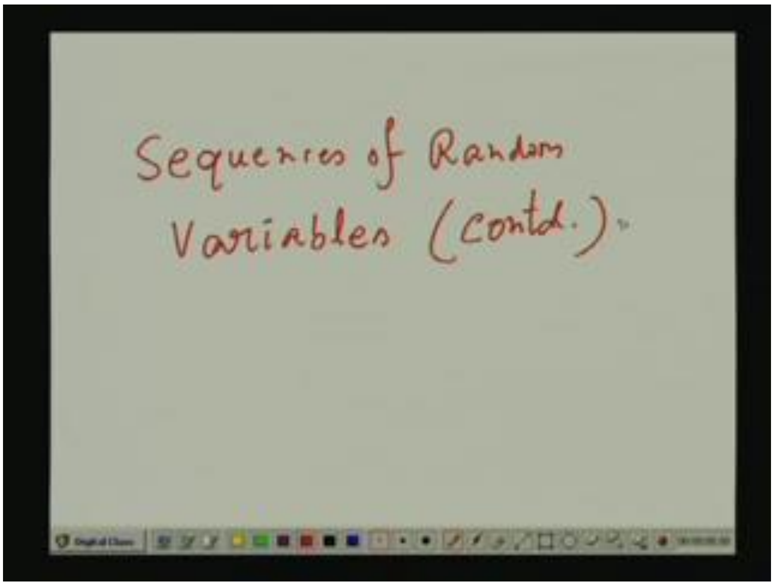
Probability and Random Variables
Prof. M. Chakraborty
Department of Electronics & Electrical Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture - 22
Sequences of Random Variable (contd.)

(Refer Slide Time: 00:36)

So, now we have been discussing this issue of random sequences.

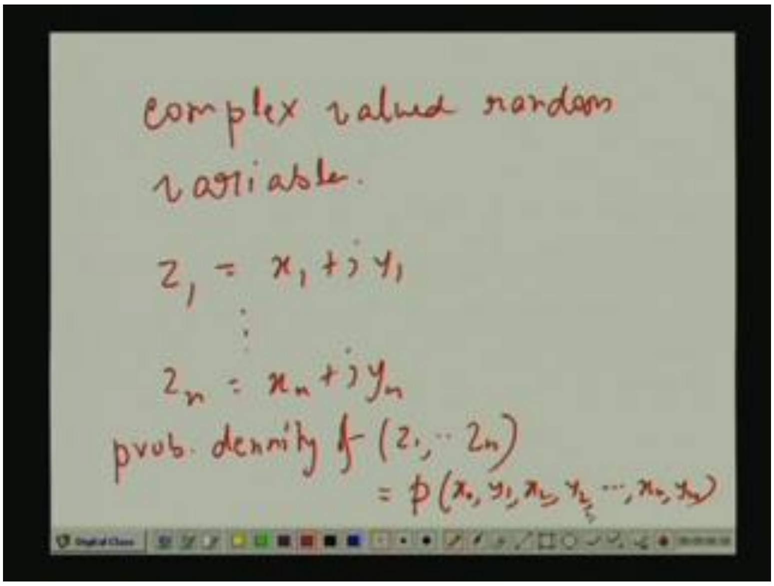
(Refer Slide Time: 00:53)



That is just for a recap. We initially started with just one random variable and function of one random variable. Then, two random variables, then one function involvement two random variables and two functions involvement two random variables. Here, we try to generalize that to a set of n random variables which are ordered as sequences x 1, x 2 up to x n, and we first consider one function of such n random variables, and then n functions of such n random variables, right.

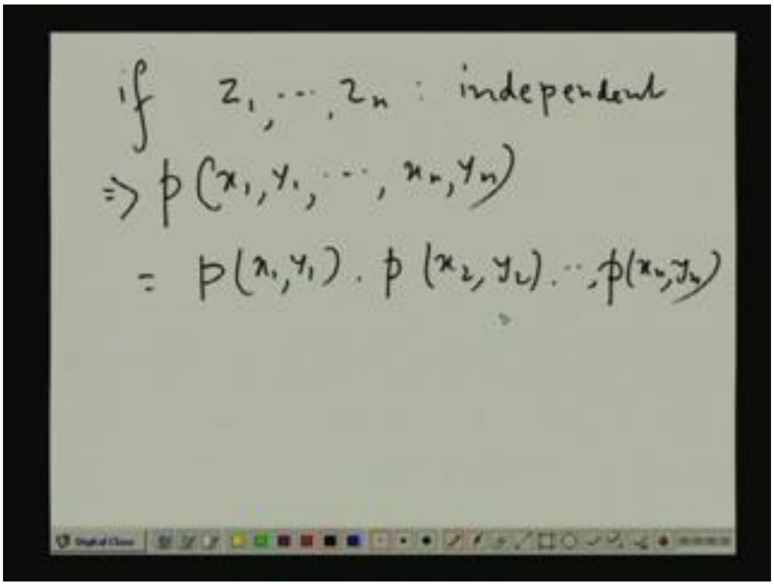
So, there were several issues which we did not discuss last time. They are all you know just analogous to those issues which we considered in the case of two random variables. So, we will just continue from where we left last time, but just one remark. We have so long used to real value random variables, but random variables could be complex valued also. Now, suppose we have got this.

(Refer Slide Time: 01:52)



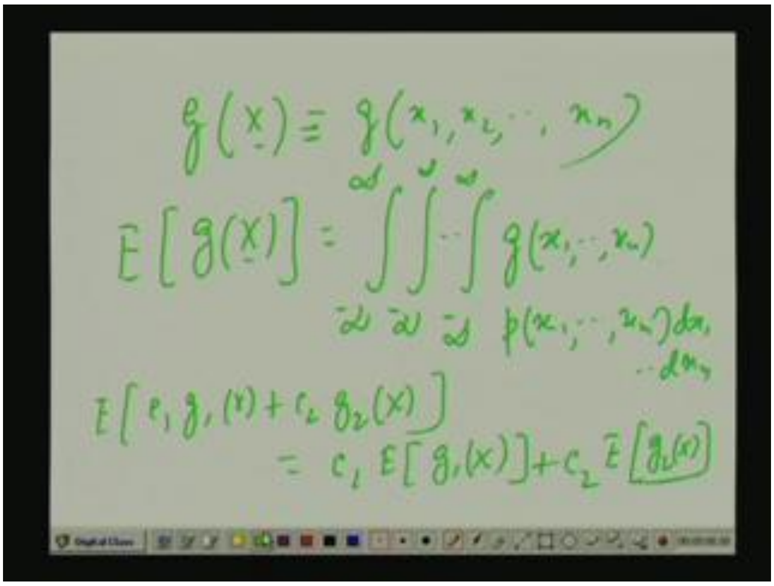
Suppose z 1 is x 1 plus jy 1 dot dot dot up to say z n x n plus jyn. So, z 1 to z n, we have got set of n complex random variables, but remember each of these, z 1 to z n, they have got two components. One is a real component and another is an imaginary component x 1 y 1 and now, x 1 also is random variable, y 1 also is random variable. Similarly, x 2 is random variable; y 2 is random variable. So, a set of n complex random variable equivalent being is a set of twice n real value random variables which means probability density of z n is same as actually a probability density function of 2 n variables x 1 y 1 x 2 y 2 dot dot dot x n yn.

(Refer Slide Time: 03:19)



Further, if z 1 dot dot dot z n, they are independent, then it means that p x 1 y 1 dot dot dot x n yn that will be like this which stands for probability density for z 1, because z 1 has the two random variables, real value random variables x 1 y 1 in it, then p x 2 y 2 dot dot dot p x n yn.

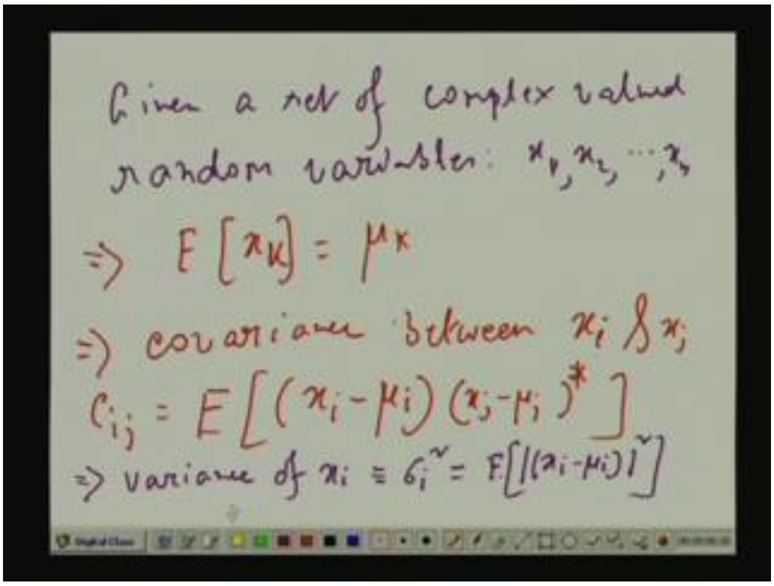
(Refer Slide Time: 04:10)



Suppose we have got single function g x. X is a vector. Basically x means actually is a function of n random variables, say x 1 x 2 dot dot dot x n. Now, by extending our previous argument, we can then say that expected value of gx is nothing but times the corresponding probability density function dx 1 dot dot dot up to dx n. Obviously, if instead of these real variables x 1, x 2 up to x n, we are a set of complex n number of complex valued random variables z 1, z 2 up to z n.

Then, instead of having n fold multiple integral, you would have had twice n fold multiple integral, and there probability density function would have been I mean this is g z 1 up to z n, and probability density function as we proved have been p of x 1 y 1 x 2 y 2 dot dot dot x n yn and we would have a dx 1 dy 1 dx 2 dy 2 dot dot dot dx n dy n. Just that generalization also as before you can see that there is a linearity that is valued here. Also, that is expected value of say sum of functions c 1 g 1 x. Say plus c 2 g 2 x is same as c 1. Expected value of g 1 x plus c 2 expected value of g 2 x, and instead of having just two terms, you can always have n number of terms. So, the linearity works.

(Refer Slide Time: 06:27)

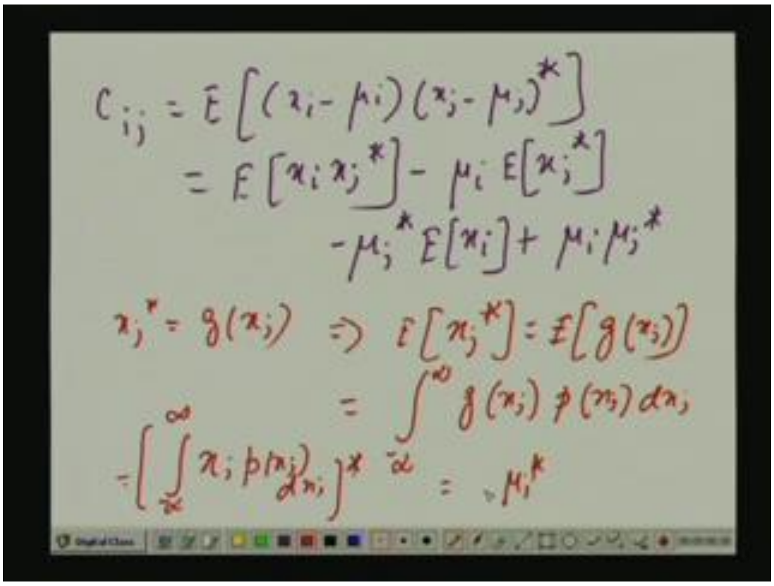


Given a set of complex valued in general, because now I wanted to be more general. So, complex valued random variables say x 1 instead of z 1, I am calling them x 1. Now, x 1 x 2 x n, then covariance we now define first say E of x 1 say is mu 1 or E of xk is mu k, where k goes from 1, 2 up to n, that is x 1 as a mean value mu 1, x 2 as a mean value mu 2 dot dot dot x n has a mu value mu n, and then we define covariance between say xi and xj as expected value. You can call it, sorry you can call it cij. This is expected value of xi minus mu i times.

Now, the difference xj minus mu j and a complex conjugation here. Earlier we have dealt with similar things, but since that time we were dealing with only real value complex variables. We have more simplified expression for this covariance, where there was no complex conjugation, but the more general case, this is the definition of covariance. Why this complex conjugate is necessary? Simple because if now you are interested in the variance of say xi, say in variance of xi, now we all know that variance denotes power expected value of the power that is average power. So, average power has to be real and non-negative.

So, if you put a star here, that will make the variance that is expected power real. How? You can easily see variance of xi equal to say sigma i square is what expected value of xi minus mu i times again xi mu i star which makes it this because complex number and its own conjugate multiplied gives rise to the mod square, and as mod square means it is real function, right. It is happening, because of this star. So, this is a variance, this is a covariance. As before this covariance expression, you can simplify like before I repeat again here.

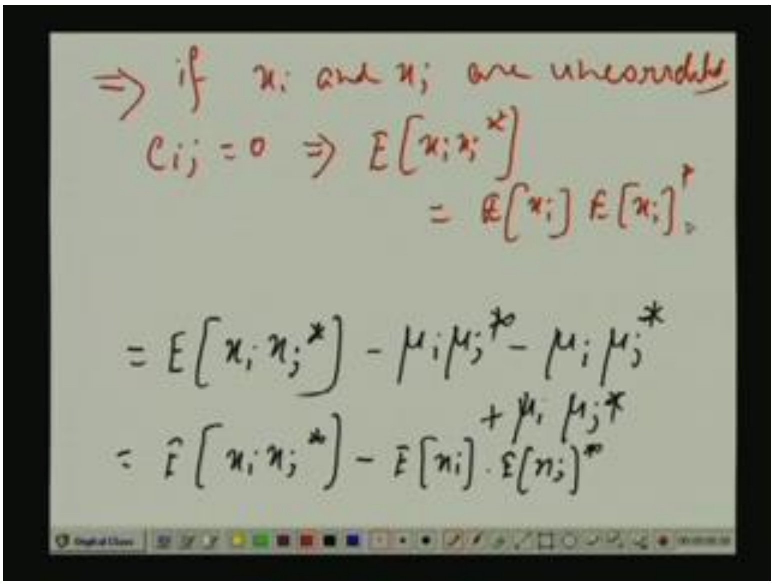
(Refer Slide Time: 10:32)



Our cij was I am just repeating. You can now bring it up one term is E xi xj star. Then, minus mu i E if this cross term xj star minus mu j star E xi plus mu i mu j star. Now, you see you have got a quantity like expected value of xj conjugate. What is that? Now, we know that you can write xj conjugate as a function of xj given xj, this function evaluate its conjugate. So, xj star is conjugate is a function of xj which means this simples expected value of xj star is nothing but expected value of gxj, and we have already seen this is nothing but gxj times pxj dxj, right and there is an integral gxj. It is a star. I replace this by xj star. Pxj is a real quantity.

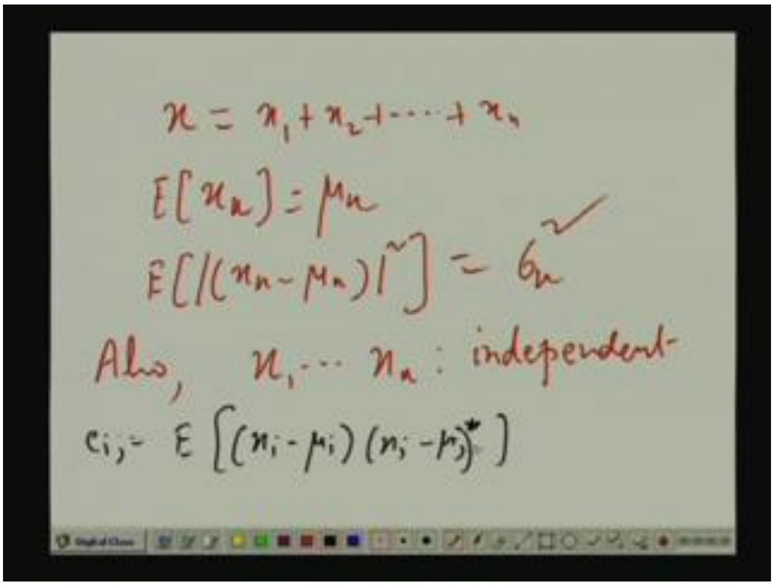
So, you can also write this like minus infinity infinity xj pxj dxj, and then star. That is first you replace gxj by xj star, but you see that other quantity that is pxj and dxj, they are real valued. So, you can take the integral first, and then conjugate. What is this integral? This is nothing but expected value of xj that is mu j. That is mu j star. So, I will use this fact here.

(Refer Slide Time: 13:22)



So, I continue this as before Exi xj star. In fact, this is called the correlation minus mu i, and you have seen while back, this is mu j star, and then minus mu j star mu i plus again mu i mew j star. So, essentially it becomes E of xi xj star minus. After simplification we only have 1 mu i mu j star which means Exi which is mu i and Exj star which is mu j star which means this implies if xi and xj are uncorrelated, then as before cij is equal to 0 meaning Exi xj star is simple Exi Exj star. Either you can take xj star, and then excepted value of it or E of xj, then conjugate. Either it is same. As an example, suppose you have got, one more thing. If x 1 to x n, there are statistically independent, then obviously p of x 1 dot dot dot x n is nothing but p of x 1 into p of x 2 into p dot dot dot pxn, right.

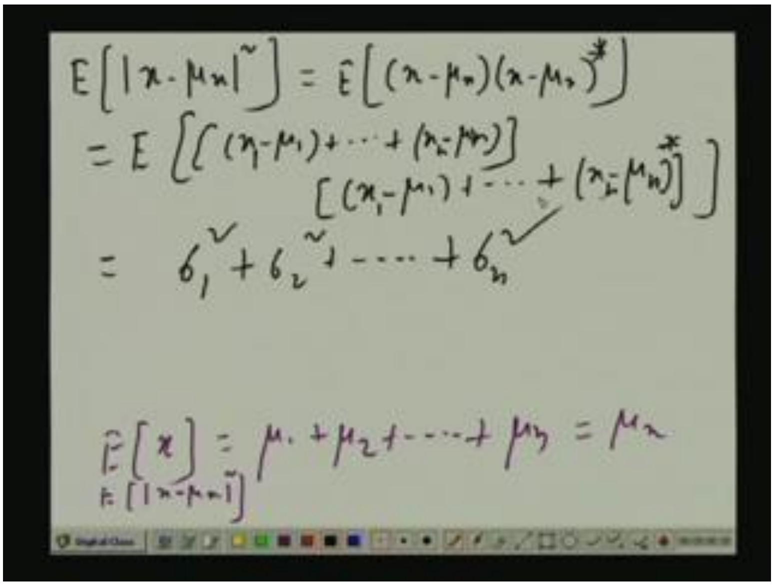
(Refer Slide Time: 15:58)



Now, suppose we are given this that X is E of xk is mu k and variance E of xk minus mu k mod square is sigma k square. Then, firstly also x 1 to xk, they are independent. We all know and it is easily seen here that if two random variables are independent, they are also uncorrelated. Then, the covariance will be 0. For instance, if you just have doubt about it, where see that suppose say xi xj as before xi minus mu i xj minus mu j star expected value, these are covariance. I am saying that if xi and xj, they are statically independent, then they are uncorrelated. That means the covariance will be 0. That is easily seen here. After all if you take the expected value of this, that means, you take this quantity as a function of two random variable xi and xj.

So, to find the expected value, you have to multiply by its joints probability in this function p of xi xj, but since xi and xj are statically independent, you can write p of xi xj as pxi into pxj and then integrate. So, one integral will be with respect to this xi minus mu i multiplied by p of xi integrated. That will be give rise to 0 because xi as a mean mu i. Similarly, the other integrals also give rise to 0. So, if they are statically independent, the correlation is covariance is 0 and we said that they are uncorrelated. So, we have given here the fact that x 1 to x n, they are n number of mutually; I mean they are n number of mutually independent random variables which means they are uncorrelated also.

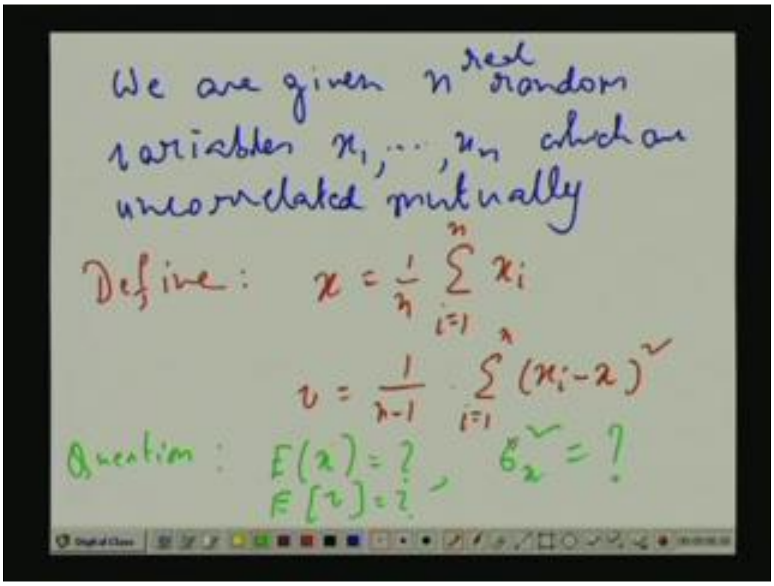
(Refer Slide Time: 18:28)



So, then what happens to E of x? Obviously, E of x 1 plus E of x 2 dot dot dot E of x n using the linearity. So, it is nothing but sigma 1, sorry mu 1 plus mu 2 plus dot dot dot mu n. How about the variance E of this? I say mu x. How about mod x minus mu x square? As you know x is x 1 plus x 2 plus dot dot dot x n which means we can write I replace x by x 1 plus x 2 plus dot dot x n mu x also by mu 1 plus mu 2 plus dot dot mu n. Then, that means, you can write it as simple here. First x minus mu 1 plus dot dot x minus mu n and on the other side, x minus again mu 1 plus dot dot dot dot x minus mu n, but star.

This is x, sorry this is x 1 to x n x 1 to x n. That is I am replacing x by x 1 plus x 2 plus dot dot x n mu x by mu 1 plus mu 2 dot dot mu n. So, taking x 1 with mu 1 x 2 mu 2 x n with mu n. Now, if you take the product, the cross products are 0 because the variables x 1 to x n, they are given to be statistically independent which means they are uncorrelated. So, the covariance is 0. Only the direct products that is x 1 minus mu 1 primes x 1 minus mu 1 star expected value only, that is non-zero and that gives a variance of x 1. Similarly, the other one gives variance of x 2. So, that means this gives rise to what variance of x n x 1, that is sigma 1 square, then sigma 2 square plus dot dot dot sigma n square. We now consider a very interesting example which uses these concepts.

(Refer Slide Time: 21:52)



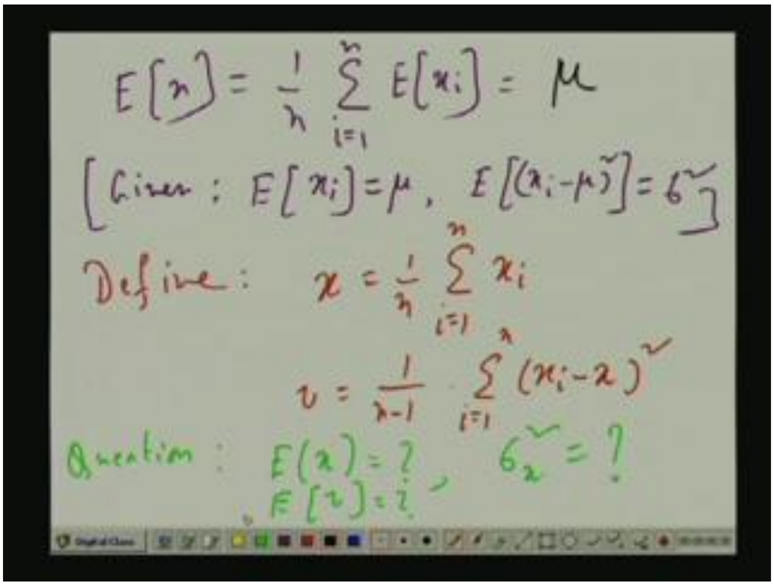
Suppose we are given means as before x n which you can say uncorrelated mutually, then define that is simple sample average of this n random variable. Just add them and divide by the total number of variables involved that is n, and this is new random variable call it is x tilde. Similarly, define random variable say v as 1 by n minus 1 times xi minus, this x tilde square.

Question is what is E x tilde equal to what? Then, what is the variance? Just a minute. What is sigma x? May be instead of x tilde, it is better to remove the tilde just because there is no x anywhere. So, I remove x tilde. So, x is the sample average. X is an average of those n random variables x 1 to x n, and v is 1 by n minus 1 times this summation xi minus that average x average random variable. Mind you x also is random variable. So, xi minus x whole square, and we are assuming, I forgot to mention that all the variables are real here. They are all real variables that is n real random variable.

So, question is what is E of x? That is this new random variable x which is formed by averaging the n random variables. What is the expected value? That is the mean value or expected value of x. Then, what is its variance? Sigma x square, sorry equal to what and the other one what is the expected value of this. What is Ev? I am not interested in squaring it up because then these terms will be rest to the power 4 and it is not of use to me here. So, just what is the expected value of Ev? It is some sort of average square, square of what deviation of xi from the average random variable.

Just take that, square it up, add for all the variables and divide instead by n here, I am dividing by n minus 1. So, that is the difference. So, it is not exact averaging. Exact averaging would have meant division by n, but here I am purposefully dividing by n minus 1. So, these two things you have to find out.

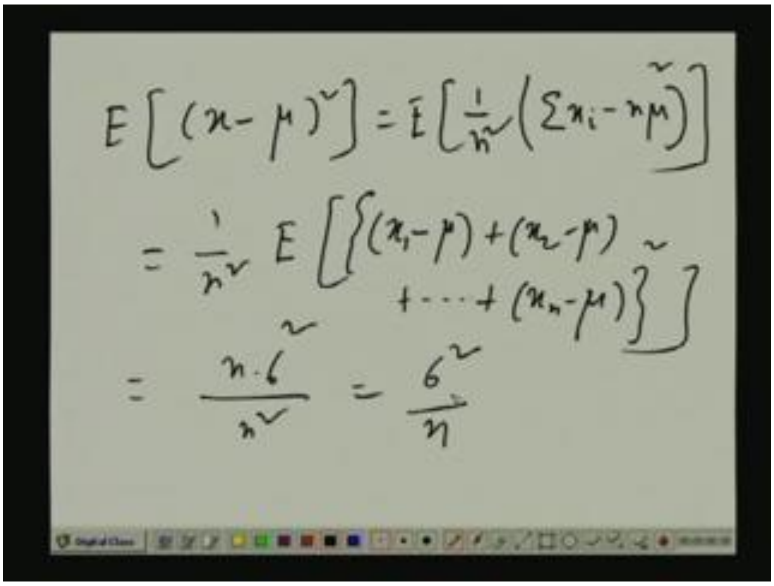
(Refer Slide Time: 27:07)



First one is very simple. E of x using the linearity. Obviously, 1 by n summation E of xi. So, E of xi, what is E of xi? I forgot to mention that given for all random variables, they have the same mean mu, ad same variance for all of them. Since they are real value, I am not putting any mod. So, sigma square is given. I forgot to mention this that we are given a set of n uncorrelated, that is mutually uncorrelated real valued random variables x 1 to x n, where each of them has same mean constant mean mu and constant variance sigma square. So, that is same for all the random variables.

Then, by just adding, and then averaging I found a new random variable. My question is what average value for that expected value of that. Then, each random variable is I found the difference between that. Each random variable with the average, what tilde square sum and average not exactly because that would have mean dividing by n. I am dividing by n minus 1. Then, we found this v. What is the expected value of v? That is the question. So, you first start with expected value of this x which obvious is in the linearity, and this formula expected value of x will be 1 by n summation. I can put the expectation operator inside the summation and E of xi for all i is equal to mu. So, mu times n n n cancel. So, this will be equal to mu. So, x as the same average same expected value as each of the random variables has. Then question is what the variance of x is?

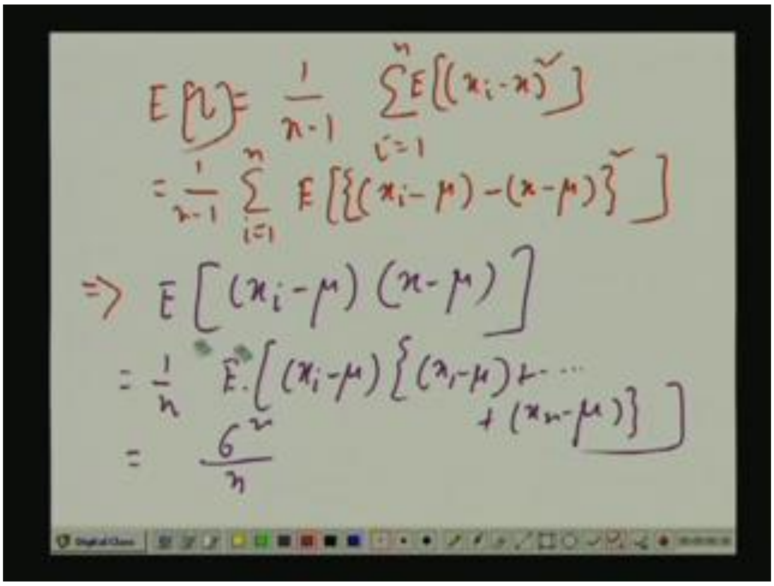
(Refer Slide Time: 30:04)



That means, that is now we have to find out what is this quantity. Mean value of x mu, we have already seen. So, what is this quantity? Now, what is x? We know that x is 1 by n summation xi i equal to 1 to n, right and now we replace this x here. So, that means, you can write it like this. E of you can take 1 by n square summation xi within bracket minus n times mu whole square. Very simple, we replace x by this. Why? We want to take 1 by n outside the brackets. So, it becomes 1 by n square. So, mu gets multiplied by n which means you can take 1 by n square outside because it is not random expected value of again xi minus mu. Sorry you have to put it this way.

X 1 minus mu plus x 2 minus mu plus dot dot dot x n minus mu whole square. In the whole square, the cos terms will be 0 as before because of uncorrelatedness. So, this will give rise to just variances for each of them and each of them as a same variance sigma square. So, there will be total n times sigma square. So, n into sigma square will come on the top n square here. So, you get sigma square by n. So, our second question is answered.

(Refer Slide Time: 32:47)

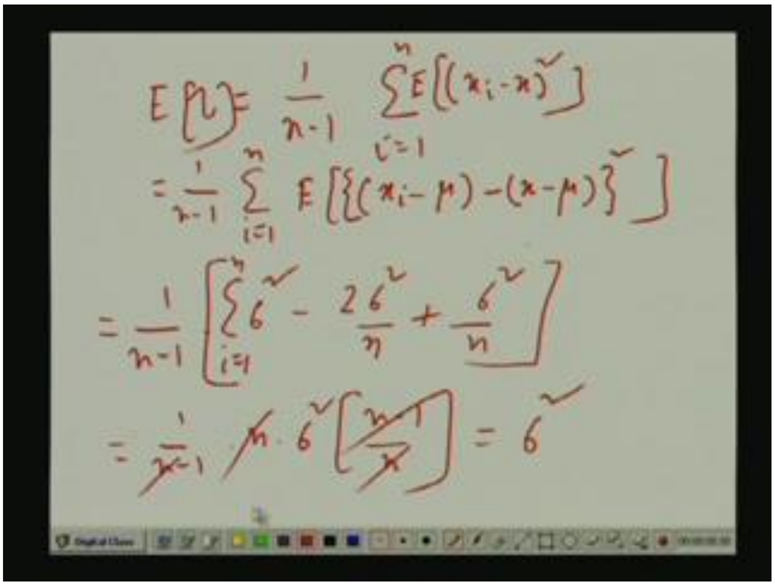


Now, the third question is so that I rewrite again. We formed the variable v. Question is what is equal to what? For that we already know that mean of xi is mu mean of x is mu. We first consider this thing E of what is this xi minus mu. That is the covariance between xi and x, xi minus mu and x minus mu because xi also as mean mu x also as mean mu. So, the covariance between any xi and x, what is that? Let us find it out.

Now, as before we replace x by its expression, so 1 by n summation xi minus mu 1 by n comes out and you can write this as x 1 minus mu plus dot dot dot dot x n minus mu. Once again all the terms will be 0 expected for the case, where from here I have got xi minus mu. So, only square of that will remain other terms like xi minus mu times x 1 mu. Expected value will be 0 because x 1 and xi, they are uncorrelated and likewise. Only when we have got a term xi minus mu from here, so square of that will come up expected value which is nothing but the variance for xi, and variance of xi we all know is a constant independent of i which is sigma square, right.

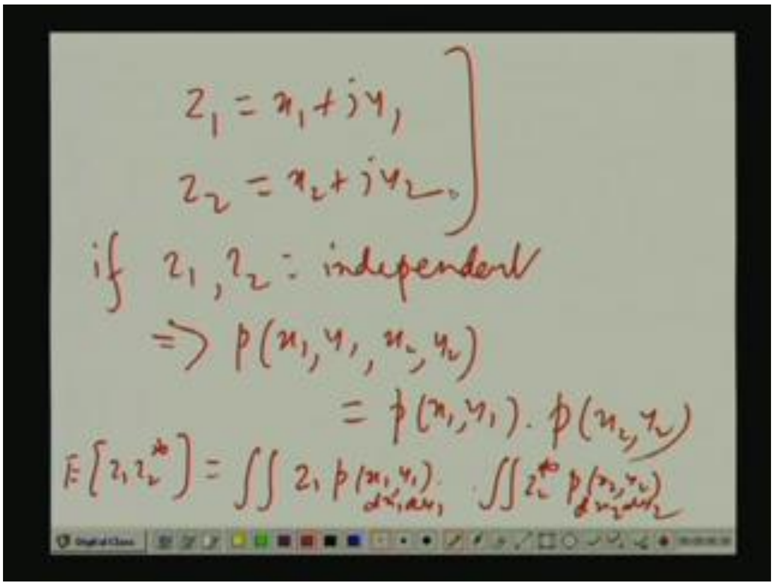
So, this will become nothing but sigma square n. So, we now come to this question. What is Ev? So, if we apply E here using linearity, we will have this and you can write here 1 by n minus 1 and same summation i equal to 1 to n E, but inner quantity you can write like this square. Now, here there will one term xi minus whole square expected value which will give x to variance of xi which is sigma square you know. Similarly, another term will have expected value of x minus mu whole square which is the variance of x, and we have already seen variance of x also be square by n, and there is a cross term twice expected value xi minus mu x minus mu. There is covariance between x xi and x which is also sigma square by n. We have just now seen.

(Refer Slide Time: 36:57)



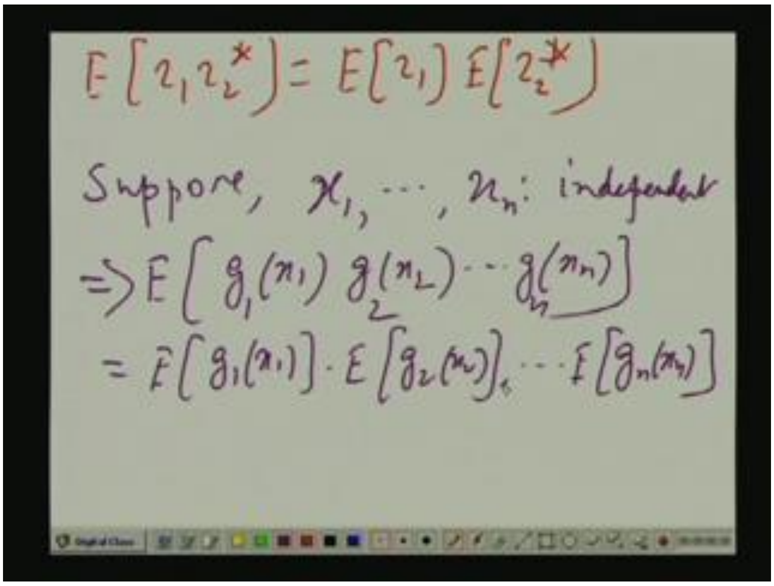
So, if we use those figures here, first is sigma square, then minus twice sigma square by n from the covariance term, the cross term, and then from the variance of the other one sigma square by n, and this sigma square. Now, we are summing that is a summation, right. Then, that is the summation of n. So, that means now what is it? If you take sigma square common, this is n times sigma square. If you take common sigma square by n minus twice sigma square root by n, so you are left only minus sigma square by n taking this sigma square common. So, you get n minus 1 by n and this n, this n cancels. You are left with just sigma square. This is answer. Few definitions when the two random, when the random say x 1 to x n, they are independent. We have seen they are uncorrelated. Obviously, same applies for the complex random variables also.

(Refer Slide Time: 38:40)



For instance, if say z 1 is x 1 plus jy 1, z 2 is say x 2 plus jy 2 and we say that z 1 and z 2, they are statistically independent. That means that if z 1 z 2 are independent, this will imply that probability density z 1 z 2 which is nothing but some px 1 y 1 x 2 y 2 will be same as x 1 y 1 which goes for z 1 times px 2 y 2, and if this is given, then for complex variables also complex random variable also, they will be uncorrelated. That is E of z 1, z 2 star will be Ez 1 into z 2 star. That is very easily seen, that is E of z 1 z 2 star will be nothing but z 1 z 2 star multiplied by the joint density integrated, but joint density can be written like this. It can be broken like this. So, there will be one on one side, we have got z 1 times px 1 y 1 dx 1 dy 1, and again z 2 star px 2 y 2 dx 2 dy 2. Obviously, this will give rise to Ez 1 and this will give rise to Ez 2 star. So, that means I write separately here.

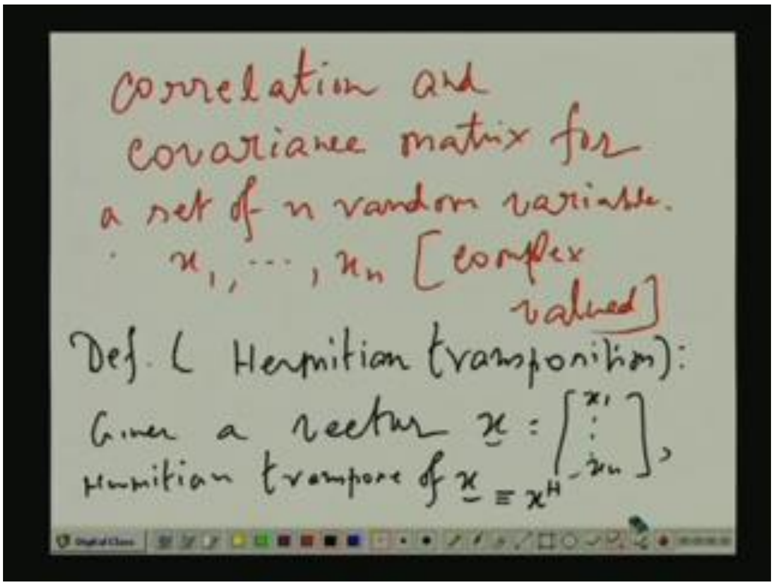
(Refer Slide Time: 40:28)



Here we see easily. Finally, suppose sorry suppose are independent. That means if you take instead of x 1, a function gx 1 instead of x 2, a function gx 2 dot dot dot a function gx n is same as after all. This whole product is again another function you calculate g 1 sorry. You can write g 2 gn different function. X 1 as a function I mean I take a function g 1 which works on x 1. Then, a function g 2 which works on x 2 gn works on x n, and I consider the products which is basically a function g of x 1 x 2 x n. Expected value will be what this product multiplied by the joint density integrated, but if they are independent, joint density is nothing but product of individual and marginal densities.

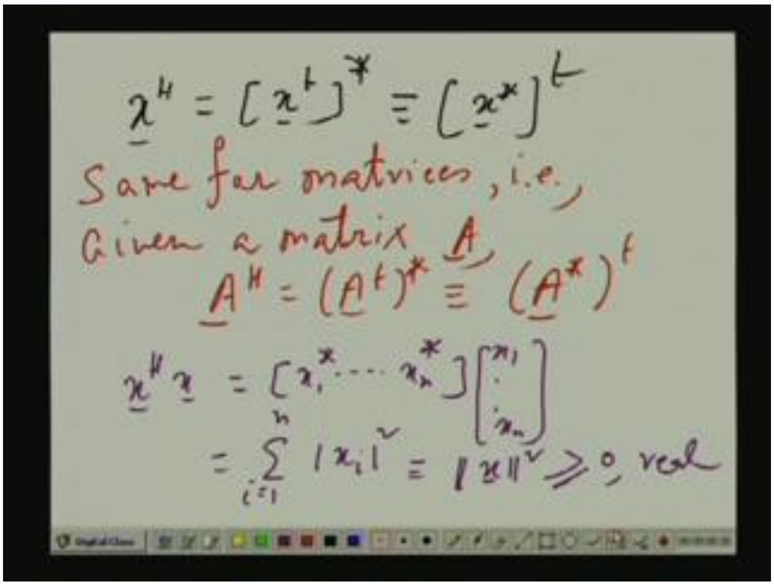
So, you just separate out the integral one with respect to x 1. Thus, g 1 x 1 times probability density of x 1 integrated same for x 2 same for x n. So, this will give rise to first one. Integral give rise to the expected value of that is g 1 x 1, then expected value of g 2 x 2 dot dot dot expected value of gn x n. Now, I come to a very important topic that is correlation matrix and covariance matrix for a set of n random variables in general complex value to random variables in general complex valued.

(Refer Slide Time: 43:18)



I need to define you know certain things like what is meant by hermitian transposition, definition of hermitian transposition given a vector x as say x 1 dot dot dot x n hermitian transpose which we denote by x hermitian. This will be given by actually by conjugate transpose. You have to transpose the vector, and then conjugate all the elements. It is not ordinary transposition. Ordinary transposition means just transpose the elements row vector becomes column vector or column vector becomes row vector likewise, but here we have to do one more step, that is conjugate the elements also.

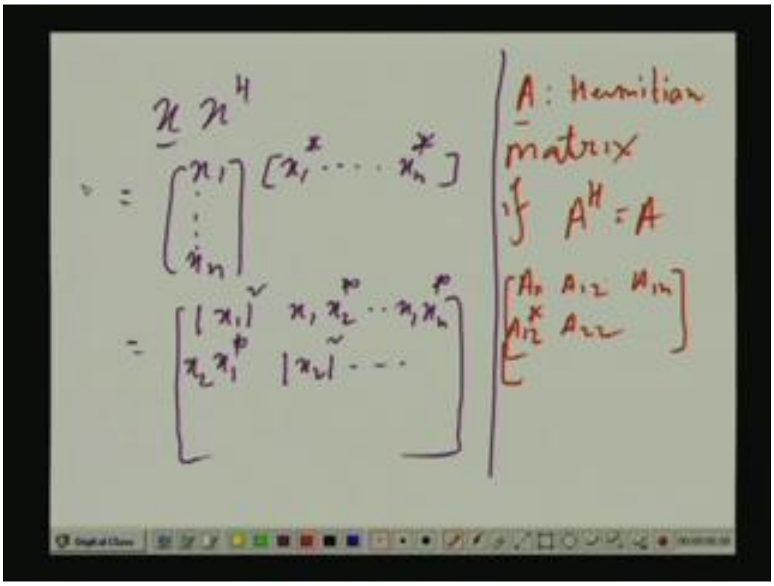
(Refer Slide Time: 45:53)



So, xH is nothing but you can either write x transpose. So, if it is a column vector, it becomes a row vector and vice versa, and then complex conjugation or equivalently first conjugate and transpose is called hermitian transpose, and is denoted by xH. Similarly, for matrixes, same for matrixes, that is given a matrix A. AH is nothing but either A transpose. Then, conjugate restrict the matrix transpose. You get another matrix. Then, conjugate all the elements or equally you first conjugate all the elements, and then transpose either we will get the same thing. Also, my notations if any symbol, any letter as a bar or underscore, then it is either a vector or a matrix.

Now, if it is written using capital letter, then it is a matrix. If it is written using lower case letter, then it is a vector, but both have an underscore that difference differentiates it or differentiates from scalars. One more thing you will see x hermitian x, this will what X is a column vector. So, it becomes a row vector transposition, and then conjugates it, and this is original x. So, basically it means mod of x 1 square, then mod of x 2 square dot dot dot mod x n square, all that which is also called non square of the vector, and this is always greater than equal to 0 and real.

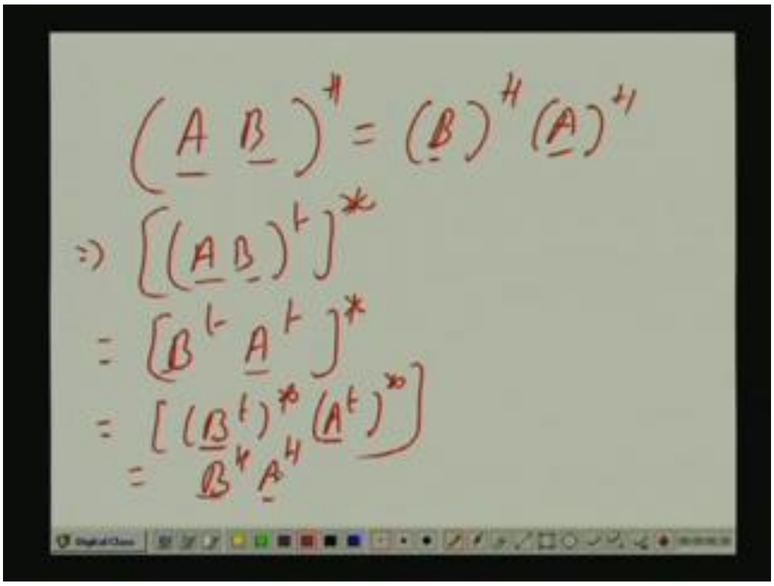
(Refer Slide Time: 48:44)



On the other hand how about x xH? This will not become a scalar. Previous xH x was a scalar number because xH was a row vector, x was column vector. So, product was a scalar, but now here we have x as a row column vector as before x 1 dot dot x n and obviously, xH will be first is transpose which is row vector, and then conjugate it. So, x 1 star dot dot dot x n star which means we have trans like this. Mod x 1 square, then x 1 star x 2 star dot dot dot x. Sorry I make a mistake x 1 x 2 star. Then, x 1 x n star, then again x 2 x 1 star, and then mod x 2 square likewise.

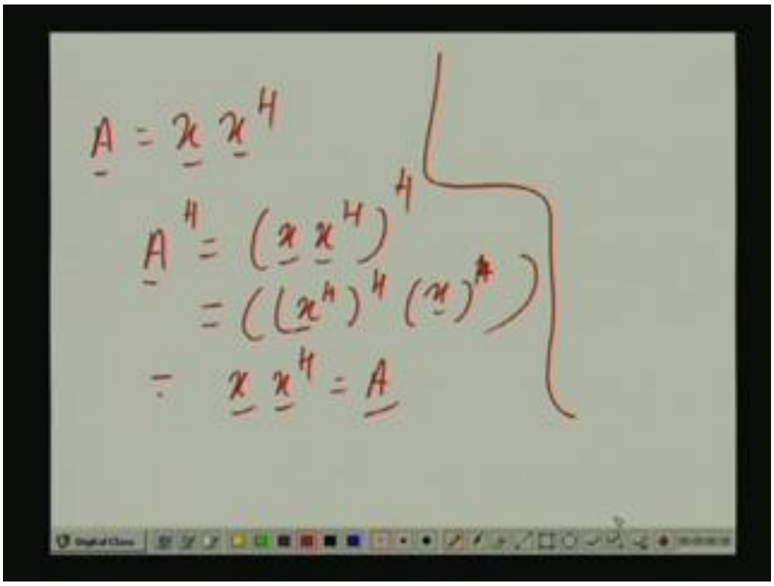
You see the first row and second column element is x 1 x 2 star and second row first column element is there conjugate of this x 2 x 1 star, and this works. This matrix actually is called a hermitian matrix which means if you have trans like say A 1 1 A 1 2 dot dot dot say A 1 n. If I transpose it, A 1 2 comes here, and then conjugate. So, this place will be A 1 2 star. This remains A 2 2 and likewise. Now, this matrixes are hermitian matrixes, but before that you see one thing.

(Refer Slide Time: 51:36)



A B hermitian is B hermitian A hermitian. This is not difficult to see after all AB hermitian means you take AB transpose conjugate, and AB transpose is B transpose A transpose conjugate and conjugate of products means product of the conjugate. You have to take the conjugate on each of them, right. So, that means, B transpose conjugate A transpose conjugate which is nothing but B hermitian A hermitian.

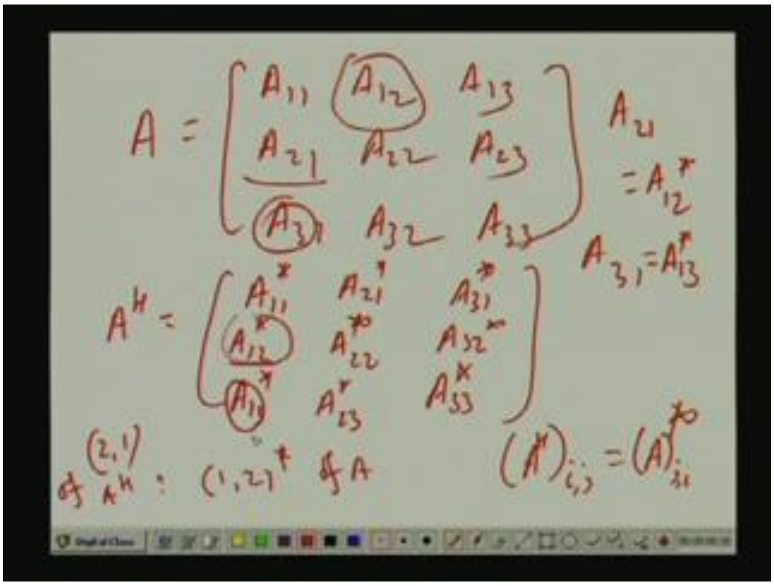
(Refer Slide Time: 52:27)



Now, previously I considered a matrix of this form. You can call it A. My claim is this is a hermitian matrix, that is A, and its hermitian transpose, they are same. Hermitian transpose means I mean you take the transpose. So, I, j th element where i is the row, j is the column on transposition that goes to the j, i th position, and then you conjugate that also. So, after hermitian transposition j, i th position gets filled by the ij th element of the previous matrix with the conjugate, that is a hermitian transposition. So, you get back the same matrix, and then we call it hermitian matrix.

Now, you can see one thing. This hermitian because what is AH, what is the transposition H, right and AB hermitian means hermitian x hermitian, right. You can see hermitian of a hermitian transpose. That gives you the original vector. After all this was a column vector by hermitian transposition. I mean in a row vector with conjugate, I mean elements conjugated. Again taking the hermitian transpose, so it again becomes a column vector and conjugation disappears. So, you get back x and here as before xH which is the original vector matrix A. So, a matrix which on hermitian transposition gives you back itself, then it is a hermitian matrix.

(Refer Slide Time: 54:33)



That is suppose you take A 1 1, A 1 2, A 1 3, A 2 1, A 2 2, A 2 3, A 3 1, A 3 2, A 3 3, and then what is AH transposition will not change the diagonal elements. 1 1 remains in 1 1 only. Thing is that they get conjugated. Hermitian transposition means A 1 2 comes here with a conjugate, and A 2 1 goes here with a conjugate, A 3 1 comes here with a conjugate, A 1 3 comes here with a conjugate, A 2 3 comes here with a conjugate, A 3 2 comes here with a conjugate. Like these two matrixes are same, then A will be called a hermitian matrix. That means firstly, the diagonally elements A 1 1 must be A 1 star which means it must be real valued element.

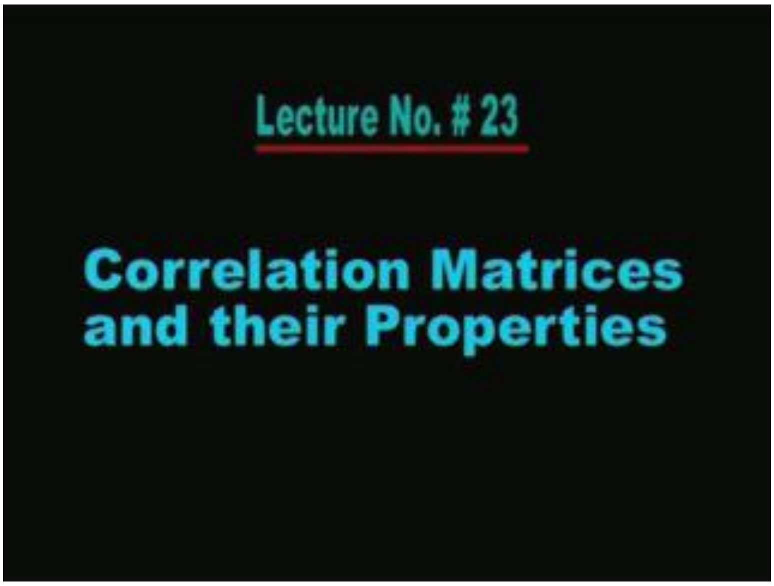
So, the diagonal similarly for A 2 2, similarly for A 3 3. So, for a hermitian matrix diagonal entries are real and other entries are conjugate symmetric of each other, like 2 1 th element of AH, that is second row first column. This is nothing but 1 2 th element of that is 2 1 th of AH is same as 1 2 th star of A. So, 2 1 th element is this and 1 2 th element of A is this. If you put a star here, you get back this element. That means AH this matrix it is ij th element is nothing but A matrix. It is ji th element star. That is a hermitian transposition.

Now, if A and H are same, that means, A 2 1 and that is we should have A 2 1 is equal to A 2 1 should be just a conjugate of this A 1 2 star. A 3 1 should be A 1 3 star because this and this are to be same and likewise. So, the transmission of the hermitian matrix ij th element. That means I think I continue from here in the next class because some more properties are to be discussed because this hermitian matrices, they give rise to what is called the correlation and covariance matrices. So, that is all for today. We will continue from here in the next class because time is up.

Thank you very much.

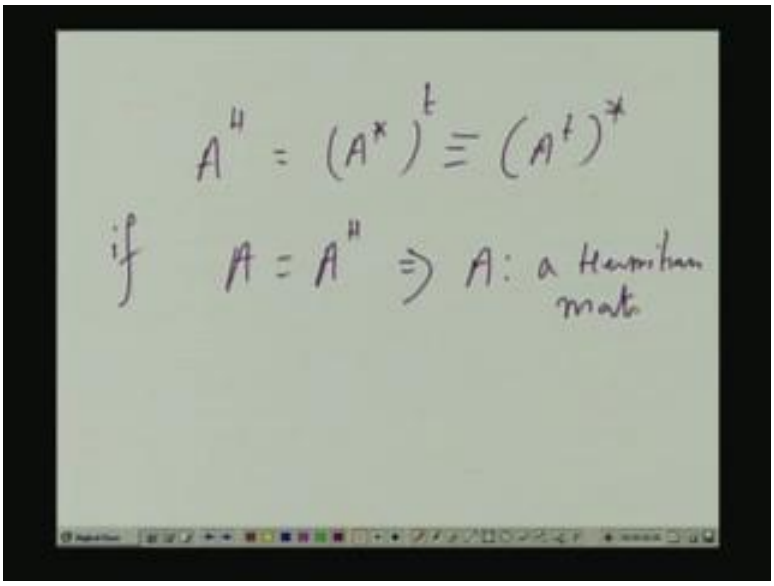
(Refer Slide Time: 57:34)

(Refer Slide Time: 57:39)



So, in the last class, we were discussing these correlation matrices and in that connection I talked about what is called hermitian matrices and hermitian transposition and things like that. So, today we will continue from there.

(Refer Slide Time: 58:06)



So, just to recall what we did last time given a matrix say A, we define a hermitian transposition that is AH which actually is nothing but A conjugate that is complex. Take the complex conjugate of each entry of A and their transpose which is also equivalent to doing in the other way. First take the transposition of A, and then a complex conjugate. That is very simple. So, if A is not a matrix, but just a vector, say column vector, then its hermitian transposition is what we first transpose. So, it becomes a row vector, and then takes the conjugate of each element. Similarly, if A is a row vector, there hermitian transposition will be a column vector with all the original elements are complex conjugated. So, it happens if A is a hermitian, and then A is called a hermitian matrix.