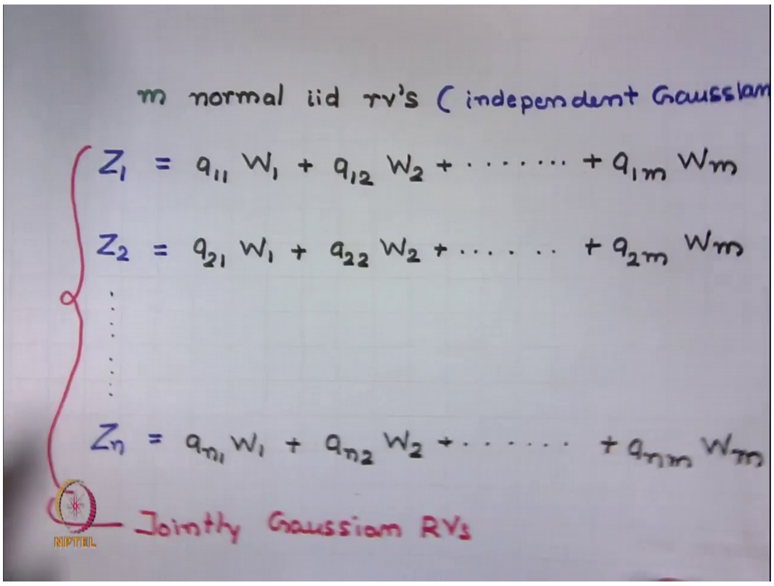
Principles of Digital Communication
Prof. Abhishek Dixit
Department of Electrical Engineering
Indian Institute of Technology, Delhi

Lecture – 14
Random Variables and Random Processes: Gaussian Random Process (Part – 2)

Good morning. Welcome to new lecture in the unit of Random process and today in this lecture, we will continue our discussion about Gaussian Processes. So, let us first recap the main ideas that we discussed in the last lecture.

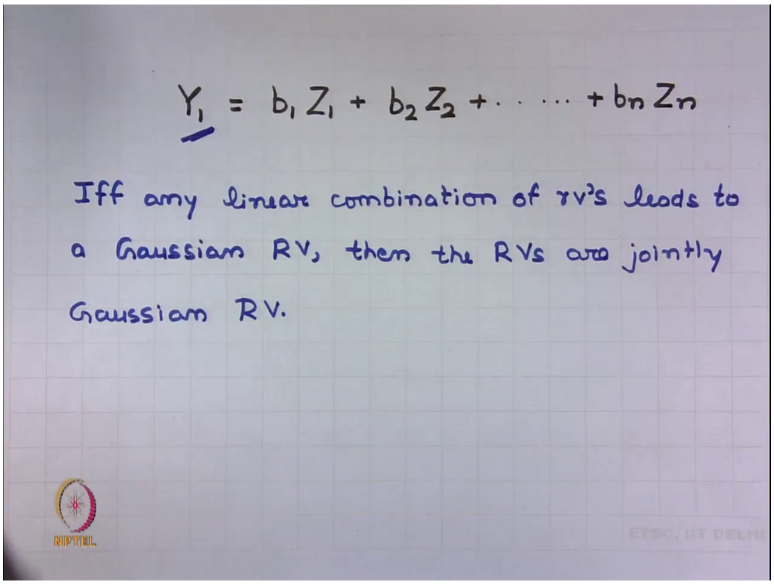
(Refer Slide Time: 00:32)



So here, comes the first big idea that we discussed. So, we said that if we have m iid normal random variables. So, W 1 W 2 and W m are m normal iid random variables and if I make different linear combination of these normal iid random variables, then what I end up with are Jointly Gaussian Random Variables. So, Z 1, Z 2 and Z n are Jointly Gaussian Random Variables.

So, we have defined Jointly Gaussian Random Variables as the random variables which are obtained by different linear combinations of common underlying set of normal iid random variables. Also note that even though we have been developing our model assuming these W 1 W 2 and W m as normal iid random variables, but we could have also assumed them to be independent and Gaussian ok. This assumption is just for the simplicity. So, that one big idea that we have seen in the last lecture.

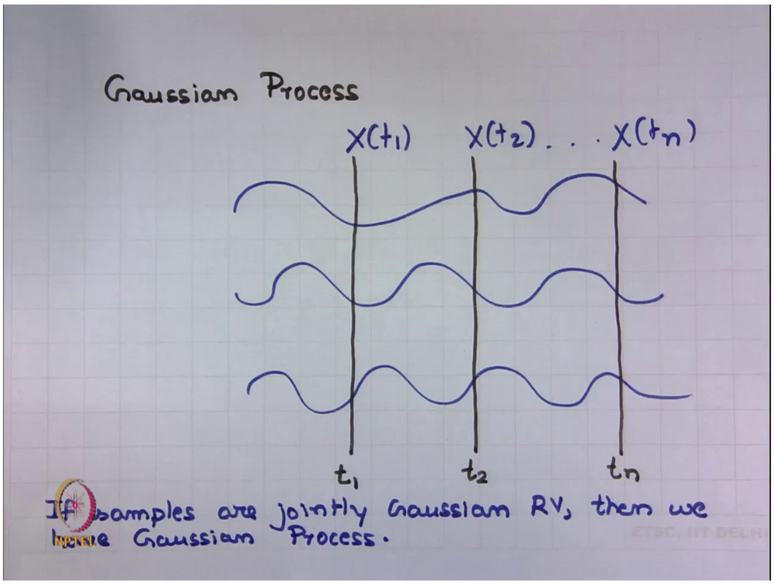
(Refer Slide Time: 01:41)



We have also looked the second way of defining these jointly Gaussian random variables and this is that if and only if any linear combination of random variables. So, these Z 1, Z 2 and Z n are random variables and if any linear combination of these random variables leads to a Gaussian random variable.

So, Y 1 has to be a Gaussian random variable, then the underlying random variables are jointly Gaussian random variable ok. And we have also seen the equivalence between the first definition and the second definition ok. So, this is for jointly Gaussian random variables. We have also defined Gaussian process and what is a Gaussian process.

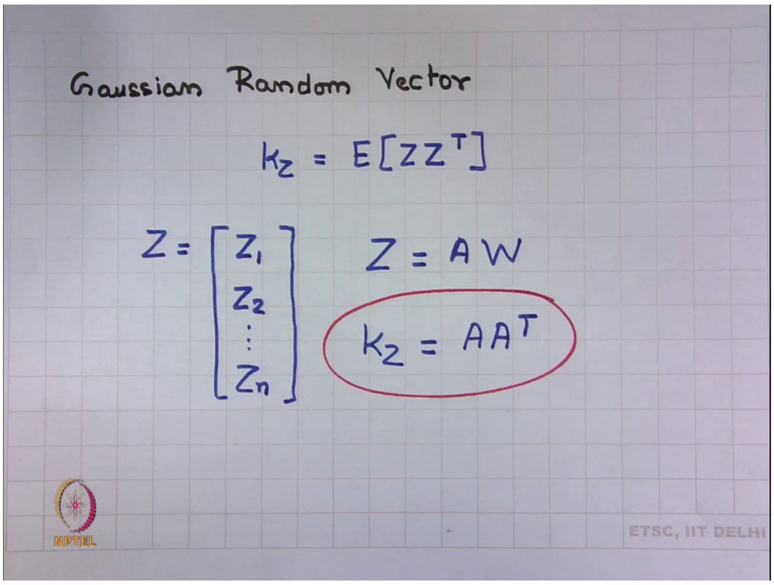
(Refer Slide Time: 02:24)



So, first Gaussian process is a random process and we know that a random process can be understood by taking samples of this random process at different time instants. For example, we have taken the samples of this random process at time instants t 1 t 2 and t n. Once you sample a random process at different time instants, we end up with different random variables ok.

So, in this case let us say we have obtained n random variables and if these random variables are jointly Gaussian random variables; then the underlying process is a Gaussian process. Thus, we can define Gaussian process as a process whose samples are jointly Gaussian random variables ok.

(Refer Slide Time: 03:14)

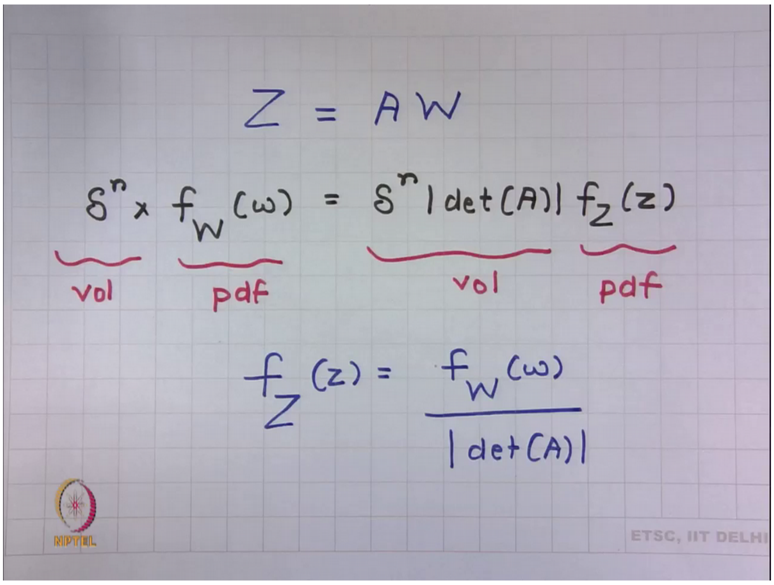


So, a lot of definitions flowing in, we also defined Gaussian random vector. So, what is the Gaussian random vector? It is a vector whose elements are jointly Gaussian random variables. So, for example, here Z is a vector and the elements of Z of Z 1, Z 2 and Z n and as we have already seen that these are jointly Gaussian random variables. So, that is a Gaussian random vector. We have defined the covariance of a Gaussian random vector which is given by expected value of Z into Z transpose.

Remember that in last lecture and in today's lecture, we always assume random variables to be with 0 mean and if these random variables are with 0 mean. Then, the covariance of this random vector Z can be simply obtained by carrying out this expectation operation and what we have also showed that this is nothing but A into A transpose and what is A? A is the matrix with which we were multiplying this vector w to get vector Z ok.

So, this is one idea that we will also use that the covariance of a Gaussian random vector is nothing but A into A transpose ok. Then, we are trying to derive the probability density function of this random Vector Z.

(Refer Slide Time: 04:36)



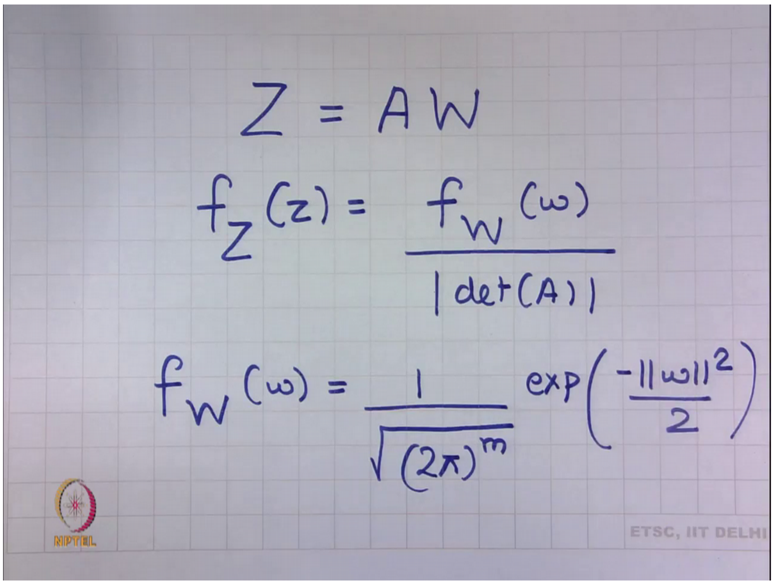
And what we have said is that the probability when you go from W space to Z space will be preserved and probability can be thought as multiplying the probability density function with the volume. So, the probability in w space is the probability density function in w space multiplied by the volume in the w space and this is the probability in the Z space and these probabilities must be same.

Because every point in W space is mapped to a point in Z space thus, the probability of finding a point in w space would be same as the probability with which you will find the point in the Z space where the points are the map of the points of the W space ok.

This we have seen in the last lecture in lot of details right and from this we derived that the probability density function of random vector Z can be obtained in terms of probability density function of random vector W divided by mod of determinant of A. We have also seen that determinant has nothing but it is the volume scaling factor which describes the linear transformation ok.

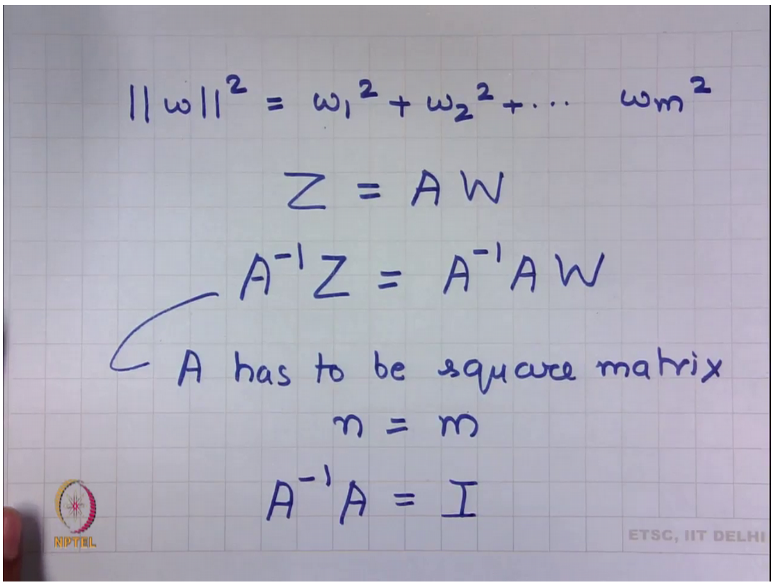
So, we will start today’s lecture from here. Let us rewrite some key equations that we have seen in the last lecture.

(Refer Slide Time: 06:09)



So, first equation is Z is A times W and the probability density function of random vector Z is probability density function of random vector W divided by mod of determinant of A ok. And we have already derived the probability density function of this random vector W. So, probability density function of random vector W, this is also what we derived in the last lecture is this thing ok, where m is the number of elements in this vector W and we have also seen that this is norm square or vector w.

(Refer Slide Time: 07:13)

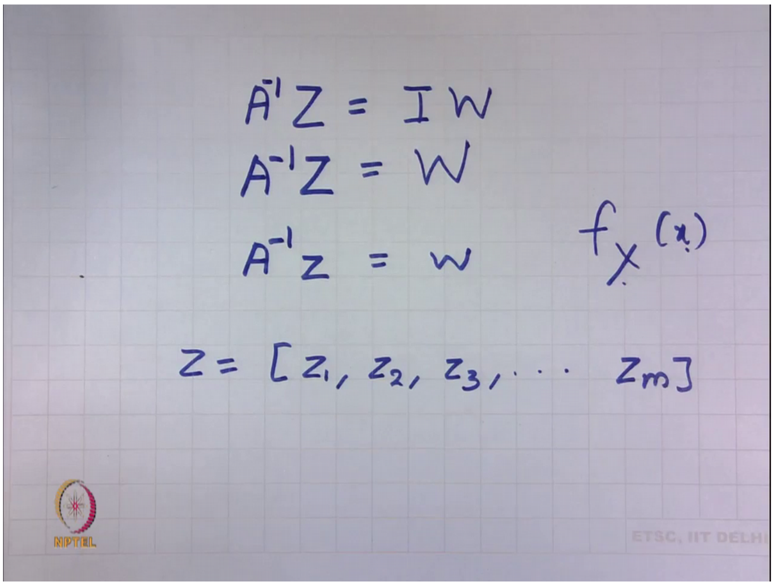


And we know that norm square of any vector can be thought as sum of squares of these elements w 1, w 2 and w m are nothing but the elements of vector w ok. So, this is how we can think about norm square of a vector ok.

So, let us now. So, this is what we have seen in the last lecture. Now let us try to develop few more equations. So, starting from Z is A times W. I can multiply this with A inverse. So, multiplying with A inverse on both side and what we are assuming is the inverse of a matrix exists. So, first thing is A has to be a square matrix and thus and n should be same as m and we will see that even if A is a square matrix, there can be a case when inverse does not exist. But let us assume at this moment that the inverse of this A matrix exists and let us also assume that A is a square matrix ok; where, n is same as m.

Now, we know that if the inverse of a matrix exists; then A inverse times A is nothing but identity matrix ok.

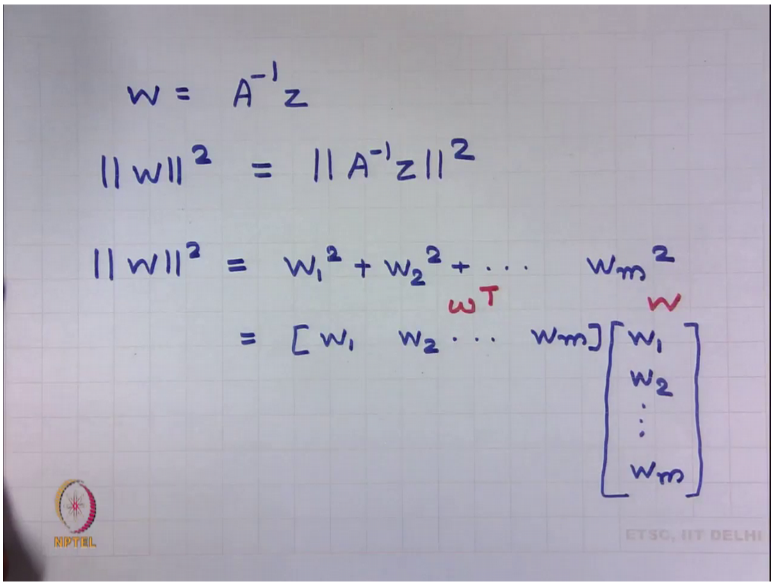
(Refer Slide Time: 09:15)



So, we can say A inverse times Z is I times w and identity times a matrix is nothing but the matrix itself. So, we know that w vector is nothing but A inverse multiplied by Z. And now, what we want to do is we want to find out the probability density function of Z in terms of probability density function of w and we have seen that vector w is nothing but A inverse times z. So, numerical value of this vector Ww would be same as A inverse times small z whereas, small z denotes that it is the numerical value of vector Z.

So, what I mean with this is z would be nothing but some vector with some numbers. So, it denotes the z 1, z 2, z 3 and z m are some numbers and Z is also a numerical value consisting of m numbers and this Z is a random vector, this z is a vector which contains some numbers right. As we have seen in the case of random variables that X is a random variable and small x is the numerical value of that random variable ok.

(Refer Slide Time: 11:13)



So, if I know that W is A inverse Z, we had norm of W square and norm of w square would be same thing as this ok. Now, what is this norm of w square? We have already seen that this is nothing but w 1 square plus w 2 square up to w m square and all these w's are small w's and this is nothing but w 1 w 2 w m multiplied by w 1 w 2 w m and this is what this is w transpose and this is w. So, remember as I have already said that the vectors are always assumed to be column vectors. So, this is a row vector. So, it has to come with a transpose. So, this is w transpose times w. So, what we are saying is simply that long square of w is nothing but w transpose times w.

(Refer Slide Time: 12:37)

$$\|w\|^2 = w^T w$$

$$\|A^{-1}z\|^2 = (A^{-1}z)^T (A^{-1}z)$$

$$= z^T \underline{A^{-T} A^{-1}} z$$

$$K_Z = A A^T$$

$$K_Z^{-1} = (A A^T)^{-1} = \underline{A^{-T} A^{-1}}$$

ETSC, IIT DELHI

Similarly, norm square of inverse z is nothing but A inverse z transpose times A inverse z and this we can solve using properties of matrix. This will be z transpose A inverse transpose A inverse into z.

Now, we have already seen that the covariance of this random vector Z is nothing but A times A transpose. So, the inverse of this covariance matrix is A A transpose inverse ok. Now, this is simply A inverse transpose times A inverse. All this follows from the properties of matrices ok. So, if you have forgotten about them it is a good idea to have a look to these properties and now something interesting has happened, we have this term here. So, this is nothing but inverse of covariance of random vector z. So, using this idea what we can say is nothing but.

(Refer Slide Time: 14:26)

$$\|A^{-1}z\|^2 = z^T k_Z^{-1} z \longrightarrow (1)$$

$$k_Z = A A^T$$

$$\det(k_Z) = \det(A A^T)$$

$$= \det(A) \det(A^T)$$

$$\det(k_Z) = \det{}^2(A)$$

$$\det(A) = \pm\sqrt{\det(k_Z)} \longrightarrow (2)$$

So, we have z transpose k z inverse times z. So, that is one equation that we will use. Second thing that we want to see is that k z is A A transpose. So, determinant of k z is nothing but it is determinant of A A transpose and this thing is same as determinant of A into determinant of A transpose and determinant of a matrix and its transpose is same. So, this is determinant square of A and so, determinant of A will be plus or minus square root of determinant of k z. So, this is the second equation that we will use. So, now let us see what we are up to? So, let me rewrite some equations.

(Refer Slide Time: 16:07)

$$f_W(\omega) = \frac{1}{\sqrt{(2\pi)^m}} \exp\left(-\frac{\|\omega\|^2}{2}\right)$$

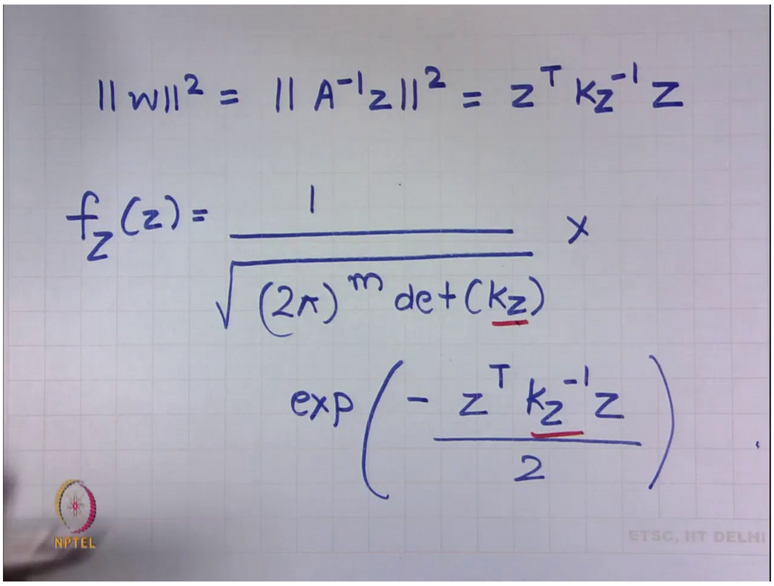
$$f_Z(z) = \frac{1}{\sqrt{(2\pi)^m}\,|\det(A)|} \exp\left(-\frac{\|\omega\|^2}{2}\right)$$

$$\det(A) = \pm\sqrt{\det(k_Z)}$$

$$|\det(A)| = \sqrt{\det(k_Z)}$$

So, we have said that probability density function of random vector w is and we know that the probability density function of random vector z would be this probability density function divided by mod of determinant of A into exponential of this thing and now we have developed two equations that determinant of A is nothing but plus or minus square root of determinant of k z. So, mod of determinant of A is square root of determinant of k Z; just one thing.

(Refer Slide Time: 17:31)

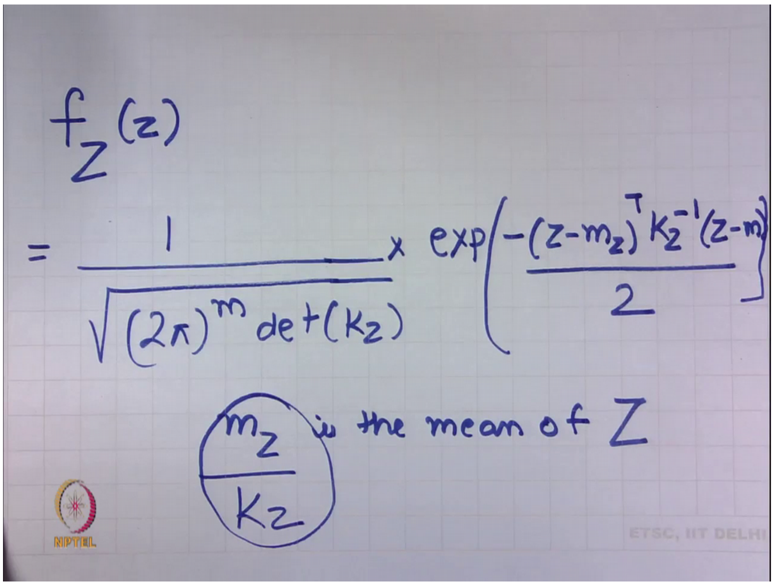


Second thing is we have seen that norm of w square is nothing but norm of A inverse z square and we have proven this is simply z transpose k z inverse times z. So, what we are up to is we trying to express this probability density function in terms of z. So, we want to get rid of w because this is a function of z. So, we want to talk in terms of z rather than in terms of w right. So, the final expression for probability density function of z will be and this is an important result. So, we have been able to derive the probability density function of random vector z and remember that this random vector z is actually a Gaussian random vector ok.

So, now we have the probability density function of a Gaussian random vector at a disposal. The few things that are interesting here that this probability density function is completely specified by this covariance function right. So, it is an important idea this probability density function is just a function of k z right. So, if you know k z, you can completely specify the probability density function of Gaussian random vector. Thus, a probability density function of a Gaussian random vector is only dependent on its covariance matrix ok. So, in this expression what we have assumed is that this random vector is with 0 mean.

Let me write the expression of probability density function which will be when this random vector has some mean. So, let me first write and then we will discuss about this.

(Refer Slide Time: 20:13)

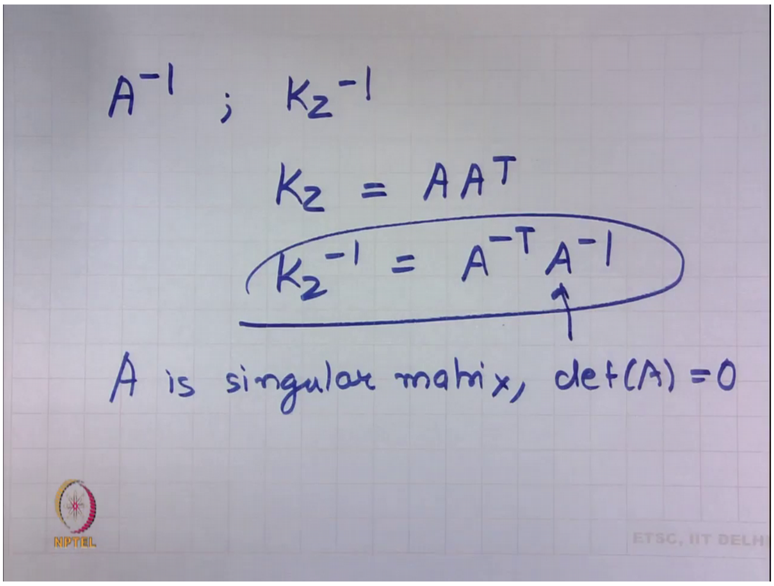


So, the probability density function of this random vector z when this is not with 0 mean will be simply. So, what we see here is what is this m z? m z is the mean of random vector z ok. So, now, if the z is not with 0 mean, the probability density function depends upon two unknowns; one is this covariance matrix and then the mean right. So, if z is not 0 mean, then this probability density function is a function of m z and k z. But overall this probability density function will only be function of these two variables right. Once you know the mean of a Gaussian random vector and a covariance of a Gaussian random vector, then you can completely specify the probability density function of Gaussian random vector right.

So, when we started talking about these random processes, we said that defining and specifying probability density function can be quite complicated, because you have to talk this probability density function for every time epochs, you have to talk for every argument and so on so forth. But this probability density function is a much simpler probability density function, just a function of mean and covariance and that is why these

Gaussian processes are easy to handle and deal with alright. So, let us talk about one more issue in here.

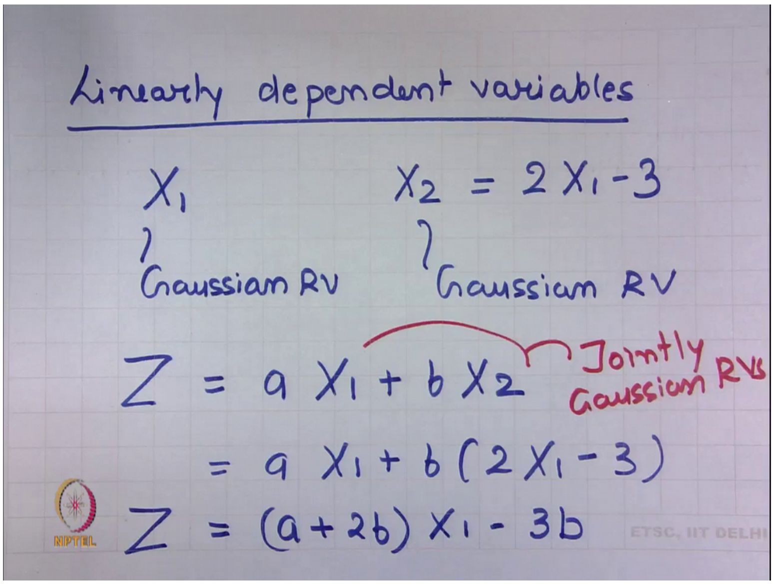
(Refer Slide Time: 22:35)



So, so far we have assumed that the inverse of a matrix exists right. This should exist and then only the inverse of this covariance matrix will also exist right. Because this probability density function is a function of inverse of a covariance matrix and if the inverse of a covariance matrix exists, what we also require is that inverse of A should also exist; remember the covariance matrix is nothing but A times A transpose. So, the inverse is as we have already seen is this thing all right. So, what we want is for an inverse of this covariance matrix to exist, then the inverse of this matrix A should also exist. What happens if this matrix A is a singular matrix ok?

So, if A is a singular matrix; that means, the determinant of A is 0 and if determinant of A is 0, then the inverse of this matrix will not exist and in that case the inverse of this covariance matrix will also not exist and then what would happen to this joint probability density function, let us see this. And to understand this in slightly more detail, let me first introduce the idea of linearly dependent variables ok.

(Refer Slide Time: 24:06)

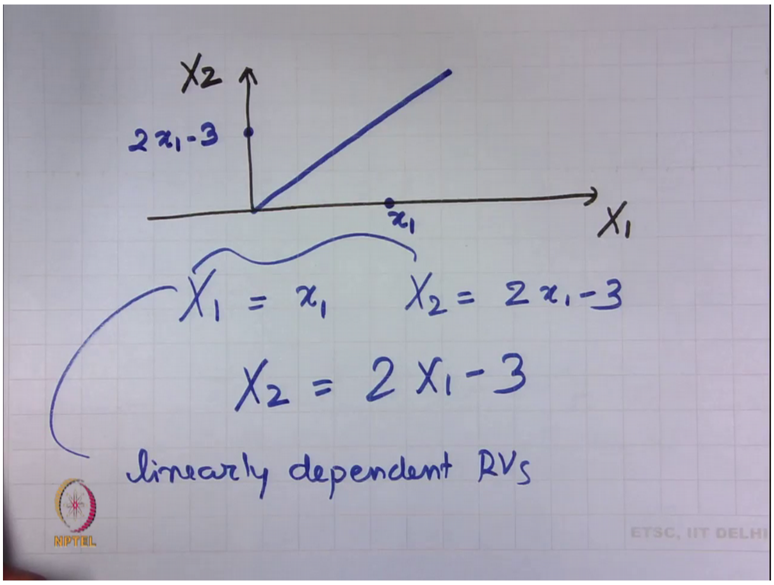


So, let us assume that I have two Gaussian random variables; X 1 is a Gaussian random variable and let me assume that X 2 is 2 X 1 minus 3 because X 2 is a linear function of this random variable X 1. This is also going to be a Gaussian random variable ok. Now let me define a random variable Z which is a linear combination of these two random variables.

Now, let us substitute the value of X 2 in this equation. We get this and simplifying this we get. Now, for any values of a and b you can easily appreciate that Z is also going to be a Gaussian random variable right, because Z is nothing but it is a linear function of X 1 and we have seen any linear function of a Gaussian random variable is a Gaussian random variable. So, whatever values of a and b you choose Z is going to be a Gaussian random variable ok.

So, now the question is are these X 1 and X 2 jointly Gaussian random variables; the answer is obviously, yes. Why are these jointly Gaussian random variables? Because any linear combination of X 1 and X 2 gives us a Gaussian random variable.

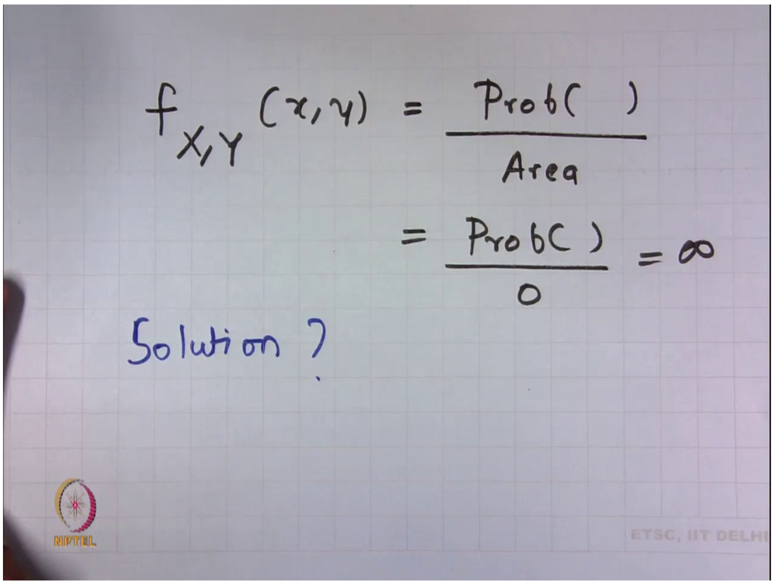
(Refer Slide Time: 26:38)



So, now, if we look at again these X 1 and X 2 after having understood that these are jointly Gaussian random variables, what kind of random variables are these X 1 and X 2? For example, if X 1 takes in a numerical value small x 1. We know for sure that X 2 will take a numerical value 2 x 1 minus 3 because X 2 is nothing but it is a linear function of X 1. It is a deterministic function of X 1. Thus, if X 1 takes a numerical value x 1 small x 1, random variable X 2 will take in a numerical value 2 x 1 minus 3 and thus, if I know what value random variable X 1 has taken, I can completely tell you the value that random variable X 2 will take.

And these kind of random variables are referred to as linearly dependent random variables ok. So, X 1 and X 2 are linearly dependent random variables. So, X 2 is a deterministic it is a linear function of X 1. So, linearly dependent on X 1 specifying X 1 will tell me everything about X 2 ok. For example, if someone says the random variable X 1 has taken a numerical value x 1, we know for sure that this random variable X 2 would take in a numerical value 2 x 1 minus 3 and thus, the probability density in this two dimension case will lie along a line. Isn't it? The whole probability density in this two dimensional case just collapses to a single line and in this case thus the joint probability density function is not defined ok. Let us see let us see why.

(Refer Slide Time: 28:51)



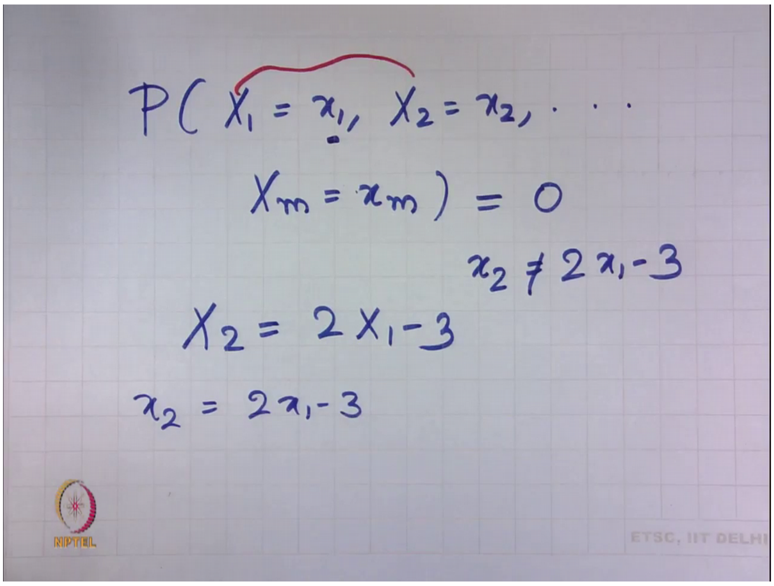
So, remember that the joint probability density function of two random variables is what? Probability divided by unit area in to dimension it is divided by area and what is the area of a line? Area of a line is 0. So, the joint probability density function is nothing but probability upon 0 and this is going to be infinite.

Thus, we see what happens is if two random variables are linearly dependent random variables, the probability density function collapses to a line and that is what we see is that the joint probability density function will not be defined, until and unless you are willing to allow impulses and delta functions ok. We are not willing to take that route because there is lot of notational mess that will be introduced. We want to keep things simple ok.

So, what we do now? So, how to get rid of this problem? So, what is the solution ok? What is the solution if we have bunch of random variables which are linearly dependent? If we have bunch of random variables which are linearly dependent, we have seen that the probability density function would not be defined ok. You can also think it in just one more way.

So, we know that the area is in this case it will be delta square times mod of determinant of A right. We have seen once you go from one space to second space, this is what the area is and we have seen that the area is 0. If the area is 0 determinant of A would be 0 and determinant of A would be 0; that means, A is a singular matrix. It is inverse does not exist ok. So, we are telling you the same thing in a different way right. So, what is the solution? Let us try to look at the solution.

(Refer Slide Time: 30:59)

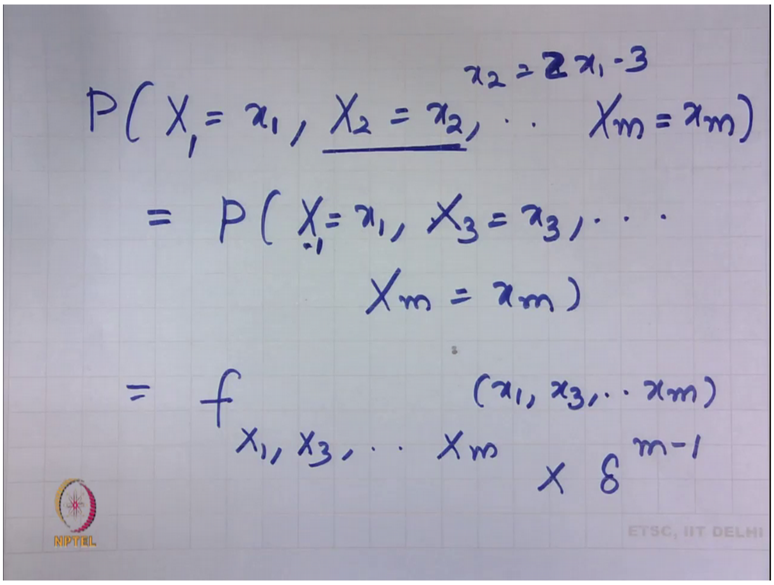


So, let us assume that I have m random variables. So, here I am assuming that these x 1 x 2 x m are continuous random variables and when I say probability that X 1 is equal to small x 1, X 2 equals to small x 2 and X m equals to small x m, I simply mean that these random variables are taking values around their corresponding numerical values. So, namely X 1 is taking a value around the small x 1, X 2 is taking a value around small x 2 and so on so forth.

Now, let me assume that X 2 and X 1 are linearly dependent ok; rest let us assume that they are linearly independent. So, once you want to talk about the probability of random variable X 1 taking in a value small x 1, X 2 taking in a value x 2, X m taking in a value x m, this probability will be 0 if x 2 is not 2 x 1 minus 3. Remember we are we are assuming as before that X 2 is 2 X 1 minus 3.

So, if random variable X 1 has taken any value x 1, random variable X 2 has to take a value 2 x 1 minus 3; it cannot take any other value. So, if x 2 is not same as 2 x 1 minus 3; then this probability is going to be 0 because these two events cannot happen at the same time. So, we know easily how to calculate this probability if x 2 is not same as 2 x 1 minus 3. What happens if x 2 is 2 x 1 minus 3; what happens then?

(Refer Slide Time: 33:10)



If x 2 is 2 x 1 minus 3 probability of X taking in a value x 1 and X 2 taking in a value x 2, where x 2 is 2 x 1 minus 3 and X m taking in a value x m is nothing but probability of X taking in a value x 1, X 3 taking in a value x 3 and X m taking in a value x m. So, what I am saying is this probability of X 2 taking in a value small x 2, where x 2 is 2 x 1 minus 2 x 1 minus 3 is 100 percent, once you know that X has taken in a value or X 1 has taken in a value small x 1. Let me put here ok.

And this probability can be computed from the joint probability density function multiplied by delta m minus 1. See what I have done? What I have done is I have removed this linearly dependent random variable because I know that once this event has happened, this event would automatically happen if x 2 is 2 x 1 minus 3, these probabilities are same. If I am interested in finding the probability for this event, this could be simply found from the joint probability density function of random variables X 1, X 3 and X m and multiplying it with the volume of that space because now only I have m minus 1 random variables because I have checked out this random variable X 2.

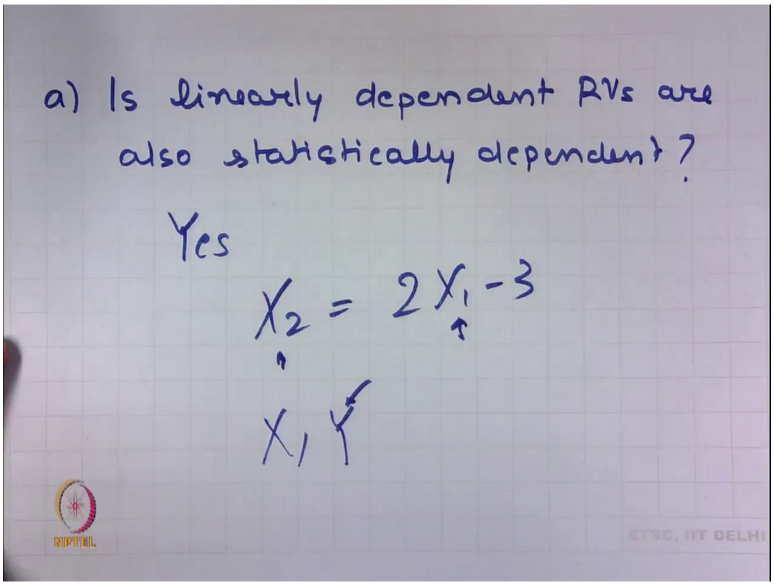
The dimension of my space has reduced from m to m minus 1 and this joint probability density function is also a specified over a lower dimensional space and this probability density function will be defined. Because now there are no linearly independent random variables here anymore; whatever random variables were dependent I have removed them because whether you include this or not, the probabilities will remain unchanged.

So, this event as the same probability as this event for x 2 taking any value to x 1 minus 3.

So, what we are doing is we are defining this probability density function over a lower dimensional space, when we are doing this we are checking out or linearly dependent random variables because of which this probability density function in original dimension was not defined and now we have this probability density function which will be always defined because this is only consisting of random variables which are linearly independent and we are multiplying it with the volume of a lower dimensional space ok. So, in short this is what we are doing.

We were having a bunch of random variables, these bunch of random variables might have some random which are linearly dependent. We check out those random variables, the linearly dependent one and then we have a probability density function or joint probability density function which we can get defined over a lower dimensional space which only consists of random variables which are linearly independent and we can get rid of this issue ok.

(Refer Slide Time: 37:12)



So, one question that is of interest is, is linearly dependent random variables are also statistically dependent? And the answer is obviously, yes because what does linear dependence? Linear dependence means that one random variable is a deterministic function of another random variable. So, if we know this random variable you know everything about this random variable. So, it is a completely deterministic dependence. When we are talking about statistical dependence, it simply means that once you know about random variables.

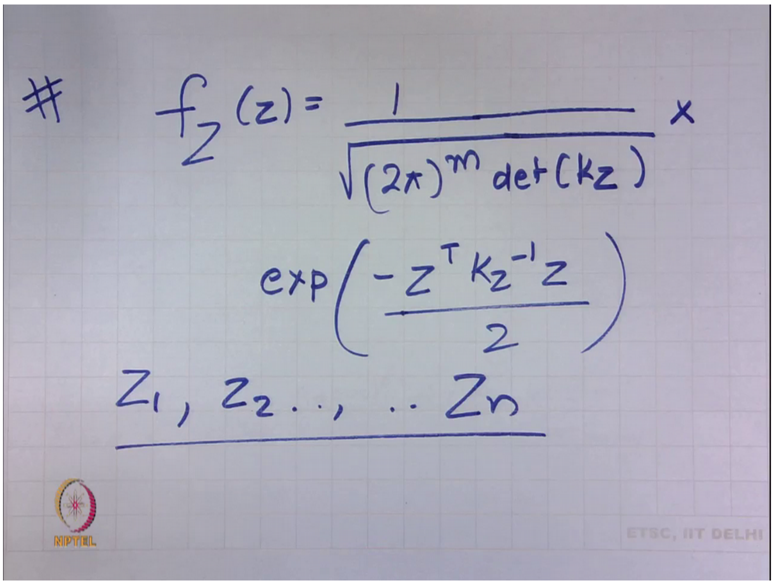
So, for example, if two random variables X and Y are statistically dependent, it simply means that if you are observing Y your information about X as improved, but you would not know everything about X ok. So, if these are linearly dependent that means, you know everything about the other random variable. Then, these two random variables will also be statistically dependent right.

But if these random variables are statistically dependent, then need not be linearly dependent. The statistical dependence is weak dependence right, it is a probabilistic dependence; whereas, the linear dependence is a completely deterministic dependence. So, statistical dependence does not imply linear dependence, but linear dependence imply statistical dependence ok.

Let us conclude this issue by really relooking the probability density function of random vector Z. So, let us use this one. So, we have already derived the probability density or joint probability density function of random vector Z and this is also a tool to define Gaussian random vector. So, what we are saying is if a random vector has such a joint probability density function, then that random vector is a Gaussian random vector ok. If it is a Gaussian random vector, they should be the probability density function.

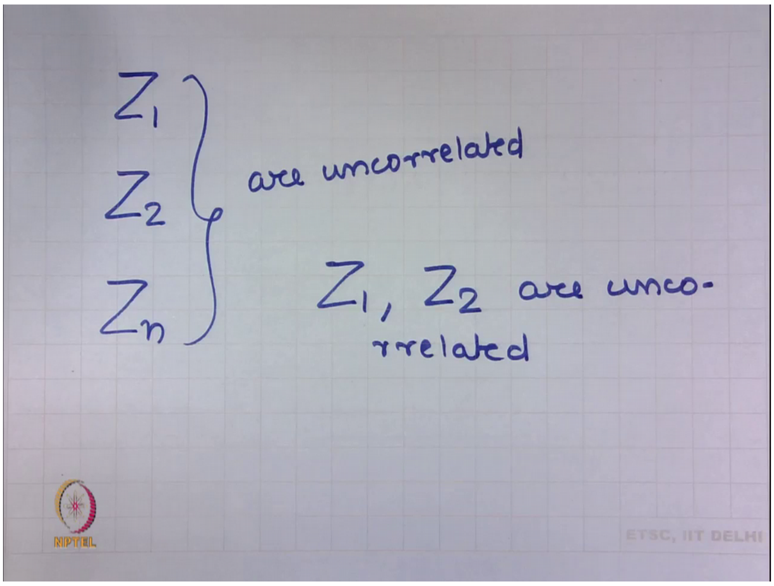
If this is the probability density function or joint probability density function of that random vector, then that random vector involved is a Gaussian random vector. So, this is third way in which we can talk about Gaussian random vector or jointly Gaussian random variables. Once we have said that this is a Gaussian random vector and then, we can also say that it is the element should be jointly Gaussian random variables ok. So, this is a third definition of jointly Gaussian random variables that means, jointly Gaussian random variables has such a probability density function ok. So, let us now try to look at the last issue and this issue is trying to understand this expression little better in some special cases.

(Refer Slide Time: 40:26)



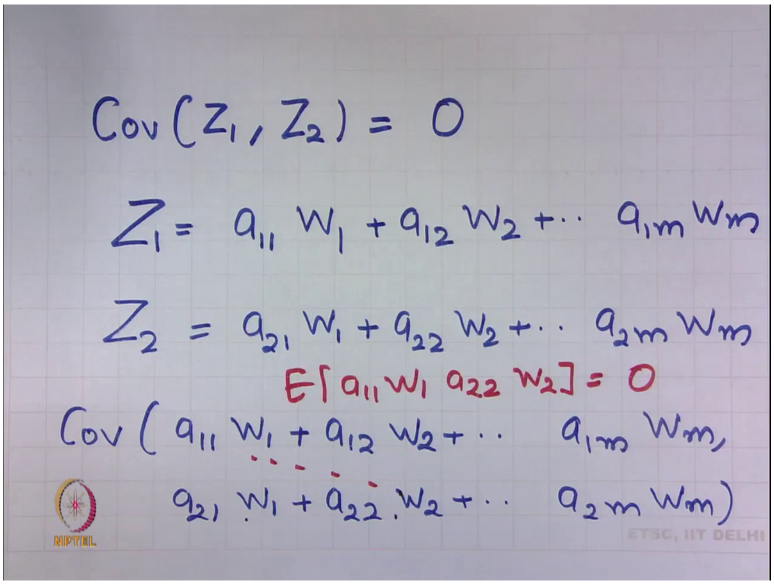
I am rewriting this time in again. So, that it gets imprinted in your head. So, this is the expression and remember that these z are small z right, they denote the numerical values rather than the vector itself ok. Now, let us try to ask yourself a question what happens if the elements of Z; that means, Z 1, Z 2, Z n. So, these Z 1, Z 2, Z n are random variables; what happens if these random variables are uncorrelated; what happens to this probability density function? So, this is what we are trying to discuss and see and we will end up with very interesting result.

(Refer Slide Time: 41:32)



So, let us start with the premises that is Z 1, Z 2, and Z n are uncorrelated and to begin thinking about this let us just start taking two random variables Z 1 and Z 2 and let us see what happens if Z 1 and Z 2 are uncorrelated. Now if these two random variables are uncorrelated, the covariance must be 0.

(Refer Slide Time: 42:11)

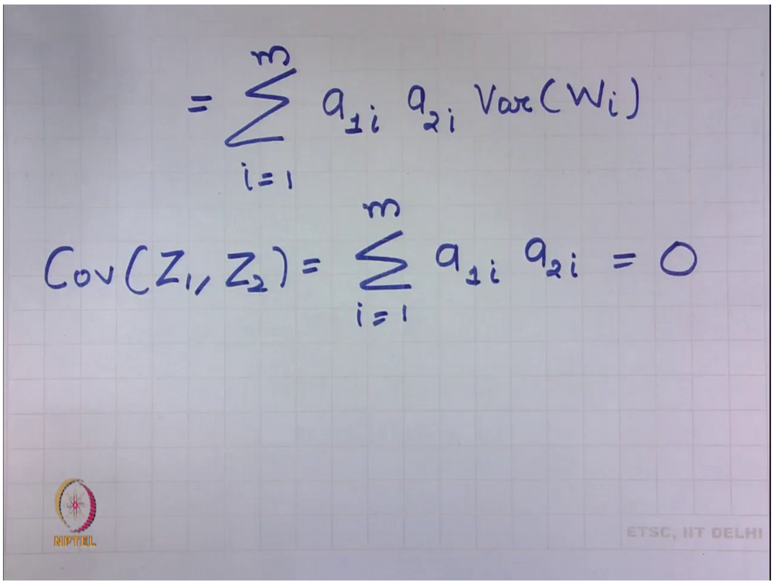


So, covariance of Z 1 and Z 2 must be 0 if Z 1 and Z 2 are uncorrelated right. This is how we define uncorrelated random variables. So, what is Z 1? Z 1 we know is a 11 times w 1 plus a 12 times w 2 and so on so forth a 1m times w m and Z 2 is a 21 times w 1 plus a 22 times w 2 plus a 2m w m ok.

Now, substituting these expansions of Z 1 and Z 2 in this equation, we get all right. Now, we know that covariance is a bilinear operation. So, we can multiply these terms with this terms and there will be some cross terms ok. I will not derive this completely; I will leave it to you to see that once you multiply these cross terms and you take the expectation after multiplying these cross terms, what we will get is applied 0.

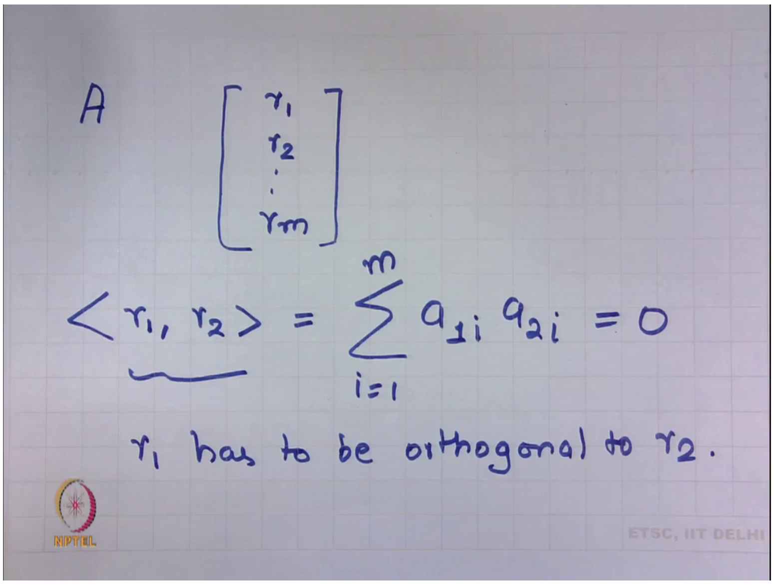
Because w 1 and w 2 are independent right and they are with 0 mean right. So, expected value let me write a 11 w 1 into a 22 w 2 would be 0 right, because they are independent and there with 0 mean. We have done it several times and now, I request you to do it yourself and think that all these cross terms will vanish ok.

(Refer Slide Time: 44:36)



So, finally, what we will have is summation a 1i a 2i variance of w i; where, i goes from 1 to m ok. Just try to prove it yourself. It should be simple and easy. Furthermore, what we know is that these random variables are also normal random variable that means their variance is 1. So, from this, we can say that covariance of Z 1 and Z 2 is nothing but summation i going from 1 to m a 1i times a 2i assuming that this is 1 and we know that if these two random variables have to be uncorrelated, this has to be 0 alright. So, let us see under which condition this will be 0.

(Refer Slide Time: 45:41)

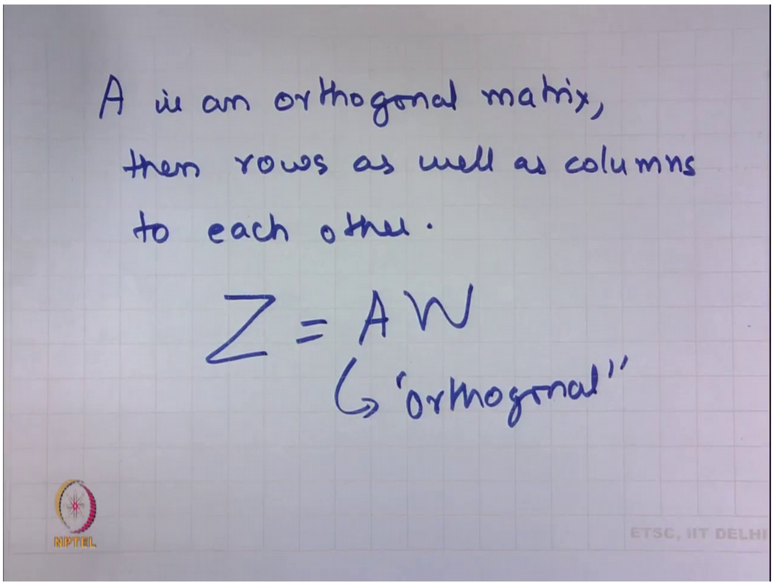


And for that let us assume that I have a matrix A and this matrix A consists of m rows and let us say that I am interested in the inner product of row 1 with row 2, and what is the inner product of row 1 with row 2? It can be obtained by multiplying the elements of row 1 with the corresponding elements of row 2 and then, summing this up. So, the inner product of row 1 with row 2 is exactly the same thing as we have derived before.

Of course, you can have conjugate if these elements are complex, but we are assuming everything to be real and those there are no conjugates and so on and so forth. So, what we are saying is if these two random variables Z 1 and Z 2 have to be uncorrelated this quantity has to be 0; that means, the inner product of row 1 with row 2 has to be 0 and this simply means that for the random variables to be uncorrelated row 1 has to be orthogonal to row 2.

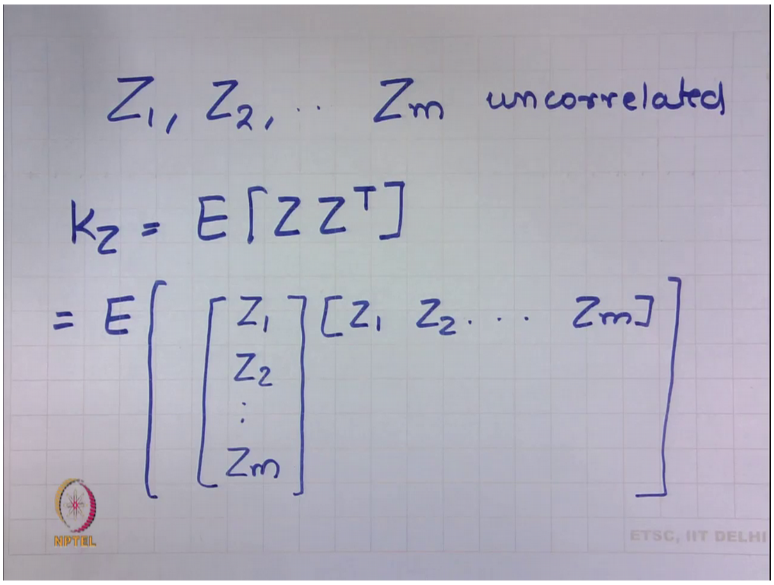
And if you want random variable Z 1, Z 2 and Z n to be uncorrelated with each other, what you simply require is that these row should be orthogonal to each other.

(Refer Slide Time: 47:30)



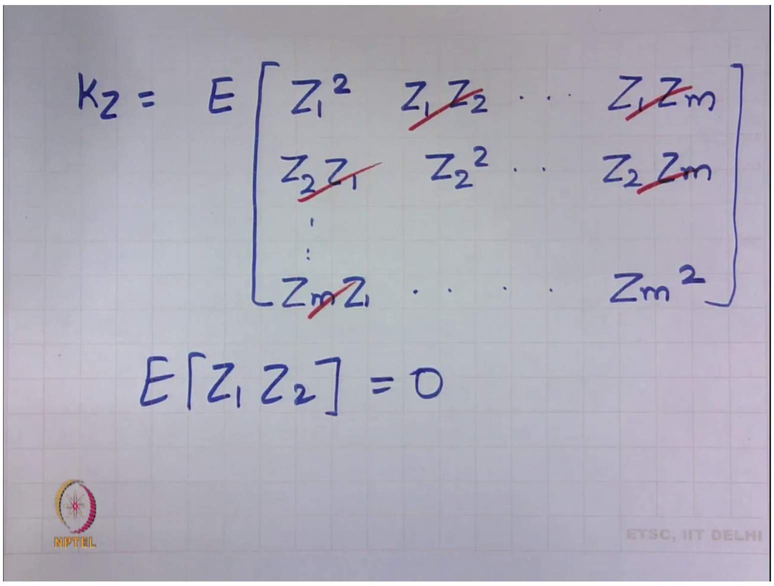
And you know that if A is an orthogonal matrix; if A is an orthogonal matrix, then rows as well as columns are orthogonal to each other. Thus, if you have this linear transformation Z is a times w and if you assume a to be orthogonal, then what you ensure is that these elements of Z which are Z 1, Z 2 and Z n or Z m they are uncorrelated ok. Now, let us see what happens to this joint probability density function, if these elements are uncorrelated.

(Refer Slide Time: 48:30)



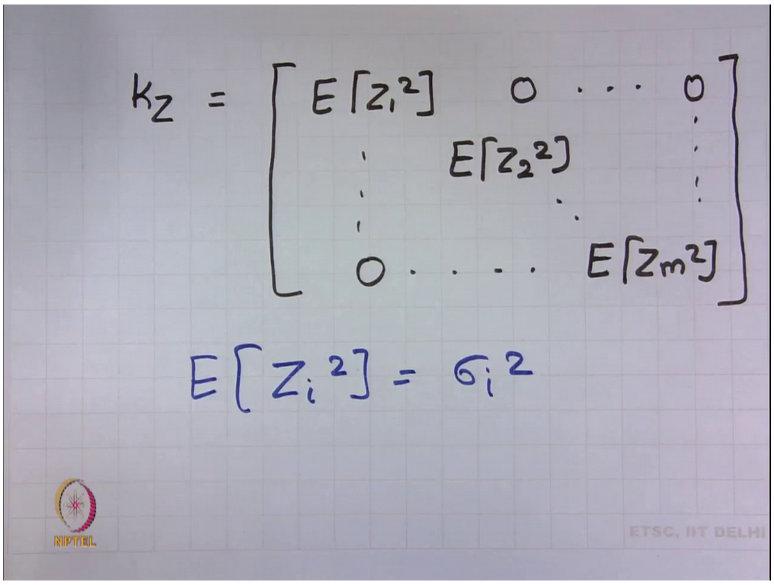
So, let us now assume that these are uncorrelated and let us see what happens to the joint PDF and joint PDF depends upon the covariance matrix. So, let us see what happens to the covariance matrix and what was the covariance matrix? It is the expected value of Z times Z transpose and so, this is expected value of Z. What is Z? Remember by notation we are using, we are assuming the vectors to be column vectors. So, this will be Z Z transpose will be this and we have done this several times.

(Refer Slide Time: 49:27)



This would be ok. Now, we can pull this expectation operator inside and we know that expectation of cross terms for example, Z 1, Z 2; this would be 0 because we know that these are uncorrelated and we also know that they are with 0 mean. So, expected value of Z 1 times Z 2 would be 0. So, all these cross terms would vanish.

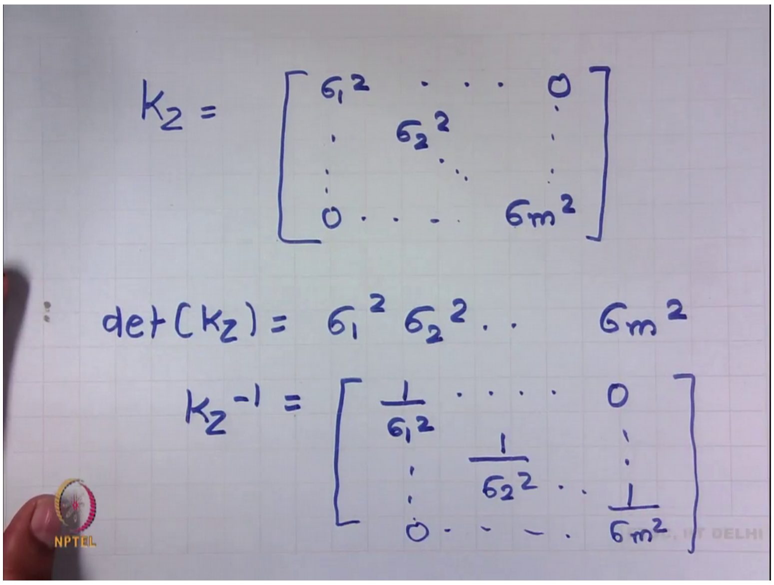
(Refer Slide Time: 50:36)



So, finally, what we have is covariance matrix will be expected value of Z 1 square it will be a diagonal matrix expected value of Z 2 square expected value of Z m square and so on and so forth, rest will be 0.

Now, let me use a notation to simplify this expression that expected value of Z i square is sigma i square ok. So, we are using this notation to talk about expected value of Z i square and hence, our covariance matrix becomes sigma 1 square sigma 2 square sigma m square right.

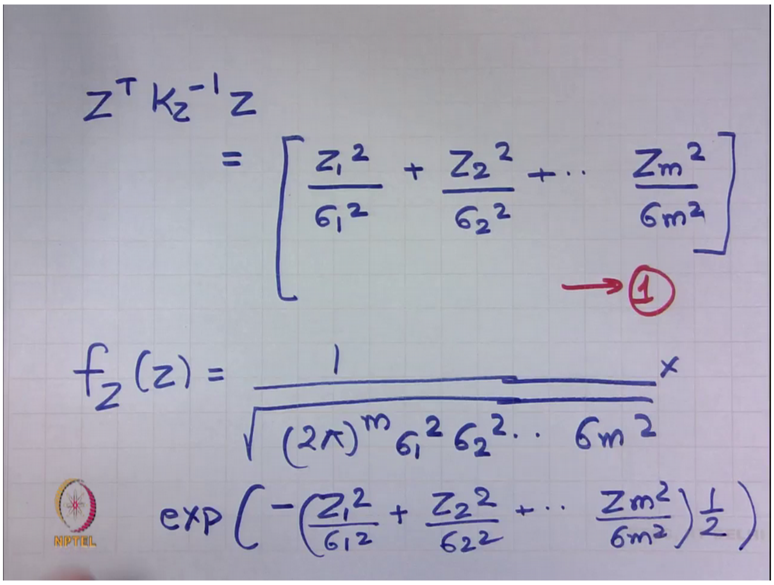
(Refer Slide Time: 51:21)



So, we have to find now the determinant of this k z and k z inverse right because joint PDF depends upon that. So, let us do the easy one first. What is determinant of this k z? Determinant of a diagonal matrix you must have seen is nothing but it is the product of diagonal elements.

So, determinant of k z will be nothing but sigma square times sigma 2 square to sigma m square; this was easy. Was an inverse of a diagonal matrix; inverse of a diagonal matrix is also easy. This is obtained by the reciprocal of diagonal elements. So, this will be 1 by sigma 1 square 1 by sigma 2 square 1 by sigma m square and everything else will remain 0. So, this is how you take the inverse of a diagonal matrix.

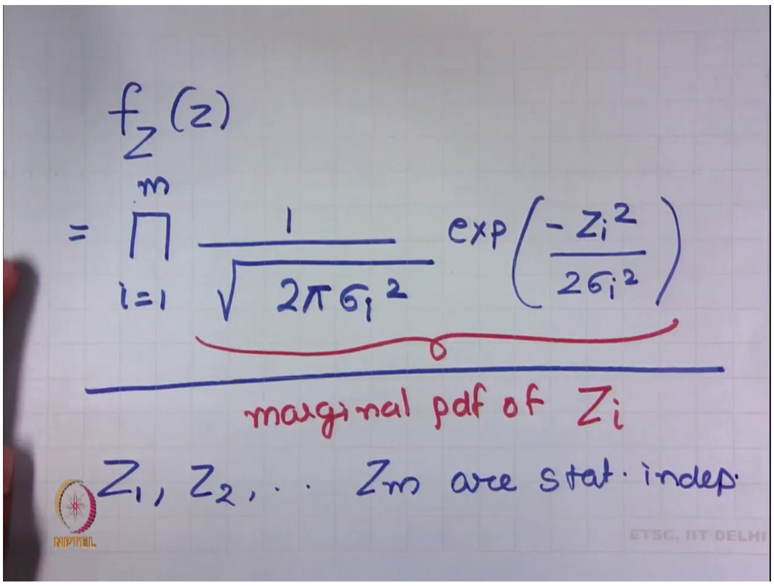
(Refer Slide Time: 52:48)



Let us now see what is Z transpose k z inverse times z, I would not derive this I leave the derivation to you because it is quite a straightforward. This will be z 1 square by sigma 1 square plus z 2 square by sigma 2 square z m square by sigma m square. So, now, is the time to plug these things. So, let us say this is expression number 1 and let us say this is expression number 2 in the expression of joint PDF.

So, we will have 1 by 2 pi raise to m determinant of k z will be sigma 1 square sigma 2 square up to sigma m square and then, I will have exponential minus z 1 square by sigma 1 square. So, all this z are small z's ok; whole divided by 2 and what you can see is that you can conveniently express this joint PDF like this.

(Refer Slide Time: 54:38)

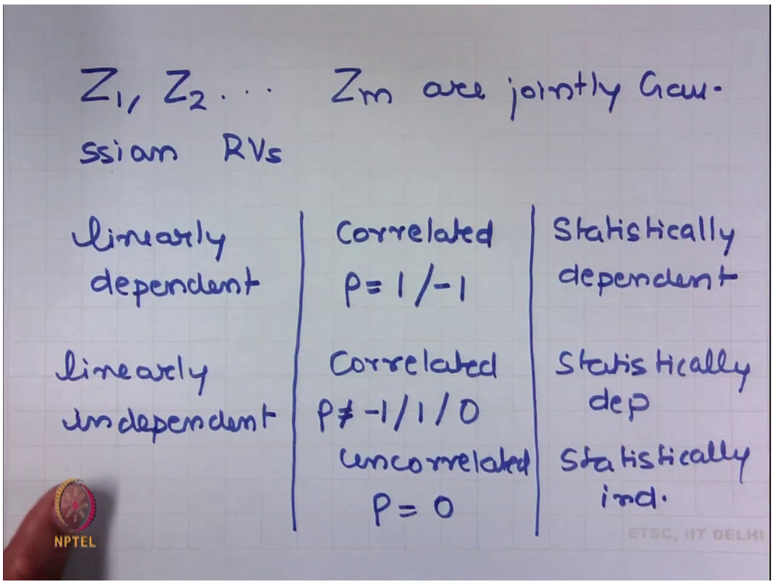


So, this denotes that I am interested in taking the product of the terms. So, I will have this is i. So, the joint probability density function of a Gaussian random vector z when the elements of this random vector z are uncorrelated is computed to be like this. And if you see this carefully what happens there is something interesting going on here. First thing that you realize that this is the marginal PDF of Gaussian random variable Z i and what we are now saying is the joint PDF is nothing but it is the product of marginal PDFs of the involved random variables and we know, when this happens when does the joint PDF is simply the product of marginal PDF.

This happens when the underlying random variables are statistically independent and thus, we see something beautiful has happened that these random variables are statistically independent random variables. So, we can say from this that Z 1, Z 2 and Z m are statistically independent random variables. So, this is a proof that when these random variables are jointly Gaussian random variables and if these random variables are also uncorrelated, then these random variables are also statistically independent; that means, uncorrelated jointly Gaussian random variables are also always statistically independent.

Remember this was not true in general. Uncorrelated random variables does not imply statistical independence, but if these uncorrelated random variables are also jointly Gaussian random variables. Then, uncorrelatedness implies statistical independence and this is something important. Let us try to finish this lecture. Let me make a small table which might help you in clearing out some confusions if they are remaining.

(Refer Slide Time: 57:37)



So, what we are saying is if Z 1, Z 2 and Z m are jointly Gaussian random variables, they might be linearly dependent. Isn't it? If these are linearly dependent, it also means that they are correlated and the correlation coefficient can be 1 or minus 1. It also means that they are statistically dependent.

Now, the case 2; these random variables can be linearly independent; if these are linearly independent, they might be correlated, where the correlation coefficient will lie between minus 1 and 1. But it cannot be minus 1 or 1 or 0 if they are correlated, they are also statistically dependent; they can be uncorrelated and then, the correlation coefficient is 0 and this implies that they are also statistically independent ok.

So, this brings us to the conclusion of today's lecture. In today's lecture, we have covered lot of important concepts. We have derived the joint probability density function of a Gaussian random vector and we have seen there that the joint probability density function of a Gaussian random vector is only a function of mean and covariance matrix. We have further seen that it might happen that the jointly Gaussian random variables involved are linearly dependent and then, the best way to sort out that issue is to consider the bunch of random variables which are linearly independent and define the probability density function over a lower dimensional space.

The third thing that we have seen is that if we take the A to be an orthogonal matrix then the elements of Z will be uncorrelated and if the elements of Z are uncorrelated, it will also guarantee that these random variables are statistically independent. Thus, uncorrelated jointly Gaussian random variables are always statistically independent.

From next lecture, we will see some more interesting things in the context of random processes. We will see what happens when this random process pass through linear time invariant systems or digital communication systems.

Thank you.