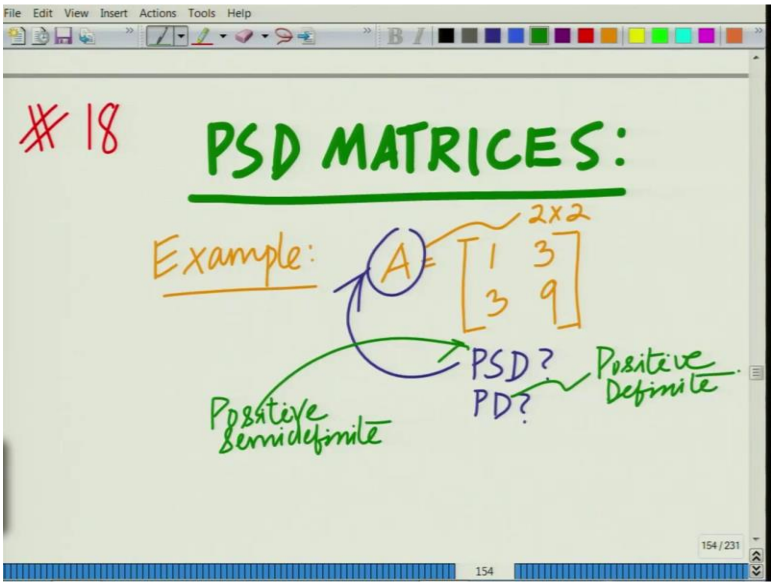
Applied Linear Algebra for Signal Processing, Data Analytics, and Machine Learning
Professor. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur
Lecture No. 18
Positive semi-definite matrices: example and illustration of eigenvalue decomposition

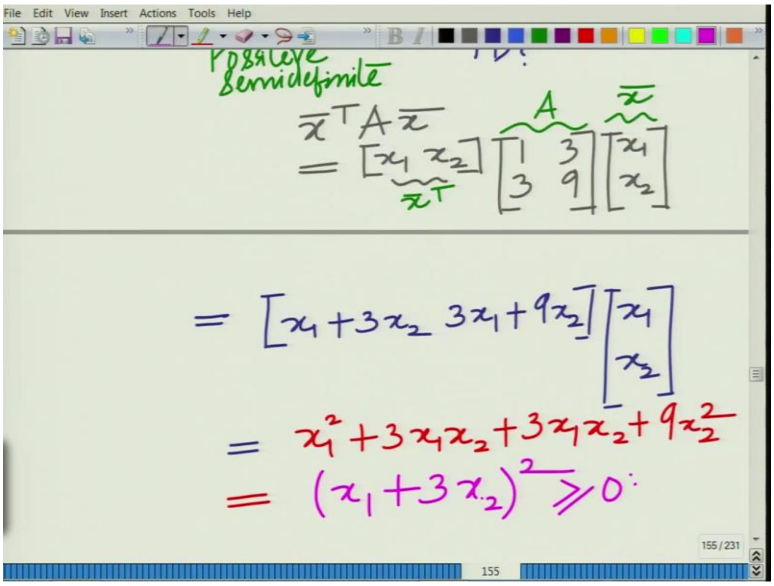
Hello, welcome to another module in this massive open online course. So, we are looking at positive semi-definite and positive definite matrices. So, let us continue our discussion.

(Refer Slide Time: 00:24)



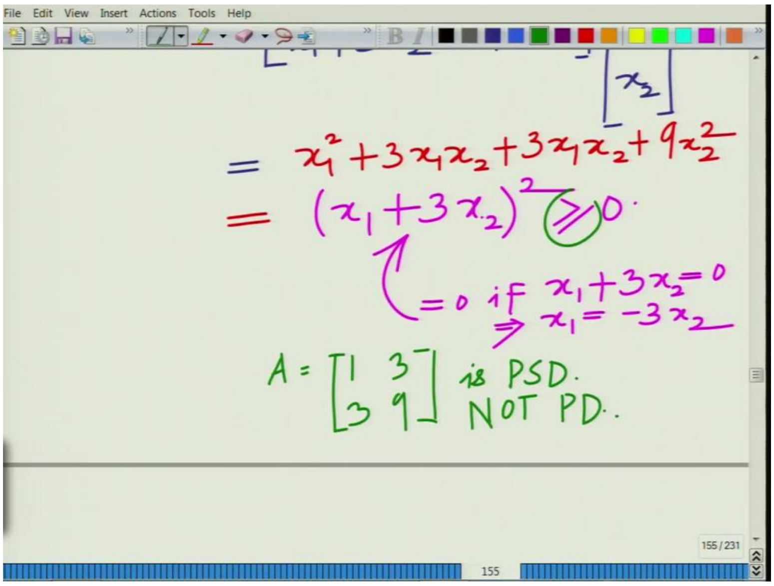
So, in fact, let us look at an example, let us look at a simple example for a PSD matrix. Consider the matrix A which is equal to 1 3 this is our 2 cross 2 matrix A, let us try to see is this positive definite or positive semi-definite, so is this PSD or is this PD? Of course, you might recall PSD essentially means positive semi-definite and PD is positive definite. So, this is either positive semi-definite or this is positive definite.

(Refer Slide Time: 1:46)



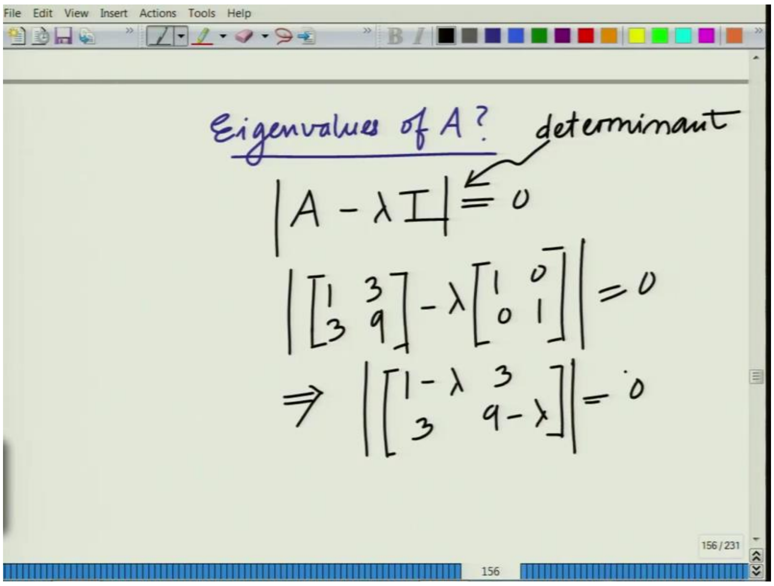
So, to do that, let us follow the first principles let us look at x bar transpose A x bar for any vector x bar. So, this is a real matrix clearly. So, we can look at x bar transpose A x bar and this is equal to x1, of course, the vector x has to be 2-dimensional since a is 2 cross 2 matrix 1, 3, 3, 9 times x bar. So, this is your x bar transpose, this is the matrix A, this is the matrix x bar this is equal to, I can write this as follows, so first multiply the row on the left by the matrix A so this will give me x1 plus 3 x2 and then you have 3 x1 plus 9 x2 times x1 x2 and that gives us, if you continue with this, that gives us x1 square plus 3 x1 x2 plus 3 x1 x2 plus 9 x2 square which is equal to, you can easily, x1 plus 3 x2 whole square which is greater than or equal to 0.

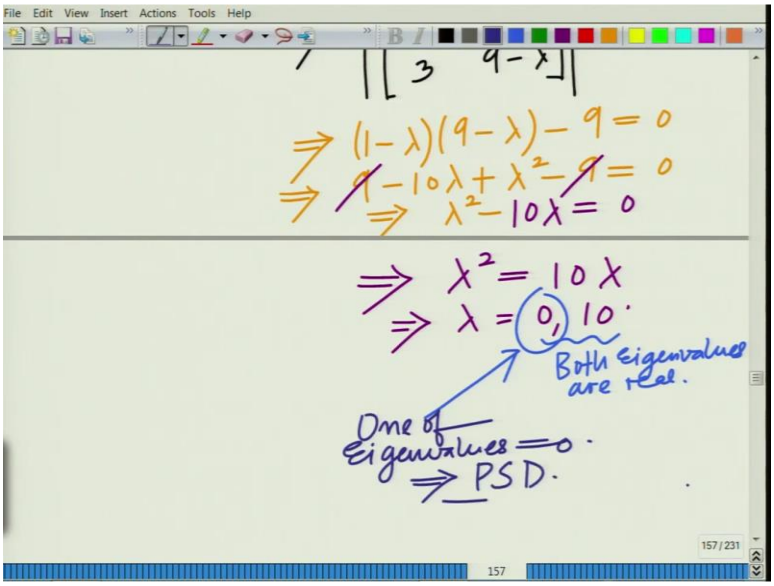
(Refer Slide Time: 03:56)



In fact, this can also be equal to 0 if x1 equals minus x1 plus 3 x2 equal to 0 or x1 equal to minus 3 x2. So, this is only greater than or equal to 0, this is not always greater than 0. In fact, it can be equal to 0 if x is equal to minus three x2 therefore, this is only positive semi-definite, not positive definite. So, in fact, since this is only greater than or equal to 0, this matrix A equals 1, 3, 3, 9 this is PSD, not PD. Since this can be equal to 0 whenever you have a vector x bar such that x1 equal to minus 3 x2. So, A is in fact a positive semi-definite matrix. Let us continue with this example 5 further.

(Refer Slide Time: 05:10)

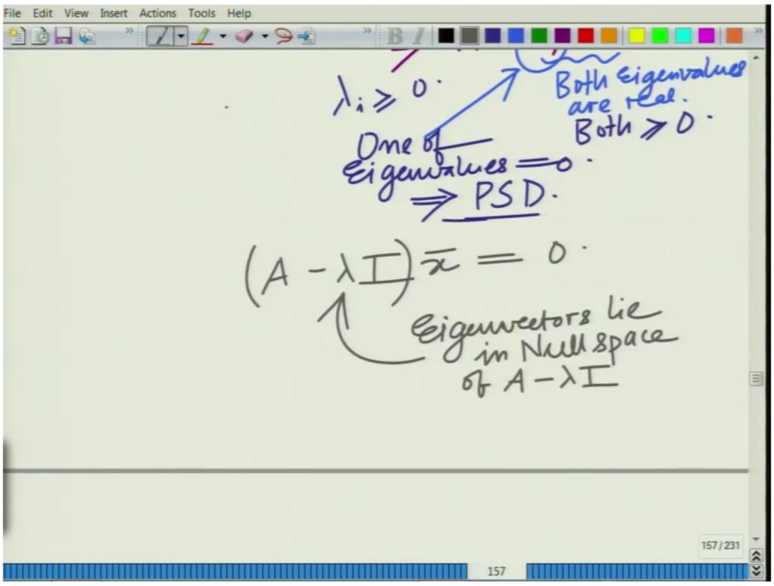


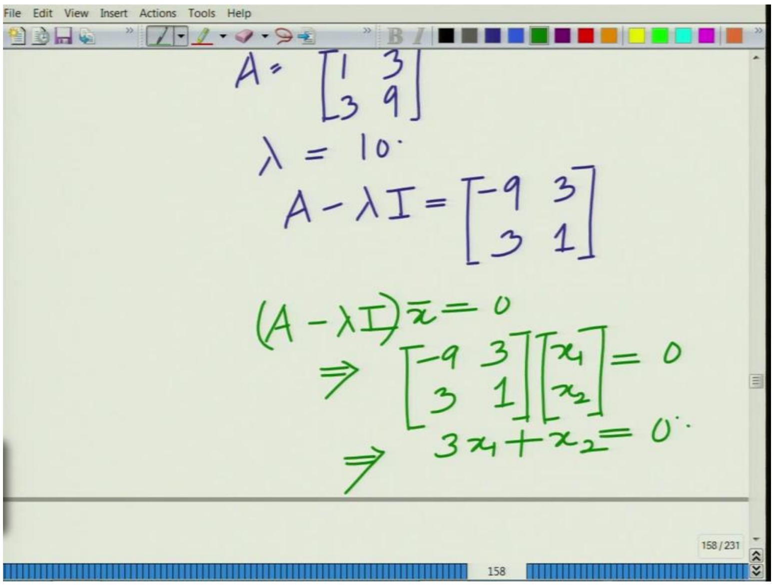


What about the eigenvalues of A? So, to find the eigenvalues remember we have to find A minus lambda I, set the determinant equal to 0 you have to set the determinant equal to 0 this is 1, 3, 3, 9 minus lambda times, of course, you have to take the 2 cross 2 identity matrix, this equal to 0, this implies that you have 1 minus lambda 3, 9 minus lambda the determinant of this must be equal to 0. This implies that 1 minus lambda into 9 minus lambda minus 3 into 3, 9 equal to 0, which implies 9 minus 9 lambda minus 10 lambda plus lambda square minus 9 equal to 0 which implies lambda square.

So, you can see the nines go away. So, you have lambda square minus 10 lambda equal to 0 which essentially implies lambda square equal to 10 lambda which implies lambda equal to either 0 or 10. And in fact, now, you can see one of the eigenvalues is 0, there is eigenvalues are not greater than 0 but eigenvalues are greater than or equal to 0. Hence, this can only be a positive semi-definite matrix. So, first interesting thing both eigenvalues are real, one of the eigenvalues is equal to 0 implies once again that this can only be a positive semi-definite matrix.

(Refer Slide Time: 08:01)



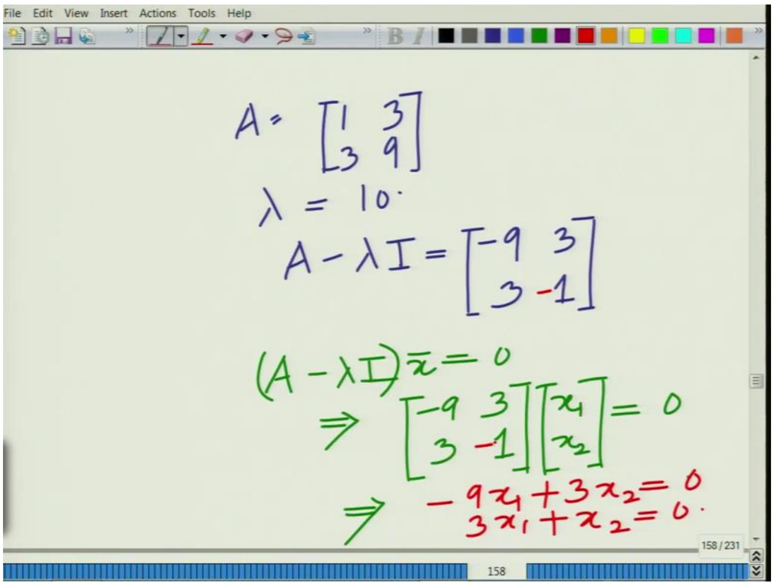


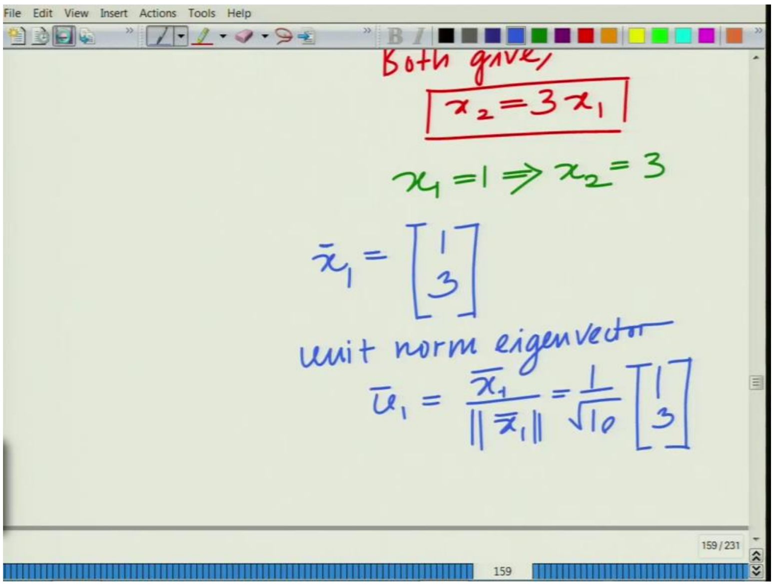
In fact, you can see lambda i eigenvalues are greater than or equal to 0, both eigenvalues are real both are greater than or equal to 0, both eigenvalues are in fact greater than or equal to 0 that is what we expect of a positive semi-definite matrix eigenvalue are positive definite matrix, eigenvalues have to be greater than 0, positive semi-definite matrix eigenvalues have to be greater than or equal to 0. In fact, in this case, one of the eigenvalues is 0. So, this is only a positive semi-definite matrix.

So, let us now find eigenvectors of this matrix and let us check the property about the eigenvectors, I hope all of you remember the property about the, of course, the way to find the eigenvectors is, eigenvectors lie in the null space of A minus lambda i. I hope all of you remember this, eigenvectors lie in the null space of A minus lambda i, what is A? You have a equals 1, 3, 3, 9, take lambda equal to 10 A minus lambda i equals 1 minus 10 that is minus 9, 3, 3, 9, minus 10 that is 1.

So, we have A minus lambda i, we have to find the vector x bar in the null space of this. So, A minus lambda i times x bar equal to 0 this implies minus 9, 3, 3, 1 times x1, x2 equal to 0 which implies you can see there is redundancy in this. So, we can only solve one equation. So, 3 x1 plus x2 equal to 0, we are solving the second equation, you can solve the first equation, you will get the same factor.

(Refer Slide Time: 10:17)



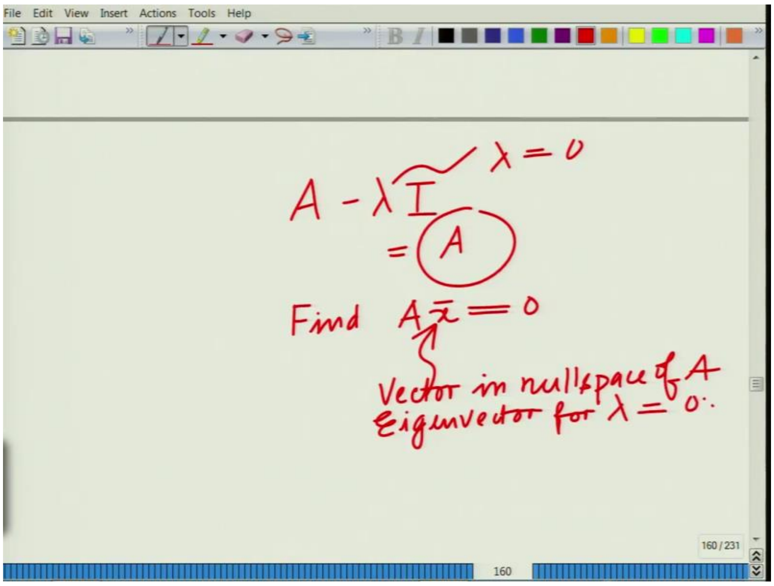


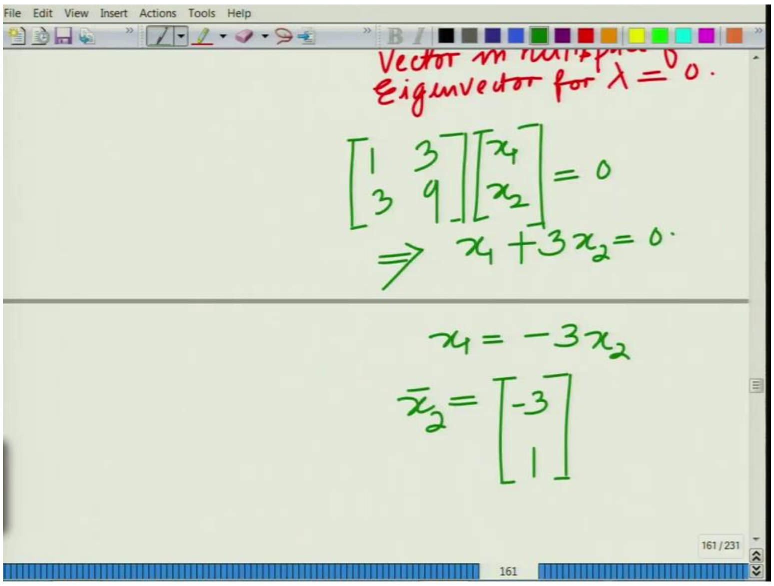
We write the first equation, let us write that I mean to avoid any confusion, minus 9 x1 plus 3 x2 equal to 0, 3 x1 plus x2 equal to 0 both give minus 9 x1, this has to be minus 1 I guess. So, A minus lambda this has to be minus 1. So, this has to be 9 minus 10, which is minus 1 both, so 3 x1 minus x2 equal to 0 both give x2 equal to 3 x1, this is the property the eigenvector has to satisfy.

So, if you set x2 equal to 1, x2 or x2 equal to, or x1 equal to 1, this implies x2 equal to 3. So, you have a free parameter that is x1 and you can, and that naturally is the case because the eigenvalue, eigenvector is invariant on scaling. So, you can scale the eigenvector by any constant and it will still remain an eigenvector and that is exactly reflected here where you can see there is a free parameter you can choose x1, x2 gets determined automatically.

So, the eigenvector x bar equal to, what will the eigenvector be x1 equal to 1. Now, the unit norm eigenvector I can call this as the eigenvector. Now, for the corresponding unit norm eigenvector I have to divide this by the norm. So, the unit norm u1 bar let us call this as x1 bar, u1 bar equals x1 bar divided by norm x1 bar which is equal to 1 over square root of 10 1 comma 3.

(Refer Slide Time: 12:43)



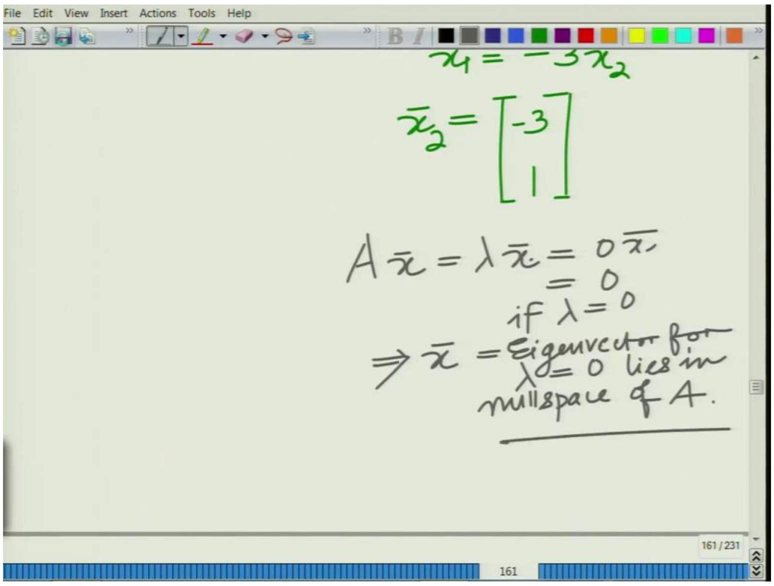


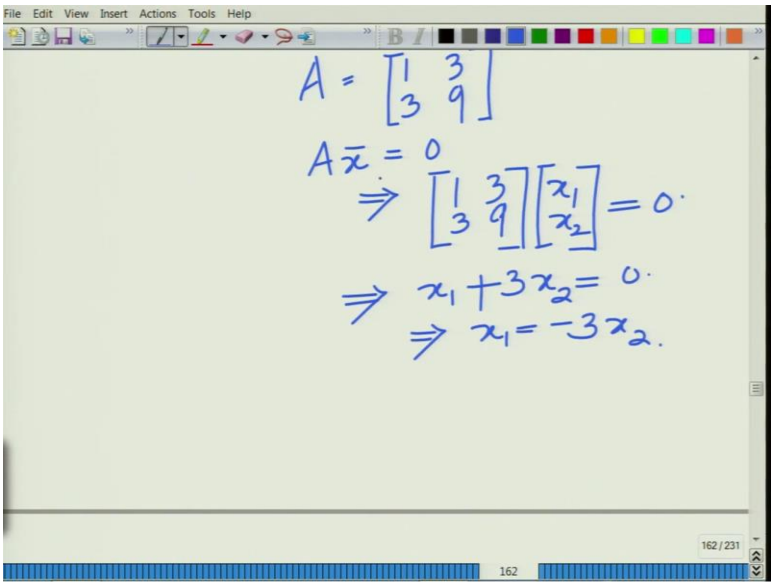
Now, similarly, if you look at A minus lambda I corresponding to lambda equal to 0, this will simply be A. Now, to find the eigenvector corresponding to lambda equal to 0, we have to find the null space, find A x bar equal to 0 that is vector in null space of A. Since lambda equal to 0 this gives eigenvector corresponding to lambda equal to 0 and this you can find as follows, you have 1, 3, 3, 9 times x1, x2 equal to 0 this implies x1 plus 3 x2 equal to 0.

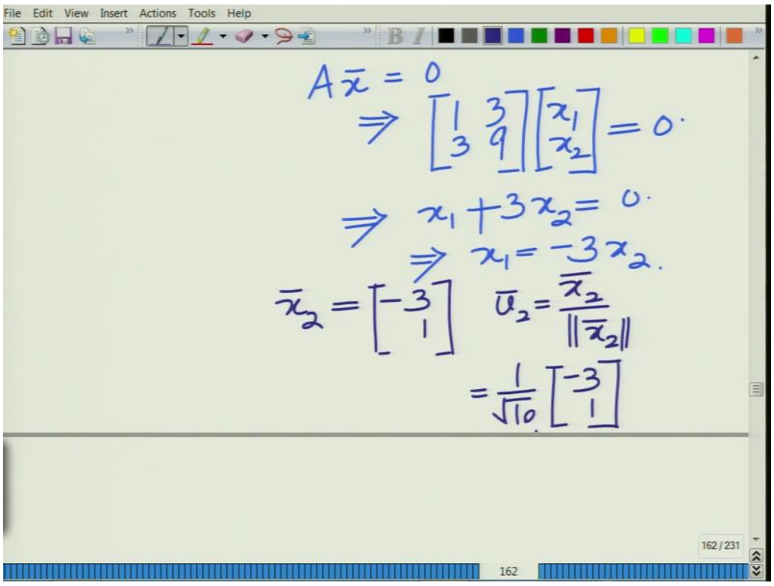
You can solve again the same thing as the previous case you can solve the second equation it will give you the same thing which implies x1 equals minus 3 x2 there is again a free parameter in this case you can say x2 is a free parameter or x1 is a free parameter whatever it is, so, if you set one it determines the other, so x1. So, if you set x2 equal to 1 x1 will be equal to minus 3, so that is we can call this as or x2 bar this is eigenvector.

In fact, this is also the vector, this is also a basis for the null space of the matrix A, you can say this also corresponds to an eigenvector of the matrix, this eigenvector of the matrix A corresponding to the eigenvalue 0, which is essentially the same thing as saying A x bar equals 0 times x bar which is 0 and that essentially says that x bar is in the null space. So, I hope that is also clear.

(Refer Slide Time: 14:43)



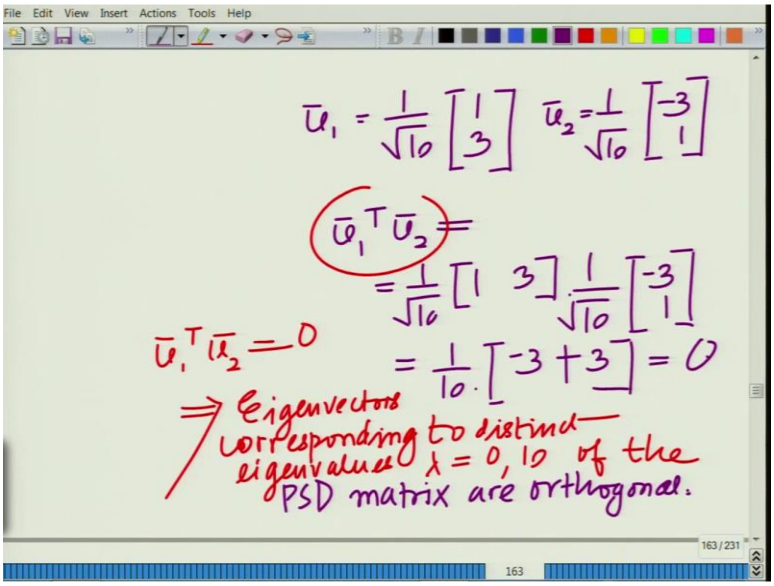




So, A x bar, if you look at A x bar equal to lambda times x bar equal to 0 into x bar is equal to 0, if lambda equal to 0 which means x bar eigenvector for which is x bar which is the eigenvector for lambda equal to 0 lies in null space of A. So, it is a straightforward result I mean it is not very difficult to see and now if you look at A equal to 1, 3, 3, 9 and if you set out to find A x, solve A x bar equal to 0 this implies 1, 3, 3, 9, x1 into x2 equal to 0 and this implies if you look at this x1 plus 3 x2 equal to 0 which implies x1 equals minus 3 x2. I think this is essentially what we have already found.

So, this is the eigenvector corresponding and x2 bar equal to essentially that is what we have so, then the eigenvector is corresponding to this is x2 bar equal to minus 3 comma 1. So, we have x2 bar equals minus 3, 1 and we have u2 bar which is x2 bar divided by norm x2 bar and this is 1 over square root of n minus 3 comma 1.

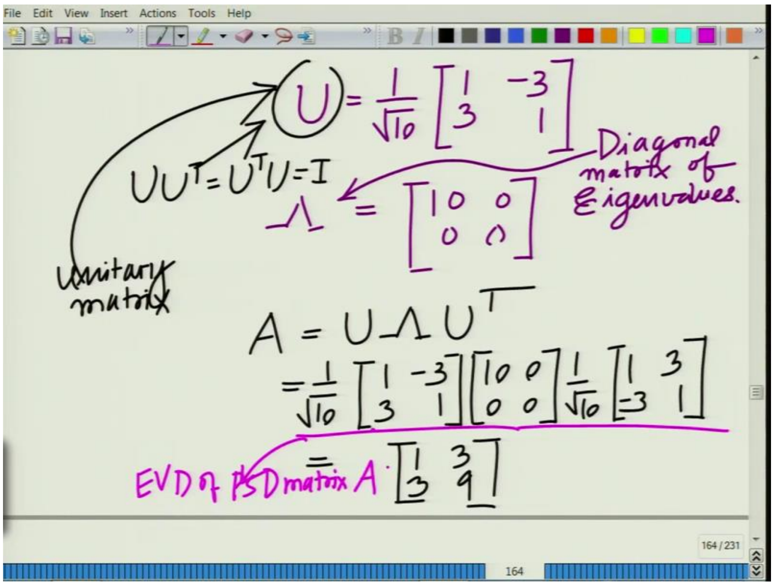
(Refer Slide Time: 16:57)



Now, let us look at these two eigenvectors u1 bar equals to unit norm eigenvectors u1 bar equal to 1 over square root of 10, we have I guess 1, 3 and u2 bar equal to 1 over square root of 10 minus 3, 1 it is not very difficult to see that u1 bar transpose u2 bar equals 1 over square root of 10 1, 3 into 1 over square root of 10 minus 3, 1 and this is nothing but 1 over 10 into 1 into minus 3, minus 3, plus 3 which is equal to essentially 0.

Therefore, what you can see very easily is that u1 bar transpose u2 bar equal to 0 which again verifies the property that the eigenvectors corresponding to distinct eigenvalues of the positive semi-definite matrix. In this case lambda equal to 0 comma 10 of the PSD matrix are orthogonal that is eigenvalues eigenvectors corresponding to the distinct eigenvectors that is u1 bar equal to 1 over square root of 10 1 comma 3 and u2 bar are equal to 1 over square root of 10 minus 3 comma 1 these are orthogonal.

(Refer Slide Time: 19:02)



And now, we follow the matrix of these unit norm eigenvectors. So, we have unit norm orthogonal eigenvectors, we have U equals 1 over square root of 10 that will be 1, 3, minus 3, 1 and the diagonal matrix of eigenvalues, in this case, is lambda which is equal to, it is not very difficult to see, 10, 0. This is the diagonal matrix of eigenvalues. In fact, U you can see is a unitary matrix UU transpose equal to U transpose U equals identity, this U you can see contains orthonormal columns.

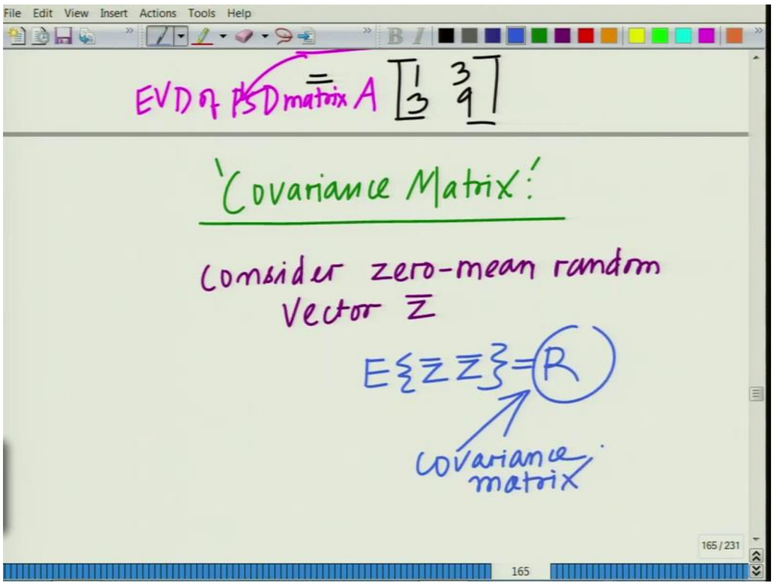
So, this is unitary and therefore, A you can clearly see can be written as U lambda U transpose which is 1 over square root of 10, 1 comma 3, minus 3 comma 1 times lambda which is essentially 10, 0, 0, 0 times U transpose which is 1 over 10, 1 comma 3, minus 3 comma 1 and if you multiply this you can see you will get back A which is 1 comma 3. So, this is the eigenvalue decomposition of the PSD matrix that is essentially U sigma U transpose that is the eigenvalue decomposition of the PSD matrix A.

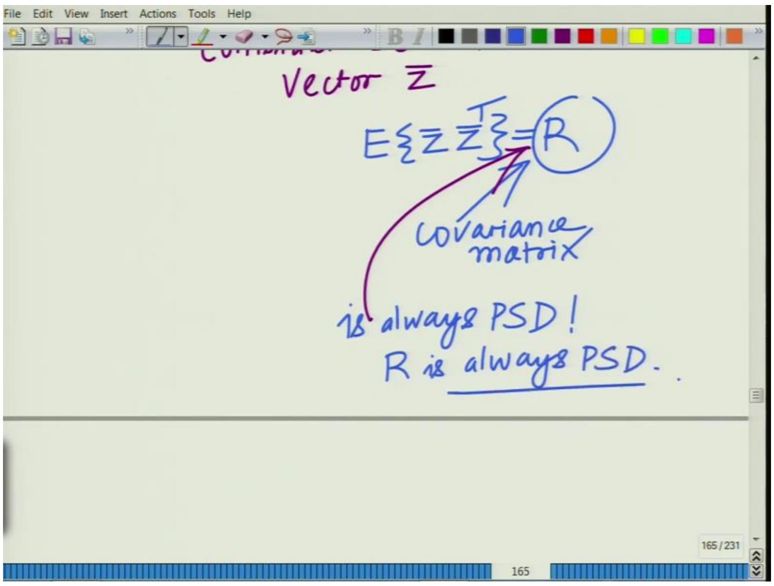
So, although this is a very simple example, it is simple 2 cross 2 example illustrates several interesting points, essentially, how do you find the eigenvalues of the PSD matrix and from that, you find that one of the eigenvalues is 0. So, the in fact is a positive semi-definite matrix, then you find the eigenvectors, then you find the unit norm eigenvectors. In fact, you realize that the eigenvectors are orthogonal because these are the eigenvectors corresponding to a PSD matrix for distinct eigenvalues, the eigenvalues are also greater than or equal to 0 they are real and greater than or equal to 0.

And finally, once you construct the unitary matrix U corresponding comprising of the orthonormal that is the unit norm and orthogonal eigenvectors as its columns, then you will realize that U is a unitary matrix UU transpose U transpose U is identity and the eigenvalue decomposition is in fact given as U lambda U transpose. See, normally, for a normal square matrix is given as U lambda U inverse but for the positive semi-definite matrix it is given as U lambda U transpose because U transpose itself is U inverse.

In fact, in this case, it is given as U times lambda times U transpose where lambda is the diagonal matrix comprising of the eigenvalues the first one is the lambda 1 0, lambda 1 is 10 and the next one lambda 2 equals 0. So, the diagonal matrix will be lambda is a diagonal matrix with the diagonal elements given us 10 0. So, it is a very interesting example very simple, but very interesting clarifies a lot of, so the simple examples as they are clarify a lot of interesting ideas and clarify I mean, they illustrate very clearly several ideas.

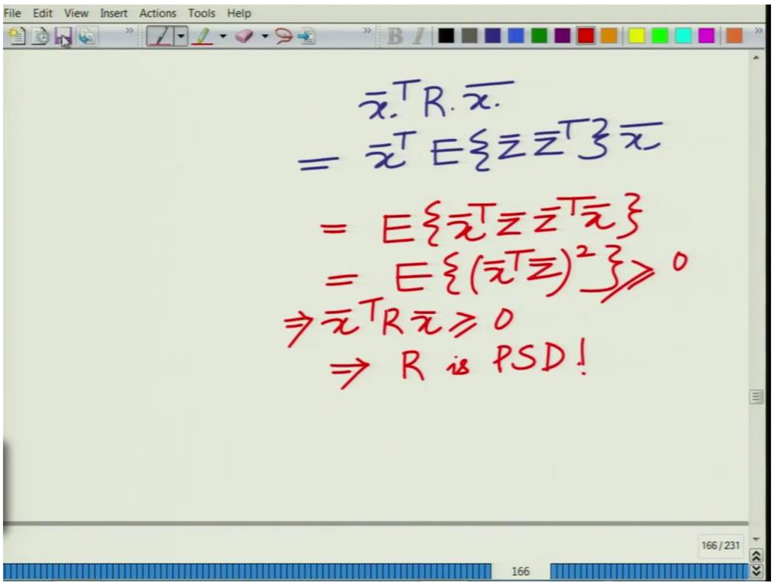
(Refer Slide Time: 23:17)

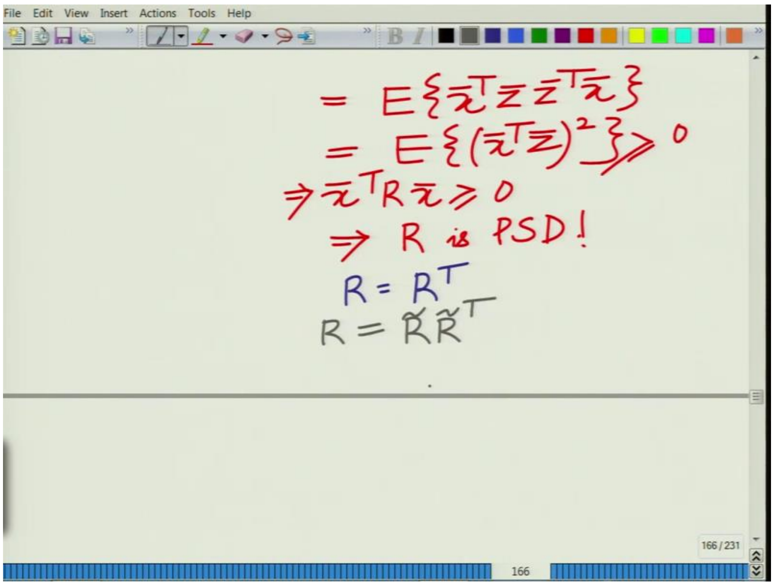




Now, let us look at another interesting application. Remember I told you positive semi-definite matrices arise very frequently in linear algebra random vectors, why is that the case, and for that you have to look at again the covariance matrix that is consider, let us consider a zero-mean random vector with covariance, this R we know is called or is termed as the covariance matrix, and this has to be expected of Z bar Z bar transpose. Now, turns out that R is always, the covariance matrix is always a positive semi-definite matrix that is R is always any covariance matrix is always PSD, how, why is that?

(Refer Slide Time: 25:05)



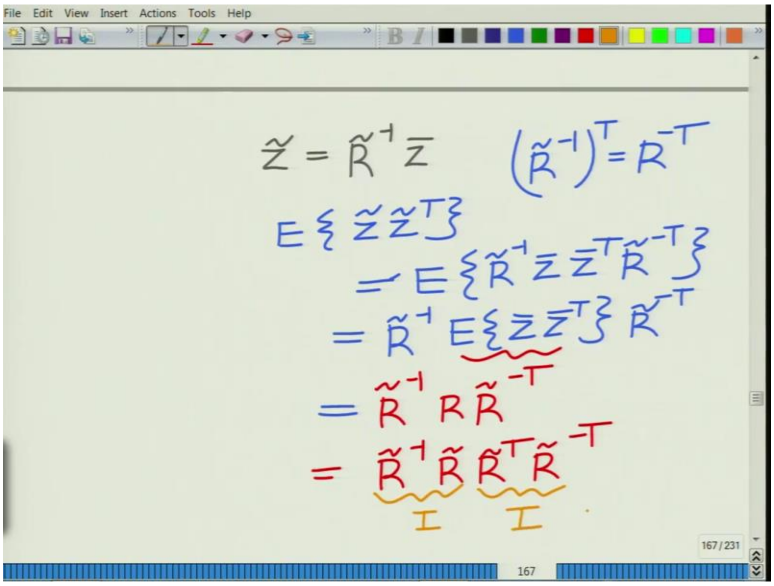


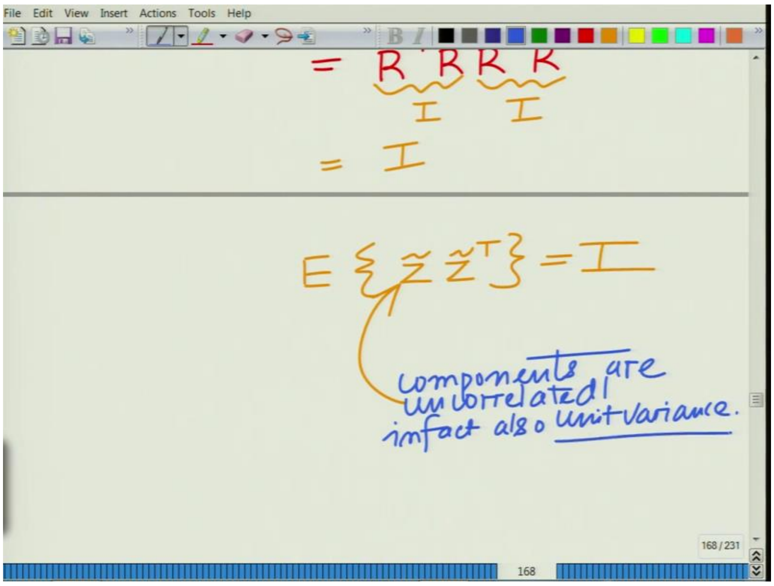
Now, let us take any vector Z bar, let us perform Z bar transpose R into Z bar this will be, let us perform x bar transpose R into x bar is equal to x bar transpose expected value of, R is expected value of Z bar Z bar transpose times x bar. Now, take the x bar inside this is expected value of x bar transpose Z bar Z bar, x bar which is equal to expected value of x bar transpose Z bar whole square.

And now, this is the expected value of x bar transpose Z bar whole square which is a non-negative quantity therefore, this is always greater than or equal to 0. So, this implies x bar transpose R x bar is always greater than or equal to 0 for all x bar, this implies by definition R has to be a PSD matrix because x bar transpose R x bar is always greater than or equal to 0 it being expected value of x bar transpose Z bar whole square.

Therefore, any covariance matrix is always a positive semi-definite matrix and this has very interesting applications. So, first of all, it is a symmetric matrix. So, we have R equal to R transpose. Secondly, if you look at this R can be decomposed because it is a positive semi-definite matrix, R can be decomposed as R tilda R tilda transpose, remember we saw this is the Cholesky decomposition R can be decomposed as R tilda R tilda transpose because it is a positive semi-definite matrix.

(Refer Slide Time: 27:07)



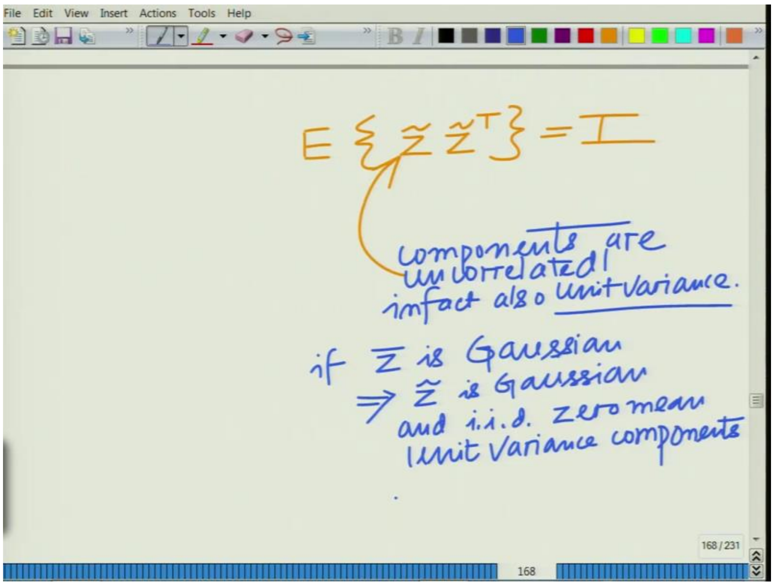


Now, let us look at what happens to our tilda inverse Z bar, let us call that as Z tilda Z tilda equals R tilda inverse Z bar then we have Z tilda Z tilda transpose is expected value of R tilda inverse Z bar Z bar transpose R tilda, which is now if you take our tilda outside, this is R tilda inverse expected value of Z bar Z bar transpose R tilda, this has to be R tilda inverse transpose which I am going to write it R tilda minus T.

So, note R tilda inverse transpose is what I am writing and what is frequently also written as R tilda minus T. And now, this is equal to R Tilda the inverse expected value of Z bar Z bar transpose this is R R tilda of minus T which is nothing but R tilda inverse into R is R tilda R tilda transpose into R tilda minus T and now, you can see this R Tilda inverse into R tilda. This is identity R tilda transpose R tilda inverse transpose is the identity.

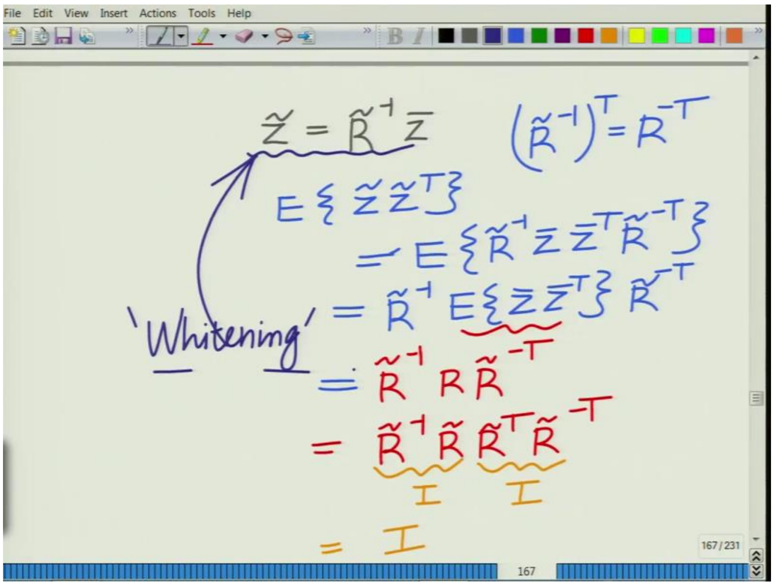
So, this is essentially equal to identity. So, expected value of Z tilda Z tilda transpose equals identity which means, Z tilda contains uncorrelated components. Remember we said whenever the diagonal whenever the covariance matrix is identity, the components are uncorrelated because off diagonal elements are 0. So, components are uncorrelated in fact unit variance. In fact, there are also unit variance this is the diagonal elements are 1. In fact, also, unit variance.

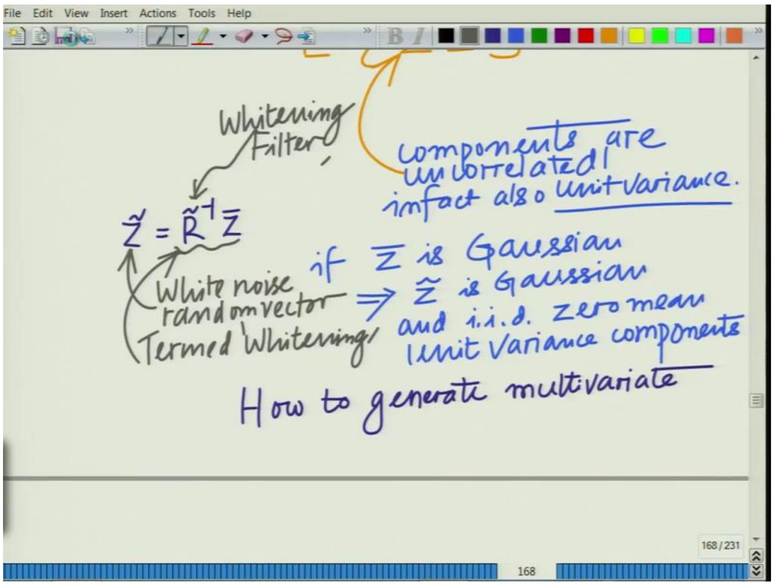
(Refer Slide Time: 29:31)

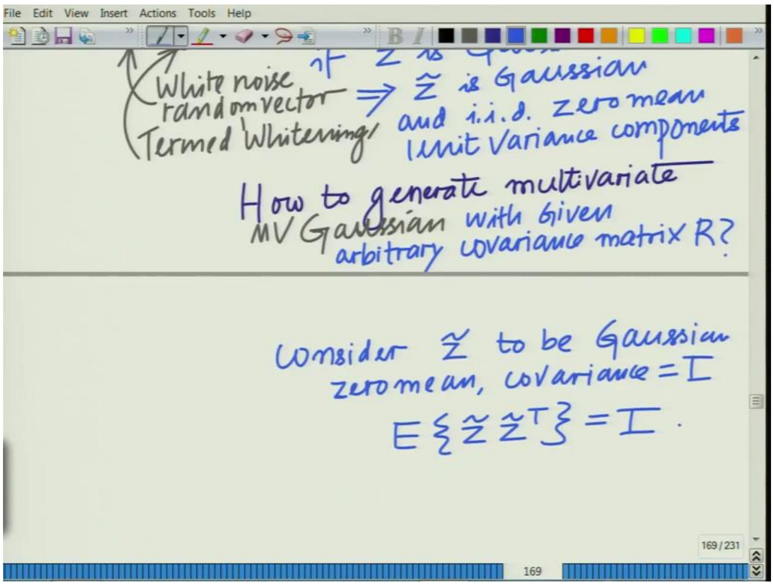


And in fact, if Z is Gaussian, if Z bar is multivariate Gaussian, this implies Z tilda is Gaussian and for Gaussian uncorrelated implies independent Gaussian and i.i.d. zero-mean unit variance components because the tilda the covariance matrix is identity. So, the components are uncorrelated, which for the case of Gaussian also means that they are independent and, of course, since the diagonal elements are 1 it means the variance of each component or each element of this vector Z tilda that is equal to 1.

(Refer Slide Time: 30:43)







The other way out, now, if you want to generate, how to generate multivariate Gaussian, this process is an interesting name. So, this process is termed as whitening. So, this process where you have Z tilda is equal to R inverse Z this is termed as whitening. Why is this termed as whitening? Because it results in white or uncorrelated components because Z tilda contains uncorrelated components.

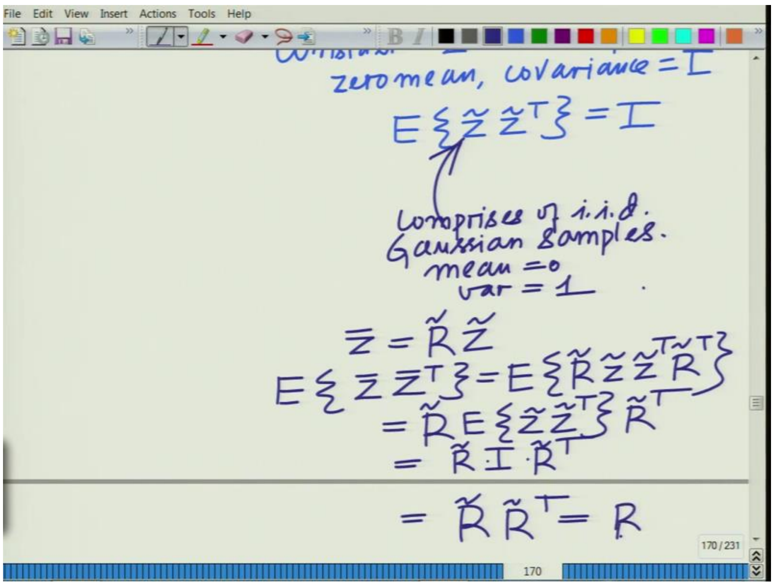
So, if you go back to your knowledge of random processes, when you have a random process in which the different samples are uncorrelated there is autocorrelation is the impulse function if you look at the power spectral density it becomes flat and we call that as a white random process. So, similarly, you have essentially because the covariance matrix is identity. This is a white I can think of this as a white random vector of white noise and therefore, this process that is R tilda inverse into Z bar which results in this white random vector, white noise random vector, this is termed as a whitening filter or this process termed as whitening.

So, Z tilda equals R tilda inverse into Z inverse this Z tilda is a white noise random vector, and this process can be termed as a whitening. And in fact, this R tilda inverse is termed as the whitening filter. This process R tilda inverse and this is termed as a whitening filter. Now, how to generate Gaussian or multivariate Gaussian with given arbitrary covariance matrix R? So, we have looked at the whitening, now, the analog of that in fact you can think of it as the inverse problem that is given a noise process, colored noise process Z bar with covariance matrix are we produce the white noise process.

Now, the question is the other way around, given a white noise because in most simulators you can generate very easily Gaussian samples which are independent and unit variants for instance even MATLAB the standard function Rand n that gives you Gaussian random variables with mean zero-unit variants.

Now, if you construct a vector, how do you generate a vector with a given arbitrary covariance matrix? And that is again very simple consider Z tilda to be Gaussian zero-mean covariance equals identity, that is expected value of Z tilda Z tilda transpose equals identity because most simulators will give you independent identically distributed Gaussian samples of mean 0 and variance 1.

(Refer Slide Time: 35:21)

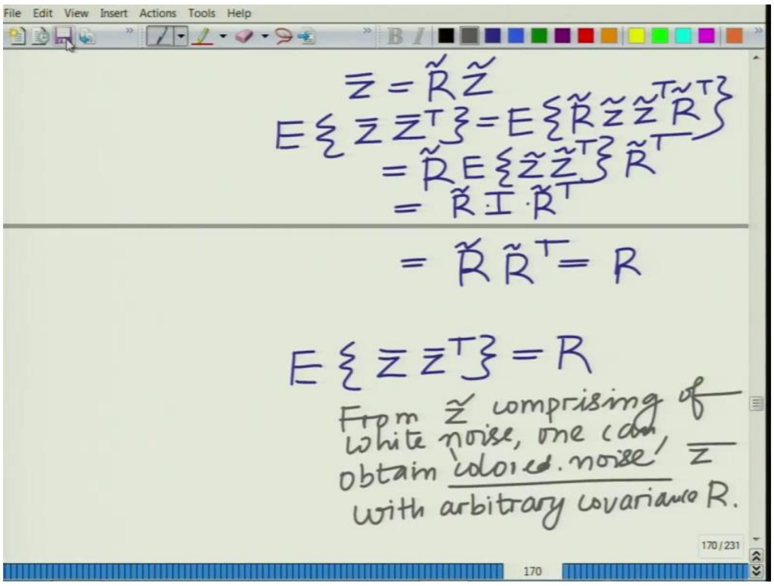


So, essentially Z tilda comprises of i.i.d. Gaussian samples mean equal to 0 variance equal to unity. Now, how do we generate Z bar with given covariance R? Very simple, consider Z bar equal to R tilda Z tilda. And now, if you look at expected value of Z bar Z bar transpose equal to R tilda this becomes equal to R tilda expected value of R tilda Z tilda Z tilda transpose R tilda this equals R tilda expected value of Z tilda Z tilda transpose R tilda transpose this is equal to R tilda.

The expected value of Z tilda Z tilda transpose identity. So, this is R tilda, so this becomes equal to R tilda R tilda transpose which is equal to R. So, now given a Gaussian random vector Z tilda which contains i.i.d. components with mean 0 variance unity, you are now generating a Gaussian random vector Z bar zero mean and arbitrary covariance.

So, this is a coloured noise so, from coloured noise you can come to white noise, from white noise you can go to coloured noise, and that the properties of the positive semi-definite covariance matrix that is R which can be decomposed as R tilda into R tilda transpose that plays a very important role here and this is in fact plays a very important role in entire signal processing. So, it allows you to go from coloured noise to white noise, from white noise allows you to generate coloured noise.

(Refer Slide Time: 37:23)



So, x therefore, expected value of Z bar Z bar transpose equal to R and so, that you have now coloured the noise, so from Z tilda comprising of white noise, one can obtain coloured noise Z bar with arbitrary covariance R. So, from Z tilda comprising of white noise one can obtain coloured noise Z bar with arbitrary covariance matrix that is R. So, these are essentially the way various interesting because we looked at a simple example and we looked at a very interesting practical application of the properties of positive semi-definite matrices.

As I already told you positive semi-definite matrices arise very frequently because these are the covariance matrices have multivariate probability density functions and these are the covariance matrices as well as sample covariance matrices that is when you evaluate the estimates of these covariance matrices naturally any estimate of the covariance matrix also has to be positive semi-definite.

So, the covariance matrices, as well as the estimates of these covariance matrices, are positive semi-definite they have many interesting properties, we have seen that eigenvalues are real they are greater than or equal to 0, eigenvectors corresponding to distinct eigenvalues they are orthogonal and in fact, you can have a square root or Cholesky decomposition of these that is any positive semi-definite matrix R or positive definite matrix R can be expressed as R equal to R tilda into R tilda transpose for a complex matrix it will be R tilda into R tilda Hermitian.

So, these are very, very interesting properties and these arise everywhere in linear algebra, signal processing, machine learning, image processing, these covariance matrices arise everywhere wherever you have noise, wherever you have random vectors, positive semi-definite, positive definite matrices arise everywhere. So, this is a very, very important concept. Please go through this again and understand it thoroughly. Thank you very much.