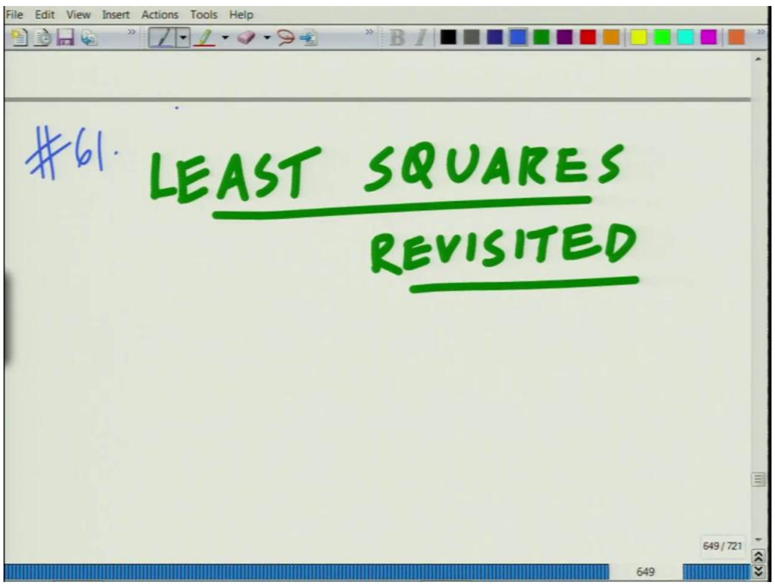
Applied Linear Algebra for Signal Processing, Data Analytics and Machine Learning
Professor Aditya K Jagannatham
Department of Electrical Engineering
Indian Institute of Kanpur
Lecture 61
Least Squares Revisited: Rank Deficient Matrix

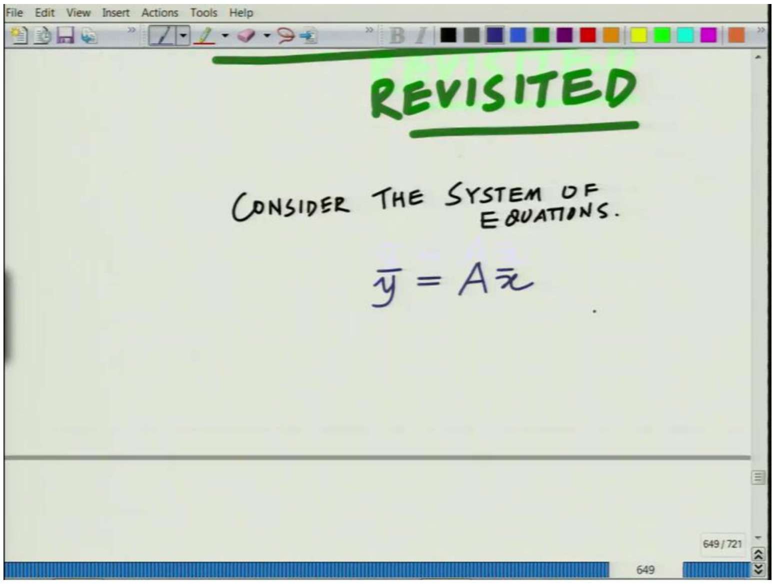
Hello, welcome to another module in this Massive Open Online Course. So, in this module let us start let us, again revisit the Least Squares problem and discuss and try to elaborate on it a little more by looking at some special cases or certain special properties of the least squares.

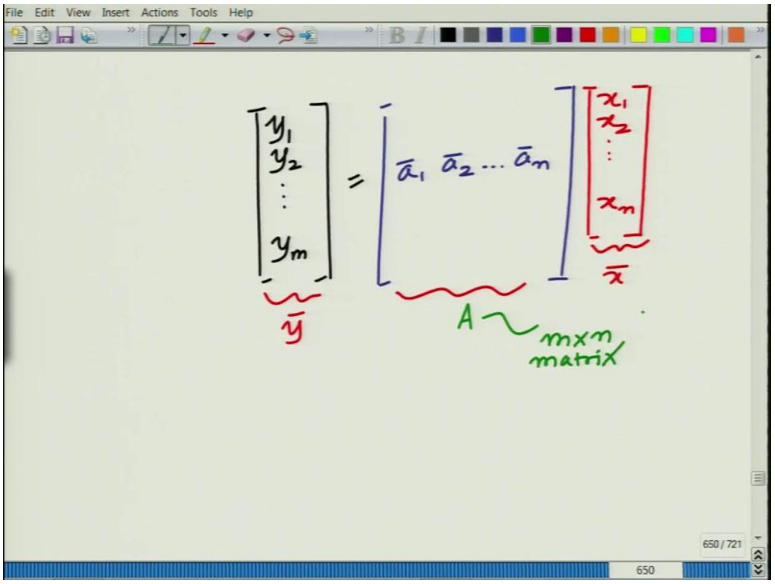
(Refer Slide Time: 00:38)



So, let us again take a relook. So, I would like to revisit the least squares problem and understand this better especially looking at some interesting cases. Now, you remember and remember the least squares is a very popular framework has several applications in wireless communication, signal processing, machine learning, and everywhere essentially, linear regression and so on and so forth. So, let us take a deeper look at it.

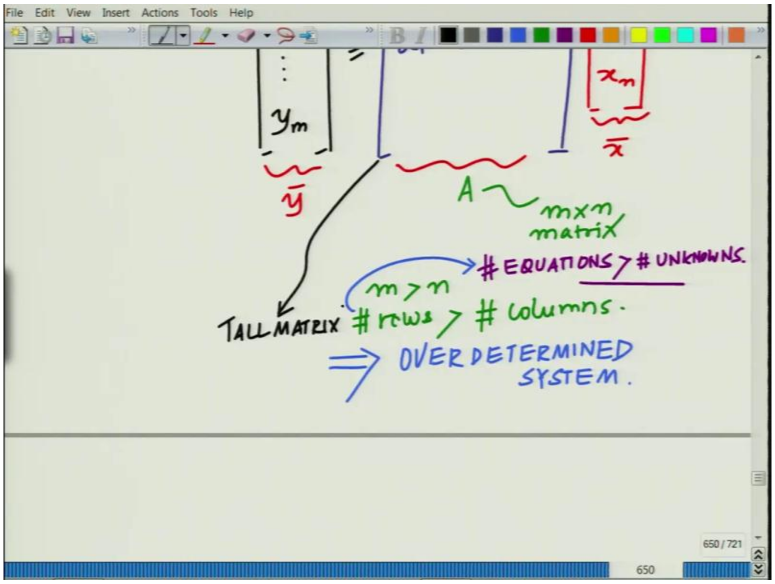
(Refer Slide Time: 01:24)

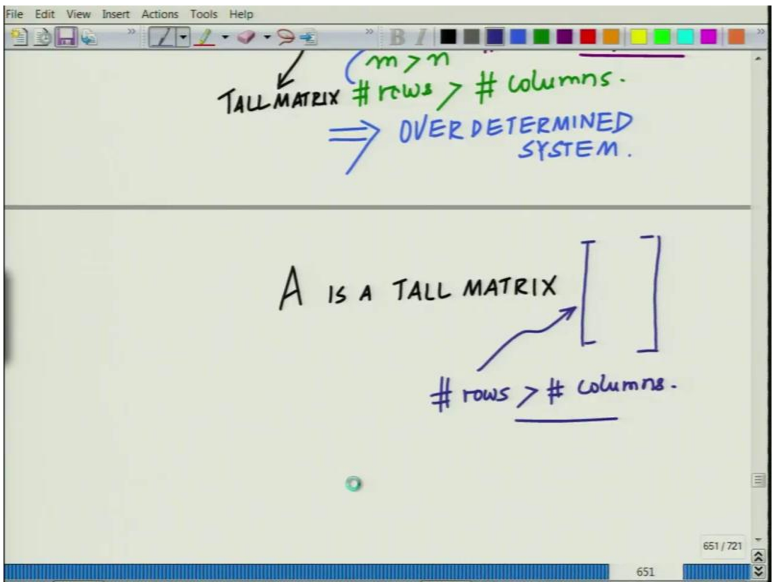




Now, remember the least squares problem is essentially where we have the system of equations. So, we have the system of equations y bar equal to A x bar, which if you look at it essentially I can write it as follows, I can write the vector y 1, y 2, y m this is equal to the matrix A which can be represented in terms of its columns a 1 bar, a 2 bar, a n bar times the vector x 1, x 2 up to x n. So, this is essentially your vector y bar, this is essentially your matrix A, and this is the vector x bar, and this we are saying is essentially your m cross n matrix, this is your m cross n matrix.

(Refer Slide Time: 03:08)



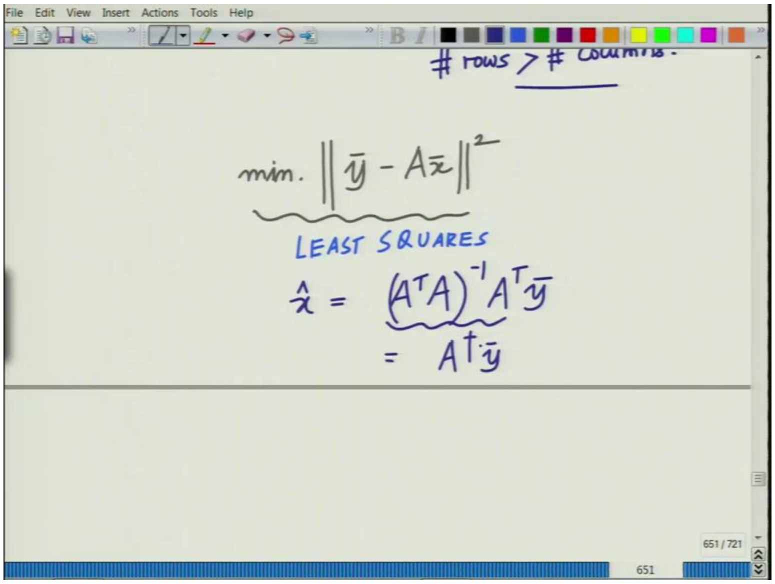


And, for the least squares problem we have m greater than n that is number of rows greater than number of columns implies this is an over determined system. So, this is an over determined system of equations. Because, essentially when the number of rows greater than is greater than the number of columns it essentially means that the number of variable, number of equations which is m is greater than the number of variables that is n.

So, this essentially implies, if you look at this essentially implies that number of equations is greater than number of unknowns, number of equations is greater than the number of unknowns is over determined system and also since the number of rows is greater than number of columns is A informally colloquially this is also known as a tall matrix. So, A is a tall matrix.

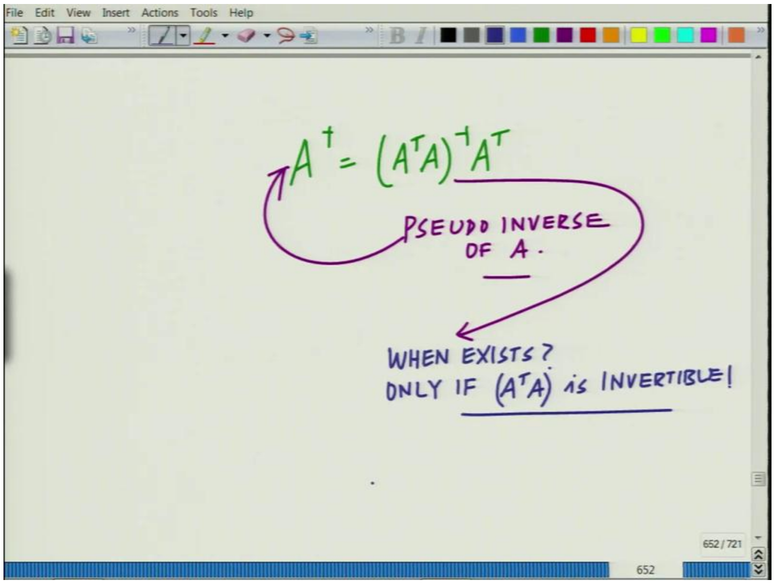
A is tall matrix which essentially, looks something like this that is essentially, this essentially means that number of rows is greater than the, number of rows is greater than the number of columns. And now, essentially what we say is to solve this system of equations remember typically for such an over determined system a solution does not exist therefore, we instead saw look for the x that minimizes the error that is we form.

(Refer Slide Time: 05:38)



Look at the error, y minus A x bar which when we look at the norm of the error and we look at the norm square and we minimize this and this is the least square error and this is essentially what we term as the least square solution, which is essentially the least square norm the implication of the meaning of that is we are looking at the least square norm solution and remember the solution to this is given as, now this is essentially the key x equal to A Transpose A Inverse A Transpose y bar and this quantity we also the pseudo Inverse of A we denote this by A Dagger A y bar.

(Refer Slide Time: 06:45)

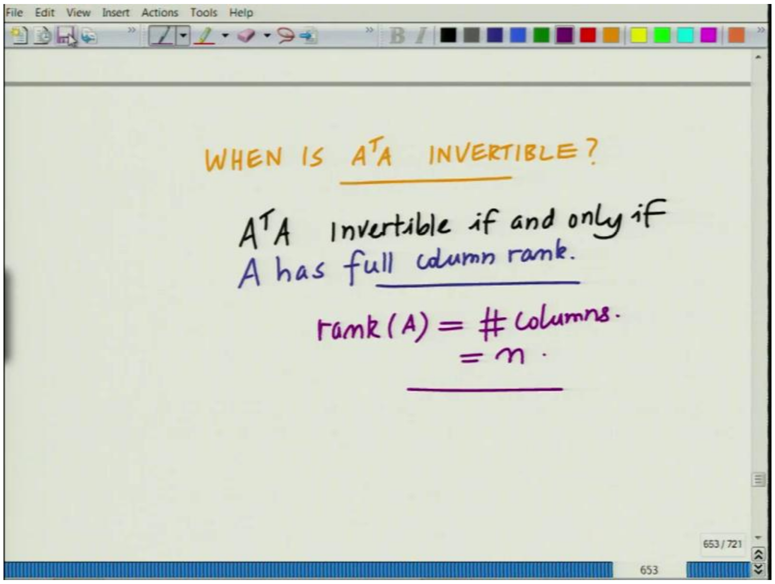


So, this is essentially you might remember this quantity A dagger which is essentially equal to A Transpose A Inverse A Transpose y bar this I am sorry, this is A Transpose A Inverse A Transpose this quantity this is termed as the Pseudo Inverse of A.

Now, the point is when does this exist? When can you write it in this fashion, remember we have not looked into that, this is only when A Transpose A, when does this exist? Now, the question that we want to ask is when it does it exist, if A Transpose only if that is only if, A Transpose A is, when this exists only if this quantity A Transpose A is invertible.

Now, A Transpose A is invertible when is? So, therefor we are asking the question when is A Transpose A invertible? When can we write it in this? Can we always write it in this form? That would be good, but turns out that we cannot write it in this form. So, always so we would like to understand under what conditions is this expression valid?

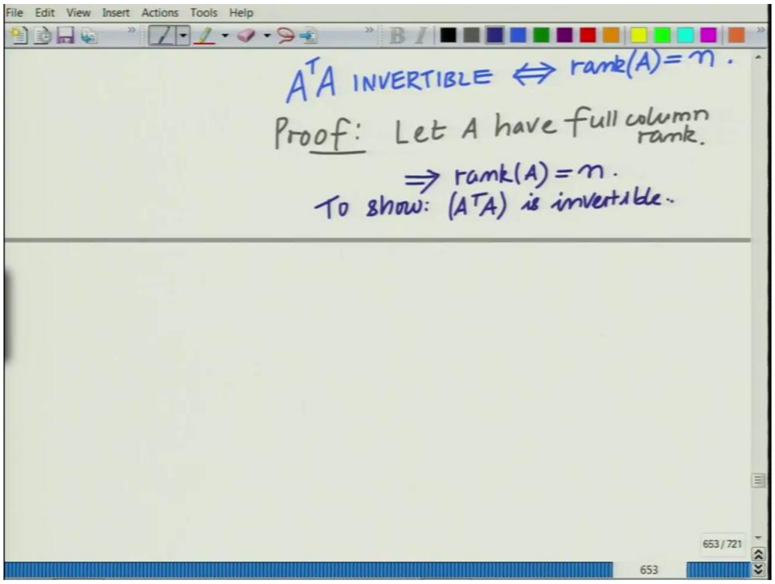
(Refer Slide Time: 08:37)



So, when is A Transpose A invertible? So, that is the question that would like to ask, when is A Transpose A, when is this quantity A Transpose A? And the answer is A Transpose A is invertible if and only if A the matrix A, if and only if A has full column rank, what is the meaning of that? Meaning of that is rank of A equals the number of columns equal to n.

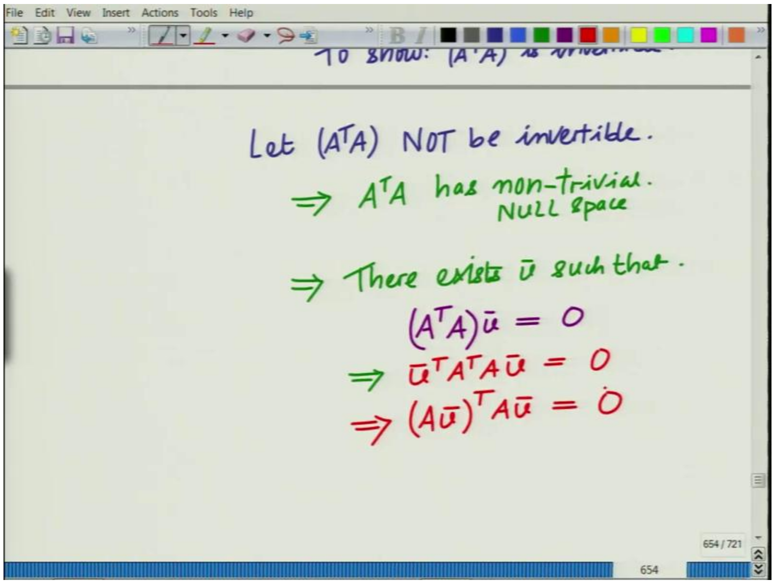
So, A Transpose A is invertible only if, remember A has n columns, m rows, m is greater than remember rank is less than equal to minimum of m comma n but, m is greater than n. So, minimum of m comma n is n. So, rank has to be less than equal to n. Now, if rank if A is actually equal to n that is if it has full column rank, then this matrix A Transpose A is invertible if and only if, that is A has full column rank this is invertible, this is invertible implies A has full column rank.

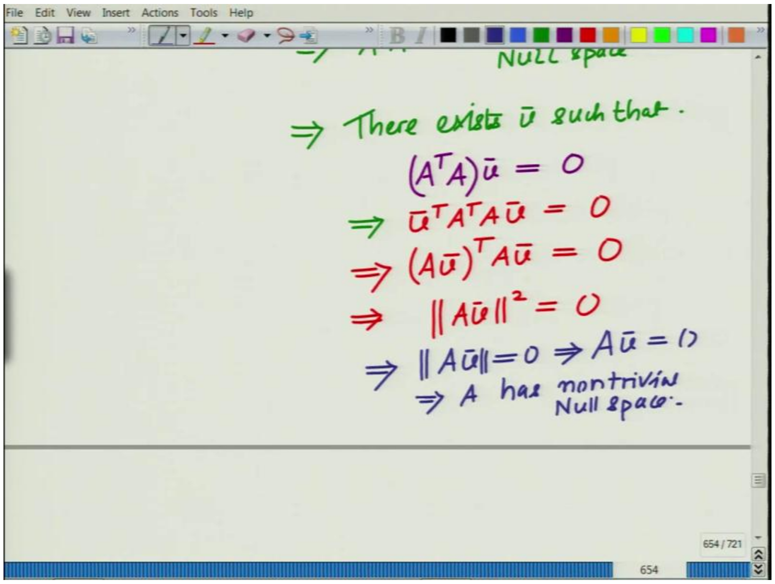
(Refer Slide Time: 10:42)



So, if you state it succinctly you can say A Transpose A invertible if and only if rank of A equal to n consider A Transpose A. Now, if A Transpose A does not have. Now, let us demonstrate this let prove us it. Let us, look at a proof of this. Now, let A have full column rank. Let us, start with one way let A have let A have full column rank. Now, if A has full column rank this implies, this implies rank A equal to n, equal to n. Now, to show A Transpose A is invertible.

(Refer Slide Time: 12:11)



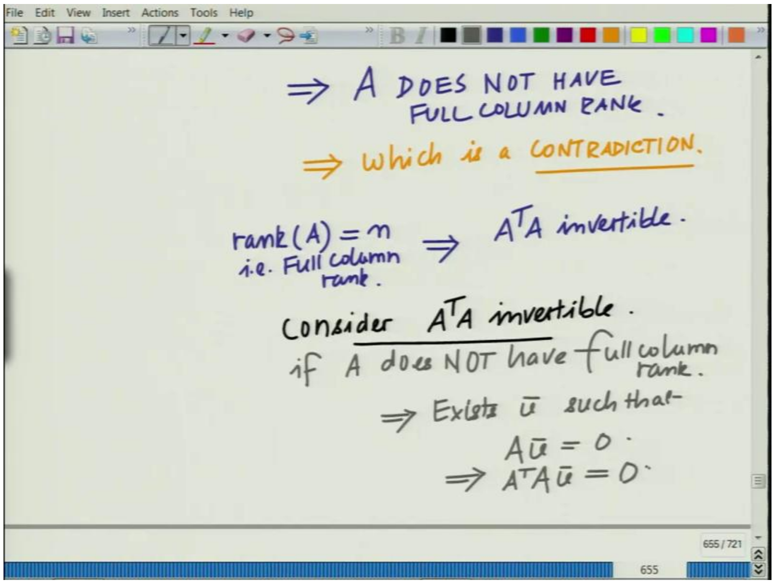


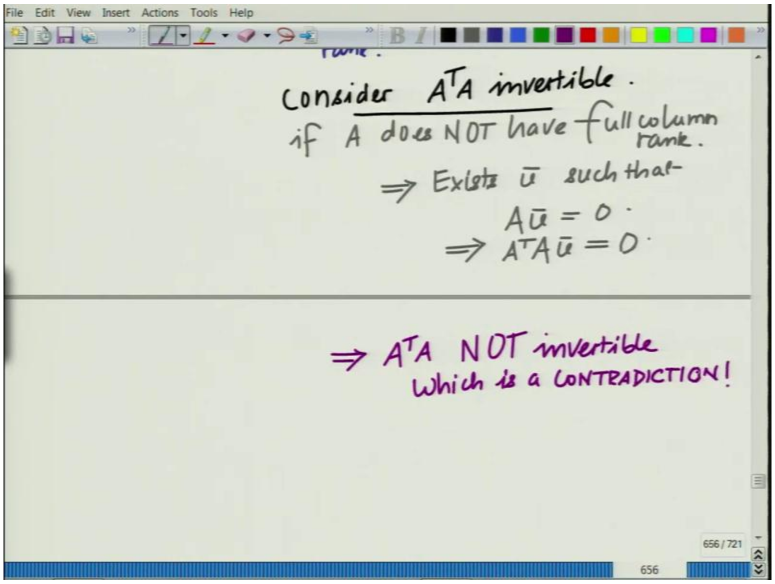
Now, let us assume A contradiction let A Transpose A not be invertible. Let, A Transpose A not be invertible this would imply that A Transpose A is singular, this would imply that there exists A vector u bar such that, A Transpose A into u bar equal to 0 remember it has a non-trivial null space that is the meaning of A Transpose A not invertible.

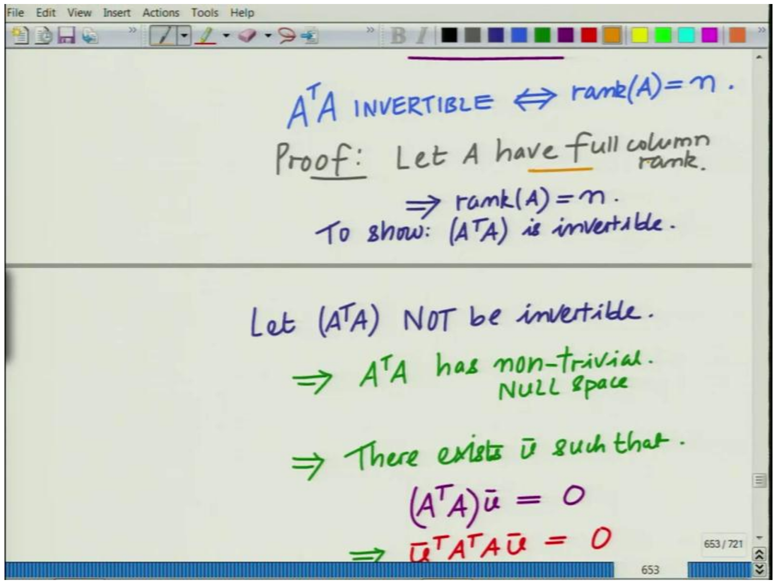
So, implies A Transpose A has non-trivial null space, this implies there exists a vector u bar. Such that, A Transpose A into u bar equal to 0. Now, multiply this on the left by u bar this implies that u bar Transpose A Transpose A u bar equal to 0, this implies that A u bar Transpose A u bar equal to 0.

But, vector Transpose itself is nothing but, the norm square of the vector, this implies that norm square of A u bar equal to 0 this implies therefore, since the norm square norm square is 0 this implies the norm is 0, this implies A u bar, since the norm is greater than equal to 0, this implies that the norm is equal to, A u bar equal to 0, this implies A is has non-trivial null space, this implies A has non-trivial null space.

(Refer Slide Time: 14:37)





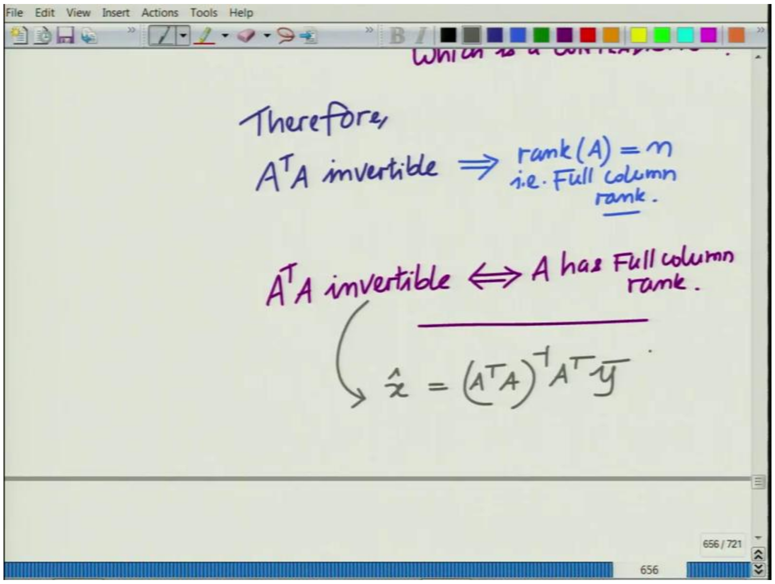


This implies, essentially implies A does not have full column rank, this implies A does not have full column rank, this implies that A which is a contradiction basically, which is a contradiction. Why is this a contradiction? Remember, we started with the assumption that A has a full column rank, we started with the assumption that A has the full column rank. So, A has full column. So, A has full column rank implies.

So, rank A equal to n i.e. full column rank this implies A Transpose A is invertible. Now, let us say A does not have full columns. Now, let us say A Transpose A is invertible. Now, let us say. Now, consider the other way consider, A Transpose A, let us look at this the other way around. Now, let us say A Transpose A invertible if A does not have full column rank, this implies there exists u bar such that, A u bar equal to 0.

This implies multiplying it by A Transpose we have A Transpose A u bar equal to 0 implies A Transpose A implies A Transpose A, implies A Transpose A is not invertible, A Transpose A not invertible, which is a once again a, which is a contradiction, which is once again this is a contradiction.

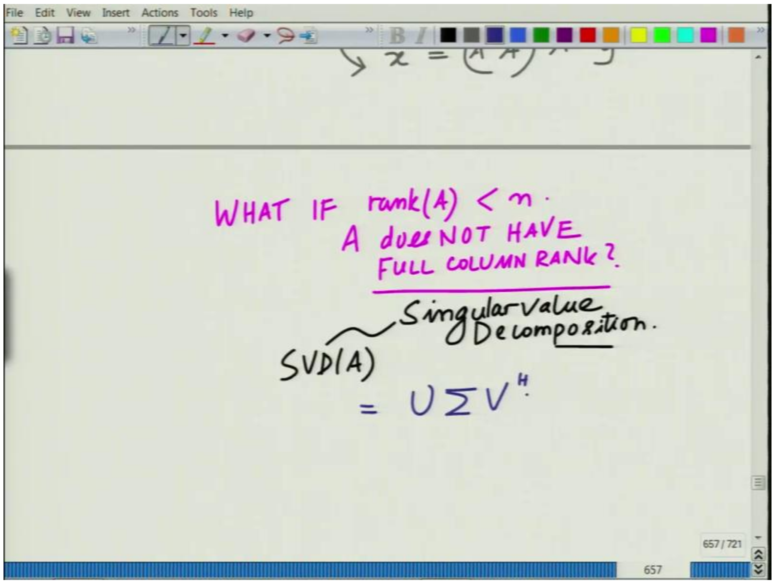
(Refer Slide Time: 18:04)

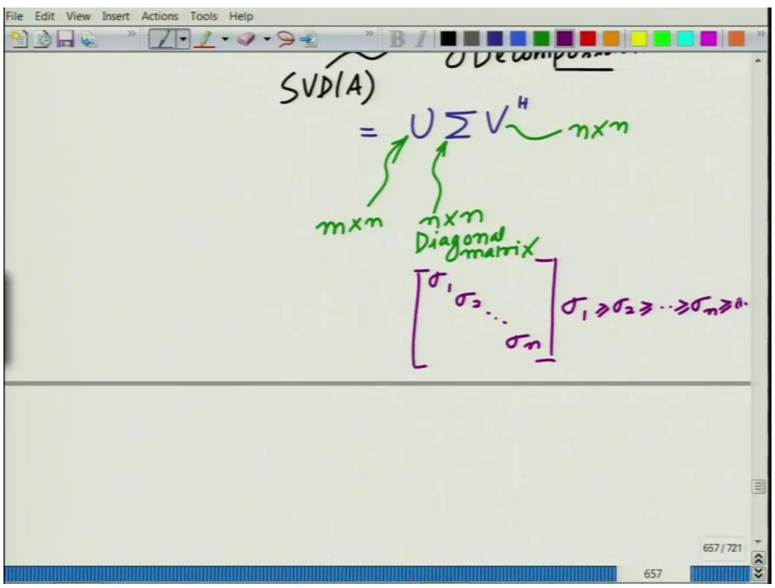


Therefore, we must have. Therefore, A Transpose A invertible. So, therefore A Transpose A invertible must imply A, must imply that rank A equal to n that is full column rank. So, therefore what we have is that A Transpose A is invertible if and only if A has full column rank. So, this completes the proof.

So, essentially what this implies is? A Transpose A is invertible if and only if A has full, A has full A Transpose A is invertible if and only if A has full column rank. So, that is when you can write and when A Transpose is invertible and therefore, in such a situation when A Transpose is invertible we have x hat equal to A Transpose A Inverse A Transpose times y bar. So, that is essentially what it is. Now, what if A Transpose is not invertible that is A is not full column rank what then? So, what if A is not full column rank?

(Refer Slide Time: 20:14)

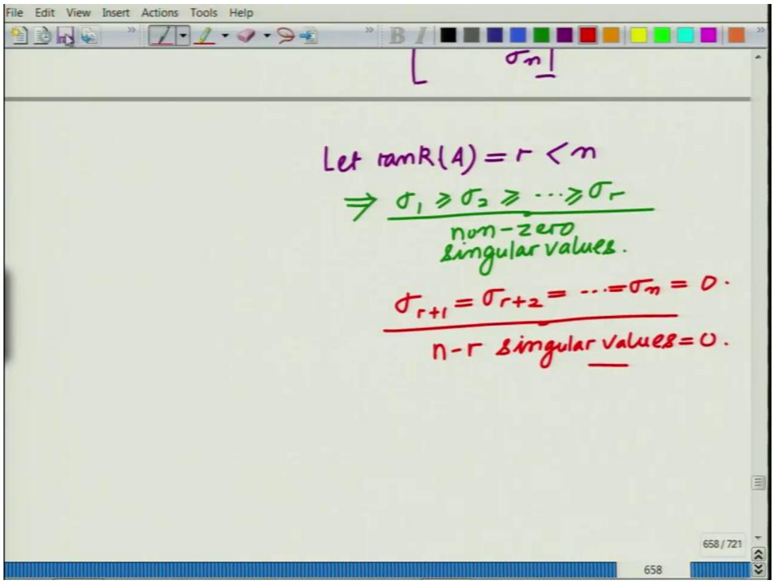




Now, what if that begs the question? What if rank A is less than n, that is, A does not have, what if A does not have full column rank? Then we start with the SVD of A, remember SVD stands for the singular value decomposition, we start with the SVD of A this is U sigma V Hermitian, where if you remember this matrix sigma, remember. Now, remember for a tall matrix this is given as an m cross n matrix, this V will be an n cross n matrix and sigma will be an n cross n diagonal matrix.

So, this will be an n cross n diagonal matrix of the singular values, you have sigma 1, sigma 2, so on up to sigma n and these are non-negative arranged in decreasing order that is essentially your properties of the singular value decomposition. And now, what happens is since A is not full column rank, if A is not full column that implies that some of the singular values are 0 remember the rank of the matrix equals the number of non-zero or the number of positive singular values.

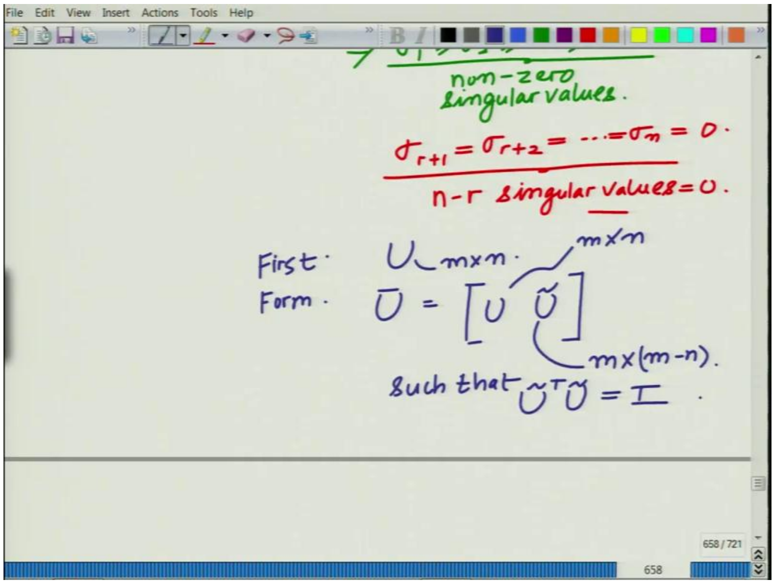
(Refer Slide Time: 22:29)



So, essentially, let rank A equal to, since we are considering A not rank deficient let, rank A equal to r that is less than n, this implies sigma 1 greater than equal to sigma 2 greater than equal to sigma r these are non-zero singular values.

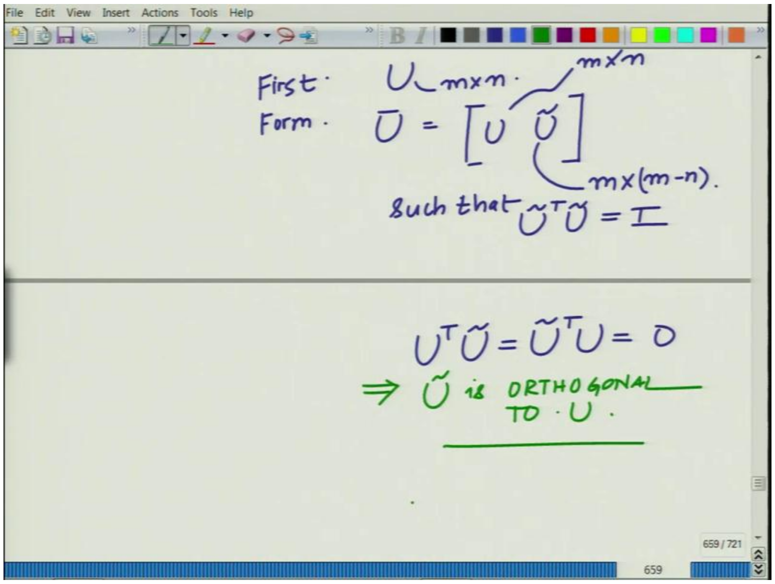
And then, we have sigma r plus 1 equals to sigma r plus 2 so on until sigma n these are 0, these are n minus r, you have n minus r singular values equal to 0. And now, how do you solve this least squares problem. So, let us start out, let us set the model for today in this module and we will continue with this in the subsequent module.

(Refer Slide Time: 24:00)



Now, the key to this is first let us expand this. First we have U this is of size m cross n. Now, we form the matrix U bar which is the matrix U and you have U Tilde. Now, this is m cross n, this we can make it as m cross m minus n. So, you are adding m minus n such that, U tilde is also orthonormal U Tilde Transpose U tilde equals identity, you can always find U Tilde because, you can basically you can give any m dimensional space, you can find a set of m orthonormal basis vector. So, you can already have n you are finding m minus n moreover, and put them together you have this matrix U bar.

(Refer Slide Time: 25:18)



So, U Tilde U Transpose is identity and further U Transpose U Tilde equals U tilde Transpose U equal to 0 that is U Tilde orthogonal to U. So, essentially what we are going to do at this point, so let us stop here in this module. So, now we have constructed this matrix U. And now, how we are going to use this singular value decomposition to solve this least squares problem when A is not full column rank is what we are going to discuss in the next module. Thank you very much.