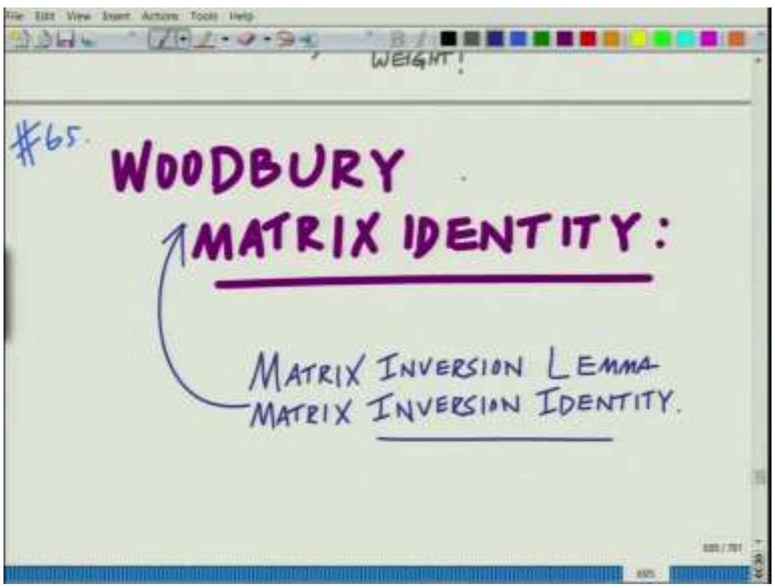
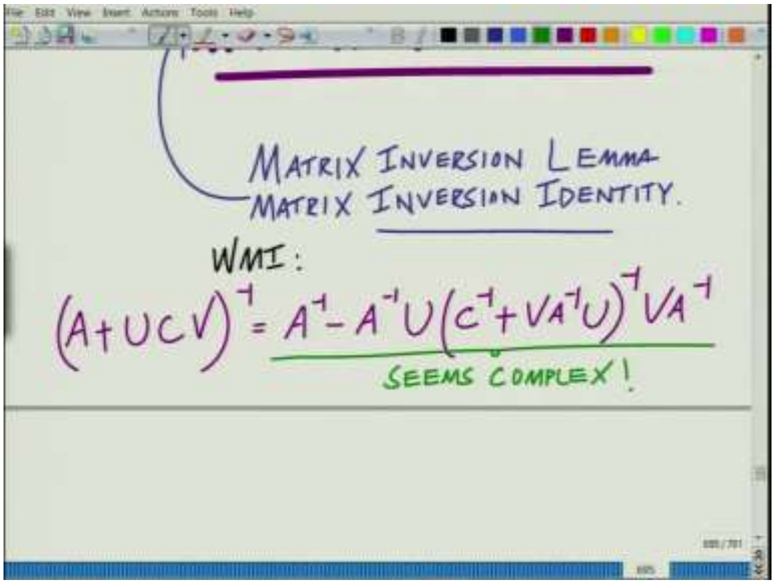
Applied Linear Algebra for Signal Processing, Data Analytics and Machine Learning
Professor Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology Kanpur
Lecture - 65
Woodbury Matrix Identity – Matrix Inversion Lemma

Hello, welcome to another module in this massive open online course. In this module let us look at a very useful matrix relation or a matrix inversion identity, which is also termed as the Woodbury Matrix Identity which makes the evaluation of matrix inverses very easy or significantly reduces the complexity of matrix inversion in certain cases. So, let us try to understand this.

(Refer Slide Time: 00:45)



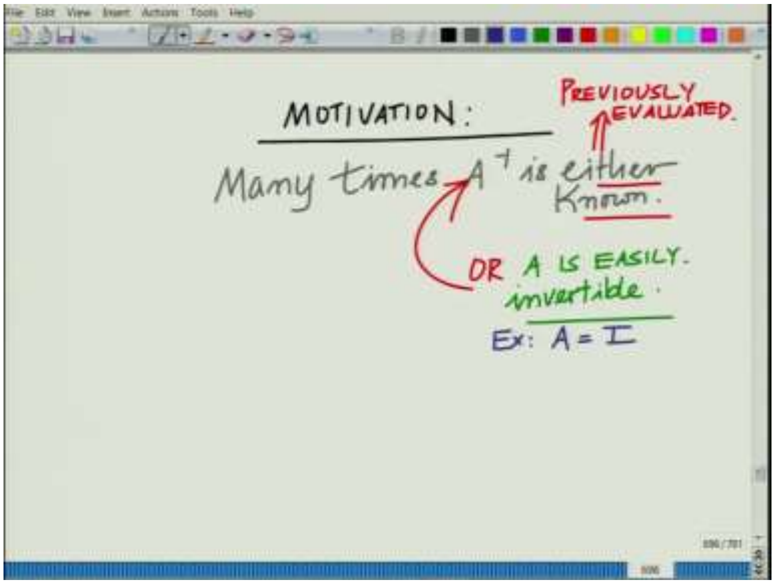


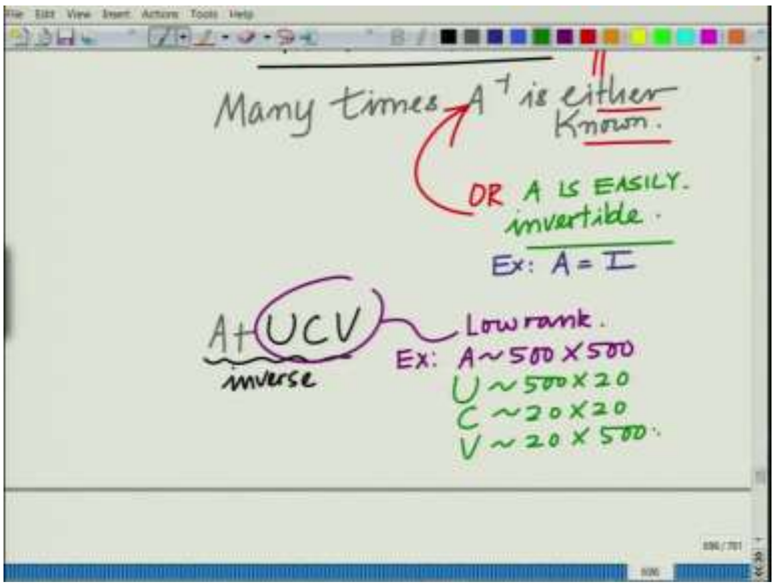
So, we are going to discuss the Woodbury matrix. So, we are going to discuss the Woodbury matrix identity. And what is this matrix, this is also by the way known as the matrix inversion lemma, I mean there are many names for this, this is also known as the matrix inversion lemma or the matrix inversion identity or you know, so many other are the matrix inversion rule property, so on. So, this essentially but these are the two popular names that is the Woodbury matrix identity, Woodbury matrix identity and the matrix inversion lemma.

Let us first state it then try to understand this application then also prove it So, the Woodbury matrix identity abbreviating it as this, Woodbury matrix identity this can be stated as, if you look at any compatible matrices U or, A plus U C V inverse this can be evaluated as, or let me just write it a little A plus U C V inverse can be evaluated as A inverse minus A inverse U times C inverse plus V A inverse U inverse V A inverse.

So, this is the expression for the matrix inversion identity or the Woodbury matrix identity. And at first glance it seems as if the matrix, the identity is more complicated than the inverse itself, it seems like the left is very compact. Well, the right is sort of really long, messy and seems rather complex to evaluate. But this really what we have said in the beginning is that this really simplifies the evaluation of the matrix inverse in certain cases. Now, what are these cases, let us try to understand this. So, if you look at this at first glance, this seems complex, I mean if you just look at it as is, it seems complex.

(Refer Slide Time: 03:49)



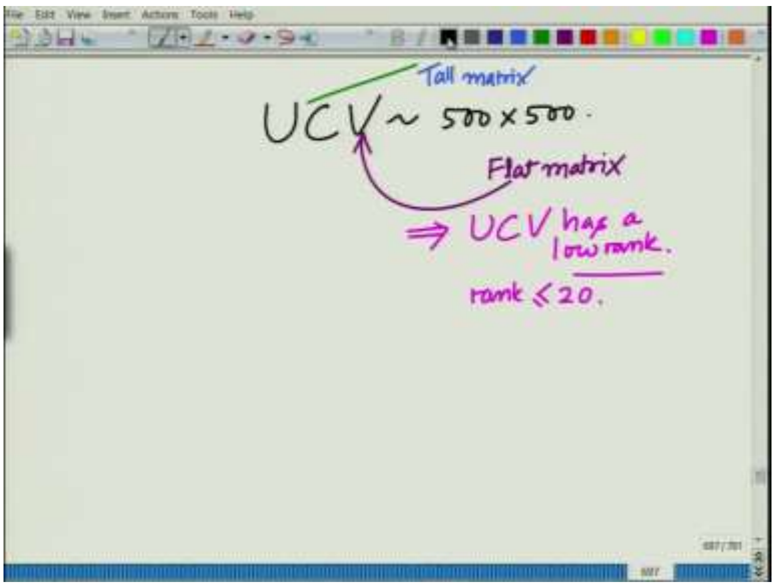


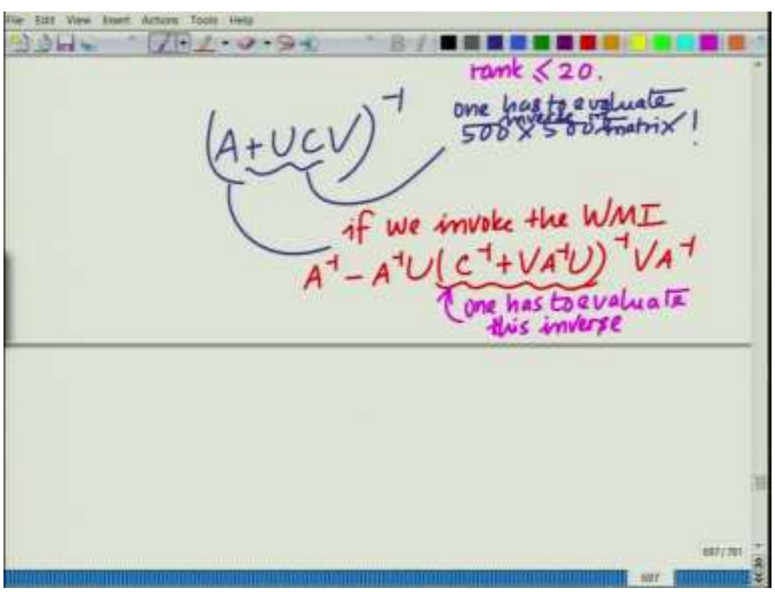
So, let us understand what is the motivation for the matrix inversion identity. The motivation is as follows, several times, many times many or many times A inverse A is either known that is has already been evaluated, is either known or either known implies essentially, previously evaluated, this has been previously evaluated or what happens is A is easily invertible, A is readily invertible.

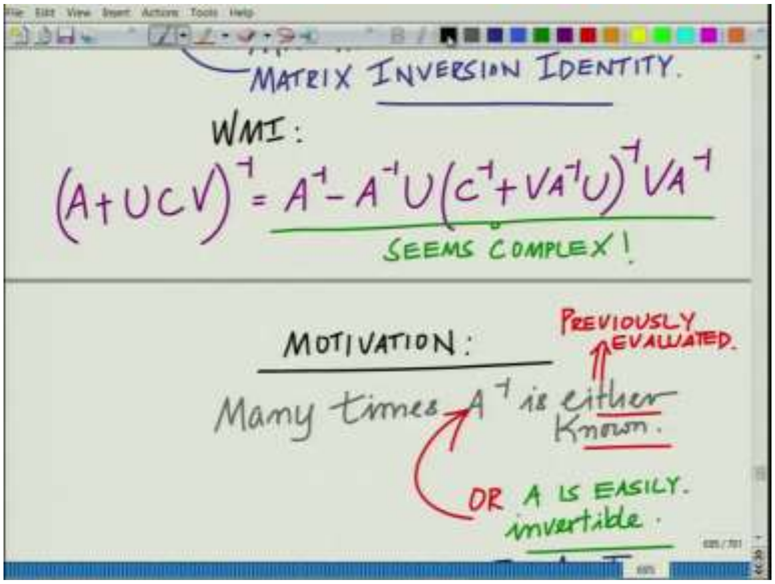
For example, when we say, for example, let us say A equals identity matrix, very simple example. A equals an identity matrix, one can compute the inverse of A very easily or lets a easy unitary matrix, we know U transpose is the inverse of you, so so on and so forth. So, now, what happens in such a scenario, now let us consider, now remember we are talking about this sum A plus U C V and we are talking about the inverse of this sum, the inverse of this quantity.

Now, we want a specifically consider scenarios where U C V has a low rank that is the key. U C V has a lower rank for instance, what do we mean by that, let us say example, let us say A equals of 500 cross 500 A is of size not equal to A is of size 500 cross 500 U is of size 500 cross 20 and such large matrices arise very frequently. C is of size 20 cross 20 and V is of size 20 cross 500.

(Refer Slide Time: 06:30)





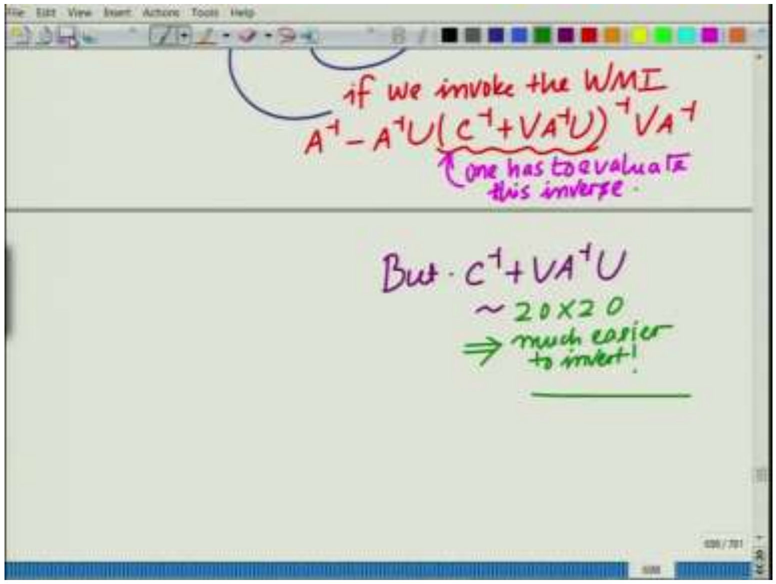


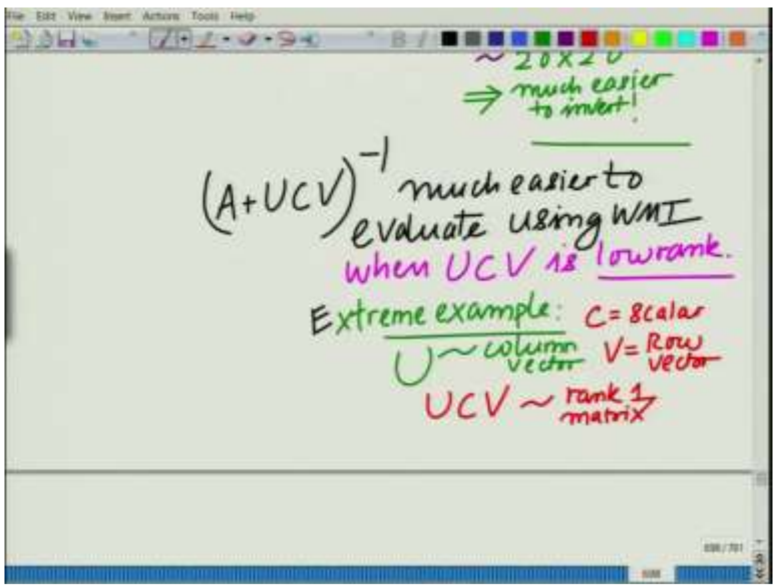
Now, what you can see is the size of UCV this is of size 500 cross 500 but, because U is of size 500 cross 20 that is a tall matrix, U is a tall matrix, U is a tall matrix, while V if you look at V, V is a flat matrix used a tall matrix where V is a flat matrix. So therefore, this implies that U C V is very low rank, this basically implies U C V has a low rank. In fact, rank of U C V is less than equal to 20.

Now, in such a scenario what happens is, now again go back to the matrix inversion identity what you will see is, rather than inverting, now if you look at A plus UCV inverse , now if I want to compute A plus U C V inverse, now this is 500 cross 500. So, one has to evaluate the inverse of a 500 cross 500 matrix, if you evaluate the inverse of A plus U C V.

On the other hand, using the, now if we use on the other hand, if we use, if we invoke, if we invoke the Woodbury matrix identity, then remember, let me just write that again. So, that is this quantity inverse minus A inverse U times C inverse plus V A inverse U inverse V A inverse. Now, if you look at this, I have two only, one has to only evaluate one has to only evaluate this inverse, one has to evaluate this inverse, evaluate this inverse.

(Refer Slide Time: 09:21)





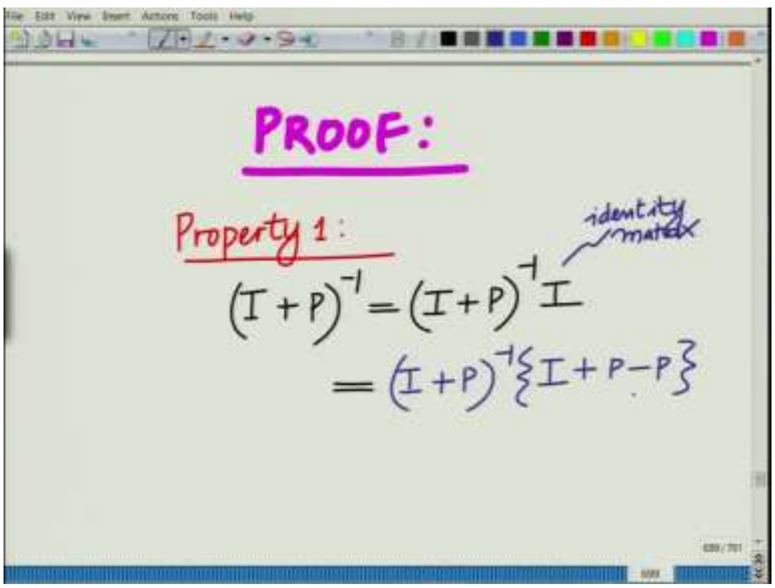
But, but if you look at this, but C inverse plus V A inverse U, what is the size, this is you can clearly see, this is 20 cross 20. So, implies much, so this is much easier to invert. That is the point. So, C inverse plus V inverse U because it is 20 cross 20 which is much lower than the 500 cross 500 matrix that we had initially. So, this is a much easier to invert.

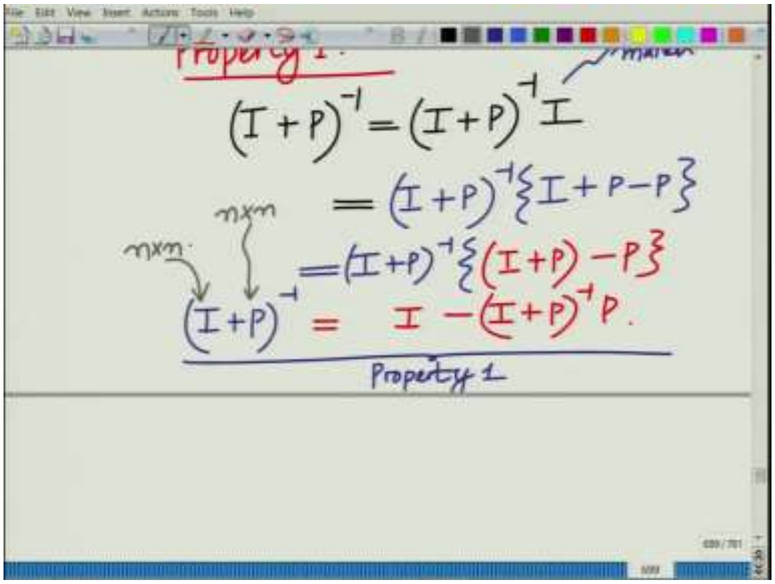
Now, therefore, the Woodbury matrix identity, WMI can significantly simplify matrix inversion where U C V is low rank. And A inverse is already known or easy to evaluate and such sequence such, such scenarios occur very frequently in practice especially in signal processing and machine learning.

So, A plus U C V inverse, this is much easier to evaluate using Woodbury matrix identity when U C V is a low rank matrix, I mean that is I think something that is important. In fact an extreme example of that is where U is a single vector and V is a row vector then it becomes a, and C is a scalar quantity and then it simply becomes rank one matrix.

Example, extreme example of this is, so you have U equals column vector and V equals row vector U C V and the C is a scalar, C equals scalar. U C V has basically, this has a rank one, this is a rank one. Size is 500 cross 500 but ranked is only 1 because U is a column vector V is a row vector. And such scenarios, it becomes very easy to employ the matrix very convenient and it is very attractive to employ the matrix inversion identity for matrix inversion.

(Refer Slide Time: 12:27)



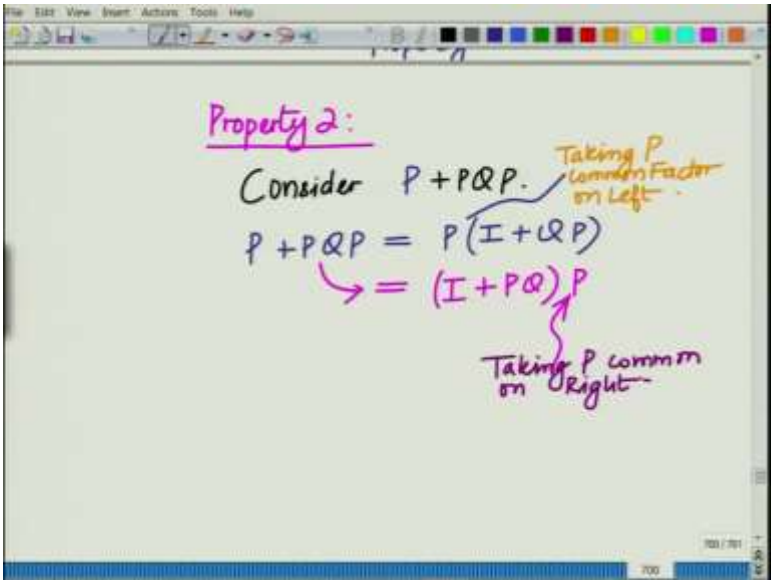


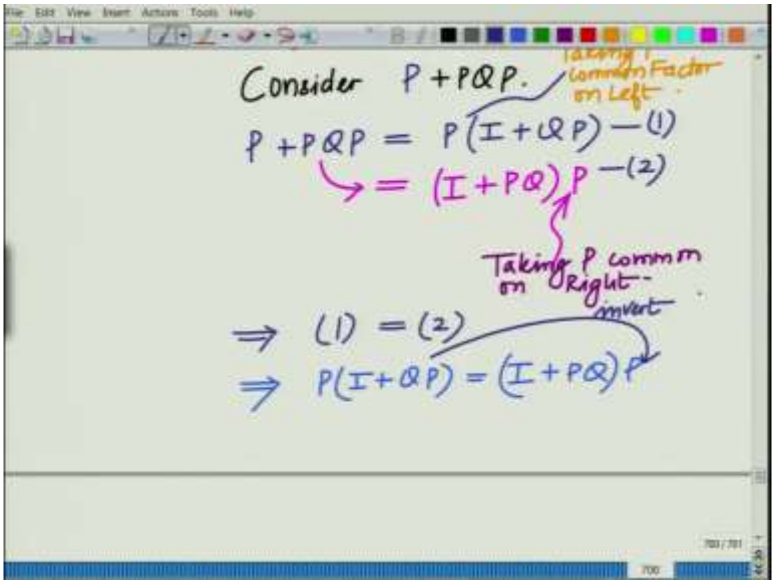
Now, let us look at the proof of this, I think this a very important property. So, let us try to understand this let us use, let us look at the proof of this in brief and it's rather informative and it is a very useful exercise to understand the proof of the matrix inversion identity because that by itself is a learning experience.

So, let us try to first establish two properties that we are going to use in the proof let us simply call them property 1. For property 1, let us consider this matrix I plus P inverse I can write this as I plus P inverse times identity, which i can now write as i plus P nothing has changed this is this is simply the identity matrix, which is i plus P inverse into now I am going to add and subtract P now add and subtract P.

Now, this gives me, now combine these i plus P inverse into i plus P minus P now i plus P inverse into i plus P this is identity minus i plus P inverse into P. So, this is your expression for i plus P inverse. And this is essentially what we are calling as the property and naturally for this, this holds when P is a square matrix. Let us say this is n cross n, this is the n cross n identity matrix I mean all the matrix sizes have to be compatible. So, we are considering square matrix P and identity matrix I of the same size. Fantastic, fantastic.

(Refer Slide Time: 14:53)



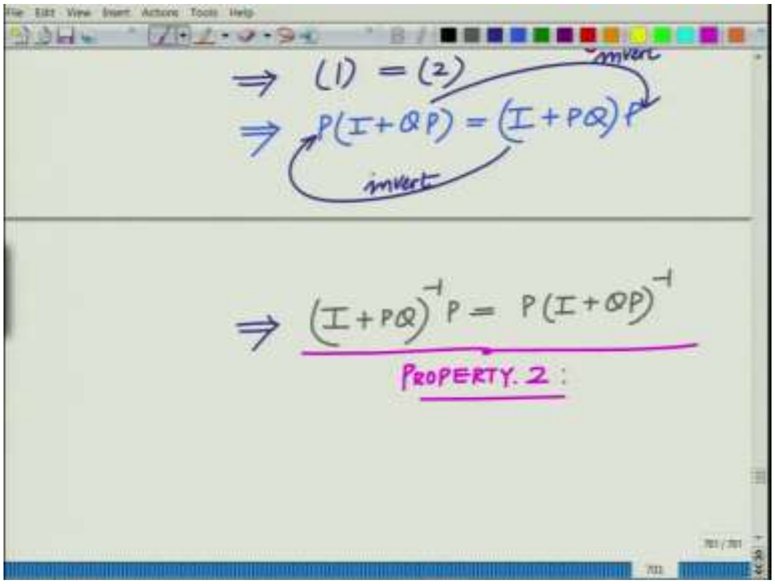


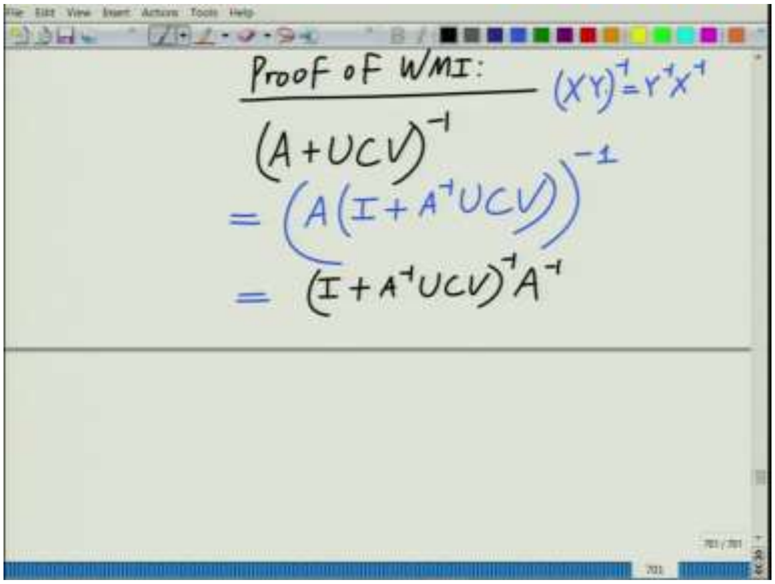
Let us now look at the next property that is property 2. Remember these are the properties that we will need to prove the matrix inversion lemma. Consider this quantity I plus P Q P. Now, I can write i plus P Q P, not i plus P Q P but rather P plus P Q P. I can write this as taking P common on the left, I can write this as P into I plus Q P.

So, this is by via taking the P as the common factor on the left, P as common factor on the left. And I can take this also write this as now taking P as a, as a common factor on the right, I can also write this as I plus P Q into P. So, this P plus P Q P I can write it two ways, this is I into taking P common on the right I can write as I plus P Q into P.

So, this is obtained via taking P common on right. And this basically implies that now, these two things are equal this way so because both are equal to P plus P Q P. So, let us say this is 1 and let us say this is 2, so both are equal. So, this implies essentially that you have 1 equals 2 which basically implies that P into I plus Q P equal to I plus P Q into P. Now, invert this and take this over here. So, that is multiplying by I plus Q P inverse on the right on both sides and I plus P Q inverse on the left. So, you invert and take this over here.

(Refer Slide Time: 17:32)



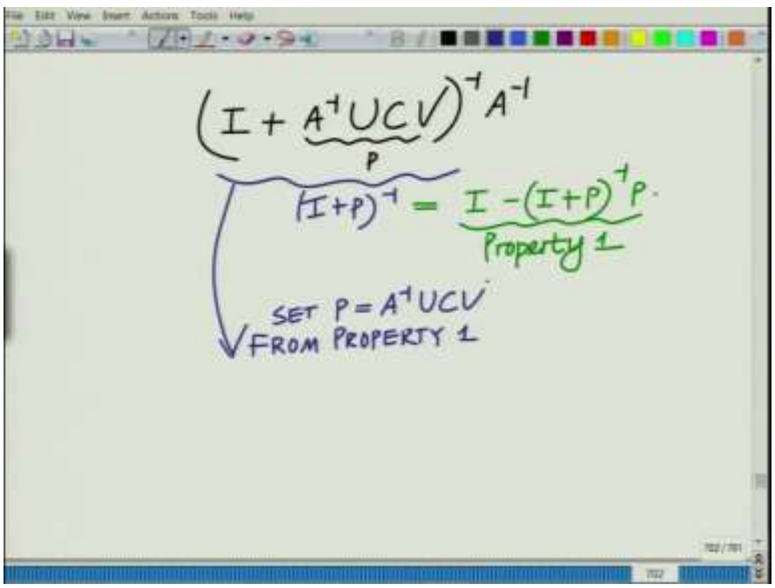


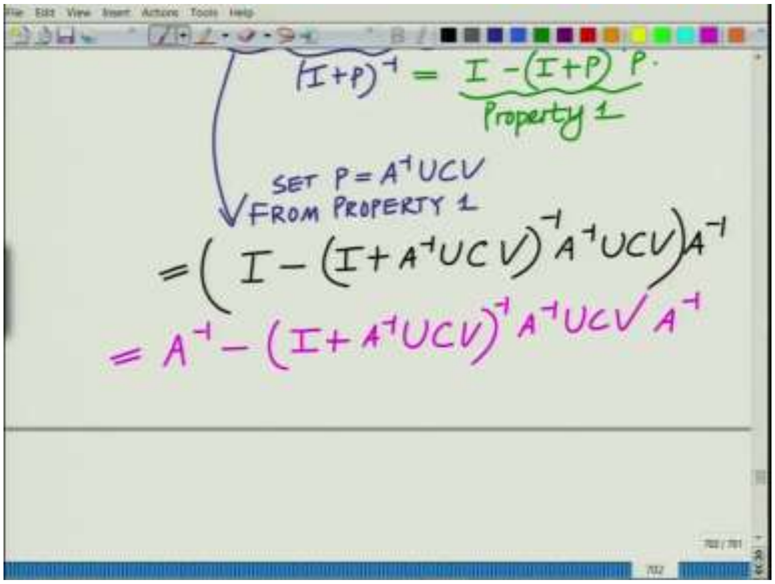
This implies now that I plus P Q inverse into P equals P into I plus Q P inverse. So, that is essentially the property that we have to see P into I plus Q P inverse. So, that is essentially the property number 2. These are the properties, so we will use these two properties to prove the matrix inversion lemma or the Woodbury matrix identity.

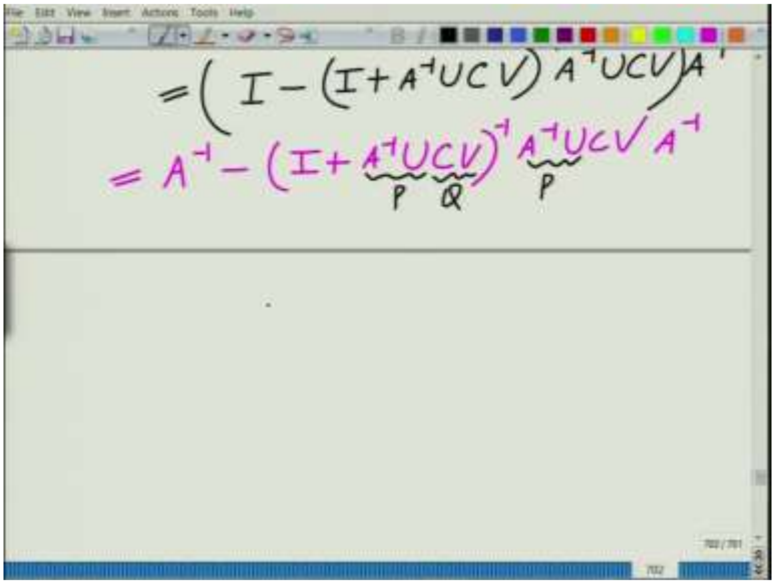
Now, let us prove the Woodbury matrix identity. So, now, we prove or let us say simply say proof of WMI. Now, proof of WMI, now we start with, so you A plus A plus U C V inverse this will be equal to let us take A out So, I can write this as A into I plus A inverse U C V inverse. Now take the inverse inside. So, this becomes, we use the property, we use the property X into Y or two invertible matrices this is Y inverse times X inverse.

So, use this property here that becomes I plus A inverse U C V inverse into A inverse, where we use the property X into Y inverse is Y inverse into X inverse, so that is essentially we have now simplified this in the first step as follows, first taking A common and then moving inside the inverse inside the bracket.

(Refer Slide Time: 20:04)



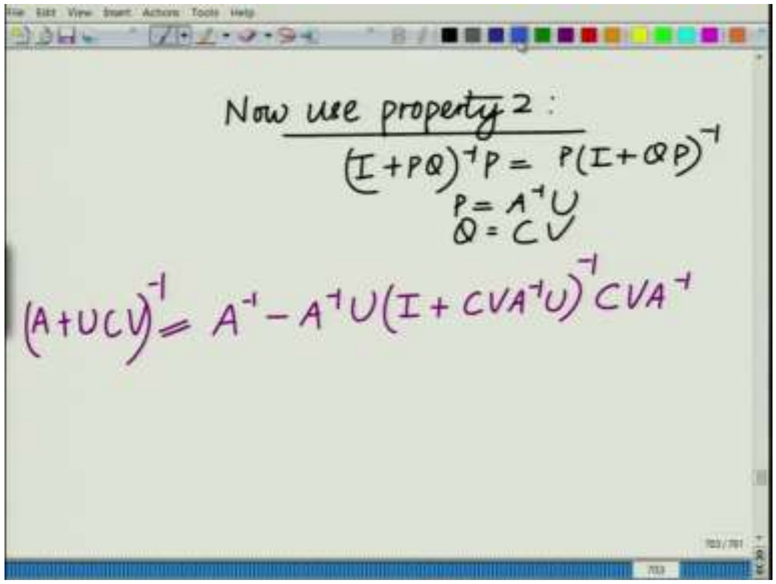


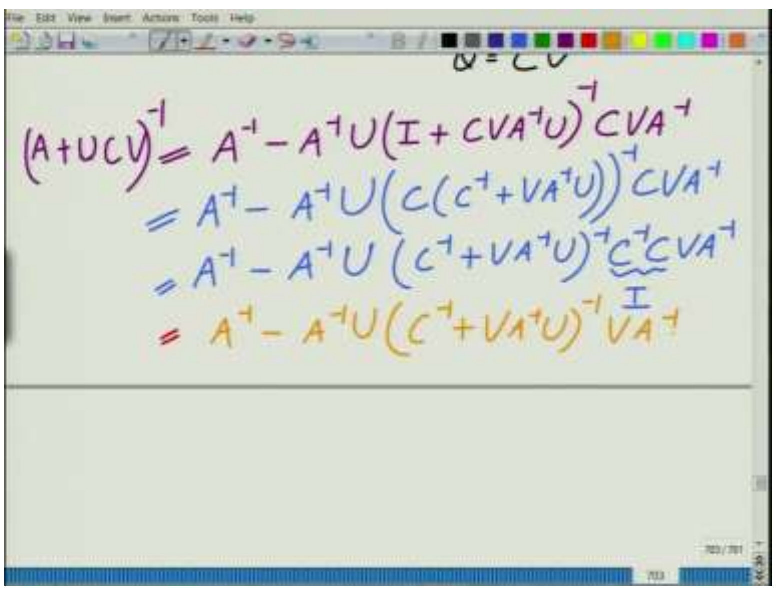


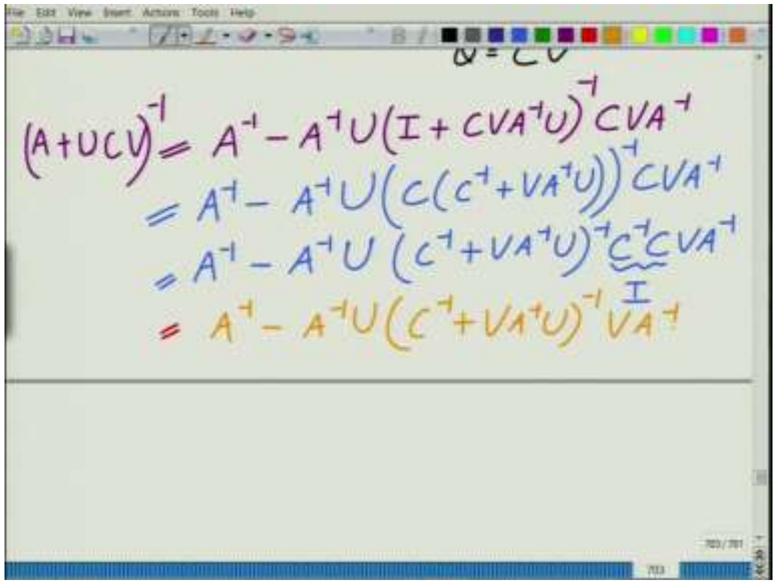
Let us now continue this, now we are going to use the property 1. So, we have this i plus A inverse U C V inverse into A inverse. Now use the property 1 treat this as P. So, this becomes essentially you are I plus P inverse which from property 1 equals I minus I plus P inverse into P. So, this is the essence of your property, this is the essence of your property 1.

So, from property one this becomes, from property one, from property 1, after setting or we set, so set P equal to A inverse U C V. From property 1 this reduces to, this becomes in the brackets I minus I plus P which is A inverse U C V inverse P which is A inverse U C V. And outside remember you already have an A inverse, that remains. And therefore, now I can write this as A inverse minus I plus A inverse U C V inverse A inverse U C V U C V into A inverse. Now, we use the property 2, so now use property 2 by setting this quantity as P, this quantity C V equals Q and this quantity equals P.

(Refer Slide Time: 22:40)







So, now use property 2, recall what is property 2, I plus P Q inverse into P equals P into I plus QP inverse and we set P equals A inverse U Q equals C V. And therefore, we have this becomes A inverse minus P which is A inverse U times I plus A inverse U I plus Q P that is C V A inverse U Q is C V P is a inverse U inverse times, times what is already there outside that is your CV into A inverse and this is what we have obtained for so far. Not to forget U for A plus U C V inverse after using the property 2. So, first we use property one in the simplification followed by property 2 and this is what we get.

And now let us simplify this further. And now the next step is easy. You simply take in the inside you can already see we all, we are almost there, this is A inverse U and inside I am going to take the C out. So, this is C times C inverse plus C C is out. So, this is V A inverse U inverse times C V inverse A once again use the property X into Y inverse equals Y inverse into X inverse.

So, this is A inverse minus A inverse U times C inverse plus V A inverse U inverse C inverse C V A inverse and C inverse C is identity and that gives you the final expression for the matrix inversion identity, which is essentially A inverse minus A inverse U into C inverse plus V inverse U inverse C inverse into C is identity what remains is V A inverse. And that completes the proof of the matrix inversion identity.

So, this is equal to A plus U C V inverse and this completes, this completes the proof, this completes the proof of the Woodbury matrix I identity. So, this completes the proof of the

Woodbury matrix identity, which as I have already told you is a very interesting property and a very convenient and handy property which arises frequently in communications signal processing, machine learning, analysis and so on.

Especially in the inversion of some of these matrices where you have the sum of an invertible matrix and a low rank matrix, which can be written as the outer product of basically, which as essentially the product of a tall matrix and a flat matrix. So, naturally such a product is going to be low rank. And this expression can be very conveniently used to simplify and compute such an inverse with very low complex. So, let us stop this module here, and we will continue our discussion in the subsequent modules. Thank you very much.