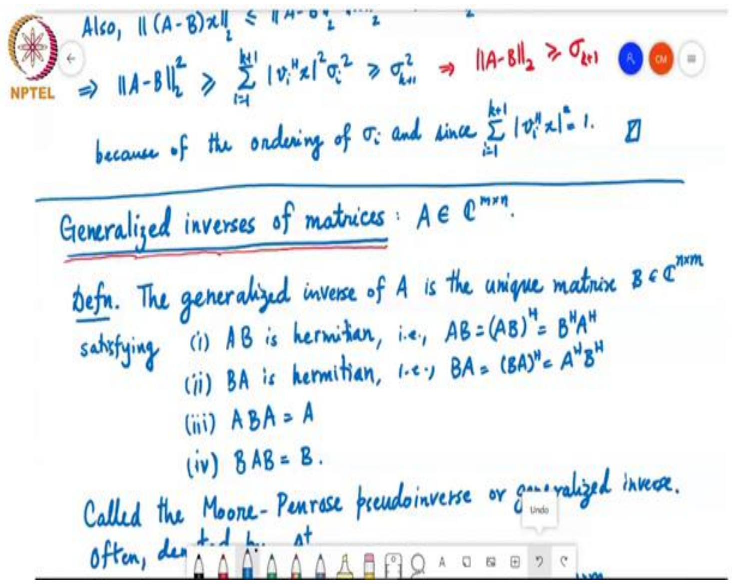
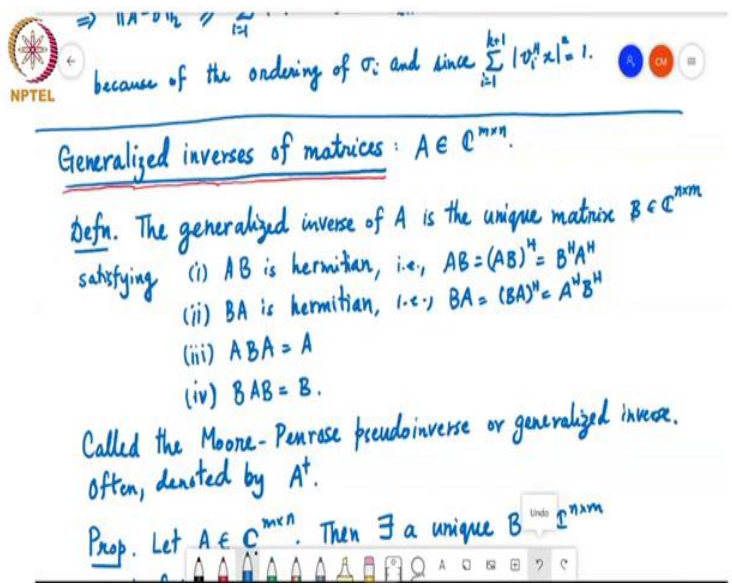
Matrix Theory
Professor Chandra R Murthy
Department of Electrical Communication Engineering
Indian Institute of Science Bangalore
Generalized inverse of matrices

(Refer Slide Time: 0:19)



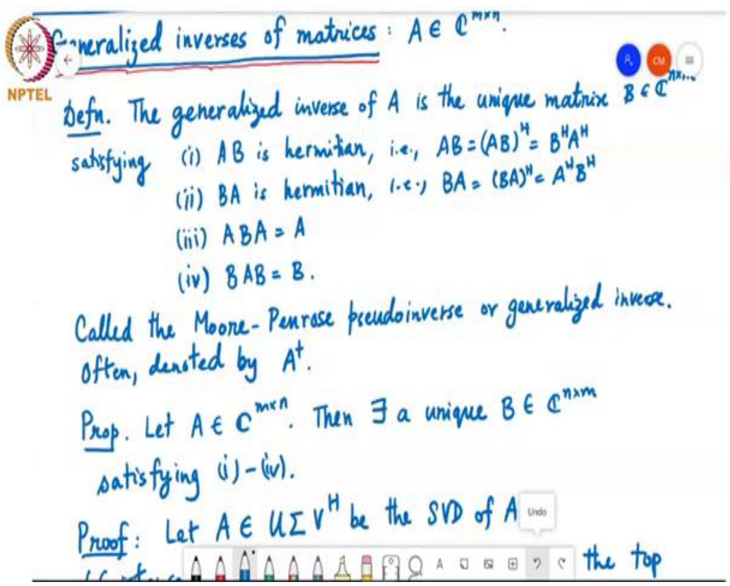


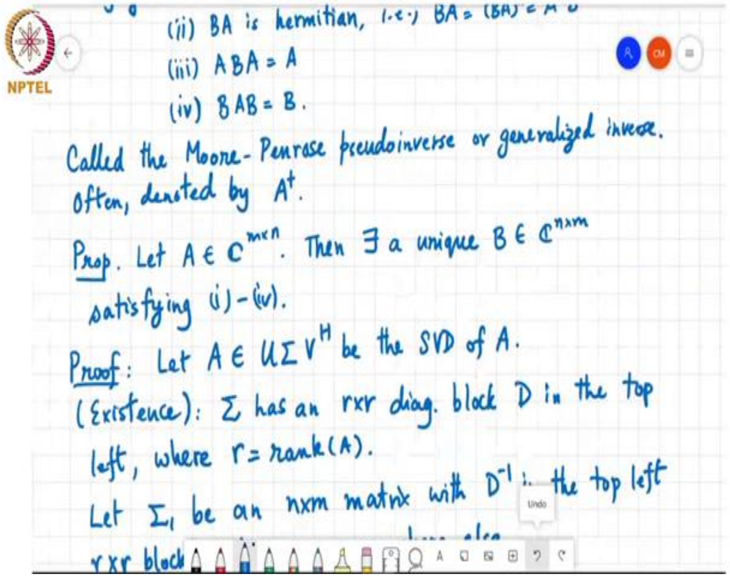
So, the next topic is generalized inverses of matrices. So we know that if a matrix is square and non singular, we can always invert it. But in order to solve systems of linear equations and even for many other problems, we will need to invert matrices that are rectangular of size m by n. So, for that we use this concept of generalized inverses. And these generalized inverses have various properties that we can also study.

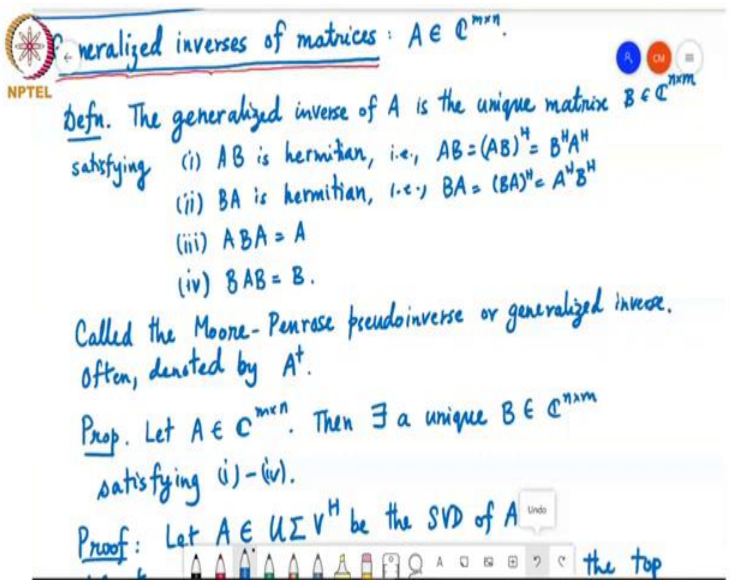
So, we will start with the basic definition, the generalized inverse of a matrix A size m by n is the unique matrix B of size n by m. So, note that the dimensions are reversed, it is of size n by m, satisfying four properties. The first property is that AB is a Hermitian symmetric matrix, that is AB equals AB Hermitian, which, of course is B Hermitian A Hermitian.

Similarly, BA is a Hermitian matrix that is BA is equal to BA Hermitian, which is equal to A Hermitian B Hermitian. Then, A times B times A equals A, B times A times B equals B. So, the matrix B that satisfies these four properties is a unique matrix and is called the generalized inverse or the Moore Penrose pseudo inverse of this matrix A, can be often denoted by A dagger. So, the fact that this is a unique matrix that satisfies all these four properties is the essence of this proposition.

(Refer Slide Time: 2:02)







So, let A be a matrix of size m by n, then there exists a unique matrix B, which satisfies 1 to 4. So, what is saying, it is in fact saying two things. One is that it is always possible to satisfy these properties 1 to 4. And there is only one matrix that satisfies all four of these properties, very interesting proposition.

So, in other words, you can find a pseudo inverse for any matrix A, no exceptions. And that pseudo inverse that you find there is no other one. It is a unique one. So here is how the proof goes. So again, we start with the singular value decomposition of the matrix A u sigma v Hermitian.

So first, we will talk about existence that is to show that there exists a matrix B, which satisfies these four properties.

Student questioning: Sir.

Professor: Go ahead.

Student: Sir, that mean that any (())(3:12) is a square matrix of not full rank, there exist this pseudo (())(3:17)?

Professor: Can you repeat your question?

Student: If A is a square matrix of rank lesser than full rank, I mean, not full rank. Does pseudo inverse exist for that also?

Professor: Always.

Student: Okay, thank you.

Professor: Yeah. So you can find a pseudo inverse of any matrix square, rectangle, anything. It always exists, and it is unique. You cannot find two of them. It is a very intriguing property.

Student: Sir one more question.

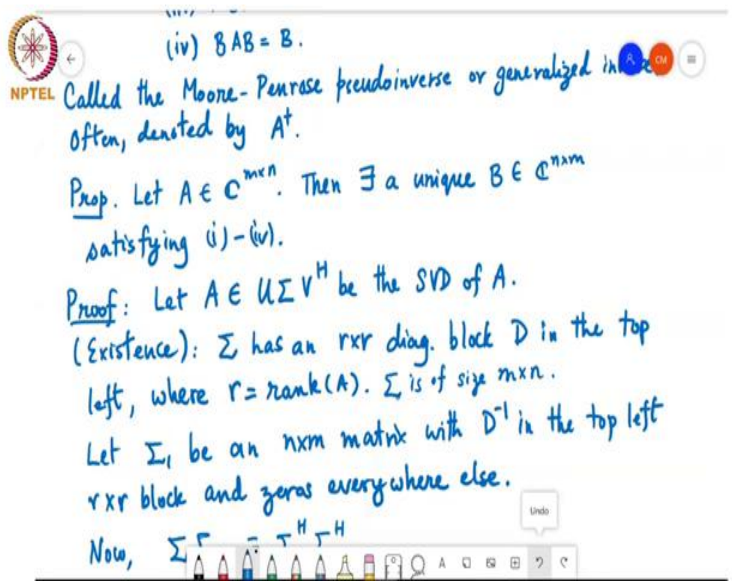
Professor: Yeah.

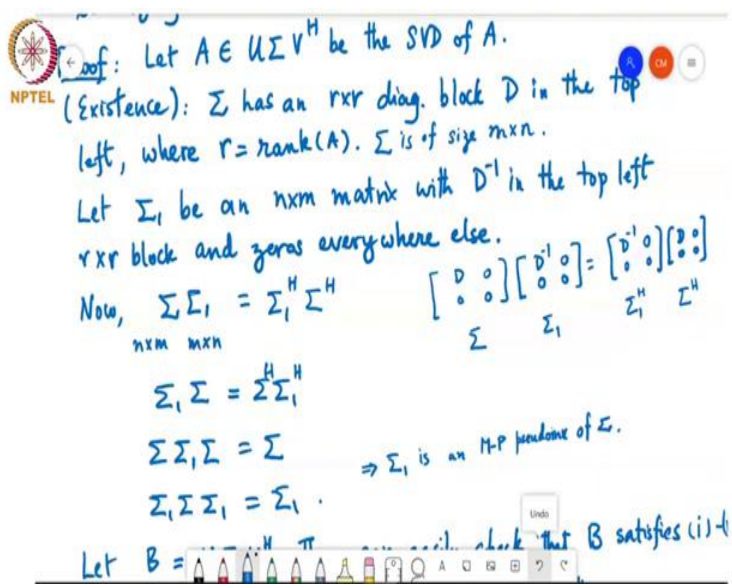
Student: Does the pseudo inverse equal to inverse?

Professor: Yes. So we will see that actually, that when, if the matrix A is square and invertible, the pseudo inverse equals the inverse. In fact, we can see something about that right away. Suppose A was a square matrix, and it was invertible. And if B was equal to A inverse, then if I take AB, AB is the identity matrix.

Yes, of course, it is Hermitian. Similarly, BA is the identity matrix. Of course it is Hermitian. If I do ABA, that BA cancel with each other, and I will be left with A, if I do BAB, AB cancel with each other, I will be left with B. So the normal inverse also satisfies all these four properties when A is square and invertible, that is the reason why it is called a generalized inverse, it is not telling you something new when you go to square matrices that are invertible, it is the same inverse that you will get if you were to invert the matrix.

(Refer Slide Time: 5:15)





So, this sigma here has an r cross r diagonal block D in the top left, where r is equal to the rank of A. So that is our standard SVD. Let sigma 1 be an n by m matrix, with D inverse in the top left r cross r blocks and everywhere, 0s everywhere else. So, this is of size m by n. And sigma 1 is of size n by m. So, the opposite size. What we will first show is that sigma 1 is the Moore Penrose pseudo inverse of sigma.

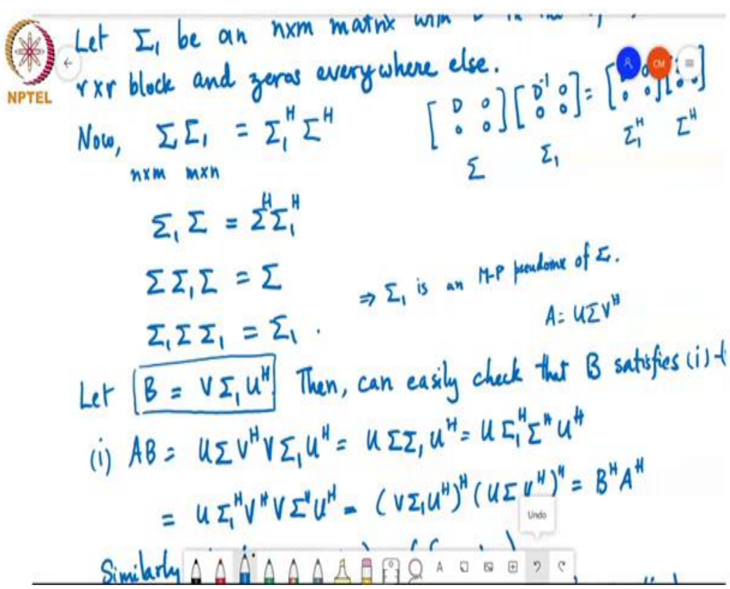
So, if I take sigma sigma 1. That is equal to sigma 1, so this you just have to write it out what it looks like. So, this will be like a D with 0, 0, 0 here, and then this sigma 1 will be D inverse. These 0s are all of different dimensions, this is a size n by m, this is a size n by n. And so I will get the identity and 0, 0s. And that is exactly equal to sigma 1 Hermitian times sigma Hermitian, which is to say that the Hermitian of this matrix is equal to itself.

So, sigma sigma 1 is Hermitian so it satisfies property 1, sigma 1 sigma is equal to sigma Hermitian sigma 1 Hermitian. So, this also satisfies that this matrix is a Hermitian symmetric matrix, sigma times sigma 1 times sigma equals sigma. So, I will just write this like this, sorry. This is actually a bad way to write it, because this is sigma, this is sigma 1, this is sigma 1 Hermitian. And this is sigma Hermitian. So, they are all of different dimensions, but you have, so, basically the 0s are all of different dimensions, but by just matching up the dimensions, you can verify that this is actually true.

And similarly, sigma sigma 1 times sigma is equal to sigma and sigma 1 sigma times sigma 1 equals sigma 1. So, basically sigma 1 is the more Penrose, for now, since I have not said shown uniqueness, I can say that it is an more Penrose inverse of sigma.

Student: Sir, is he the Penrose who got the Nobel price?

(Refer Slide Time: 8:48)



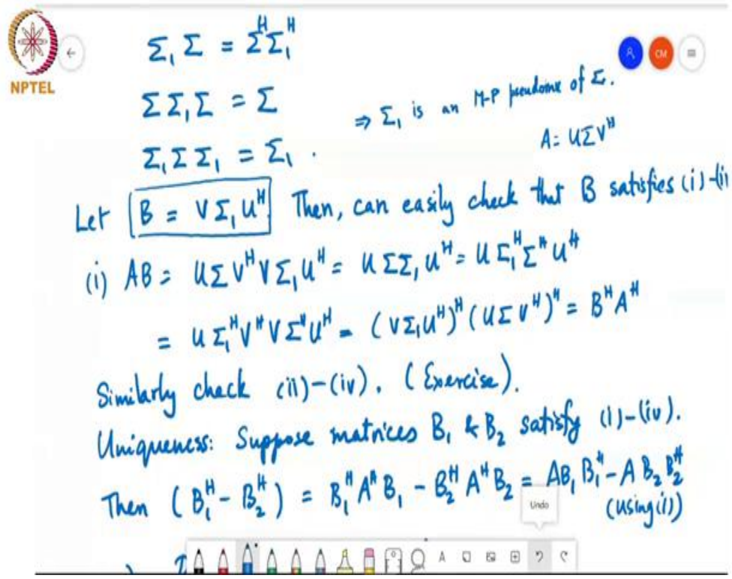
Professor: So, coming back to, coming back to our proof. Suppose B, now okay, so let us define B to be v times sigma 1 times u Hermitian. So, what is B? If I have the SVD of A, I will write it here A is equal to u sigma v Hermitian. All I have done is to take the nonzero r cross r block in sigma, invert that and put that as a matrix sigma 1 of size n by m pre multiplied by v post multiplied by u Hermitian. Interesting matrix, it is easy to obtain once you have the singular value decomposition of A.

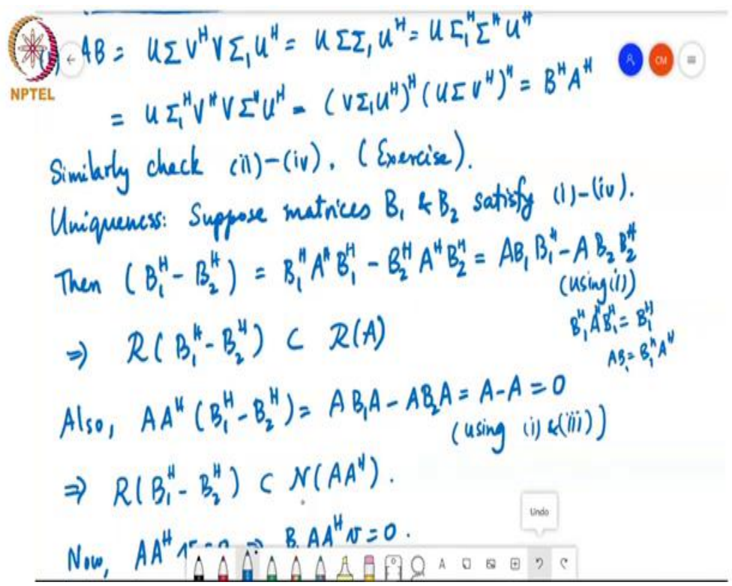
Then we can easily check that B satisfies properties 1 through 4. That is, for example, if you look at AB that is u sigma v Hermitian v times sigma 1 times u Hermitian, just substituting for A and B, but we v Hermitian v is the identity matrix. So it is sigma sigma 1 u Hermitian and sigma sigma 1, sigma sigma 1 is equal to sigma 1 Hermitian sigma Hermitian. So, it is sigma 1 Hermitian, sigma Hermitian times u Hermitian.

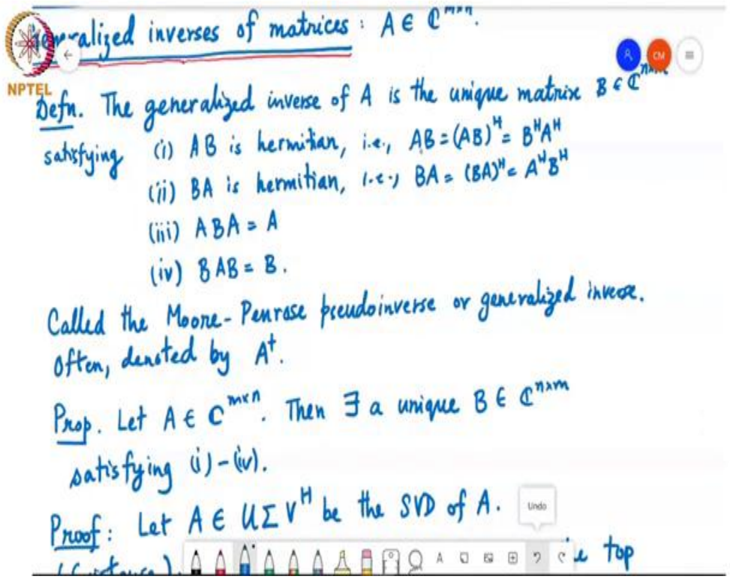
And now I can re insert a v Hermitian v in between these two matrices and write it as u sigma 1 Hermitian v Hermitian, v sigma 1 Hermitian u Hermitian. And this itself is basically B Hermitian, if B is this then B Hermitian is u sigma 1 Hermitian times v Hermitian. So, that is B Hermitian and this quantity here is just A Hermitian, A is u sigma v Hermitian. So, A Hermitian would be v sigma Hermitian u Hermitian. So, this is equal to B Hermitian A Hermitian which is the same as AB whole Hermitian.

So, AB is a Hermitian symmetric matrix, in a similar way you can verify properties 2, 3, 4. And so, this matrix B that we defined here it satisfies the four properties required by the Moore Penrose pseudo inverse.

(Refer Slide Time: 11:01)







So, we now know that there exists a Moore Penrose pseudo inverse. Now, the uniqueness is a little more complex. So, suppose there are two different matrices B1 and B2 that satisfy all four properties, 1 to 4. Then if I look at B1 Hermitian minus B2 Hermitian that is the same as B1 equals B1 AB1.

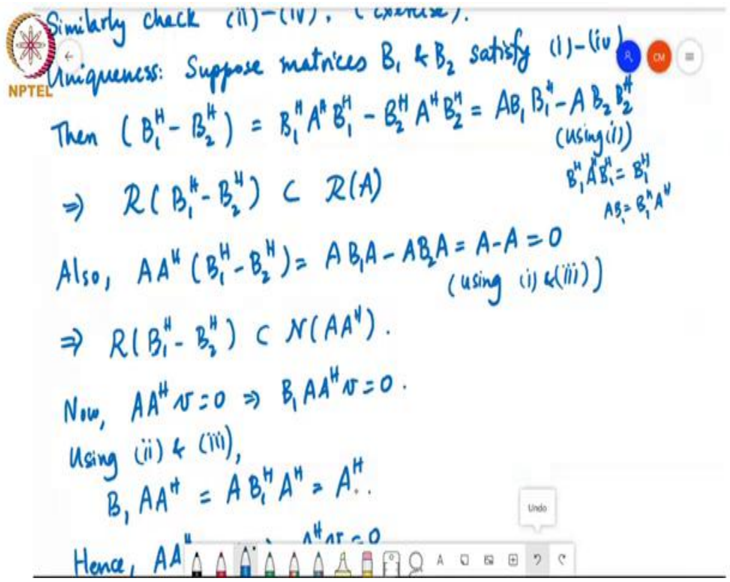
So, for example, we know that so if I take the Hermitian so B1 Hermitian is same as B1 A B1 Hermitian. So, B1, sorry B1 Hermitian is B1 A Hermitian B1 Hermitian, so B1 Hermitian A B1 Hermitian minus B2 Hermitian A B2 Hermitian So, that is the difference between B1 Hermitian and B2 Hermitian. But by property one AB is equal to B Hermitian A Hermitian.

So, I have B1 Hermitian A Hermitian which is equal to AB1. So, A B1 B1 Hermitian minus A B2 similarly, for this A B2 B2 Hermitian. So, this difference is equal to A B1 B1 Hermitian minus A B2 B2 Hermitian. So, you can see that this is equal to A times B1 B1 Hermitian minus B2 B2 Hermitian.

So, the range space of B1 minus, B1 Hermitian minus B2 Hermitian is a subset of the range space of A because you can write it as A times some matrix, so, it will have to these columns will have to lie in the column space of A. Similarly, using these properties 1 and 3, if you look at AA Hermitian times B1 Hermitian minus B2 Hermitian you can write this as AB1 A minus AB2 A, but AB1 A equals A and AB2 A equals A. And so, that is A minus A which is 0 matrix.

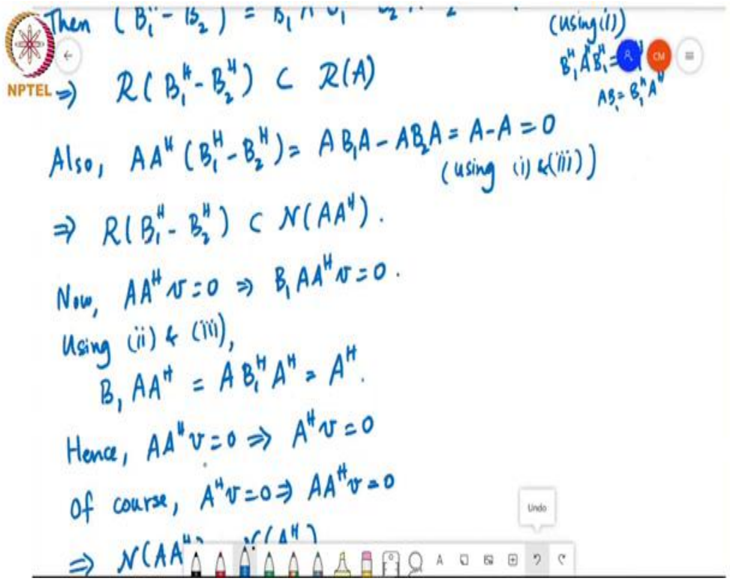
And so, what that means is the range space of B1 minus B2, any vector belonging to the column space of this matrix will lie in the null space of AA Hermitian. So, this span of B1 Hermitian minus B2 Hermitian lies in the range space of A, it also lies in the null space of AA Hermitian.

(Refer Slide Time: 14:19)



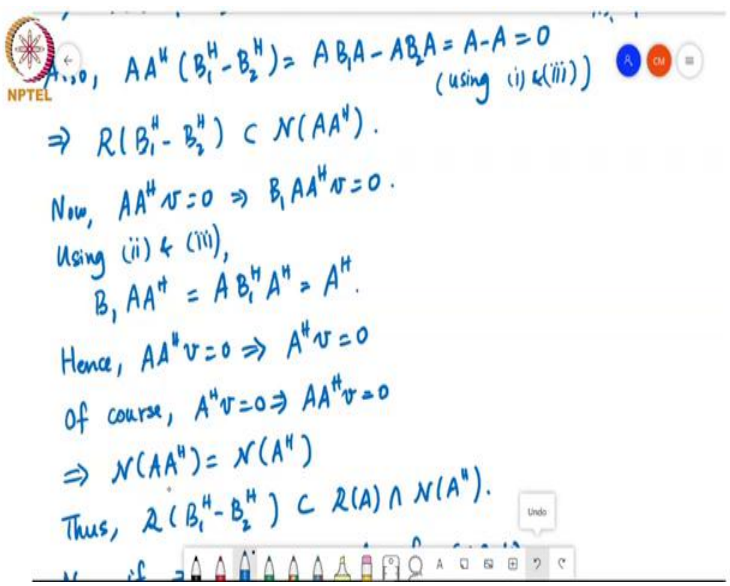
Now, if a vector lies in the null space of AA Hermitian that means AA Hermitian v equals 0, which then means that just B multiplying by B1, B1 times AA Hermitian v equals 0. Now, you should verify this, that again I am just repeatedly using these properties. It means that I mean this, these properties mean that if I look at B1 AA Hermitian, this matrix here that is the same as AB1 Hermitian times A Hermitian, which is equal to A Hermitian.

(Refer Slide Time: 14:57)



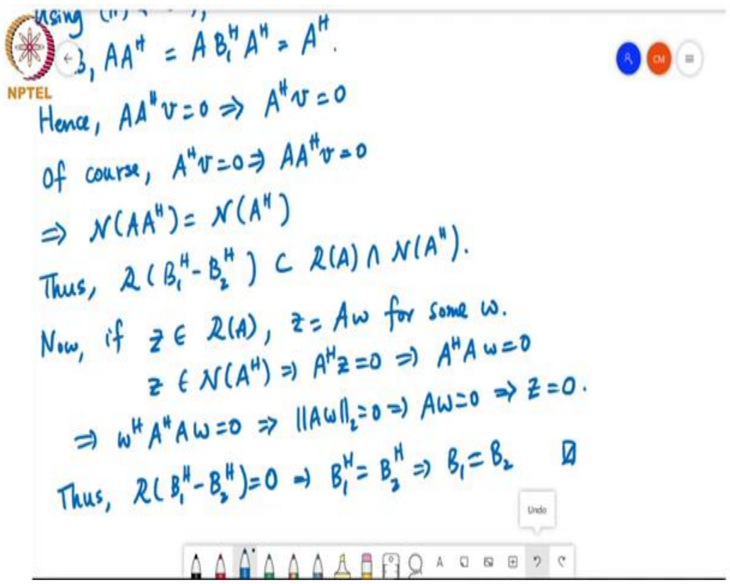
So, AA Hermitian times v equals 0 then implies then implies that A Hermitian v equals 0.

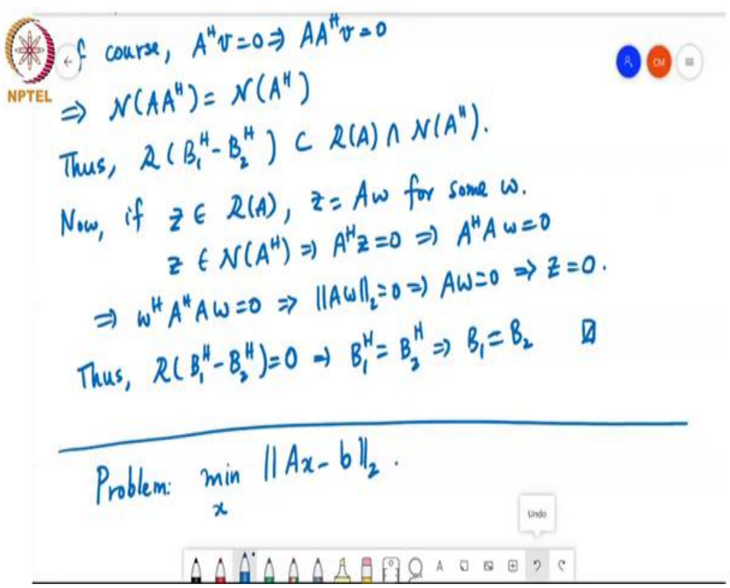
(Refer Slide Time: 15:15)



So, that means that the null space of AA Hermitian is equal to the null space of A Hermitian itself. Almost done.

(Refer Slide Time: 15:27)





So, the upshot of that is that the range, so these two are the same and the range space of B1 Hermitian minus B2 Hermitian is a subspace, subset of both the range space A and null space of A Hermitian. Now, if you recall that this is actually the null set, then you are done. So basically, that is what the rest of the proof does, is that if there is a Z, which lies in the range space of A, then Z can be written as a linear combination of the columns of A for some W.

And similarly, if Z also belongs to the null space of A Hermitian, then A Hermitian Z equals 0. And substituting for Z, I have A Hermitian A times W equals 0, which means that W Hermitian times A Hermitian AW equals 0, which means that AW equals 0, so but Z equals AW, so Z must be equal to 0.

So, this intersection space contains only the 0 vector. So the range space of B1 Hermitian minus B2 Hermitian is 0, which in turn means that B1 Hermitian equals B2 Hermitian, or B1 equals B2. So that concludes the proof. So that is all I have for today. The next class, which is the last class of the course, we will briefly discuss least squares problems.

So, the what we are working up towards is to figure out how to solve Ax equals B. And in particular, in the case where there may not be an exact solution. So those are called least squares problems. In other words, we will ask, what is the, how do we find a solution to the problem minimize with respect to x, Ax minus bl2 norm and so all these pseudo inverses basically will show up there and help us solve this problem. So that is it for today. And we will meet again on Friday.