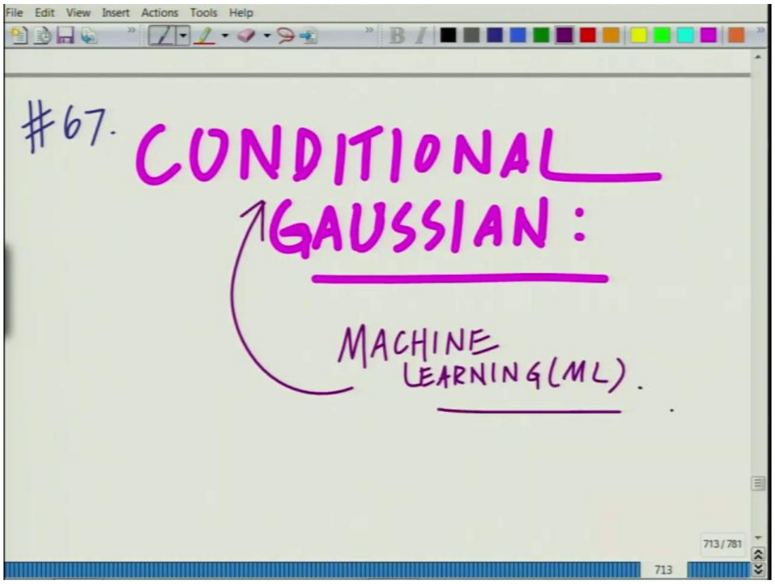
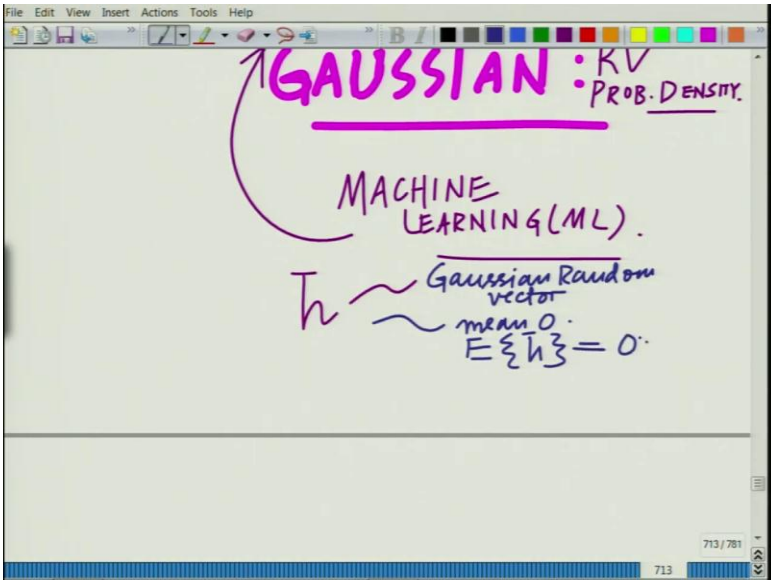
Applied Linear Algebra for Signal Processing, Data Analytics and Machine Learning
Professor Aditya K Jagannatham
Department of Electrical Engineering
Indian Institute of Technology Kanpur
Lecture 67
Conditional Gaussian Density – Mean

(Refer Slide Time: 00:34)



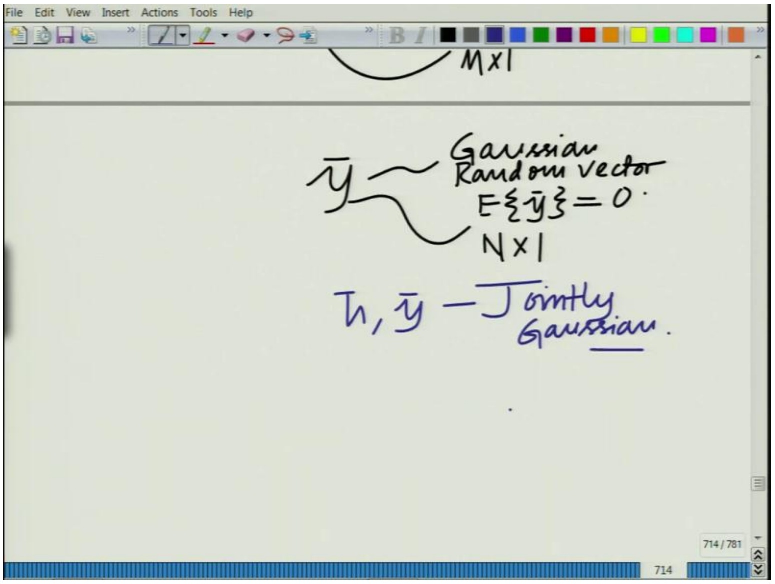
Hello, welcome to another module in this massive open online course, in this module, let us start talking about another very important and interesting concept might sound a little abstract to begin with, but I will start I will explain the motivation behind this as we go on. And this is the concept of the conditional Gaussian, this is what we call as the conditional Gaussian which arises very, this is the conditional Gaussian which again, if you are a person who is working in machine learning, data analytics and so on this arises very frequently for instance, ML is one area where this would arise significantly, this is what is called the conditional Gaussian.

(Refer Slide Time: 01:17)



You can say this is the conditional Gaussian RV or random variable or the conditional Gaussian probability density function that is the PDF and so on. Now, the genesis of this problem, let us set the stage for this problem, even before we talk about how to evaluate it. The background of this problem is as follows. So, we have a Gaussian random vector h bar. So, this is essentially your Gaussian random vector. To make things simple, let us make it mean 0, like what we usually have, we are going to talk about the non 0 mean case later, so this is expected value of h bar this equals 0.

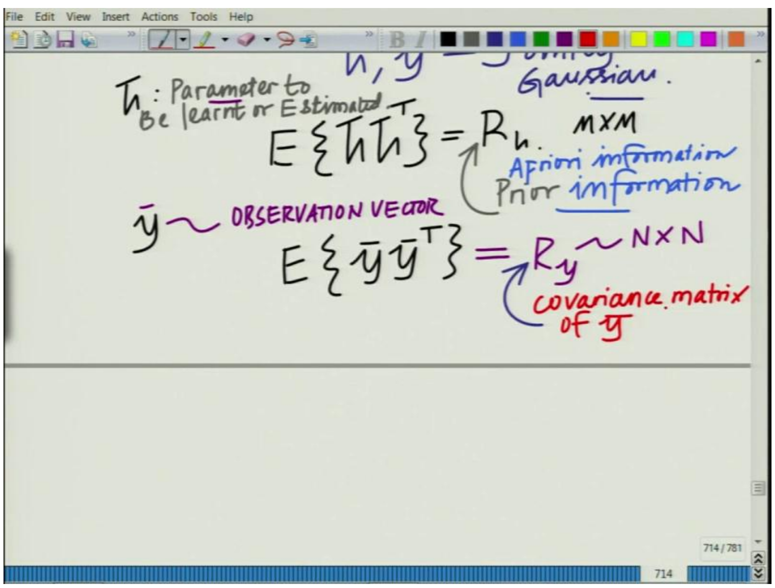
(Refer Slide Time: 02:16)

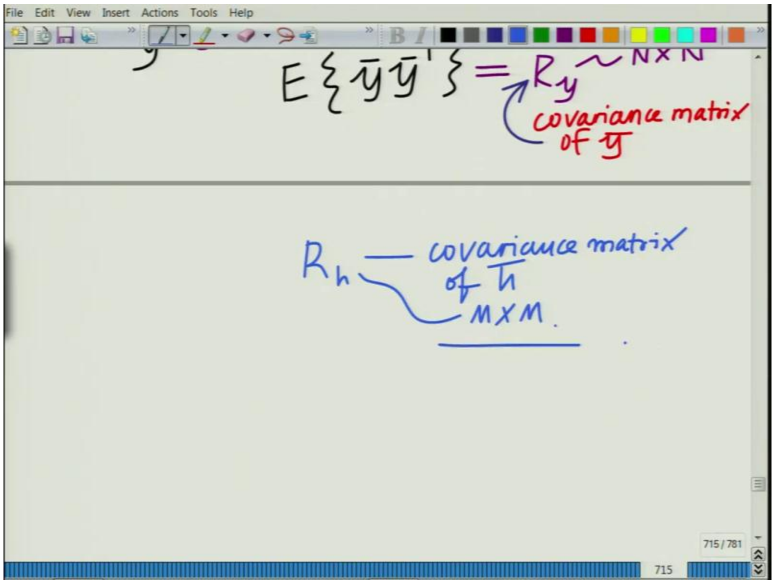


And then we have another Gaussian random vector y bar. So, this is another Gaussian random vector again, for simplicity, let us also make this 0 mean that is your expected value of y bar. So, we have 0 mean Gaussian random vector h bar, we have 0 mean Gaussian random vector y bar, let us assign some sizes to them, let us say this is an M cross 1 vector, let us say this is an N cross 1 vector.

And the point is this h bar and y bar these are jointly Gaussian, not just Gaussian, these are jointly Gaussian. So, h bar and y bar these are jointly Gaussian, multivariate random these are jointly multivariate Gaussian that is jointly Gaussian random vectors. Therefore, we have to talk also about other statistical quantities such as the covariance and the cross covariance of these quantities.

(Refer Slide Time: 03:23)



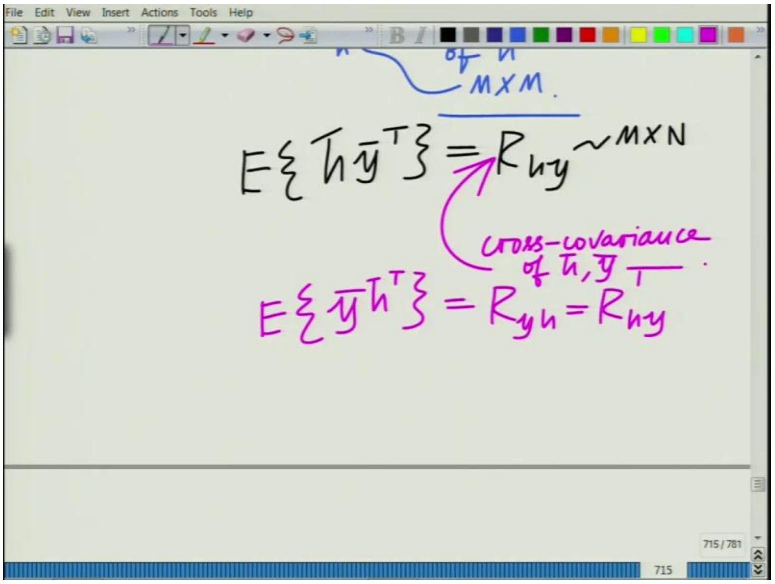


So, what is the covariance matrix? Let us define the covariance matrix expected value of h bar h bar transpose this can be defined as Rh naturally this will be an M cross M matrix. Interesting this also what we are going to term as the prior this also what in many things, such as just for instance machine learning estimation, we also call this as the prior that is eventually when we are going to talk about how to determine h bar, this is the statistical information that is available in the beginning of this process is the prior or the a priori information. So, this is also known as the prior information or this is the a priori information, something that is important in the context of machine learning, statistics and estimation and so on.

We also have the observation let us call y bar as the observation. So, this is your y bar y bar transpose. So, let us say so, this is the following nomenclature, this is the parameter vector that we have to learn, parameter to be learned or this is the parameter to be learned or essentially estimated, so that is essentially what we are talking about, but learned or estimated from where? Learned or estimated using the observations. So, this is your observation or the observation vector.

So, this is the parameter h bar this is y bar is observation vector which is 0 mean and the covariances are y and this is basically since y is N cross N y bar is N cross N this Ry this is the covariance matrix. So, this is the covariance of h bar, covariance matrix of h bar. So, this is the covariance matrix. So, I am going to write, so this is the covariance matrix of y bar and Rh this is the covariance matrix of h bar which of course, as we already said this is an M cross M matrix.

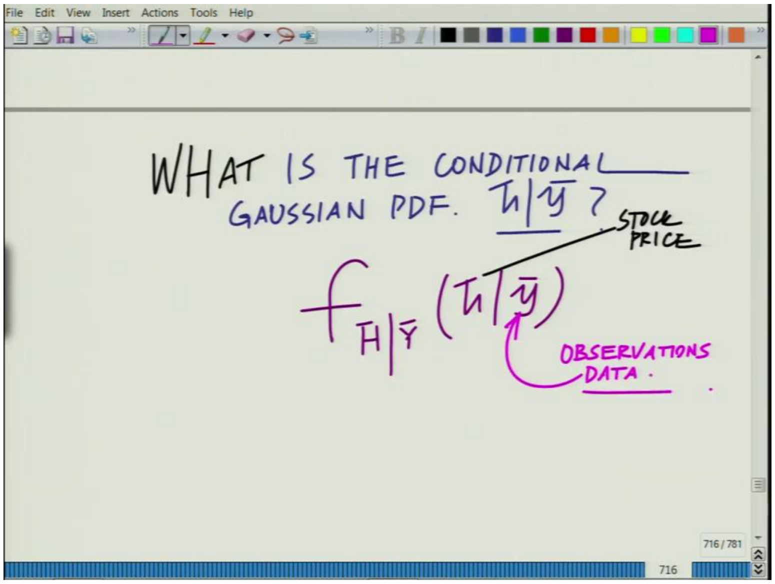
(Refer Slide Time: 06:29)

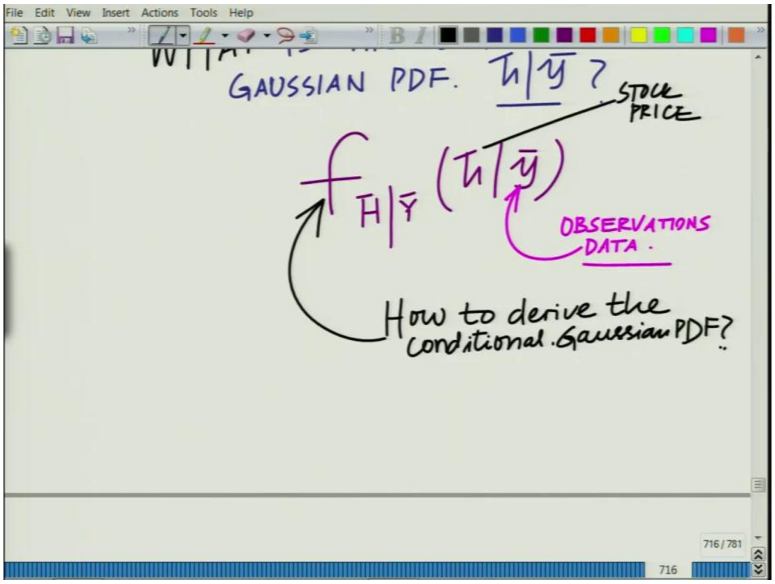


Now, we have the covariance matrix of h bar and we have the covariance matrix y bar, we need another quantity which is the cross covariance which relates y bar to h bar which basically characterizes the interplay between these random vectors that is the cross covariance matrix, that is expected value of y bar into h bar transpose or let us start with h bar into y bar transpose this is what we call as Rhy this is an M cross N matrix and this is the cross covariance of h comma y, this is the cross covariance of h bar comma y bar and similarly, you can define also the other one which is expected value of y bar h bar transpose this is Ryh which naturally this is equal to Rhy transpose, so these are the cross covariance matrices.

Now, the interesting thing, the question that we want to ask and the question that often arises in practice is now we have this random vector h bar we have this Gaussian random vector y bar, in fact jointly Gaussian random vectors h bar and y bar, what is the conditional probability density of h bar given y bar? Of course, we know the conditional probability density h bar given y bar will also be Gaussian, but what are the parameters of this conditional Gaussian probability density function? That is an interesting question.

(Refer Slide Time: 07:59)





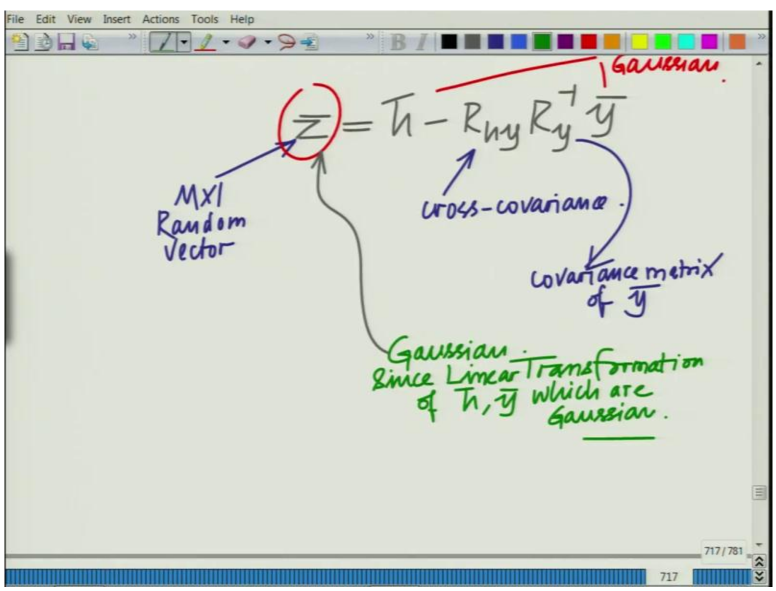
Now, what is, this is the, what is the conditional Gaussian PDF? That is h bar given y bar, what is the conditional. That is, if we were talking about the PDF, remember this is the conditional Gaussian PDF, f of H bar given Y bar, f of h bar given y bar this is a conditional, remember this is a conditional (Gaussian), this is a conditional probability density function. So, the question is, what is this conditional Gaussian PDF? That is, why is this important? This is important in machine learning and data analytics because remember, y bar is observation vector and you are trying to learn this parameter h bar that can be something such as, for instance, the stock prices, or the inventory of automobiles or basically the weather tomorrow.

So, this is h bar is something that is unknown and something that is random like a stock price and y bar basically comprise of the observations of the data that you know, how do we determine h bar given y bar, what is, how does the observation y bar affect the randomness or the pattern corresponding to h bar and we want to learn that. So, this conditional, so, this can be for instance, in your machine learning this can be random something like stock price. So, this can be your stock price and this can be for instance, something like your observations. This can be your observations or data.

So, how does this, what can you what prediction can you make about the stock price having this observation vector y bar and that is where your conditional probability density function has a very important and this basically also has a role to play in estimation like for instance you have the signal at the receiver, you observe the signal, what can you say about the channel or what can you say about the data that is essentially the concept of estimation, same thing in data analytics analysis.

So, this conditional probability density function has a very important role to play, how do we derive this. Now, to derive this, so basically the question that you want to ask is how to derive the conditional probability, how to derive the conditional Gaussian PDF and for that, we will follow the following procedure.

(Refer Slide Time: 11:01)

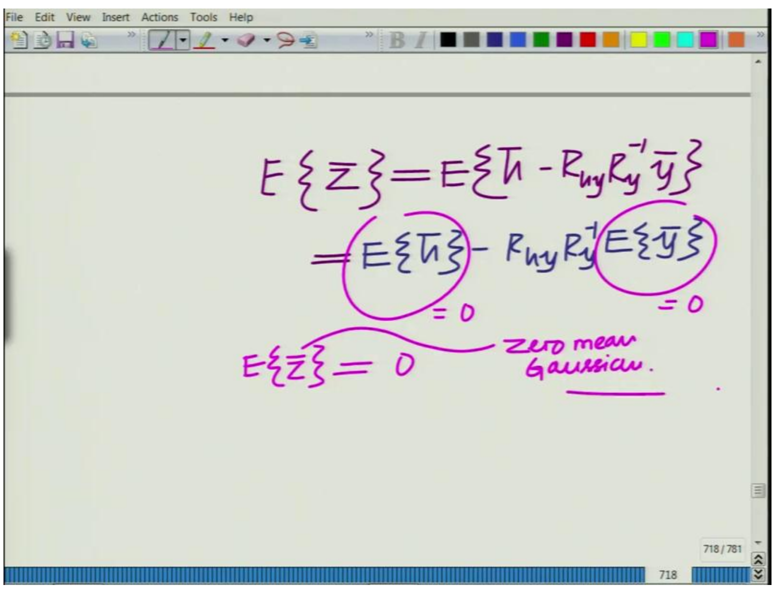


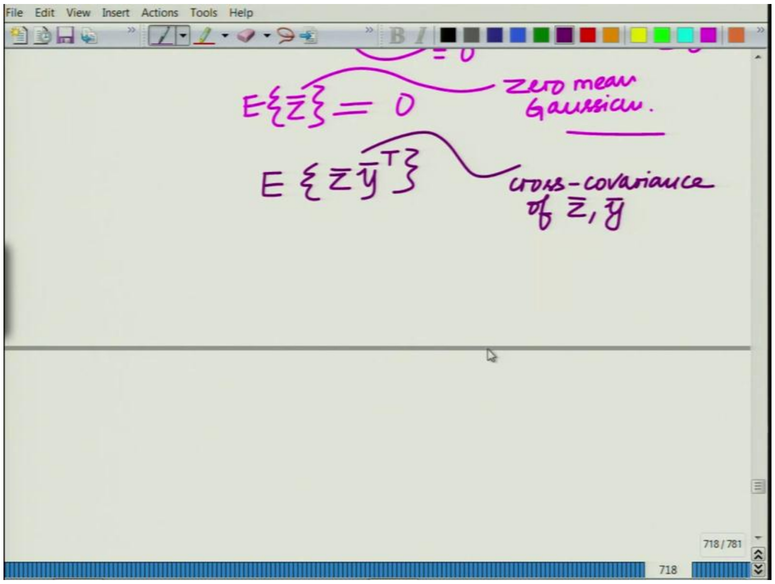
Now, let us define a new random variable to do this, let us define a new random variable z bar equals h bar minus Rhy into Ry inverse into y bar. So, this is a new random vector that we have defined, this will also naturally be M cross 1, this is a random vector.

Now, this we have already seen this is the cross covariance and this is the covariance matrix of y Ry and now what you can also see is these are Gaussian, you have the h bar y bar these are Gaussian and now z bar now naturally z bar you can see is a linear transformation of h bar and y bar. So, linear transformation of Gaussian random variables, random vectors gives another Gaussian random vector.

So, z bar is also interestingly the first major conclusion you can draw this is also Gaussian, why is this Gaussian, since linear transformation, since this is a linear transformation of h bar comma y bar which are Gaussian or jointly Gaussian, which are Gaussian, so z bar is Gaussian that is the first thing that we must have. So, gives z bar is a Gaussian random vector now, what can you say about this quantity z bar.

(Refer Slide Time: 13:17)

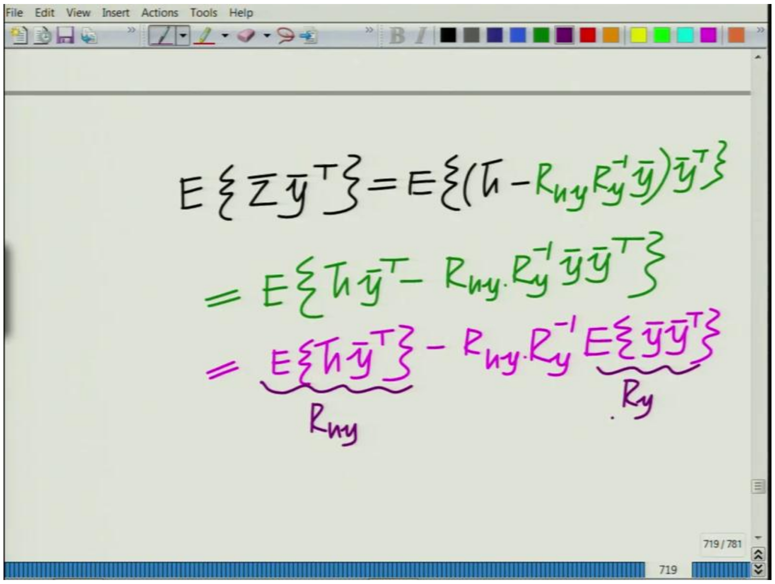




Now, let us start with the expected value or the mean of z bar, what is the expected value of z bar. So, expected value of z bar if you look at this expected value of z bar this is expected value of well h bar minus Rhy Ry inverse y bar this is the expected value of h bar minus Rhy into Ry inverse expected value of y bar and this is equal to the expected value of h bar this we can say this we are assumed these quantities to be 0 mean, so this is 0, this is equal to 0.

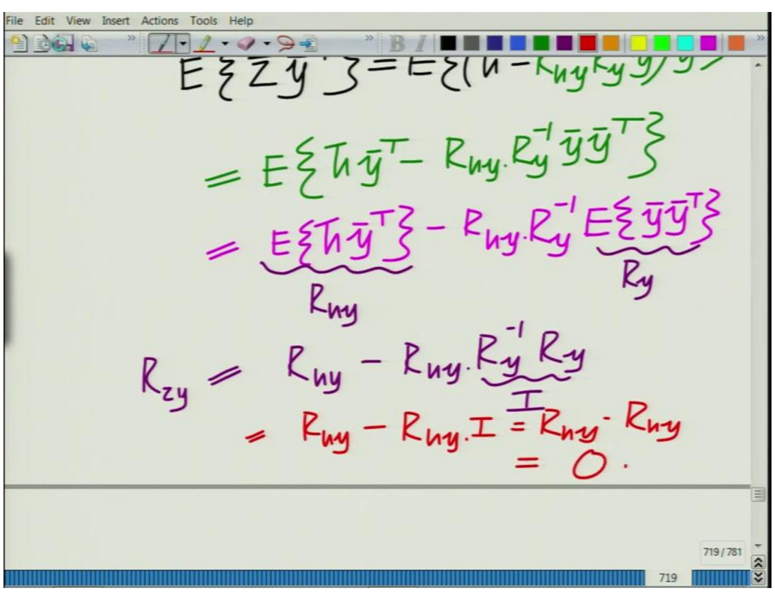
So, this is also equal to 0 expected value of z bar this is also equal to 0, so z bar is 0 mean Gaussian. So, z bar is also a 0 mean Gaussian random vector, that is what we have established about this quantity is a first of all it is Gaussian because it is a linear transformation of h bar and y bar and now the expected value is also 0. Now comes the interesting question, what can we say about the expected value of z bar y bar transpose that is the cross covariance? What can we say about the cross covariance of z bar y bar?

(Refer Slide Time: 15:14)



And this will be, if I can simplify this, not very difficult to derive that is expected value of we have the expected value of z bar y bar transpose, which is the expected value of we have the z bar which is your h bar minus Rhy Ry inverse y bar times y bar transpose which is equal to the expected value of h bar y bar transpose minus Rhy into Ry inverse into y bar y bar transpose which is equal to the expected value of h bar y bar y bar transpose minus Rhy into Ry inverse times the expected value of y bar y bar transpose this quantity is Rhy, this quantity is Ry the covariance matrix of y.

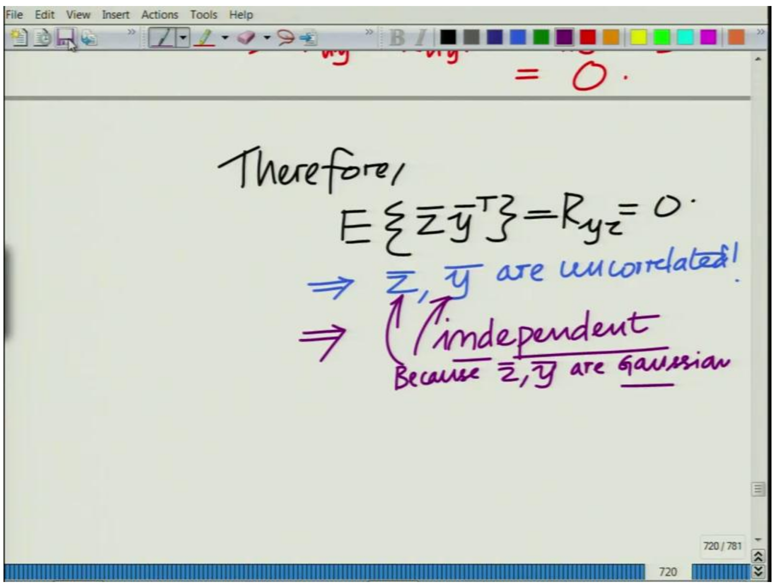
(Refer Slide Time: 16:35)

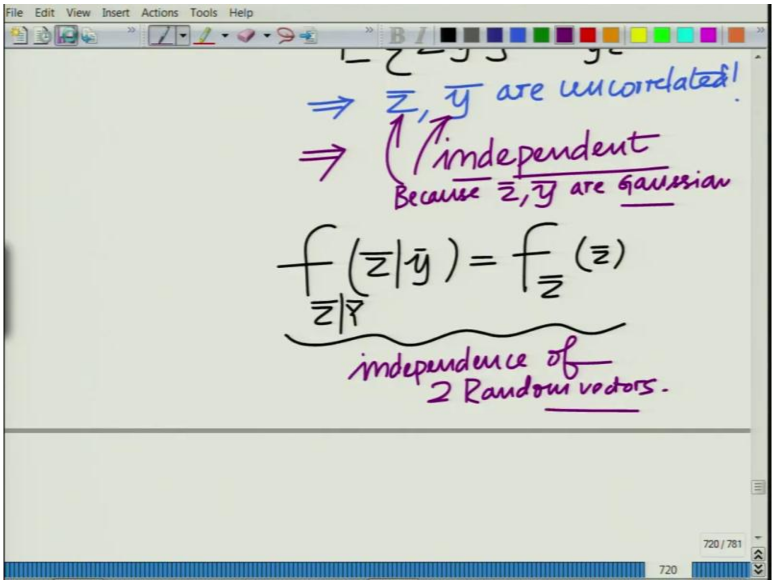


Therefore, you will observe something very interesting, if you call this as Rzy the cross covariance. This is Rhy minus Rhy into Ry inverse into Ry of course, this quantity is identity. So, essentially what you get is this is Rhy minus Rhy into I which is needless to say Rhy minus Rhy this is equal to 0. Therefore, the interesting thing that you observe is that the cross covariance expected value of z bar into y bar transpose that is equal to 0.

That has an interesting implication of z bar is Gaussian, y bar is Gaussian and these cross covariance is 0 which implies that these are uncorrelated. Now, remember for Gaussian, if the two random variables are uncorrelated two random vectors are uncorrelated they are essential, they are basically independent that is a conclusion that you can make only for the Gaussian.

(Refer Slide Time: 17:46)





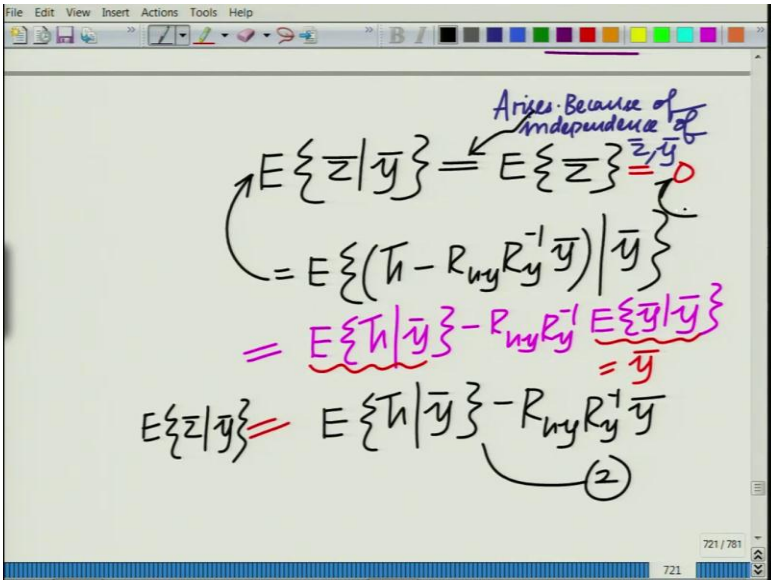
Therefore, very interestingly, therefore, expected value of z bar y bar transpose equal to Ryz equal to 0 what this basically implies is that z bar y bar are uncorrelated. What this implies as z bar y bar are uncorrelated and this implies that z bar y bar these are independent now, this is not true for all this implies that these are independent because z bar comma y bar are Gaussian.

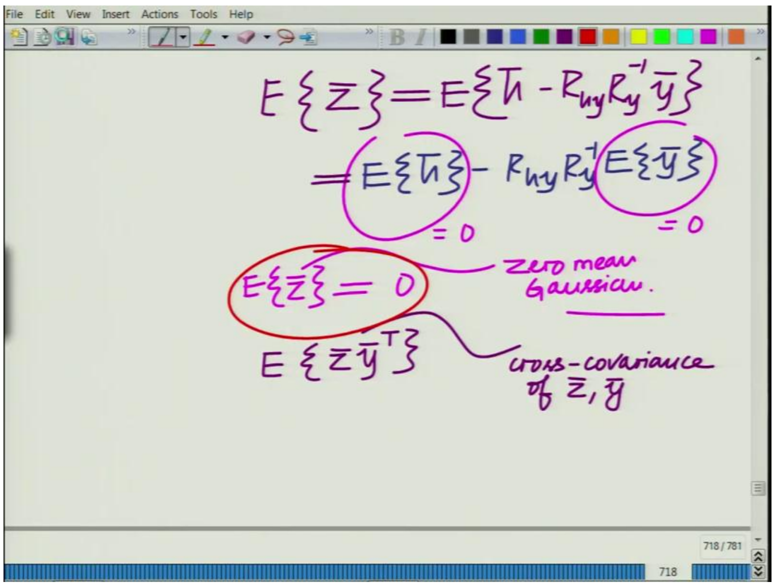
This is not always true, it is not always the case, but because z bar and y bar are Gaussian random vectors, since z bar and y bar are uncorrelated, this also implies that z bar and y bar are actually independent which means that the PDF of that is the conditional PDF of z bar given y bar the technical definition for this is z bar given y bar equals the conditional PDF of z bar. This is the definition of independence of 2 random vectors.

This is the technical definition for the independence of 2 random vector. So, because these are uncorrelated, these are also independent because these are uncorrelated. These are essential also independent. Now, remember, but remember, this is not our goal, our goal is not to determine the conditional probability density function of z bar given y bar, but this is only a tool that we implying to determine the conditional PDF of h bar given y bar.

So, let us not lose sight of that, that is our eventual goal, to determine the expected value of covariance and we know it is Gaussian, to determine, wanted to determine the expected value and the covariance of h bar given y bar that is the conditional PDF we want to derive. And we are using z bar as a tool for that.

(Refer Slide Time: 20:20)



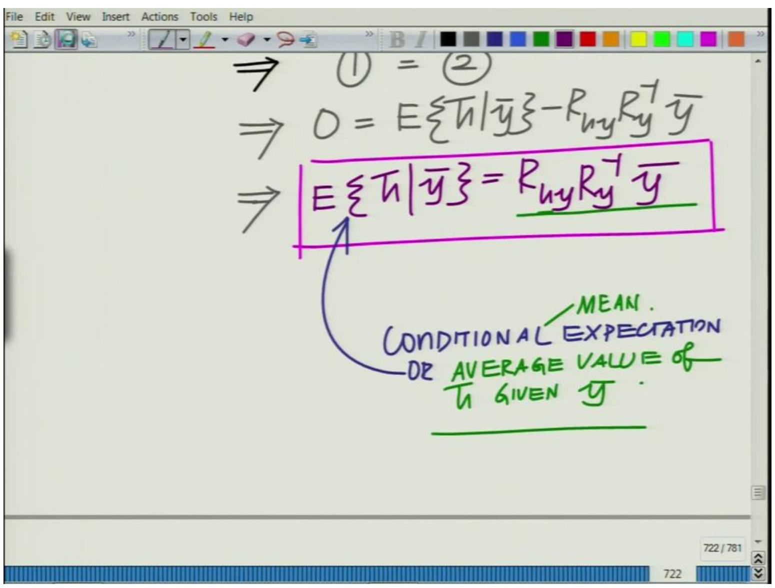


Now, let us start with that to (pro), so we have established that z bar and y bar are uncorrelated and since they are Gaussian they are independent. Now, let us proceed towards our goal. Now, let us start with expected value of z bar given y bar. Now, this is equal to expected value of simply z bar. Why does this arise? Because z bar is independent of y bar. So, the PDF is basically PDF of z bar given y bar is simply the PDF of z bar which implies that the expected value of z bar given y bar is simply the expected value of z bar that is average over the PDF of z bar.

So, this arises because of independence of z bar given y bar, but on the other hand this is also equal to expected value of substitute for z bar that is h bar minus Rhy into Ry inverse into y bar given y bar which is also equal to the expected value of h bar given y bar minus Rhy into Ry inverse expected value of y bar given y bar. By the way, expected value of z bar is 0, we already know that because we have established over here, you go all the way back. And you look at this, we have established that expected value of z bar is 0. So, z bar is 0 mean Gaussian already we have already established that expected value of z bar is 0.

Now, the point is we are trying to simplify what is expected of z bar. Now, this is expected conditional expectation of h bar. Now, look at this expected value of y bar given y bar is simply equal to y bar. Because given y bar, if you ask the question, what is the PDF or what is the expected value of y bar that is simply y bar, because y bar is already given. So, now that means that this quantity is equal to expected value of h bar given y bar minus Rhy Ry inverse y bar, let us call this as our result 2, this is the same quantity expected value of z bar given y bar. But expected value of z bar given y bar is equal to 0, we call this is from the result 1.

(Refer Slide Time: 23:24)



So, this implies that the 1 result conclusion in 1 is equal to equating 1 and 2, not the numbers, but the results, this simply implies that 0 equals expected value of h bar given y bar minus Rhy into Ry inverse into y bar, which basically implies now that the conditional expectation expected value of h bar given y bar equals Rhy into Ry inverse into y bar. So, this is an important result.

This is the conditional expectation of h bar. And this has many consequences very many applications.

As I am going to describe later, so this is the conditional, the average value the conditional expectation or you can say average value of x h bar given y bar having the observation, average or the mean value, the mean the conditional mean, conditional expectation or the conditional mean the average value of h bar given however.

So, naturally you can see this is a good answer to the question having observed y bar given the observation y bar, what can you conclude about h bar, what is the mean of h bar, what is the average value of h bar and this is a fantastic result which tells us that having observed y bar the expected value of h bar given y bar the average value of h bar given y bar is given by this quantity that is Rhy into Ry inverse into y bar.

And as you can already see, this is a good estimate for the covariance matrix, this is a good estimate for the h bar, for the parameter h bar that is having observed y bar if someone asked you what is that is having observed this y bar what is a good estimate for the stock prices in h bar, this would be your answer Rhy into Ry inverse into y bar that is an estimate of the stock prices.

So, this is can provide you a lot of can as you can already see is can be very useful in many scenarios in learning and analysis and as I have already told you has several applications alright. So, this is the conditional expectation, now, of course, that does not complete, we also have to answer, we also have to determine because it is Gaussian, for Gaussian character is a Gaussian we need the mean and the expected value and also the variance the covariance matrix.

Now, we also have to characterize what is the covariance of this quantity h bar given y bar and we will do that in the subsequent modules. So we will stop here, where we have derived the conditional mean, we will derive the conditional covariance and then also extend it to non-zero mean quantities and then infer valuable insights from this process. Thank you very much.