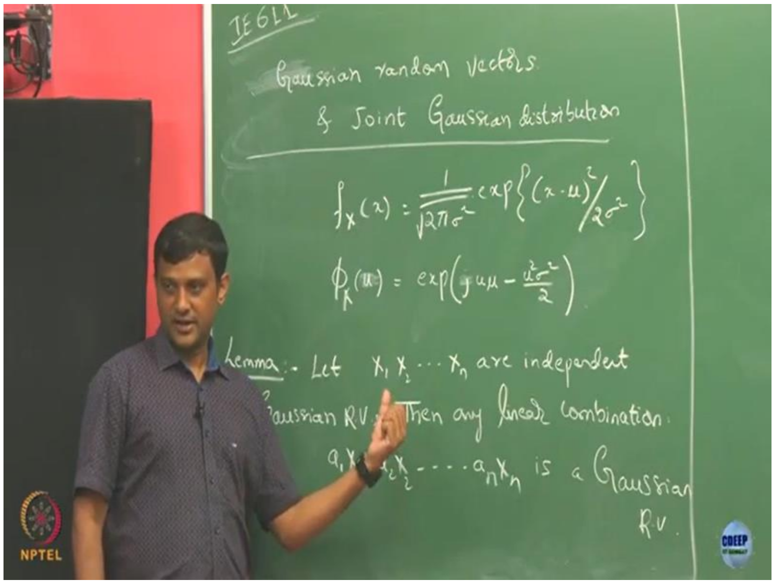
Introduction to Stochastic Processes
Professor Manjesh Hanawal
Industrial Engineering & Operational Research
Indian Institute of Technology, Bombay
Lecture No. 15
Gaussian random vectors and jointly Gaussian distribution

(Refer Slide Time: 00:18)



So, we already know what is a Gaussian distribution and what is a Gaussian random variable. And now we want to say when you have a random vector, when we are going to call it a Gaussian random vector. What is the properties it needs to satisfy? So, we already have for a single random variable, the PDF is given by, so this means I have a Gaussian random variable X that has mean mu and variance Sigma square. And what was its characteristic function? A function mu, how did its characteristic function look like? exp minus u mu, I do not know where was minus.

Student is answering: j u mu plus minus u square sigma square by two.

Professor: j u mu plus minus u square sigma square by two, this happened to be discussion. It is a characteristic function. Now, suppose I have a set of such random variables which are independent but not necessarily identical. So, what I mean by independent, we already know what I mean by independent of set of random variables, right? We know that if the joint distribution, the joint distribution splits into their marginal distributions, then we are going to call them independent. And I am going to say set of random variable is to be identical if all of them have the same distribution.

Okay, so now suppose I have a collection of random variables and all of them have and they are independent. Then it so happens that you take any linear combination or this random variables. It is still going to be again, Gaussian maybe with a new mean value and a variance. Okay, so with this, we are going to write it as lemma, so take any n random variables that are independent and all of them are Gaussian random variables. So each may have its own mean and variance different values.

Now, we are going to set, if you take the linear combination like this, you understand linear combination, right? I will just take different weights, multiply each one of them with their associated weights and then add up. Then I got this linear combination. We are saying that this is also Gaussian random variable. So, why this is true? So to see this, we are going to use the property that if a set of random variables are independent, then if you are going to look at the characteristic function of their joint distribution that also splits into characteristic function of individual random variables, right?

So that we have said, so we have said that the set of random variable are independent, if and only if their characteristic function also splits, not only their CDF splits, that also implies that their characteristic function also split. So, let us use that property. Just quickly verify this. Suppose.

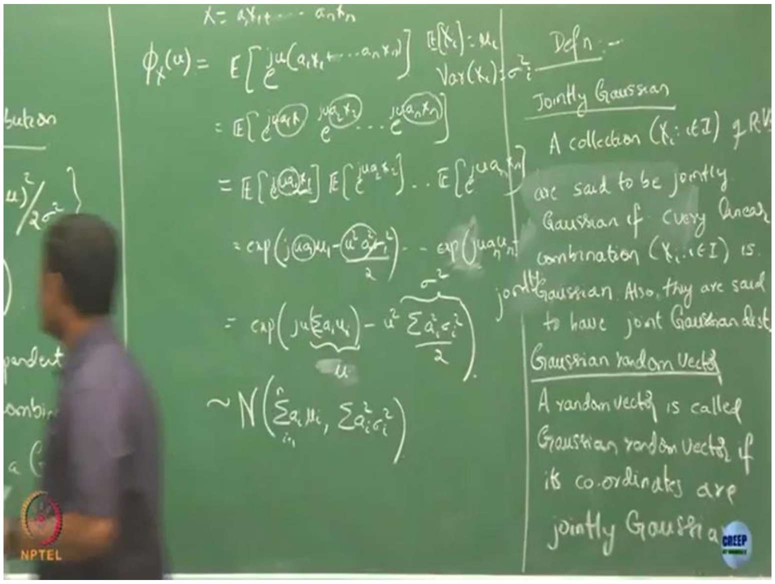
Student is questioning: Can we say that these X1, X2, X3 are identical?

Professor: No, I am not saying that. I am just saying that they are independent.

Student is questioning: But since they are all Gaussian.

Professor: Yeah, Gaussian but Gaussian they could be with different mean and variance, right? They for a given mean and variance. It is going to define different, different Gaussian distributions. It is a parameters okay.

(Refer Slide Time: 4:54)



So, now let us say I want to define characteristic function of my random variable X where X is now the joint collection of these random variables. X is a vector here now, it is not simply a scalar. So now, okay. Now, what I mean is, okay. Now, I am going to define this X not, I am going to define this X to be, this linear combination. X is now again another random variable for me it is not the vector. As I said earlier, it is just linear combination of this and I want to compute the Gaussian distributions. Sorry, the characteristic function of this random variable X.

Now, what is this going to be? It is going to be expected value of e to the power to j u and all this right. a 1 x 1 all the way up to, so this is nothing but expectation of e to the power j u a 1 X 1 into e to the power J U a 2 X 2 all the way e to the power J u an xn. Now, that X 1 X 2 they are independent. Do you think a 1 x 1 a 2 x 2 and a n x n they are also independent. If x 1 and x 2 independent. Suppose, x 1 I multiply by a 1 and X 2 multiply by a 2 will a 1 x 1 and a 2 x 2 will be also independent.

So, this random variables now are independent. Now, if they are independent now can I write this expectation as the product of expectation of each of these terms. Is they are independent, I should be able to do this. Now, x 1 is Gaussian I know, is a 1 x 1 is also Gaussian? A is a constant, right? So, what is this quantity now? This is a characteristic function of what a 1 x 1 and what it will be? So, we already know that for a, if X is a Gaussian random variable, its characteristic function is going to just look like this j you mu.

Now, so it only depends on your mean value and the variance, right? So, what is the mean value of a 1 x 1? Suppose, let us.

Student is answering: a 1 mu 1.

Professor: Say, xi, expectation of xi is equal to mu i. for ith random variable which has Gaussian has mean mu i. So, then what is this mu of this going to be? u and it will be, what will be the variance of a 1 X 1? So, if variance of let us say X square, I am going to denote it as sigma square.

What would be the variance of axi?

Student is answering: a square.

Professor: So, then what is, then u square and then I am going to replace Sigma square here by.

Student is answering: a 1 square.

Professor: 2 or you can alternatively think of this as like instead of, this as a random variable. You can think as. This is u 1, u a 1 is the point at which you want to evaluate this characteristic function. And for this random variable X. So, in that case also like it is like same like you are replaced, you are going to replace u by u a 1 here, u by u a n and u square by u square a square. But alternatively can also think of a 1 x 1 as another random variable with new mean and variance. Now, all the way to what is its value going to be?

So, now if I further simplify this, what this going to look like? It is going to look like j u summation a 1 ai mu i minus u square summation, a i square Sigma i square divided 2, right? Now, if you are not go back, and we have already said that every distribution has a unique characteristic function right. Now, you should just look at this characteristic function. What is this distribution corresponds to? What is this characters? This characteristic function corresponds to which distribution? But what parameters?

Student is answering: (())(11:07)

Professor: So, suppose let us go, if I want to map it to this. Suppose, let us say this. I have to, if this is some mu, and this is some Sigma square ok. Anyway, this is constant and right and it will be exactly in this format. So, then this is going to be a Gaussian with so this corresponds to a Gaussian with what mean and variance? So, this is going to be another, so this means my X here, which is a linear combination of independent Gaussian distributions is going to be another Gaussian distribution with mean like this and variance like this. So, what it says is that what we have just showed is any linear combination of independent Gaussian random variables is going to be another version random variable right?

So, now we are going to use couple of more definitions. So, what I mean by here is let us take a collection of random variables. You understand this notation Xi, i belong to capital I. So, I is the index set here? Here this index set is going to have finitely many values. They are said to be jointly Gaussian if every linear combination is Gaussian, and they are done if that they are said to have a joint Gaussian distribution. So what we just what what we mean here is you are, you are given a Gaussian distribution set of simply a random variable.

You shall go take any linear combination of them and they should be Gaussian. If that is the case, then you're going to call this set of random variable to be Gaussian. They are going to, they are going to call them a jointly Gaussian and they are said to have, Gaussian distribution. So, we have already said that if X 1 X. This set of Xi’s happens to be each one have to be Gaussian and they are independent. We know this is already true, right?

You take any combination of them they are again going to be Gaussian. So, so here, these are the set of, if they are independent and Gaussian they are again already jointly Gaussian, what? They are just saying if some given set of random variables happens to satisfy this property, then we are going to call them jointly Gaussian. And now when we are given a Gaussian vector, so this is a vector will have components, right?

And then we just treat those component as these components, and if their coordinates, so when I say X, X is random vector, it will have X1, X2 let us say after the Xm, then these I can treat it as coordinates of this vector, right? As we discussed in the previous class, we can treat them as these components. And if these components happens to be jointly Gaussian, then we are going to call that random vector as simply Gaussian and random vector. Okay.

So this was obviously, so here we just said collection of random variables, but I can treat all these collection of random variables as a vector, right? In which this vector constitute the components. Okay? Now, when I have a collection of random variables, how am I going to denote its distribution?

(Refer Slide Time: 15:18)

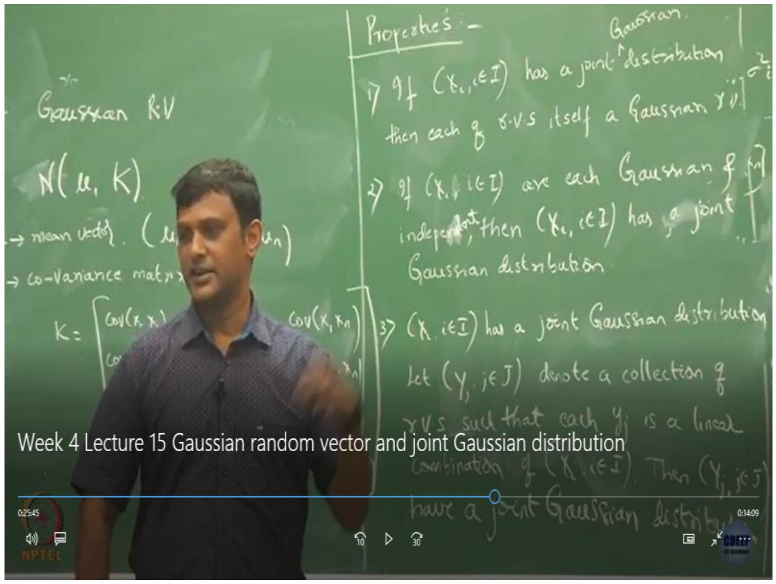
So, if I have random variable X that is Gaussian random vector. So, I am, now start using instead of random variable, I am going to use RV for random vector here, it should be clear whether I am just talking about a random vector. Or I am talking about a random. Sorry, I am just talking about a random variable or random vector. I am going to denote its distribution by mu and k here and you will see that it just like, it only depends on these parameters and you have mu is the mean vector and K is the covariance matrix.

What is mu here it is like mu 1, mu 2 all the way up to mu n. So, let us say I have components, right? X is a random vector it has some N components, let us say. Then this mean vector is nothing but the vector of individual components and what is k here in the covariance metrics? It is going to be a metrics of covariance of X 1 X 2. X1 with itself, covariance of X1 X2 to covariance of X1 X10 and similarly X2 X1 and you can write all the way up to covariance of X1 X1. So, what does this here in this covariance matrix what is this diagonal implies?

Student is answering: Variance.

Professor: So, the diagonal contains the variance of each of the components. Okay.

(Refer Slide Time: 18:32)



Now, I will just list down some of the properties, of this, Gaussian random variable and I am not going to prove any of them. You should verify all these things yourself. Gaussian random vectors or like for which we are going to say it is going to be, it has joined Gaussian distributions. Now, let us say, if Xi, I have this collection of random variable has a joint distribution, then each of the random variable itself a Gaussian random variable now. Is this true? Why is that? So, I have a collection of random variable.

I am saying that if it is jointly, it has joint, joint Gaussian distribution, then each of the random variables is itself is a Gaussian random variable. Why is that?

Student is answering: We find it like that.

Professor: Yeah, so if it is a jointly Gaussian distribution by definition we want for every linear combination, right? So, in this every linear combination, what I could do is like I hope to choose A1, A2 all the way up to An these are the weights, right? I can just choose A1 to be some non-zero value and set all the others, A2, A3 all the up to An to 0.

In this case, I have, I need to check, I need to satisfy that A1, X1 is Gaussian, but since A is simply a constant it is the better that X1 is Gaussian in this case. So and similarly, for each of the components. So, it must be the case that if the set of random variables has joint Gaussian distribution. Each one of them has, each one of them itself is a Gaussian random variable and this property we have already verified. I will just write it for the sake of completeness.

If, yeah, right now I am not saying anything, like you say you are given a set of random variables and if it satisfies. If they are jointly Gaussian distribution. For that our need is every linear combination should be Gaussian. That is the only thing, I am not the, I do not want, I am asking anything like that independent, identically distributed or anything. Just applying definition. So, suppose this Xi's are each Gaussian and independent, then Xi has a joint Gaussian distribution.

This we have the already shown, right? The second point, what we are saying is if this Xi's, this collection of random variables are such that each Gaussian and independent. Okay, did we show this or not?

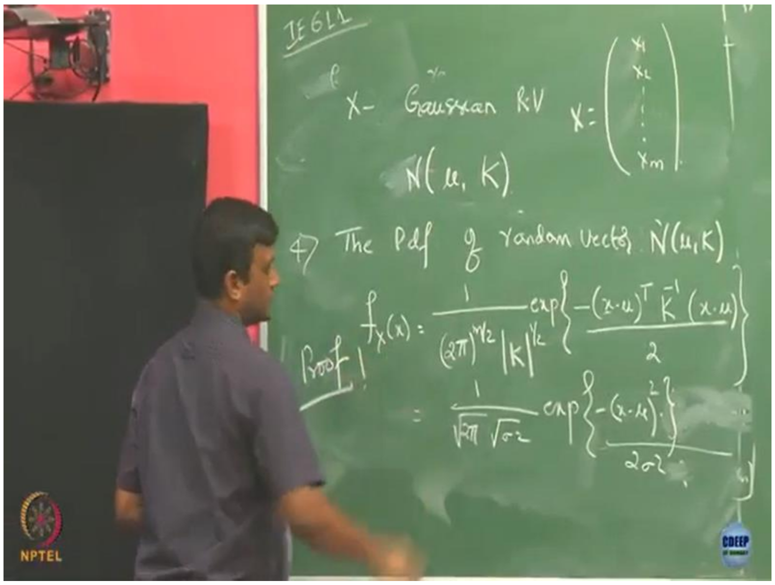
Student is answering: Yes.

Professor: If each collection of random variables they are Gaussian and they are independent and we have already shown that you take any linear combination of them that is again going to be Gaussian and then by definition it is it has it joint Gaussian distribution. The third property says, let us say Xi as a joint Gaussian distribution and you are going to construct and let this Yj be another set of random variables where j belongs to set j. So, is the statement clear here? What I am saying is take a collection of random variables that has joint Gaussian distribution. Now, what you do is you come up with another set of random variables, call them Yj, j taking again value in some set capital J where each of this Yj itself as linear combination of this Xi's.

Let us say you have X 1 X 2 all the way up to X10 you have 10 random variable that is jointly Gaussian distributed. Now, let us say you make, you take one linear combination of this random variables. We can take another linear combination like that you say, you come up with 100 linear combination of this X1 all the way up to X2. Now, this 100 random variables which I have denoted Yj. Now, this is Yj, our claim is again has a joint, we has a joint Gaussian distribution. Does it make sense?

This should be obvious that because if these Yj's are already linear combination of Xi's and now you further take any linear combination of this Yj's itself. There will another the linear combinations of Xi's. And there that is Gaussian distributed, so any a linear combination of Yj, then has to be Gaussian distributed. That is why this property already holds and we are going to call it as jointly Gaussian distributed.

(Refer Slide Time: 26:50)



So, now the, the Pdf of random vector N mu, K so let us say, I have already said that. Let us say X is a random vector and it is a Pdf. I am going to denote it as it is a, let us say I detonated it by this value, right? N mu K, where mu is the mean vector so what, how this value is going to be like. So this is just denoted, Gaussian random vector with parameters mu and K. Now, the CDF of this Gaussian random variable, Gaussian random vector is going to look like. So, I am assuming that this random vector X here as m components in this. Okay.

Just say that X is X1, X2 all the way up to Xn. So, we said that this is a Gaussian random vector where the mean vector mu and covariance matrix is K. And if that is the case, then its CDF one can write it CDF has this, what is here mod of K here means, it is going to be determinant of this matrix K. And what I mean by capital T here it is going to be transposed because this X here is a vector and mu also vector.

Student is answering: Vector.

Professor: What is this vector? This vector is the mean of all the components. And K inverse means inverse of the matrix. Okay. Now, suppose, let us see, we recover our initial Gaussian distribution for a single random variable with this formula. Suppose, X consist of only one component. Now this will, this formula will reduced to?

Student is answering: 2 Pi 1 by 2.

Professor: So, here in this case, M is going to be 1, right? This is going to be 2 pi. And in that case, what is K going to be?

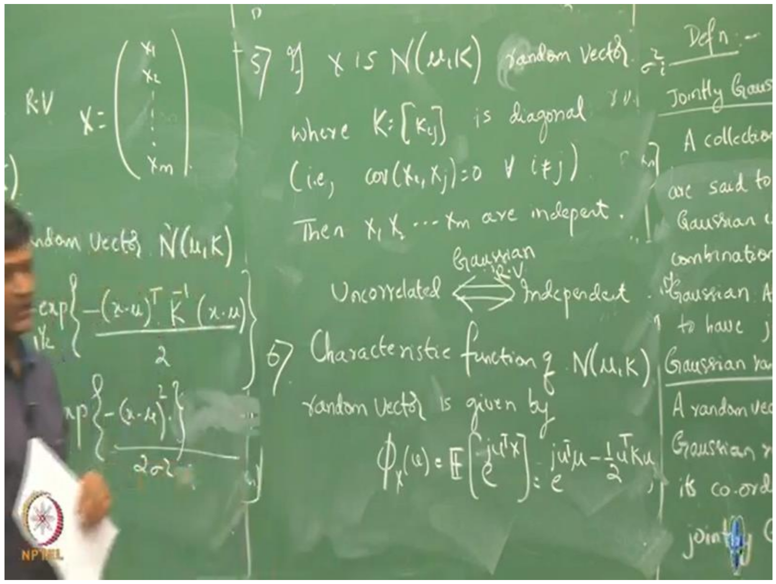
Student is answering: X 1.

Professor: So, K is going to be matrix with only one element in this. And what is that? That is covariance of X 1 with itself, right? So that is going to be variance and that is also square root. So, this is going to be this. And now for a single random variable, this is like a scalar X is a scalar and X minus mu is now just a mean of that random variable. And what is K inverse in that case? Sigma square or 1 by sigma square?

Student is answering: 1 by sigma square.

Professor: Because this is K inverse, right? That is going to be 1 by Sigma square and this is going to be X minus mu and we have 2, and this is exactly X minus mu whole square divided by 2 Sigma square that we have for a single random variable. So, please do take a look into the proof of this that is going in the book. So, this proofs comes from the eigenvalue decomposition of your vectors. Eigenvalue decomposition of your covariance matrix from that one is going to derive, do take a look into this.

(Refer Slide Time: 32:02)



Now, there is one last thing I want to say, if X is N mu K, random vector where K, is diagonal. Okay, so what we are saying is suppose you take a Gaussian random vector which has this parameter, mu and K, but this K is special here. It is such that its off diagonal elements are all 0. Only diagonal elements you allowed to be some values but off diagonals are all 0 that so off diagonals all 0 means what here?

No, in terms of covariance. When we say covariance of Xi and Xj is equal to 0. We said they are uncorrelated, right? So, all pairs of random variables are uncorrelated. If that happens then these random variables are actually independent. So, remember in the last class we said independence implies uncorrelatedness, but uncorrelatedness does not imply independence, but it so happens that for a Gaussian random variable uncorrelated also implies independence and this will provided for, this is Gaussian random vector.

This need not be true for any random vector, but provied if it is a Gaussian random vector. Then that is the case. So, you see that like if have already have a model where my random variables are jointly Gaussian distributed and they are uncorrelated, then they already implies that they are independent. Okay? So, this is like a much nicer property to have because, just by uncorrelatedness I directly get this properties of independence. Okay.

The last one. Now, how to compute the characteristic function. The characteristic function as it is said. It has some nice properties, right? That is, that is it is going to be unique for a given distribution and vice versa. So, here for the characteristic functions. So, if you have the characteristic function of a N mu K Gaussian random variable that is denoted as phi of Xu. Here X is a vector and u is also vector right? Because I am talking about random vector here. So that is going to be defined as expectation of e to the power mu transpose X.

Okay, and then that turns out to be simply e to the power j mu transpose mu. So, mu transpose, so what we are writing it as mu is for us. For us, the way they are treating it is all column vectors. Mu transpose is going to be a row vector and this mu is a vector of means which is column vector. So, u transpose mu going to be what that is just one real number, right? And this is again, and what does this quantity is going to be? So, u transpose K u, K is matrix, u is the column vector.

So, K u is going to be one column vector and then u transpose that is going to give it just one real number. So, this is a, you see that it has very much similar to what we get the characteristics of a Gaussian random variable. But in the Gaussian random vector it is just like you have to deal with vectors. So that is why this transpose and the matrix has come there. Okay.

So, this is about this Gaussian distributions also see that the Gaussian distributions has some nice properties in terms of these characteristics functions and also if they are uncorrelated directly implies independence. And this helps in many things. Where, when, when your model satisfies this, your analysis become pretty much tractable because uncorrelatedness is directly guaranteeing your independence when you have independence, all you need to worry about is the distribution of each of the random variables.

If you have that then joint distributions can be easily computed by just taking the product off this individual random variables, right? So, I do not need to really define the joint random, joint distributions there. All I have to worry about is the distribution of each of the component random variables in that case.