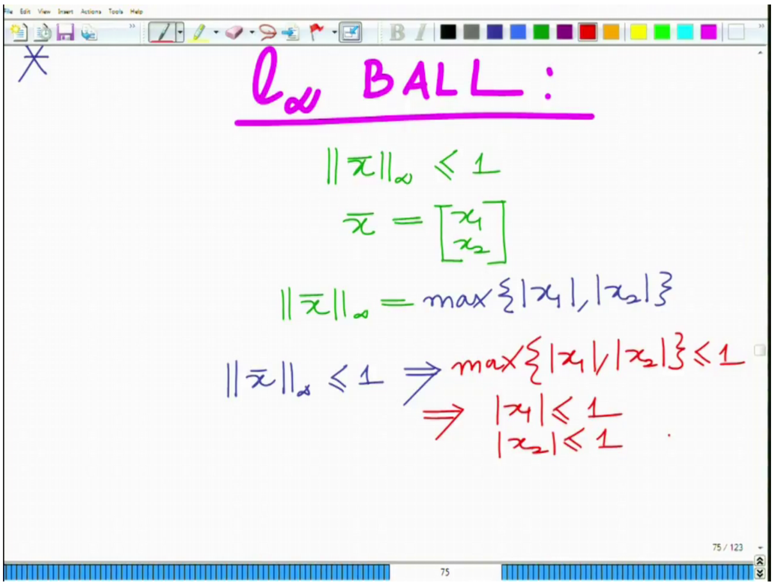
Applied Optimization for Wireless, Machine Learning, Big Data
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture – 19
norm balls and Matrix Properties: Trace, Determinant

Hello, welcome to another module in this massive open online course. So, we are looking at the l infinity ball alright; we have find the l infinity norm which is the maximum of the magnitude of the different components of a vector x bar, and let us now look at continue our discussion on the l infinity balls.

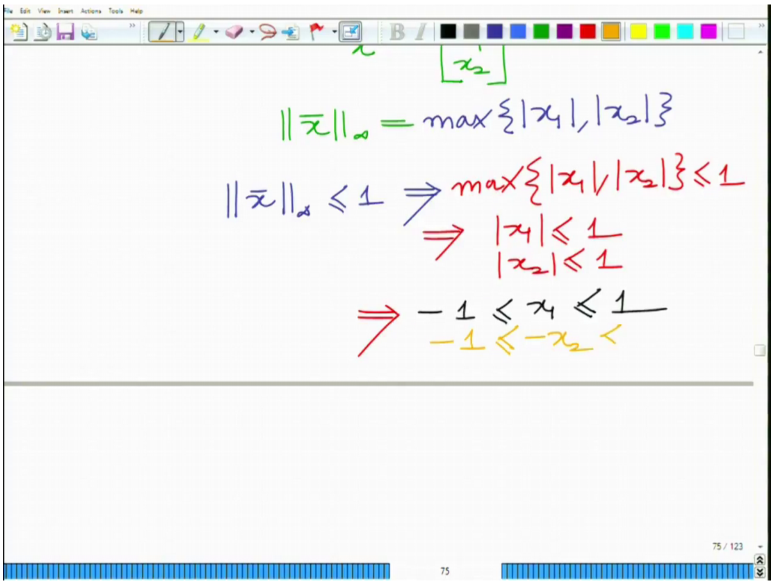
(Refer Slide Time: 00:35)



So, the we are looking at it is a l infinity ball is a simply defined as, the infinity norm of a vector less or than equal to 1. Let us consider a simple scenario 2 dimensional vector x bar which has 2 components x 1 x 2 and now the l infinity norm that is if you look at the l infinity norm that will simply be the maximum of magnitude of x 1 comma magnitude of x 2.

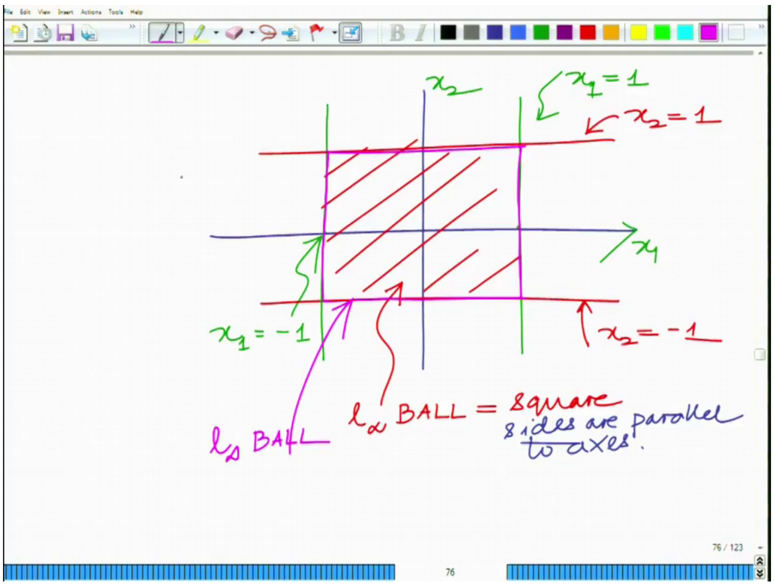
Now, l infinity norm less or equal to 1 implies this maximum of the magnitude x 1 comma magnitude x 2 less than or equal to 1 this implies that knows is a maximum of 2 quantity is less than or equal to 1, this implies that each of the quantities has to be less than or equal to 1.

(Refer Slide Time: 02:03)



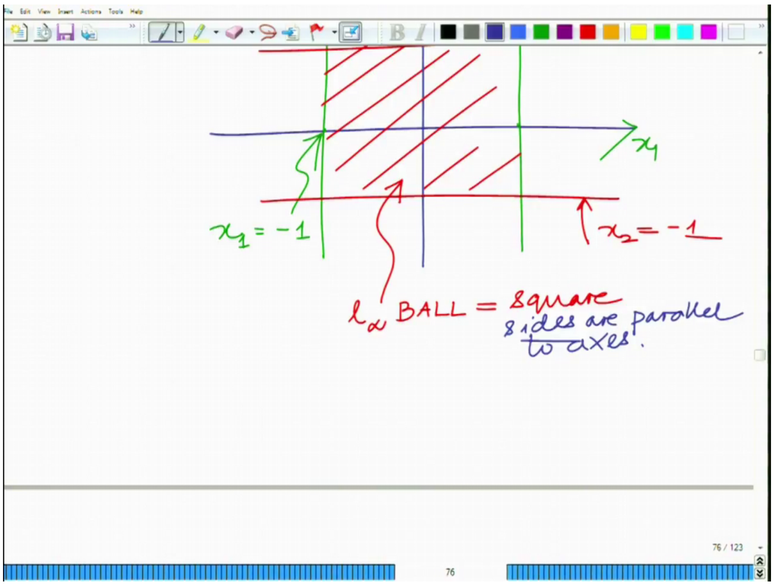
Now, magnitude x 1 less than equal to 1 this implies that minus 1 less than or equal to x 1 less than or equal to 1 that is x 1 has to lie between minus 1 and 1, and further magnitude of x 2 less than equal to 1 implies minus 1 less than or equal to x 2 less than or equal to 1.

(Refer Slide Time: 02:26)



And now so, this is the intersection of these 4 in fact, if you can see there are 4 half spaces; one is given by x 1 less than or equal to 1 that is this is the first coordinate x 1 x 2. So, let us say this corresponds to your hyperplane x 1 equal to 1 and then on the opposite side let us say this corresponds to the hyperplane x 1 equals minus 1. The strip end between denotes the region minus 1 less than equal to x 1 less than equal to 1 and similarly this corresponds to the hyperplane, this corresponds to the hyperplane x 2 equal to 1. And, this corresponds to the hyperplane x 2 equals minus 1 and now the region which is between these hyperplanes this polyhedron is in fact, what is your l infinity that is the square is the l infinity ball.

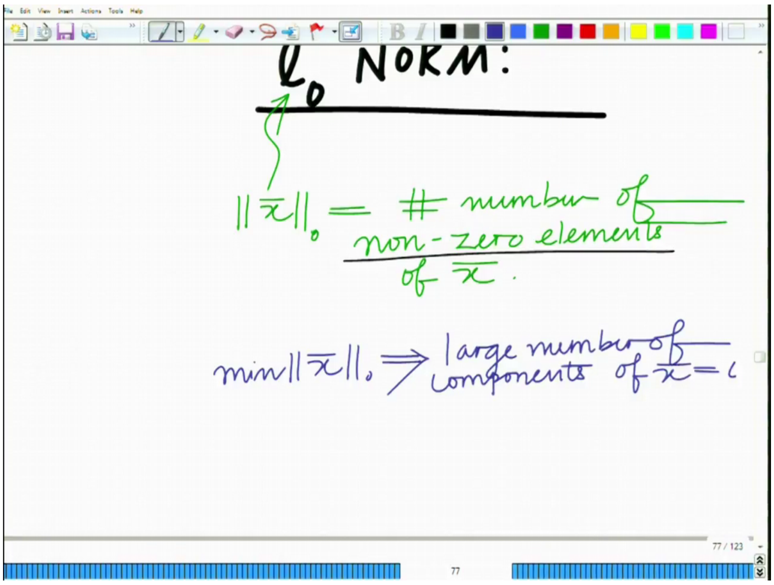
(Refer Slide Time: 03:38)



So, l infinity ball equals square this is also a square with sides that are parallel to axes; the sides are. Remember the l 1 square had the diagonals that are, if you look at the l 1 square that had the diagonals along the axes, and the l infinity square is the normal square that you would imagine which has the sides that are parallel to the diagonal sides that are parallel to the axes. So, this is your l infinity looks like a square ok.

So, this is something interesting this normally when you think of balls, you think circles and spheres, but when you look at the l 1 norm ball alright which is a tilted square and the l infinity norm ball which is essentially a square alright. So, these are also norm balls that is when you generalize the definition of a norm ball, you can derive these kinds of norm balls which are a very interesting shapes ok.

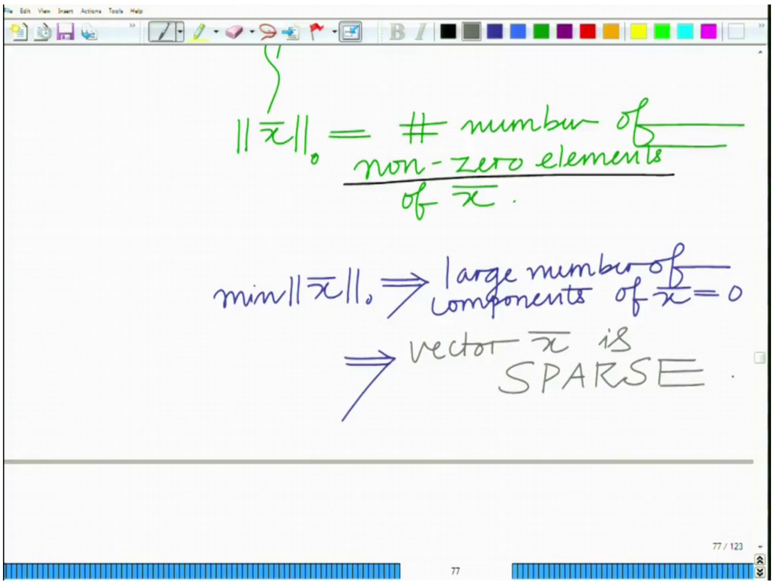
(Refer Slide Time: 05:14)



And we also have this notion of an l 0 norm, which is very interesting and well what is the l 0 norm l 0 norm and the l 0 norm you can essentially show, that if you look at a vector. What is a l 0 norm that is the definition is you can show that norm x bar 0 equals the number of non-zero; this is equal to the number of non-zero elements of x bar this is very interesting the number of non-zero elements of x bar.

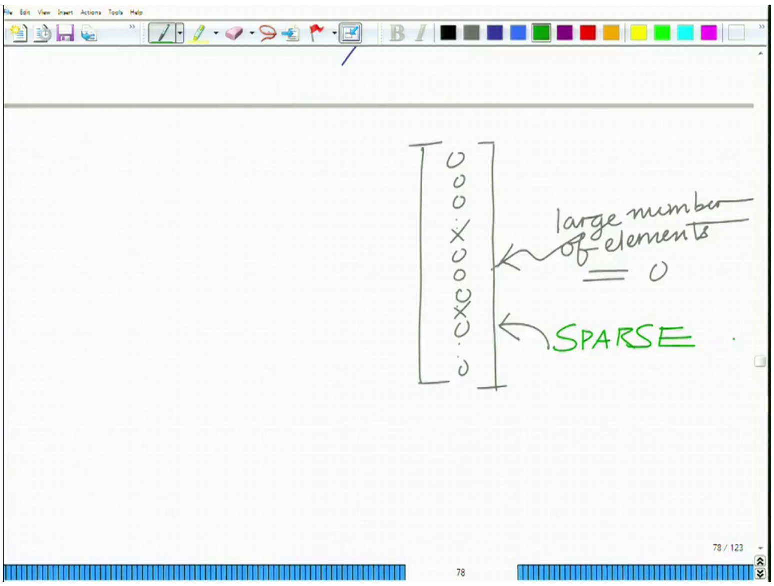
So, if you minimize the l 0 norm alright and this is a very idea, if you minimize x bar of 0 this results in a large so; large number. So, what you will observe is large number of components of x bar will be 0. So, you will get a vector x bar typically, in which a large number of its components are 0 because the l 0 norm is the number of non-zero elements.

(Refer Slide Time: 06:52)



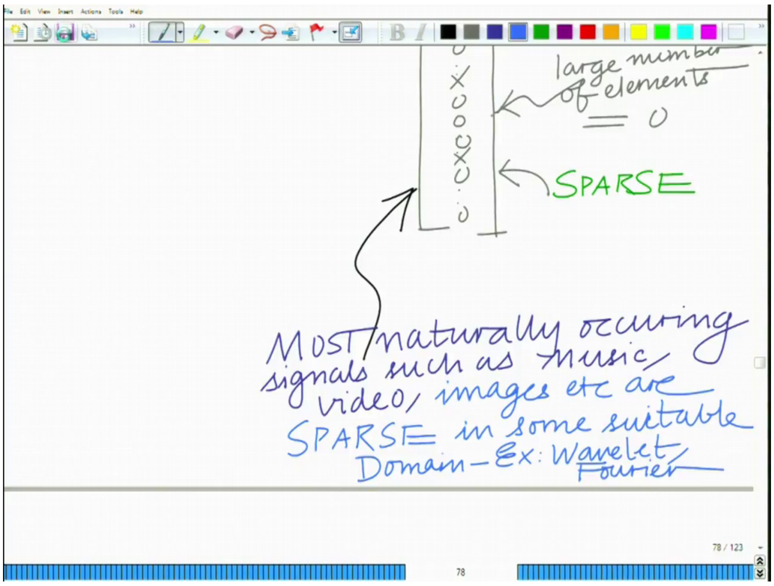
Implies and this, such a vector x bar which has a large number of non-zero elements, such a vector x bar is termed to be it has a very interesting name this is known as a sparse vector alright. Similar to what we similar to the usage a sparse sparsely populated area essentially means that, there are very few people. So, a sparse vector basically denotes a vector in which sparse it is sparsely populated with components a large only very few non-zero components, and a large number of components are a large number of components are 0.

(Refer Slide Time: 07:32)



So, it brings to mind some vector like this, which has 0 some component non-zero alright. So, what you have is a vector in which large number of elements equal to 0 such a vector is termed as a sparse vector and this is a very interesting property.

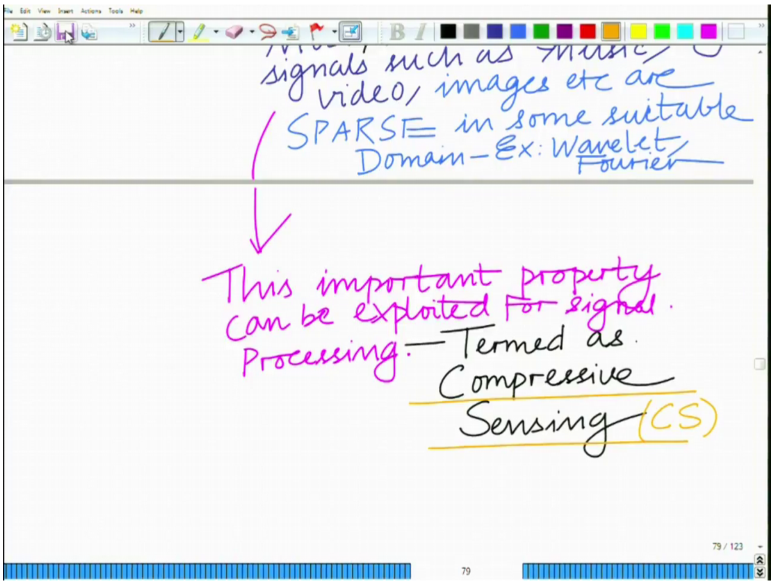
(Refer Slide Time: 08:10)



And what is interesting about this is that is you can show that most signals most naturally occurring signals such as be it either music or video or images although they are not dangerous sparse, they are sparse under some of you know some appropriate domain for instance when you look at the either the Fourier transform or the wavelet transform of these signals, they are very sparse and that is a very important property.

So, when you look at most naturally occurring signals such as either a music signal or video or images are sparse in some are sparse in some suitable domain example either the wavelet domain or the Fourier domain that is when you take the Fourier transform or basically the frequency domain, you can also say. So, they are sparse in some domain and this can be used for at this idea is a very important idea, which can be used for signal processing and to improve the performance.

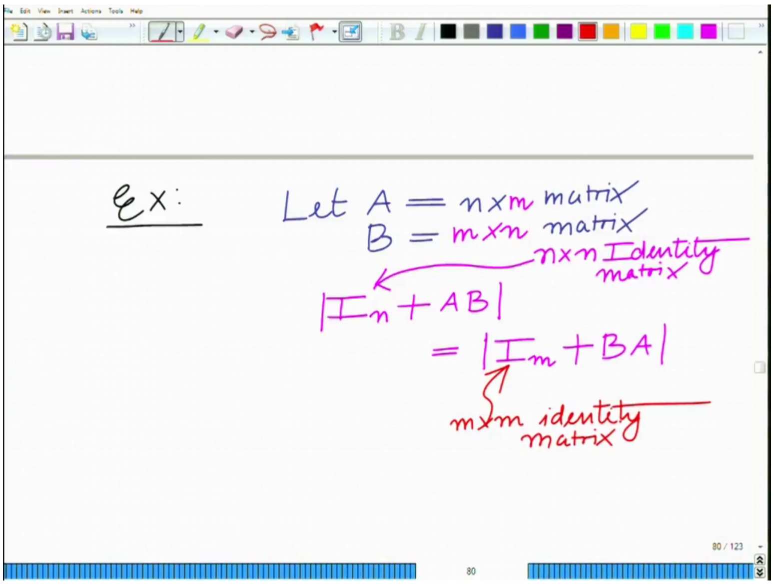
(Refer Slide Time: 09:53)



This important property that is sparsity can be exploited for signal processing. This important idea can be exploited as signal processing and this is termed as compressive sensing and one of the domains where this is exploited is known as one of the important areas, where this idea is used compressive sensing also this is abbreviated as CS. Compressive sensing this is a relatively new field in fact, a path breaking innovation I must say which has gained a lot of popularity in the recent past where in you exploit the knowledge that this vector that is this vector x bar which corresponds to a naturally occurring signal is sparse in a certain suitable domain. And, that can be used to further that can be used in signal processing and to further improve the performance alright in comparison to other schemes alright.

So, this is in fact, a break through this in fact, is a or you put its a break through paradigm alright it is a break through framework or its compressive sensing framework or the framework that exploit sparsity of the signal vectors is a break through framework that can be used for enhanced signal processing alright. So, with that let us complete this discussion, let us move on to looking at some problems to better understand the concepts. Let us start with a few problems related to the determinants and the positive semi definite property. So, I will want to start with some examples.

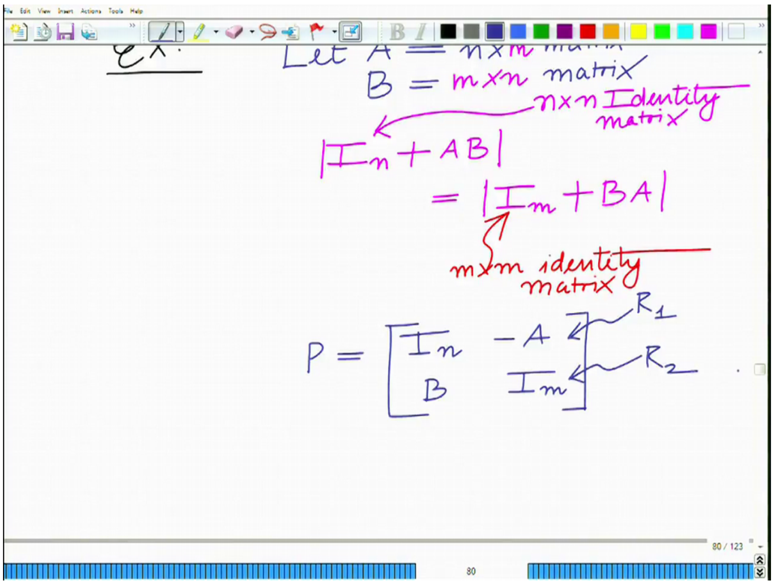
(Refer Slide Time: 12:03)



To start with let us start with some simple examples related to matrices and their properties. So, let A be an n cross P matrix and B equals P cross n matrix, the first thing we want to show is that determinant of I n plus A B equals the determinant of I m or I am sorry determinant of I n plus A B is the determinant of I p plus B A.

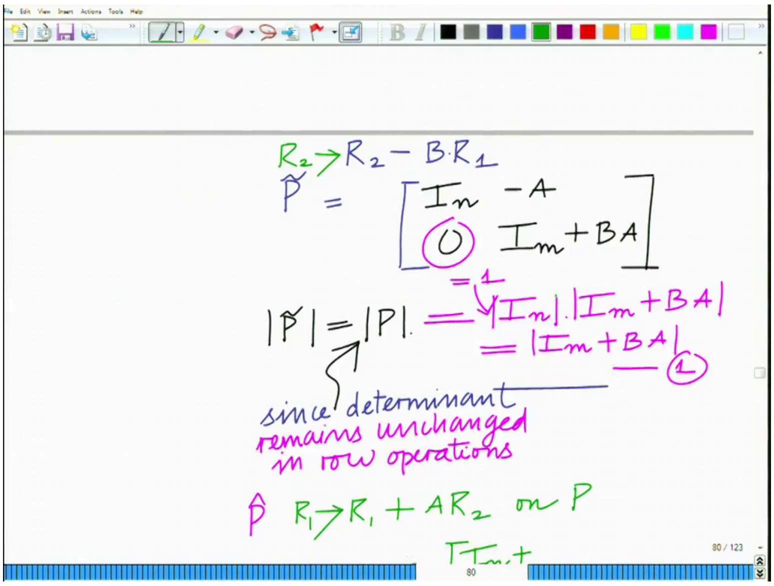
Let us make this n cross m matrix. So, I think that will make it life much more simpler. So, this B is an naturally if you are multiplying m they have to have dimensions that match up. So, this is we want to show this is an. So, I n this is the n cross n identity matrix and this on the other hand this is the m this is the m cross m identity matrix now solution is as follows.

(Refer Slide Time: 13:56)



Now, if you consider let us consider the matrix, let us consider the matrix P which is given as I n minus A, B I m. Now, I can perform row operations or in this case block row operations ok. So, the let us call this as block row 1 denoted by R 1 and this is block row 2 denoted by R 2.

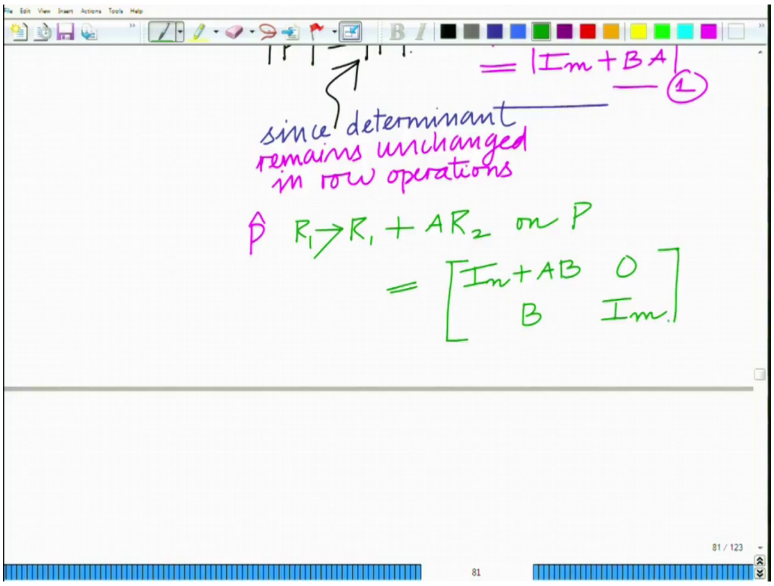
(Refer Slide Time: 14:34)



Now, let us obtain P tilde by performing R 2 that is block row 2 minus the matrix B times block row 1. So, you obtain the matrix P tilde which is basically well first row remains unchanged second row is B minus I n times B which is 0 i minus of B times minus a. So, that is I m plus B A. Now if you look at the determinant of P tilde, which is equal to determinant of P. Now this arises since row operations since determinant remember you might have seen a property of determinants before that its determinant remains unchanged in row operations. So, determinant P tilde equals determinant P now you see the lower block is 0's, which means the determinant of P is simply the determinant of I n times the determinant of I m plus B A.

Now, the determinants of I n is 1 since it is an identity matrix remains unchanged. So, this quantity equals to 1. So, this is simply the determinant of I m plus B A. So, determinant of P tilde equals the determinant of I M plus B A let us call this as our result 1.

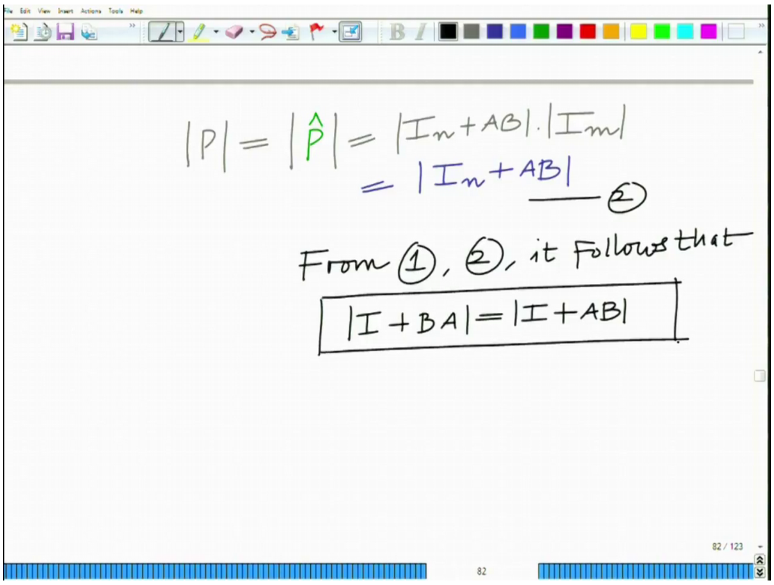
(Refer Slide Time: 16:32)



Now, let us form the matrix P hat by performing R 1 plus A times R 2 on. So, R 1 goes to let us write this as follows R 1 goes to R 1 plus A times R 2, here we have R 2 goes to R 2 minus B times R 1. So, on P and this gives the matrix you can check, this gives you the matrix well this will be I n plus A times R 2 A times B plus I m plus well let me just check this, this is n cross m matrix B times A B is m cross n; so, this B times A.

So, this will be I n or 1 plus a times R 2. So, I n plus B times A I am sorry this will be I n plus A times B and this will be I n plus A times B, this will be the other will be minus A plus A times minus A plus A times i. So, this will be 0 and the second row remains unchanged that will B and I m and now if once again.

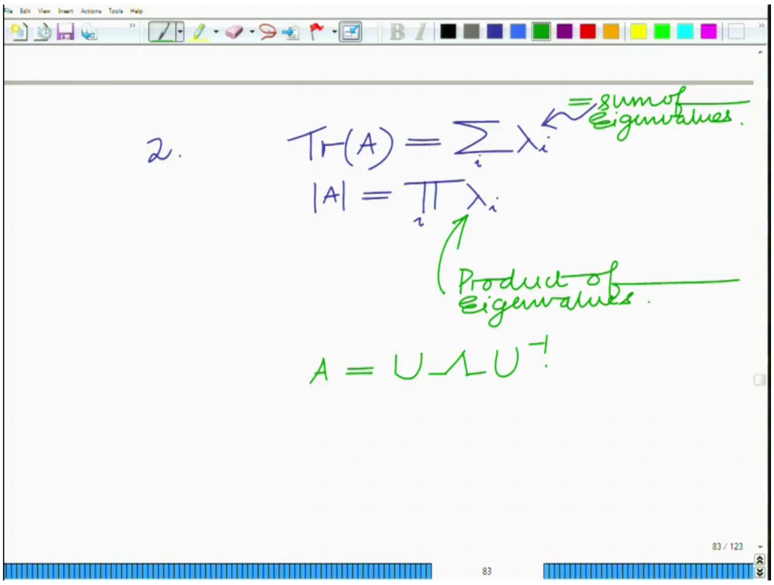
(Refer Slide Time: 18:35)



So, this is your P hat and now, if you once again formulate the determinant of P hat that is going to be simply that determinant because this is 0, that is going to be the determinant of I n plus A B times the determinant of I m.

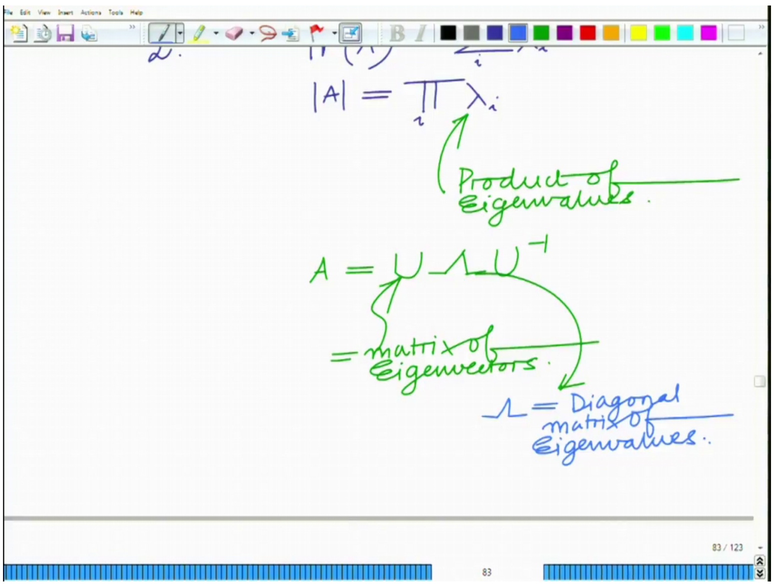
So, and by the way again once again this is equal to the determinant of P, because you are performing row operations on P determinant of I m is 0. So, this will become I n plus determinant of I n plus A B and finally, from 1 and 2 from 1 and 2 it follows that determinant of I plus B A equals determinant of I plus A B alright. So, this follows from the results in 1 and 2.

(Refer Slide Time: 19:56)



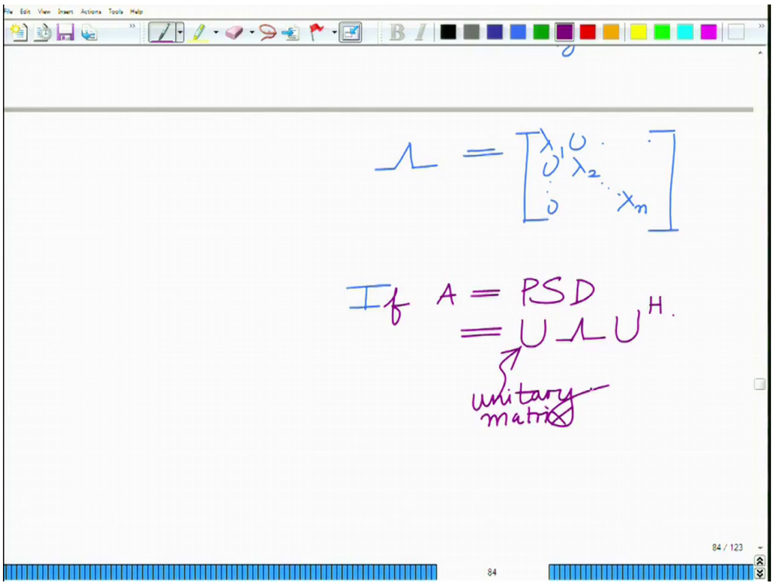
Let us look at another interesting example that is example number 2, we want to show that the trace of a (Refer Time: 20:03) square matrix A is the sum of the eigenvalues of A and the determinant of a is the product of the eigenvalues. So, this is trace which is equal to sum of eigenvalues, and the determinant is basically the product of eigenvalues ok. And this is well to do this you can start with the property the following think for a general matrix that A can be expressed as U lambda U inverse alright where U equals matrix of eigenvectors.

(Refer Slide Time: 20:58)



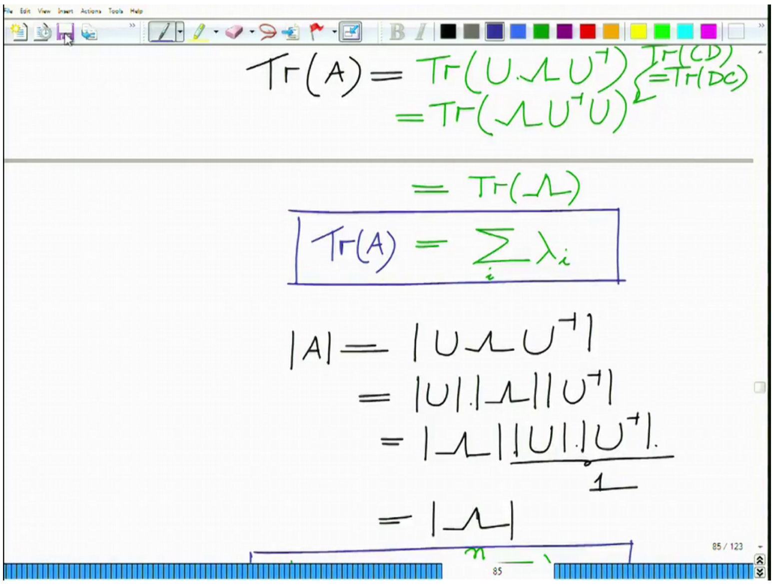
Lambda is diagonal value or diagonal matrix of eigenvalues.

(Refer Slide Time: 21:33)



Lambda equals diagonal matrix of eigenvalues and lambda has the following structure ok. If A is an n cross n matrix, lambda equals diagonal matrix of eigenvalues assumes that a equals an n cross n matrix square matrix alright. Now of course, we already seen that if a is PSD then it becomes U equal to U lambda U Hermitian. If A is a PSD matrix if A equals PSD well the above is valid for any general matrix A, this becomes U lambda U Hermitian. I am sorry because U is orthogonal U is a unitary matrix for a PSD matrix A; U is a unitary matrix and U inverse is simply U Hermitian.

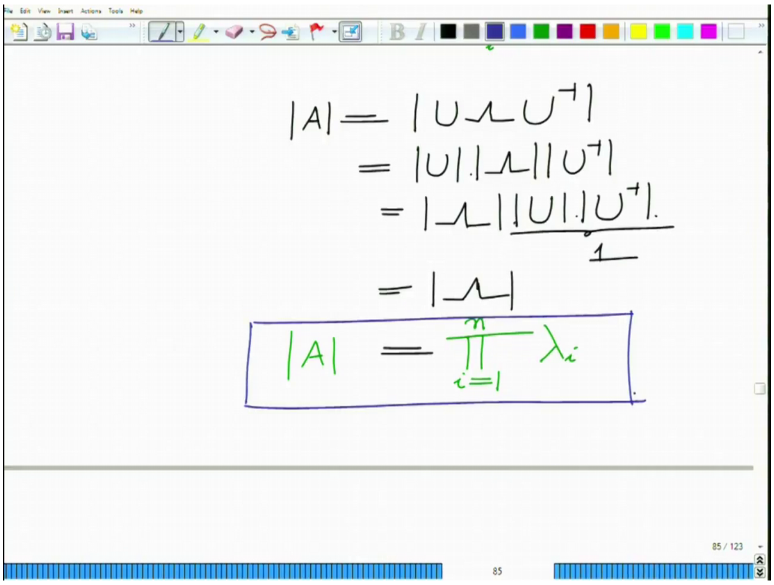
(Refer Slide Time: 22:39)



Because U Hermitian U equals U U Hermitian equals identity, which implies that U Hermitian equals U inverse ok, but this is only for a PSD matrix ok. But, we can show the above property that is the trace is equal to the trace of matrix a is equals the sum of its eigenvalues, for any general matrix and that is as follows if you consider the trace of a you use the eigenvalue decomposition and now you replace the trace as U lambda U inverse.

Now we know that the trace of A B equals trace B, that is the trace of that is if you are interchange the product of the matrices the order of the product in the trace. So, this will become trace of lambda it is very simple. So, this is become trace of lambda U inverse if you this is using the property, trace of CD just not to confuse with A equals trace of DC. Now, U inverse U is identity. So, this is trace of lambda which is the diagonal matrix of eigenvalues and this is nothing, but summation of i lambda i.

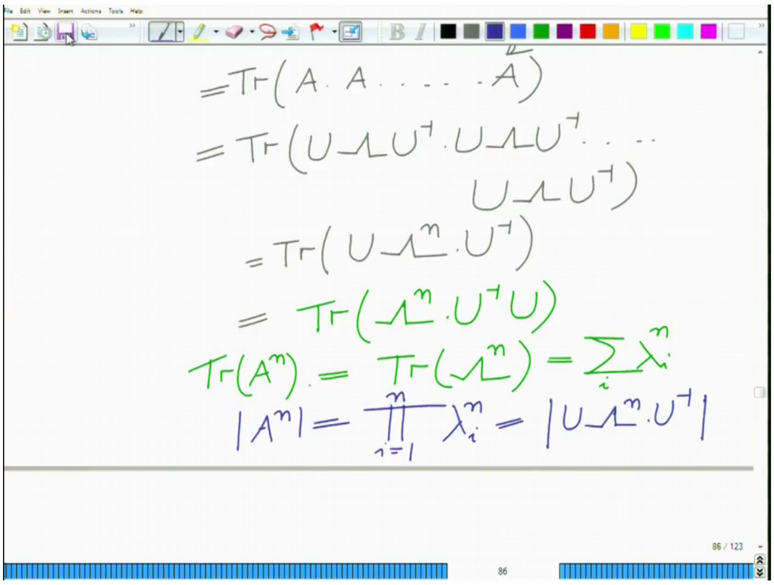
(Refer Slide Time: 24:03)



Now, similarly one can also show now if you look at the determinant of A that is the determinant of U lambda U inverse, which is now the determinant of product of matrices is the determinant of the product of the determinants, which is equal to the determinant of lambda times the determinant of U times the determinant of U inverse.

Now, if you look at the determinant of U into the determinant of U inverse, because U U inverse is identity determinant of identity is 1. So, therefore, determinant of U into determinant of U inverse is 1 this implies this equals the determinant of simply lambda which is now nothing, but basically the product of the eigenvalues and this is a very interesting property that is frequently used pi equals 1 to n this is the. So, you can say determinant of A equals product of its eigenvalues, trace is equal to the sum of its eigenvalues trace of a matrix equals the sum of its eigenvalues ok.

(Refer Slide Time: 25:32)



Now similarly, you can also exploit this property, if you look at trace of a matrix A raised to the power of n, that equals trace of A times A, the product n times we have already seen this for a positive semi definite matrix that becomes U, lambda U inverse, U lambda U inverse times U lambda U inverse, that becomes trace of I am sorry that becomes trace of well U lambda raised to the n U inverse, which is nothing, but trace of lambda raised to the power of n U inverse U which is trace of U inverse U identity. So, this is trace of lambda raised to the power of n, this is summation over i lambda i raised to the power of n that is trace a raised to the power n.

And in the similar way you can show that the determinant of a raised to the n is the product of i equals 1 to n lambda i raised to the power of n. This is nothing, but the determinant of U lambda raised to the power of n U inverse which is how we get this result ok. So, these are some interesting properties. In fact, very interesting properties that come in handy frequently during manipulation, that is a trace of a matrix a square matrix a is the sum of its eigenvalues and the determinant of a square matrix a is basically the product of its eigenvalues good.

So, will stop here and continue to this discussion looking up by looking at other examples in the subsequent modules.

Thank you very much.