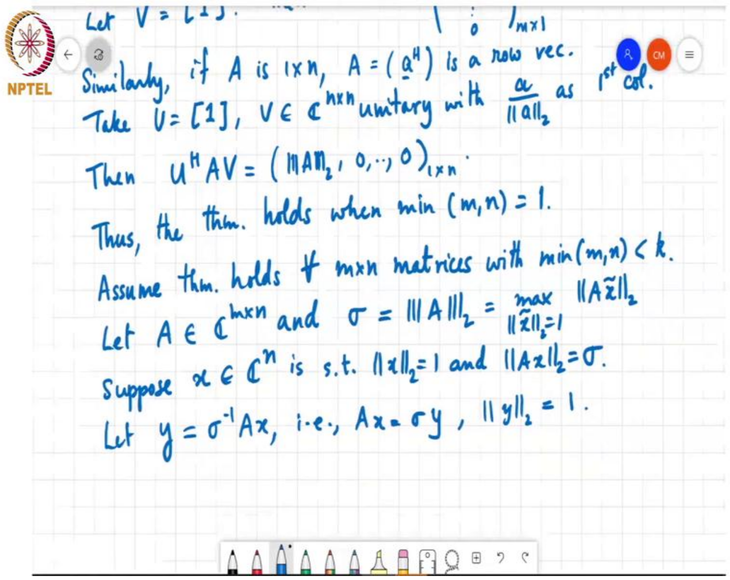
Matrix Theory
Professor Chandra R Murthy
Department of Electrical Communication Engineering
Indian Institute of Science Bangalore
Proof of Singular Value Decomposition Theorem

(Refer Slide Time: 00:14)







The very few results in this course. For example, Schur decomposition theorem is one of them, which apply to any matrix and this is also one of them, you can find the singular value decomposition of any matrix. So, the way we prove this is first we will prove using induction that there exists unitary u and v says that u Hermitian Av equals some matrix S where S is a matrix which has real non-negative diagonal block on the top left corner and 0s everywhere else.

And then we will show that S is equal to the sigma that we have defined here. So, now, so what we will do is, we will first perform the proof that there exists a u and v such that u Hermitian Av equals S where S is a diagonal matrix, which is actually a rectangular matrix. So, I mean a rectangular matrix with only the top left r cross r block being non negative and real value. We will show that by induction.

So, we will do induction on the minimum of m and n. So, in any induction based approach, the first step is to show that the induction holds for the index equal to 1. So, suppose min of m, n is equal to 1. So, that either A is either a row vector or it is a column vector. Now, if A is m by 1 then it is a column vector. So, I will call it say this vector A is a column vector. Now, we will just directly let u be in C to the m cross m be unitary with first column a over norm of a.

Now, we can without loss of generality throughout this proof we will assume that A not equal to 0, without loss of generality because, I mean, you can trivially show I mean this is trivially true if A is equal to 0 you just choose sigma equal to 0 and you can choose any unitary u and v and this will hold.

So, the result holds trivially if A is 0 matrix so, a2 here is non-zero so, it is okay to divide by the l2 norm of a. Now, it is always possible to find the unitary matrix you once you pick the first column, you just pick orthonormal columns to it and form a unitary matrix of size m by n and let v be just this matrix with a single 1. This is 1 cross 1 matrix with entry equal to 1. Then u Hermitian Av is then going to be equal to and incidentally this A2. So, this is also true I mean you can write out for this column vector you can write out and see that this is actually true.

So, then if I compute u Hermitian Av when I do u Hermitian times A, A has only one column in it and that column is in the same direction as the first row of u Hermitian and all of the rows of u Hermitian are orthogonal to the column of A and so, this is actually equal to the norm of A and then 0s everywhere else.

And this is exactly the form I wanted which is that u Hermitian Av equal to sigma and in this case sigma is of size m by n where n is equal to 1 here so, this is m by 1. Now, similarly, if A is 1 cross n then A is a row vector and I will write it let us see how do I want to write it, then I will take u equal to the 1 cross 1 matrix 1 and v to be a unitary matrix of size n cross n to with a over norm of a as first column.

Then once again u Hermitian Av so A is this A Hermitian and when I have this here the first column of v when it multiplies with A I will get A Hermitian A which is a2 squared divided by a2 so, that will be a2 and all other rows all other columns of Av will be 0 because the other columns of v are all orthogonal to this A A Hermitian here and so, this will be equal to 0, 0 and these are size 1 by m. So, again it is in the form when form that u Hermitian Av equal to sigma. So, this shows that the theorem is true when min of m, n equals 1.

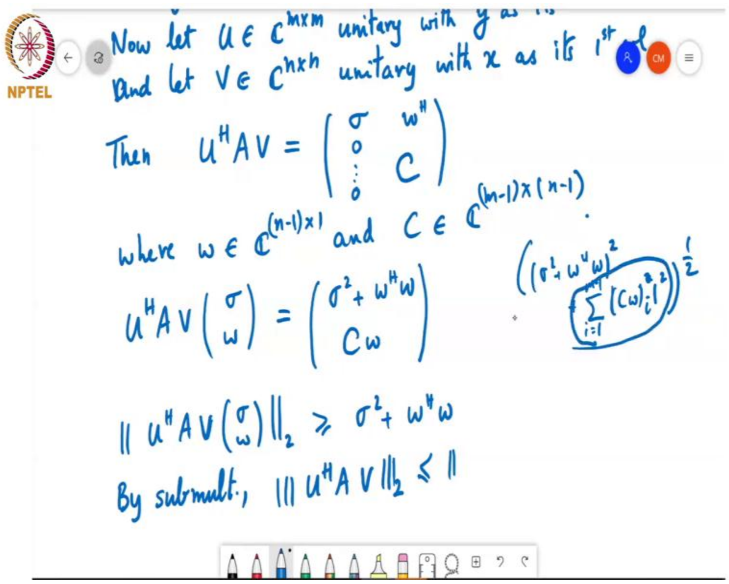
So, now we go to the induction step. So, we assume theorem holds for all m by n matrices with min of m, n strictly less than k, so min m and n up to k minus 1 and now we need to show that it holds for the k min m, n equal to k also. So, let A be of size m by n and suppose I define sigma to be the spectral norm of A, then choose I will just write what this is so that it is clear.

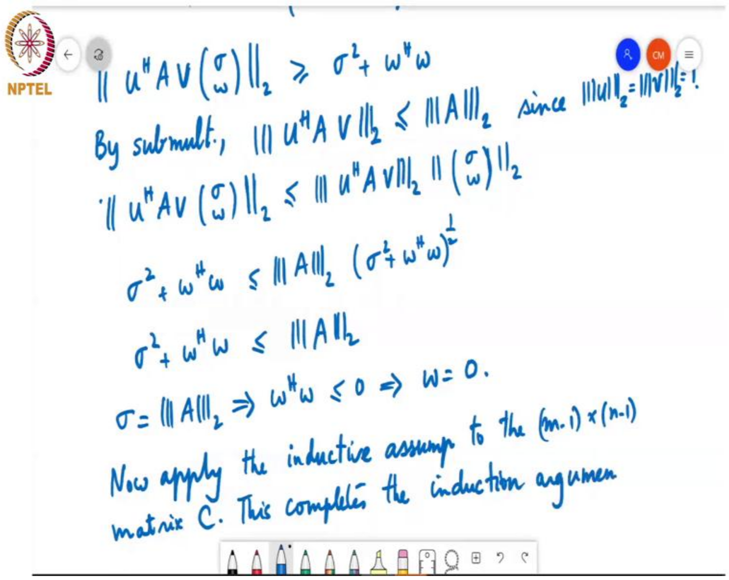
So, now, this is an optimization problem that has some solution, whatever the solution gives us as the objective function, that is, by definition equal to sigma, which means that there exists some x for which norm l2 norm of Ax equals sigma and suppose x is that particular x, so I will write this as x tilde here, then I will say suppose x in C to the n is such that x 2 equals 1 and Ax equals sigma we be back in one second.

So, suppose x is that vector that solves this one and so x is some vectors such that this is l2 norm is 1 and the l2 norm of Ax equals sigma and let y in or I will just say, let us define y to be sigma inverse times Ax. Again, because A is not equal to 0 and so this quantity is strictly positive. So, sigma is also strictly positive. So, it is okay to write sigma inverse or 1 over sigma times Ax.

So, it is again coming because we are assuming is A non-zero matrix. So then, so Ax rather i e Ax equals sigma y and also if I compute y2 I will get equal to 1 because the norm of y l2 norm of y is going to be 1 over sigma times the l2 norm of Ax and the l2 norm of Ax equals sigma. So, y2 equals 1.

(Refer Slide Time: 12:14)





Now, let u be in C to the m by m be unitary with y as its first column it is a unit norm vector, so we can choose y to be its first column and let v in C to the n cross n unitary with x as its first column. So again, x2 equals 1, so I can choose x to be the first column of this unitary matrix. Now, comes the magic. So, then, if I compute u Hermitian Av this is equal to I can write it as sigma 0, 0, some w Hermitian, and some matrix C here, where w is in C to the n minus 1 cross 1 and C is a matrix in m minus 1 cross n minus 1.

This happens because of the way we have chosen this u and v matrix you want to you should just multiply it out and convince yourself that this is in fact true. Now, if I multiply this by Av, if I multiply it by the Hermitian of this first row here, then it is the same as multiplying this by this, this vector. And what I will get then is sigma squared plus w Hermitian w as its first element, and C times w as whatever sits below that. And now if I take the norm of this, I have that the norm of u Hermitian Av times sigma w is at least equal to the square of the first entry here.

So, in other words, the actually the norm of this is this square plus the entries of this square whole square root, but if I draw up all the terms corresponding to this and then just take the square root this is what I will get sigma squared plus w Hermitian w, that becomes a lower bound on the norm of this vector here.

Now, by now we use the sub-multiplicativity we have that u Hermitian Av…

Student: Sir, this should not be in under root in the lower bound?

Professor: That is what I am saying. So, what I would do normally is I would do sigma squared plus w Hermitian w square plus sigma i equal to 1 to…

Student: Sir, I got it what was it.

Professor: You got it. So, 1, 2 I want to say m minus 1 Cw ith entry square, I need modulus over here, because these are complex valued square and then I should be taking this whole thing power half and what I am just dropping all these terms, then when I take the power half I just get this thing.

Now, this is less than or equal to the product of the norm of this times the norm this times the norm this, but these are unitary, so their norm, the Euclidean norm equals 1 or there spectral norm equals 1 and so this is just norm of A. Now, the other thing is that if I multiply if I apply the sub multiplicative on this what I have is u Hermitian Av.

Student: Sir.

Professor: Sigma w, yes.

Student 2: Sir, u Hermitian Av spectral norm will be equal to spectral norm because left and right multiplication. So, it will be less than is equal to?

Professor: Good question. So, in fact, I think actually going there, just bear with me for a minute. For now, I am just using the sub multiplicativity property to claim that the spectral norm of u emission Av is less than or equal to the spectral norm of A. So, now, if I look at the l2 norm of l2 norm of this, this is less than or equal to the norm of u Hermitian Av times the norm of this vector sigma w and so from this, we have that.

So, I could have probably done this faster if I had said that these two will be equal, but anyway, I am just following the textbook here. So, what we have from this is that sigma squared so, this quantity, this is a lower bound on this. So, I should put this over on this side. So, sigma squared plus w Hermitian w is less than or equal to this quantity which is less than or equal to this quantity and this itself is the norm of A.

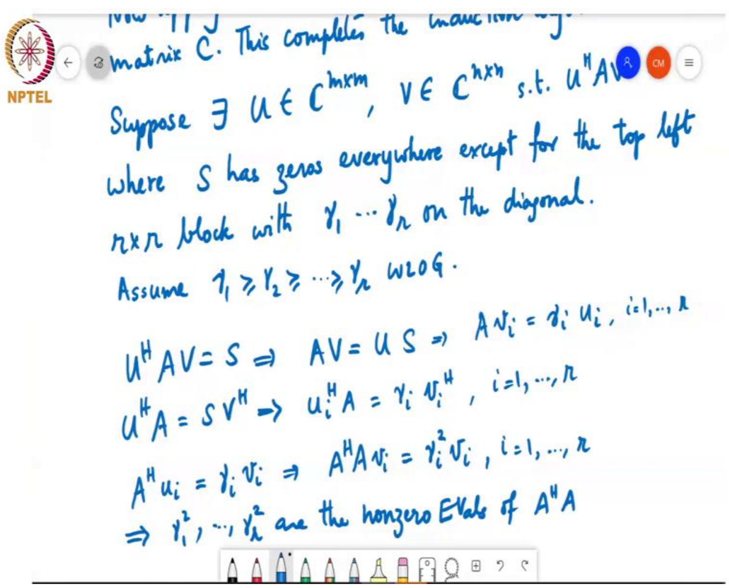
So, this itself is less than or equal to the norm of A times the l2 norm of sigma w is nothing but sigma squared plus w Hermitian w power half and so, if I just square both sides and then cancel off one of the sigma squared plus w Hermitian w I have that sigma squared plus w Hermitian w is less than or equal to the spectral norm of a.

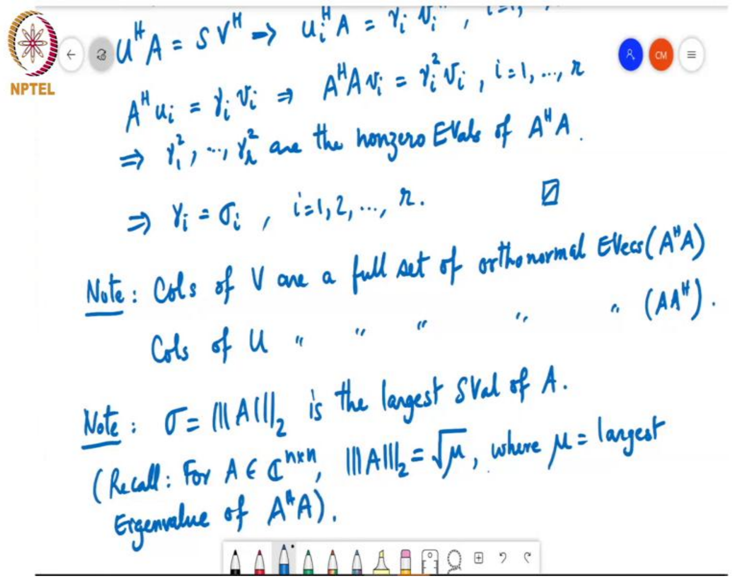
But sigma we go back to sigma, sigma is equal to the spectral norm of A. So, what that means is and so sigma so, if I substitute this here it says that w Hermitian w is less than or equal to 0 which in turn implies that w equals 0 because this is a non-negative quantity and so, if it is going to be less than or equal to 0, the only way it is possible is if w equals 0.

And so, in this matrix, whatever we wrote here, it means that this is this u Hermitian Av is now reduced to a form where you have sigma and then 0s in the first column and then 0s in the first row other than the 1 comma 1th entry and this is an m minus 1 cross n minus 1 matrix and you can now apply the inductive argument on this and so, that basically completes the induction part of the argument.

So, now apply the inductive assumption to the m minus 1 cross n minus 1 matrix C, this is similar to what we did in the case of the Schur theorem proof and so, we can and then that completes the induction argument. Now, there is one last part of the proof where I need to show that u Hermitian Av is equal to sigma.

(Refer Slide Time: 23:36)





So, suppose there exist u in C to the m by m, v in C to the n cross n and the existence of these we have shown in the, by the induction inductive argument such that u Hermitian Av equals S where S has 0s everywhere except for the top left r cross r block with say gamma 1 through gamma r on the diagonal. We need to show that these gamma 1 through gamma r are equal to sigma 1 through sigma r.

Student: Excuse me sir. Sir, how we got rid of that under root of sigma square plus w Hermitian w?

Professor: So, if you square both sides of this equation, I will get sigma squared plus w Hermitian w squared equals, so there is a square missing here and then.

Student: Yes, sir.

Professor: And then I will get sigma squared plus w Hermitian w, and I am cancelling that on the this side, so I have this thing squared and sigma equals norm of a2 so sigma squared is equal to norm of a2 squared, so this and this cancel. And so from that, I get w Hermitian w less than or equal to 0. So thanks, there is a square missing here.

So, we can assume gamma 1 greater than or equal to gamma 2 greater than or equal to et cetera gamma n gamma r without loss of generality, because if not, we can always permute the columns of u and v to make these in decreasing order. Now, u Hermitian Av equals S means that Av equals u times S, I am just pre multiplying by u. So, that A times the ith column of v is equal to gamma i times the ith column of u and this is true for i equal to 1 through r.

This is just writing out what this means because S is a diagonal r cross r diagonal sub block and everything else is equal to 0 and similarly, u Hermitian times A and my right multiplying by v Hermitian is equal to S v Hermitian which then implies ui Hermitian times A is equal to gamma i times vi Hermitian for i equal to 1 to r.

So, what this means is that if I look at let me take the Hermitian of this. So, A Hermitian times ui is equal to gamma is a real so it is just gamma i times vi, so that if I look at and ui, so, if I take this gamma i to the other side I can write this as A Hermitian A vi is equal to gamma i squared times vi i equal to 1 to r.

So, that implies that gamma 1 squared up to gamma r squared are the non-zero Eigen values of A Hermitian A. So, by definition, this implies that gamma i equals sigma i because we defined sigma i is to be the square roots of the eigenvalues of A Hermitian A for i equal to 1, 2 up to r So, that completes the proof. So, the, so, one the couple of more remarks.

So, the columns of v the full set of the orthonormal eigenvectors of the matrix A Hermitian A and the columns of u are a full set of orthonormal eigenvectors of A A Hermitian, I mean it is possible that there are multiple subsets but they this v are one such full set of orthonormal eigenvectors this by construction on how we build this proof.

And then also another consequence of this proof is that this, so, this what we defined to be sigma which is so, this is so that the spectral norm of A is the largest singular value of A. We saw that sigma is one of the singular values and it is in fact the largest singular value of A this 2 for a square matrix is equal to square root of mu, where mu is equal tos the largest of A Hermitian A. So, you can try to relate these on your own later.