Introduction To Adaptive Signal Processing

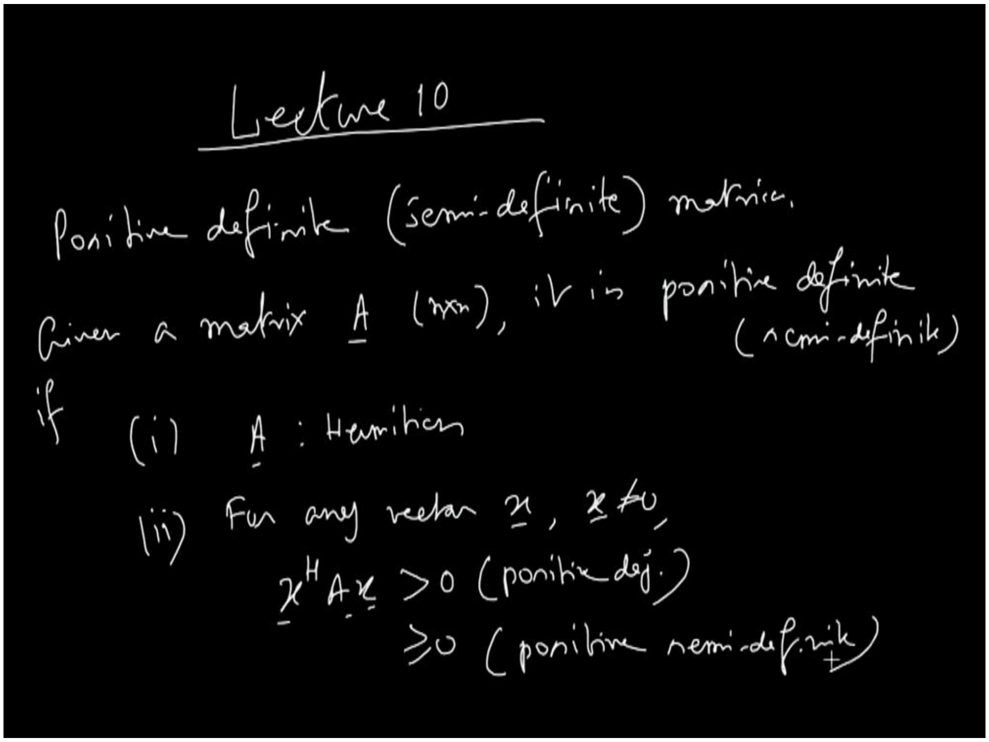
Prof. Mrityunjoy Chakraborty
Department of Electronics and Electrical Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture No # 10

Positive Definite and Semidefinite Matrices

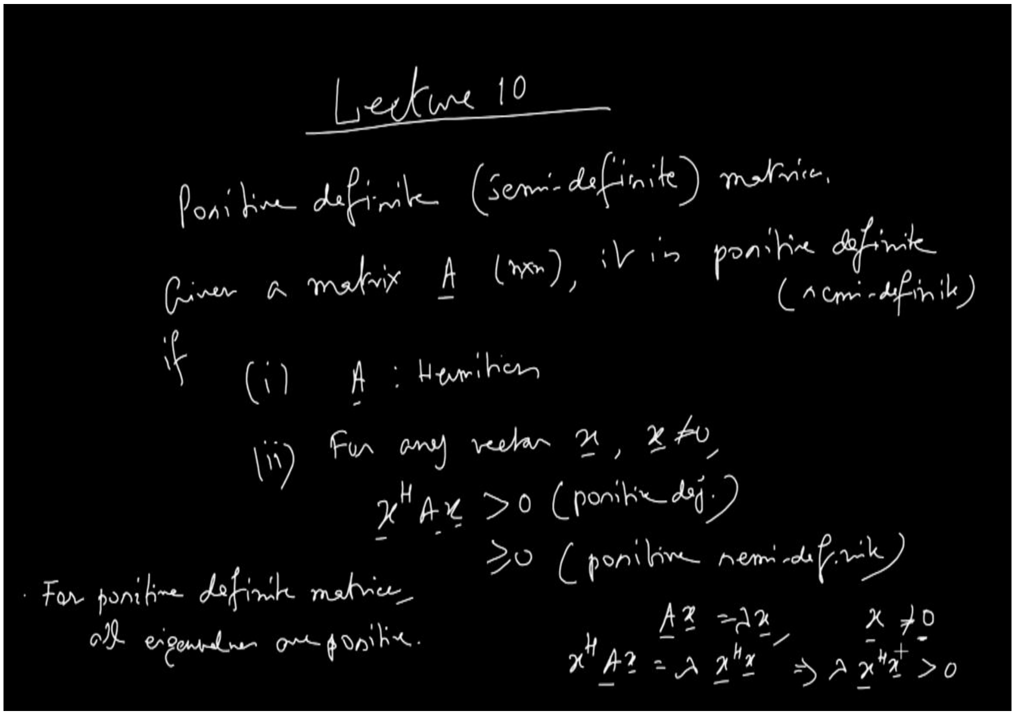
So, in the last class we discussed Hermitian matrices and their properties. A special class of Hermitian matrices play a important role in our subject that is called positive definite, sometimes positive semi definite also, matrices. Given a matrix A say N cross N it is positive definite or semi definite. If number 1 A is Hermitian. So, it must be a Hermitian matrix. Number 2 for any vector x where x is non-zero this thing xh Ax and remember in one of the previous classes I showed the derivative A is complex A has complex entries x has complex entries.

This is a scalar which is real matrix into column vector is a column vector, this is row row into column is a scalar and this is a real number this I have shown. Now for any vector x if this real number is greater than 0 then it is positive definite. and if it is greater than equal to 0 it is positive semi definite alright. This is a special class of Hermitian matrices.



Now for positive definite matrices for is a Hermitian matrix, So, eigenvalues are real. But if it is positive definite eigenvalues are just not real, they are positive, all. How suppose Ax is lambda x and since x is an eigenvector x is non-zero like x is non-zero. So, if I multiply this side left hand side and right-hand side called pre multiplication that is here if I multiply xh times Ax, xh Ax this will be lambda is a scalar. So, take it in the front and it is real xh x alright.

Now we have seen if it is positive definite for any non-zero x, xh Ax is greater than 0. And here x is an eigenvector which by definition of eigenvector is non-zero vector ok. So, for a non-zero vector x xh Ax which is real is positive. That means, lambda xh x is not only real it is greater than 0 if it is positive definite. And xh x we all know xh x will be what if x has supposed x has x1 x2 dot xn then xh x will be I mean you make it a row vector and then trans I mean there is take a transpose of this it becomes a row vector then conjugate every element that times x.

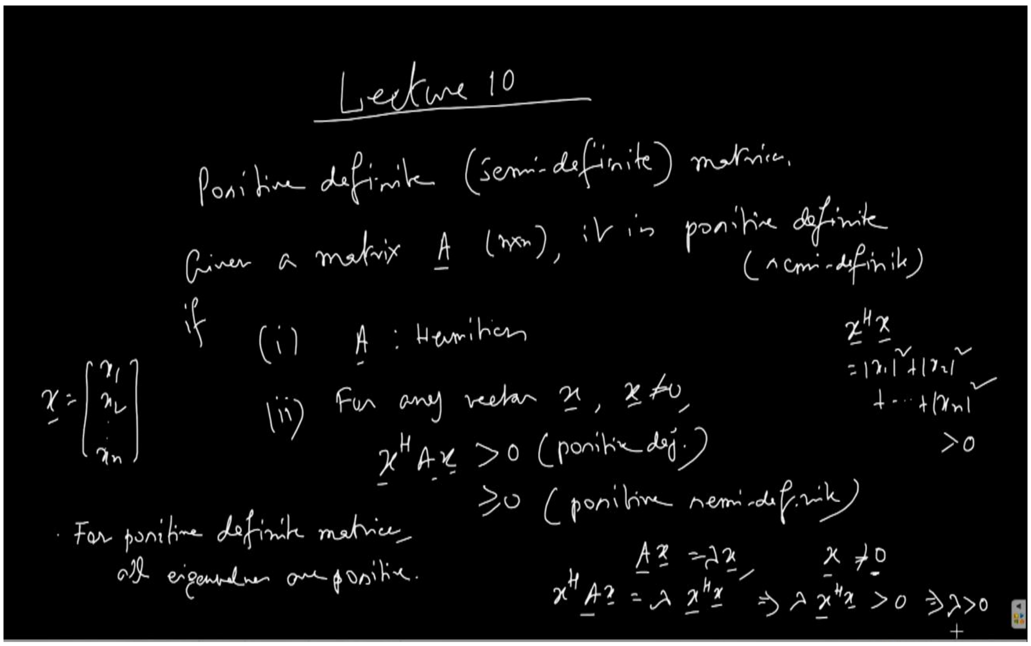


So, x1 star into x1 x2 star into x2 like that, So, it will be this we discussed earlier quite a few times and, in this case, it is said as such it is never negative real because you are taking mod of a number which may be complex but mod makes it real square. So, it is either 0 or positive it cannot be negative and they are positive all are added there is no minus there is no subtraction of two positive quantities. So, that the real thing is negative or 0. So, they are only positively contributing therefore, if it is to be 0, either it is 0 or positive. If it is to be 0 every contributor has to be 0 every term.

So, if this is 0 that means, x1 itself 0 because mod x1 square 0 means mod x1 0, and mod x1 is 0 means x1 has to be 0. So, like that. So, if it is 0 all the elements of x they have to be 0 and therefore, x must be 0 vector, but x is the non-zero vector because it is an eigenvector Therefore, it all the terms here cannot be 0. So, this is some are positive and others are 0 or maybe all are positive. So, sum together is not 0.

So, it is greater than 0. So, if it is greater than 0 this is greater than 0 this means lambda also greater than 0 because this is greater than 0. So, lambda also greater than 0. So, if it is a positive definite matrix then eigenvalues are strictly positive. If it is a positive semi definite matrix then instead of the greater than 0 we would have had greater than equal to 0 and then lambda into a positive quantity greater than equal to 0 this is positive non-zero.

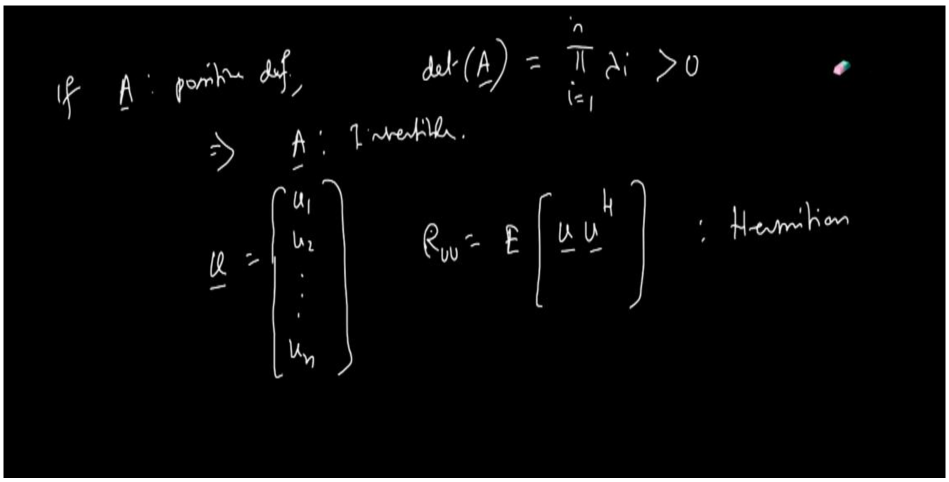
So, lambda will be greater than equal to 0. So, in case of positive semi definite matrices lambda greater than 0 greater than equal to 0 if it is strictly positive definite lambda is greater than 0.



So, if A is positive definite that means, A is Hermitian in that case we have say determinant of A for a Hermitian matrix is a product of the eigenvalues and if every eigenvalue is greater than 0, So, this product is greater than 0. Therefore, determinant is non-zero therefore, the matrix is invertible. But if A is positive semi definite then every eigenvalue is greater than equal to 0.

So, they can be equal to 0 also in that case the product which is determinant is not strictly greater than 0, it is greater than or equal to 0. So, I cannot say with 100 percent guarantee that A will be invertible because if any eigenvalue turns out to be 0 determinant is 0. So, in that case A will not be invertible. So, for invertibility A must be strictly positive definite all right. Now, consider a correlation matrix suppose I have got a vector of random variables ok or maybe I change the notation I got u random vector; these are the random variable in general complex valued.

So, form E I am not putting the subscript here you will understand uuH. So, that will be matrix and E will be operating on each. So, only the random variables present in that term will be in the subscript, otherwise, others will drop off that we know and this is called the auto correlation matrix suppose R uu all right.

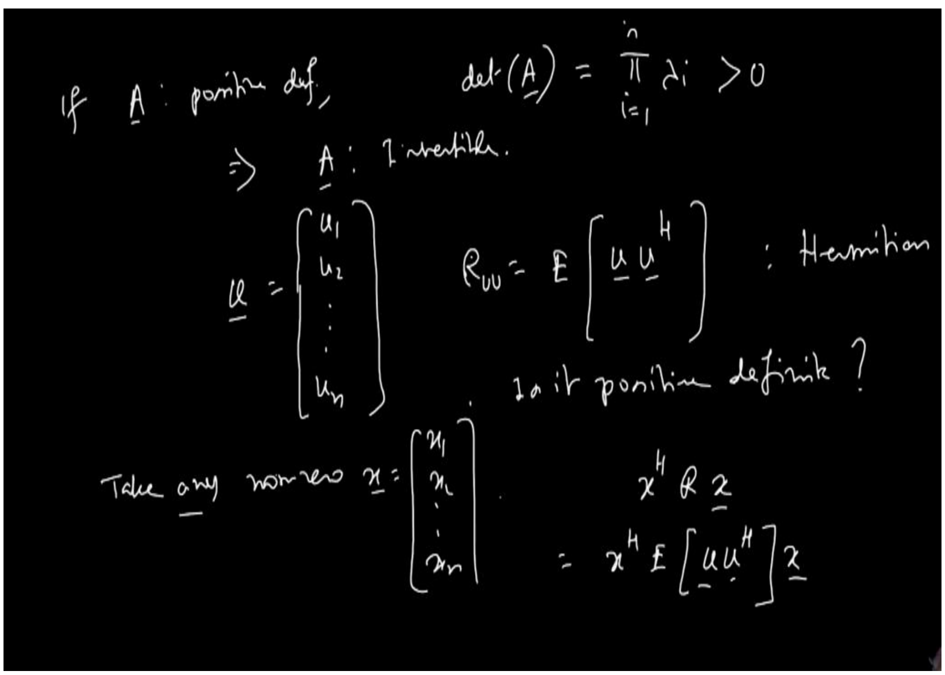


So, this is Hermitian we have already seen we have already analyzed this kind of matrices, they are conjugate symmetric i comma j and j comma ith element are conjugate of each other which makes it Hermitian. So, Hermitian we have already seen earlier.

There is correlation and covariance matrices are Hermitian. Up from covariance matrix we have shown, but in the case of covariance matrices we took the mean maybe mu1 ok and you can call this u1 prime u2 prime. So, there is again a vector and then covariance matrix becomes the correlation matrix of u1 prime u2 prime dot dot dot ok. So, there same logic applies there. There is you can apply the same logic you will find R uu also Hermitian i comma j the j comma ith element is you know conjugate of each other very easily you apply the same flow of logic you get this there is nothing here ok.

So, Hermitian is done, but is it positive definite? If positive definite we can happily say that this is invertible matrix because often we might be required to invert this matrix to solve some equations ok. Is it invertible? That is is it positive definite? For positive definite that is is it positive definite also positive definite? Positive definite this requires two condition it should be Hermitian first and there is one more condition. So, Hermitian is done. So, take any non-zero x as x1 x2 dot dot dot xn means at least one element or more than one element will be non-zero x. Then positive definiteness requires xh R I am not using the subscript Ruu, the same R, and x this would be this is real, even if R is complex, x is complex, we have seen earlier I have just a while ago told that this is real.

And if it is positive definite this should be greater than 0 not just real, greater than 0 and if it is positive semi definite this should be greater than or equal to 0. So, let us investigate this xh Rx that is xh R is EuUhx alright.

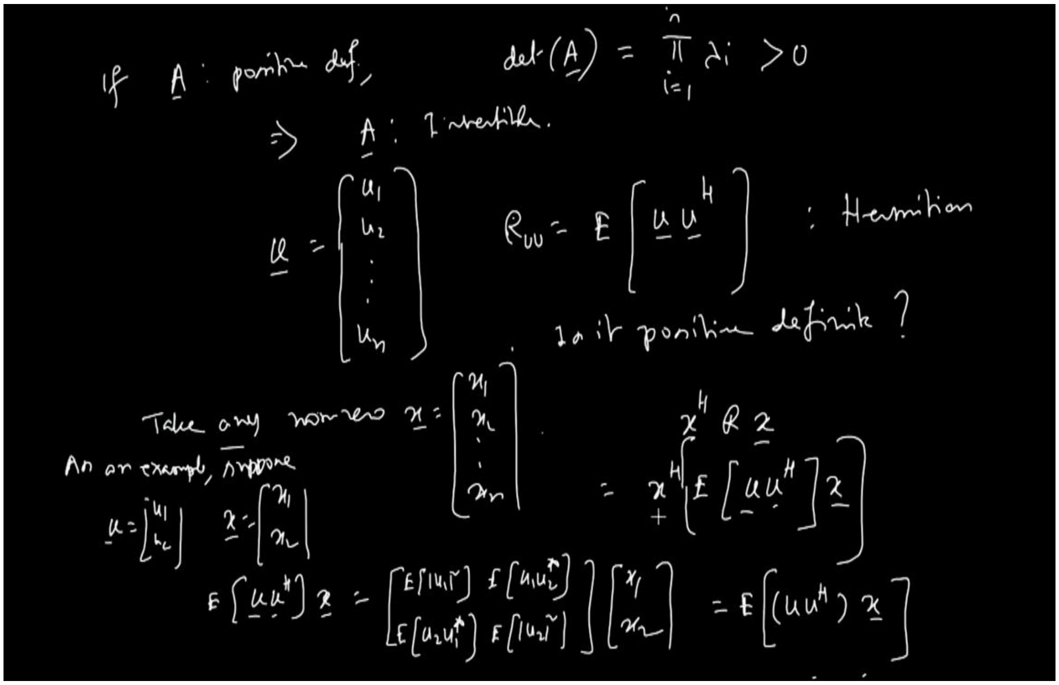


Now one thing we see here this is a matrix. This is a matrix column vector into row vector this is matrix and this is a column vector this is a row vector. But this is x is not random elements of u are random.

Ok as an example let me see how to explain to you people. Suppose just as an example u is u1 u2, x is x1 x2, as an example, suppose, this there is just two elements, small n equal to 2. In that case uH we got say this part first, you will bring this uuH, E of that, into x, will be what? uuH, u and uH means you will make it a row and conjugate and then multiply u then it will be expected value of mod u1 square u1 u2 and then u1, u1 star u2 star then u1 u2 star. Expected value u1 u2 star and then here expected value u2 u1 star and then expected value mod u2 square This into x1 x2. So, this times x1 this times x2 ok.

You can see since x1 and x2 they are not random you will get the same thing if you take E and then uUH then push the x from outside to inside the expectation. Because if you carry out uUH mod u1 square not the E operators mod u1 square u1 u2 star u2 u1 star mod u2 square and then x1 times that x2 times that added x1 times that x2 times that added then E comes in. But E of this into x1, x1 is not random. So, it will be nothing but E of only this part because this is random x1 goes out, like here this into x1. Similarly, x2 into this x2 into this quantity then E over the product will be same as E over this the random part only x2 outside as is happening here E of this part into x2 ok.

So, since x is not random from outside you can bring x here ok. So, you get one vector this into this plus this into this one vector this into this plus this into this one vector alright. This vector is a random vector, because this is, this vector is not a random vector, because after we apply expected value after we apply expected value that into x1 they are not random this into x2 not random because after expectation it is the mean average. So, it is not random.

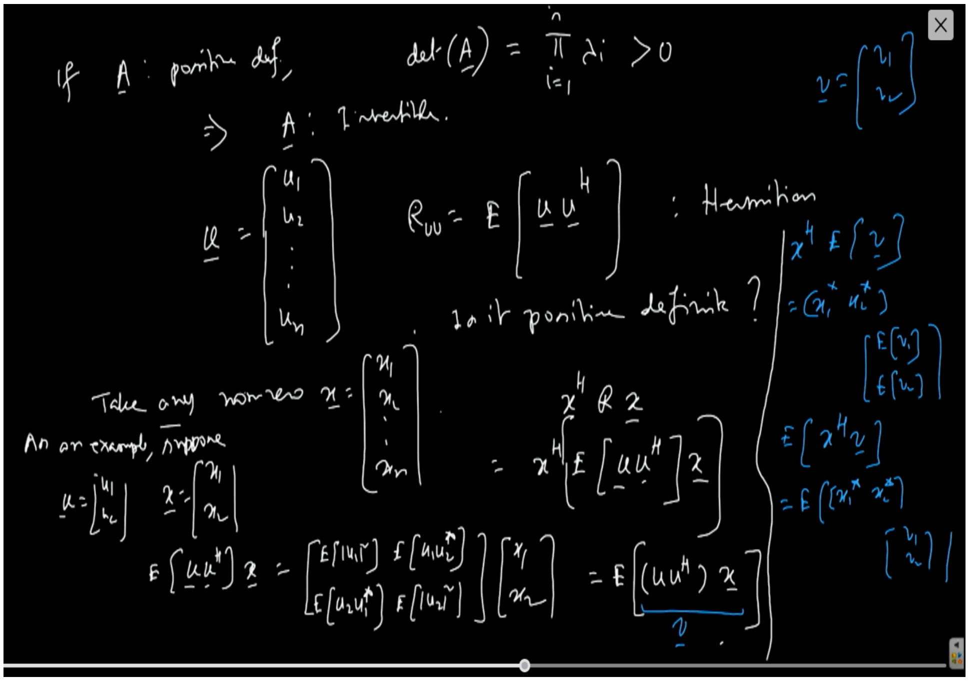


So, this part is clear ok. By the same way xh then from outside I have now xh. So, if I have xh before it I will show it can be seen that xh also can be brought inside very simple. Okay, I need some space to write there is a problem. So, this is a matrix times x is a column vector entry are not random alright. So, you can see one thing if you call this if you call this just a minute, if you call this some vectors may be small v u h at x1 x2 there expected value.

So, it is not a ev ok. So, basically u h and then x1 x2. So, mod u1 square into x1 plus this into x2 and likewise e outside. So, this vector is random vector again ok. So, now originally, I have xh times this. So, if I bring that xh and this entire thing which is E of v ok.

v is a I denote as v1 v2 2 elements, then it will be x1 h means x1 star x2 star and E v1 E v2 and we multiply this by x1 star this by x2 star. you will see we get the and then add x1 star times this plus x2 star times this and add that is what this is. But suppose I do I push xh before v, because xh is not random v then again what happens E of x1 star. So, this vector comes inside and v1 v2 as before. So, x1 star v1 E on that, but x1 star is not random.

So, it will go out like x1 star just E of v1 E of v1. Similarly, x2 star into this E on that, but x2 star is not random it will go out like here x2 star and E of v2. So, you get the same thing alright.



So, that means this will be xh Rx I can write xh Rx in general which is xh E of uuh into x. So, xh can be brought in inside x can be brought here.

So, E of xh v and uh this is xh u and uh x. x as from here gone in here xh has gone in here this is what xh u uh x. So, if uh x, u is a column vector uh is a row vector because you transpose and then conjugate. So, row vector times x, x is a column vector.

So, row into column it is a scalar. So, this quantity is a scalar. You can this x we can call it anything maybe we can call it y, y is a scalar. Then if uh x is y is a scalar if I take y star y is a scalar, scalar is also can be viewed as a 1 by 1 matrix. So, it is like yh because y is a 1 by 1 matrix first you transpose you get that y only then star, So, you get this. But yh means this side if I apply h it will be xh uhh.

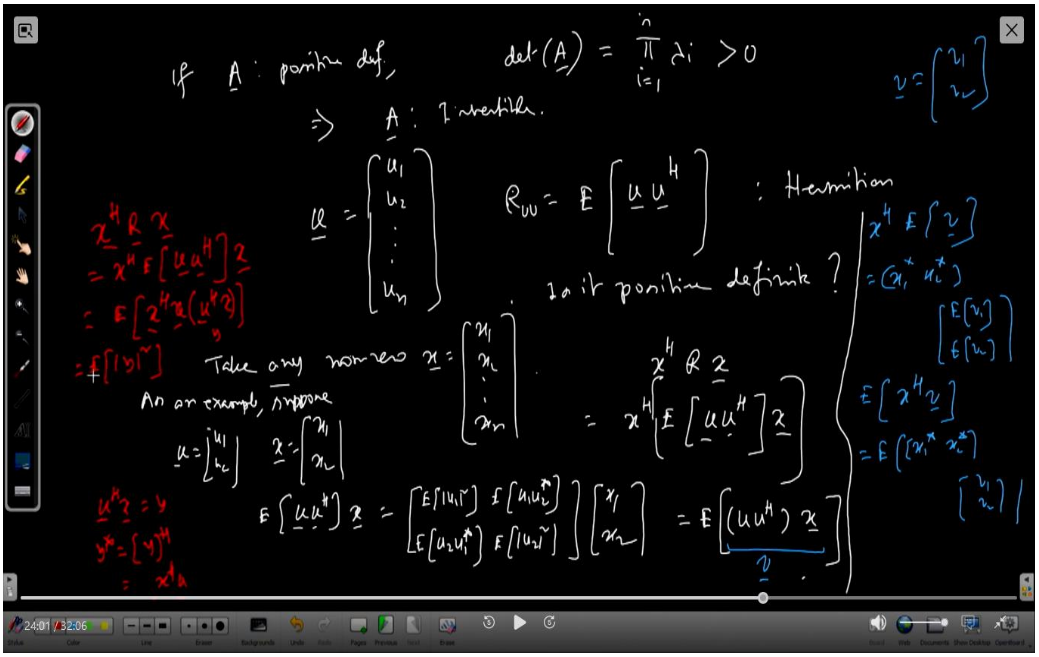
So, u alright which is my uh x sorry which is xh u. uh x I started with I called it y if I apply conjugation on y, y star which is like y hermitian I get this quantity. So, this is nothing, but E of mod y square y star this is y star this is y E of mod y square. So, this is what I take to the next page. So, if this quantity is either greater than equal to or equal to 0 that will see.

we are dealing with the special case of correlation matrices or covariance matrices. For them first condition was satisfied that this matrix is hermitian second is is it positive definite. So, I took a non-zero vector x and I found out xh rx. xh rx can be written in this kind of form R you write and then xh can be pushed inside x can be pushed inside. So, u h x I call y is a scalar because row vector into column vector.

So, this is y star E of mod y square. One thing you can see y is a random variable because even if x is deterministic vector non-random vector elements of u they are random. So, u h x will be a scalar like u1 star x1 plus u2 star x2 plus u3 star x3 dot dot dot. So, all u1 u2 u3 they are present means this entire thing is random and its conjugate also random. So, E mod y square is random because y is random and the expected value, but remember it is mod y square. So, mod y make it real and positive and then square.

So, this number can never be negative ok. This number can never be negative either 0 or positive. So, expected value can be if it is always 0 then only expected value will be 0, but if it is sometimes 0 in some experiments it is positive mod y and therefore, mod y square positive then average will be positive. So, this entire quantity will be either positive or 0. So, in general positive same in affinity. It will be 0 if mod y square and therefore, mod y turns out to be 0 in every experiment.

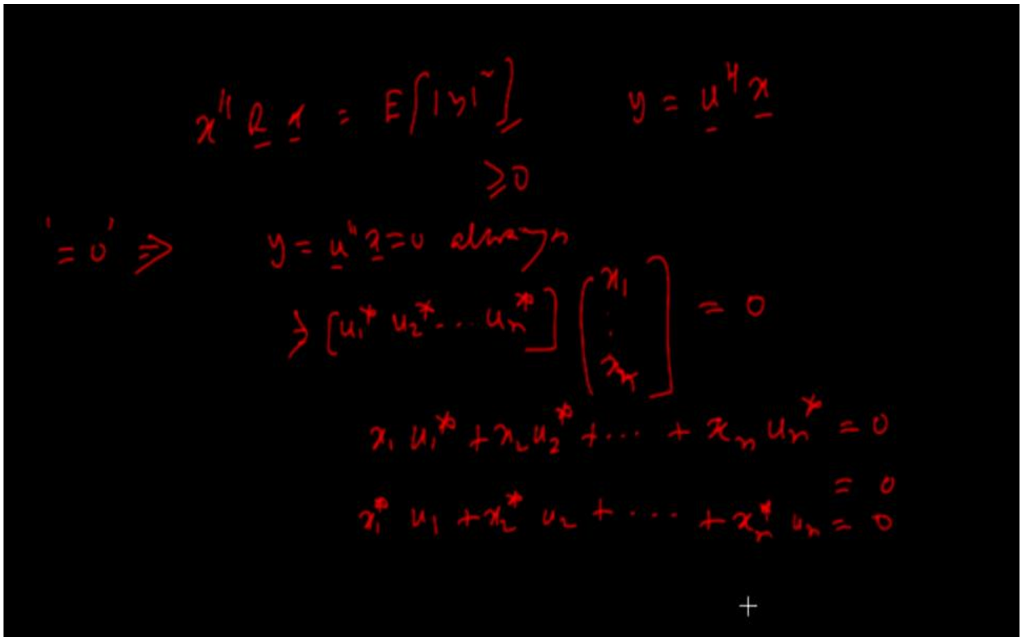
So, when you average by applying E you still get 0 that is a very special case. But if that does not happen in some experiments or maybe in all experiment’s mod y is positive or at least in some experiments positive and then you average mod y square you will get a positive quantity. So, in that case it is positive definite. So, this is either positive semi definite or positive definite.



So, with that I go to the next page. This what I had earlier x h r x is expected value of mod y square where y is a scalar and that was u h x alright it is a scalar. So, this quantity is obviously greater than equal to 0 because as I told mod of y real positive and square. So, this can never be negative average value can never be negative. If it is 0 that means, in all experiment's mod y should be 0 and this is the average is 0. Otherwise in some experiments at least it will be non 0.

So, when you average average value will be greater than 0. So, it is positive definite. So, when will it be equal to 0 that condition that equal to 0 here means equal to 0 means y equal to 0 always u h x always, but u h x is, u is a column vector u h is a row vector with conjugated elements and this is your x1 dot dot dot dot x2 equal to 0 always sorry. As I told you if this expected value has to be 0 then y must be 0 in all experiments. Then only mod y 0 and therefore, in all experiments and therefore, mod y square when averaged expected you get 0.

But this means uhx this equal to 0 which means u1 star x1 is a scalar of your choice. So, x 1 u 1 star is a scalar can be 2 3 4 anything then x 2 u 2 star and dot dot x n u n x n u n star is equal to 0. I can take complex conjugate of left hand side and right hand side right hand side is 0. So, that remains 0 and this you can say x 1 star which is again a known constant times u 1, x 2 star times u 2 and dot dot dot x n star times u n equal to 0. Suppose all the coefficients x 1 star x 2 star x n star they are non 0 and still this is satisfied.



In that case I can keep one element maybe u 1 on one side take the other ones on the right hand side. So, u 1 will be a linear combination of other variables u 2 u 3 u n. Even if not all are non 0 maybe 2 are non 0. I take the non 0 1 the other one on right hand side. So, it will be u 1 is minus x 2 star by x 1 star into u 2 ok.

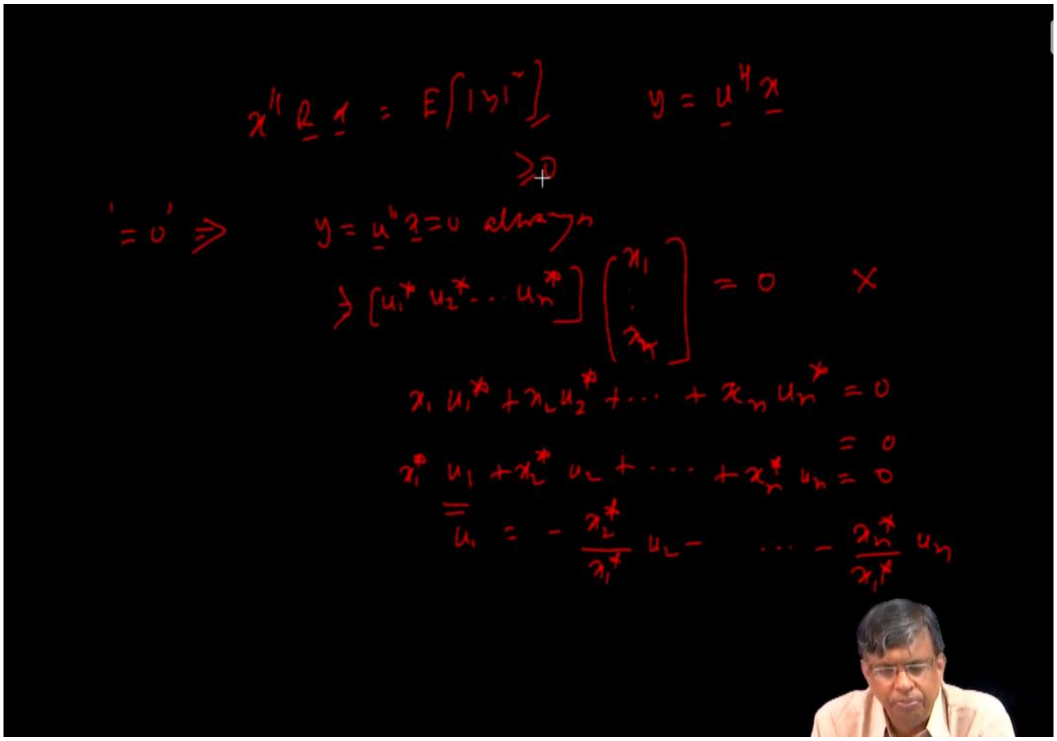
So, one random variable at least will be a linear combination of the remaining ones or part of the remaining ones. But that does not happen in practice usually because when I have got subset of random variable there is no such linear relation between them then at least one of them is a linear combination of the remaining ones. Like as I told you if I keep u 1 here and suppose all are non 0 it will be like this is non 0. So, it can go in the bottom. So, as though u 1 is a random variable, but it is basically having a linear relation with the rest.

So, if I know the rest, I know u 1 ok. That is u 1 does not have independence which is a linearly dependent that does not happen in practice most often. Random variables are random there is no linear relation underneath you know in the background there is no linear relation between them. That just like somebody speech you just suppose say something and I see the waveform I take a sample, samples are random. Because next time I see the same waveform I mean same speech I will see a different waveform sample will vary. So, at all-time point samples are dancing fluctuating, there is no linear equation relation between those samples no way.

If that happens that means, this cannot be equal to 0. So, if this is not possible there is no there is no linear relation between the random variables, they are purely random ok. They are they are independent of each other we say linearly independent of each other because I am taking a linear combination equated to 0. In that case this fellow at least is dependent on the others may be more fellows are dependent on the others that is called linear dependence. So, there is no such linear dependence within these random variables.

They are purely random like the speech example I gave. In that case you cannot have this equal to 0 with non-zero coefficients. Because when I talk of positive definite this x must be non-zero, that is at least one coefficient non-zero or all coefficients are non-zero. So, non-zero coefficients they are these equal to 0 cannot happen because that will have been linear dependence that is at least one random variable is linearly related to others ok. Even if suppose this is a vector only, I mean I want x to be I have to have x to be non-zero. So, suppose you say let only one fellow be non-zero others are 0 let x1 be non-zero others are 0 that will mean x1 times u1 star or rather x1 star u1 equal to 0 and x1 is non-zero.

So, that means u1 is a 0 random variable because if x1 star u1 is 0 and x1 is non-zero. So, x1 star also non-zero. So, non-zero into u1 equal to 0 means u1 must be 0 that is u1 is such a random variable which will take 0 value always ok. So, when you take a product with some non-random scalar result is 0 in all experiments. But in practice we do not encounter this kind of random variable where which in every trial takes only prefixed 0 value.



So, we assume that does not happen there is there is no linear relation between these random variables between any pair or more between 2 between 3 or between all n of them and none of them is a 0 random variable. That is none of them is such that it takes 0 random 0 value in all experiments which is a very practical case. In such cases this is not possible for a non-zero x ok which means in such cases it is we can assume the matrix to be R to be positive definite because it will be actually greater than 0 equal to 0 rule out alright. And in that case R will be positive definite and therefore, R is invertible. So, in our future lecture's correlation or covariance matrices like R will take them to be assume them to be positive definite and then we can write R inverse that is we will assume this inverse exists ok.

This much for this lecture and from here I will pick up in the next. Till then goodbye.