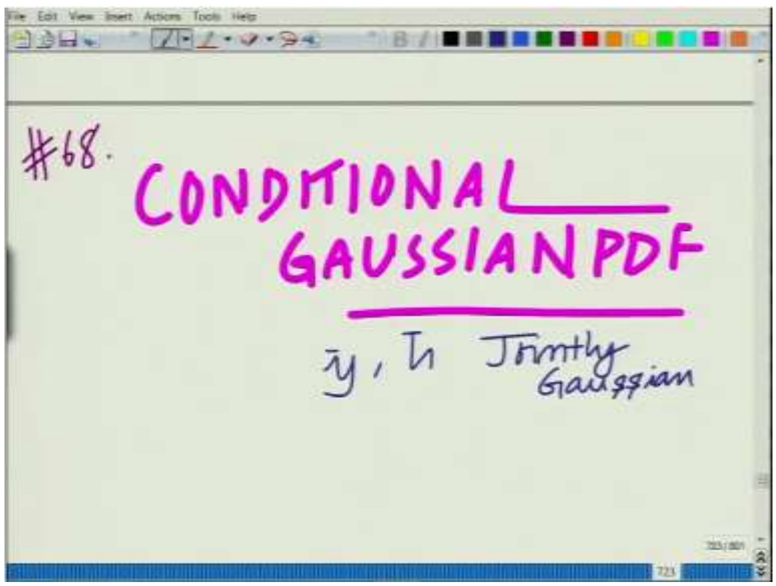
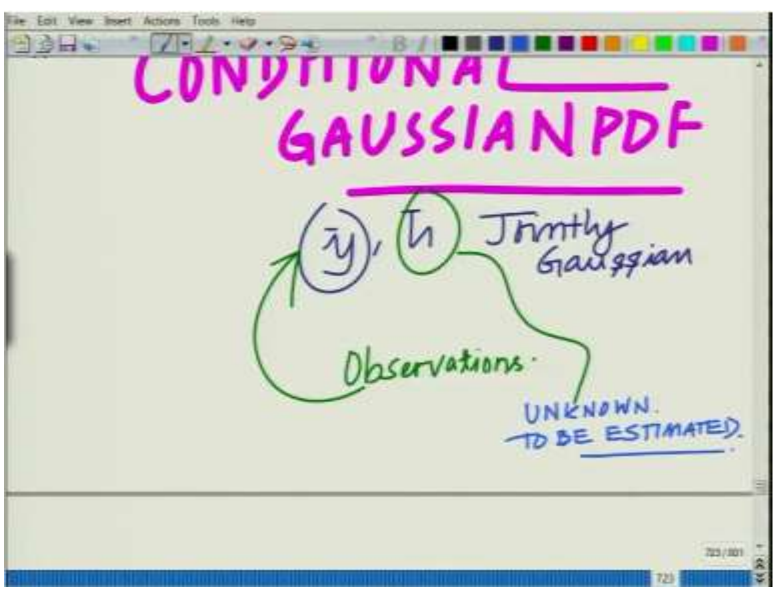
Applied Linear Algebra for Signal Processing, Data Analytics and Machin Learning
Professor. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur
Lecture No. 68
Conditional Gaussian Density – Covariance

Hello, welcome to another module in this massive open online course. So, we are looking at the conditional Gaussian probability density function and let us continue our discussion.

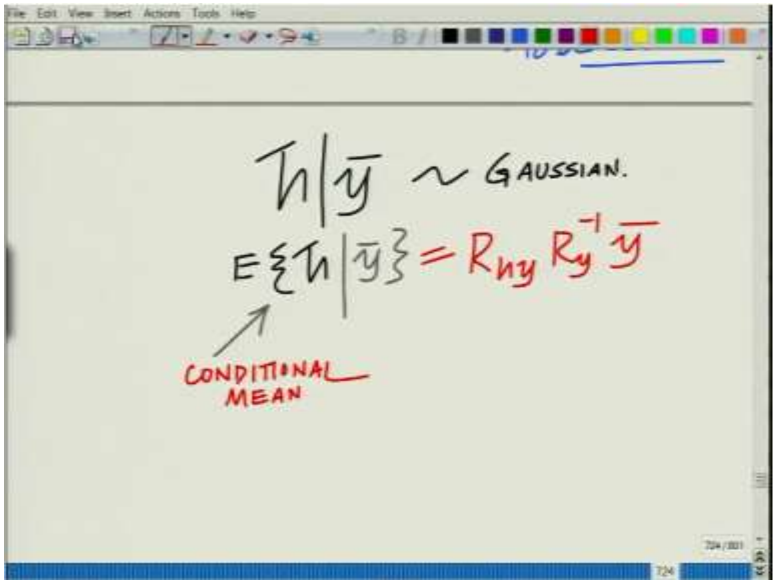
(Refer Slide Time: 0:24)

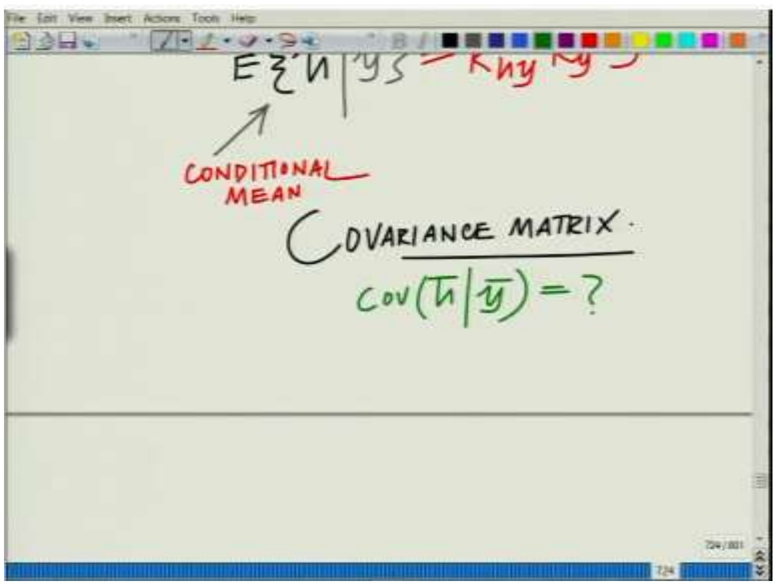




So, one of the important concepts or principles of linear algebra is the Conditional Gaussian, the Conditional Gaussian PDF and we are specifically looking at a scenario where you have two quantities y bar, h bar these are jointly Gaussian, typically you would like to consider a scenario where for instance y bar is an observation. So, we have a set of observations and h bar is an unknown to be estimated, an unknown quantity, this quantity is to be, this is an unknown quantity that is to be estimated.

(Refer Slide Time: 1:44)

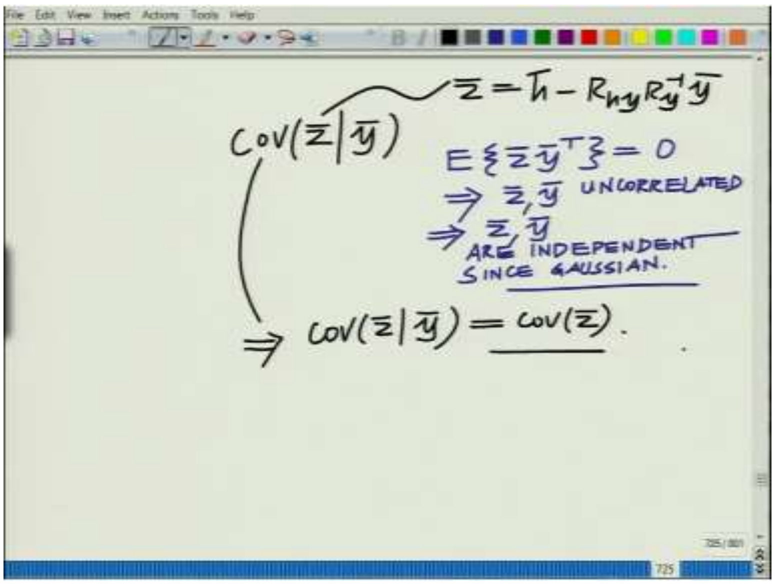




And we have seen that the conditional mean that is we have y bar or we have this probability density function, conditional probability density function h bar given y bar this is Gaussian, in fact the conditional mean that is what we established last time, if you go back and take a look at it is that expected value of h bar given y bar, this is what we call as the conditional mean, conditional mean of h bar conditioned on y bar. This has an interesting structure, this is Rhy, cross covariance of h and y times Ry inverse into y bar this is the conditional mean.

Now, in order to determine the PDF we also have to determine the covariance of h power given y bar. So, you ask the question what is remaining is the covariance, in particular the conditional covariance, the covariance matrix, in particular the conditional covariance, so we have this denoted by this covariance of h bar given y bar, what is this quantity?

(Refer Slide Time: 3:20)

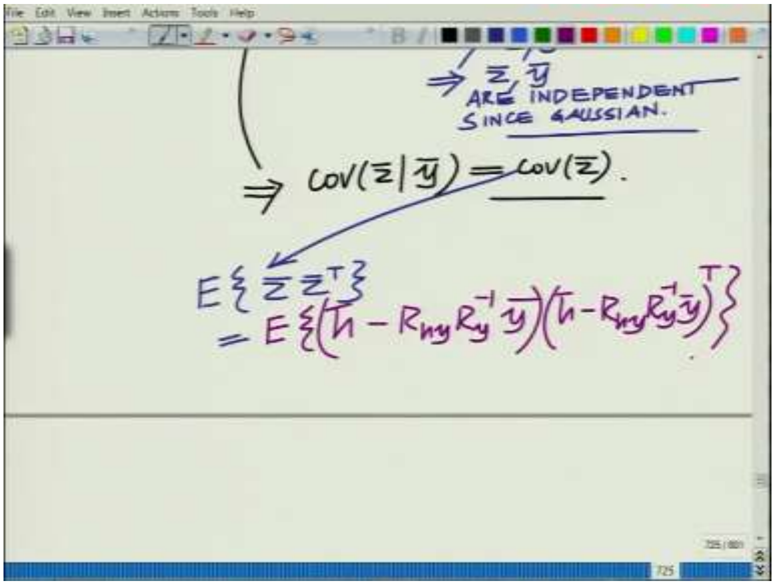


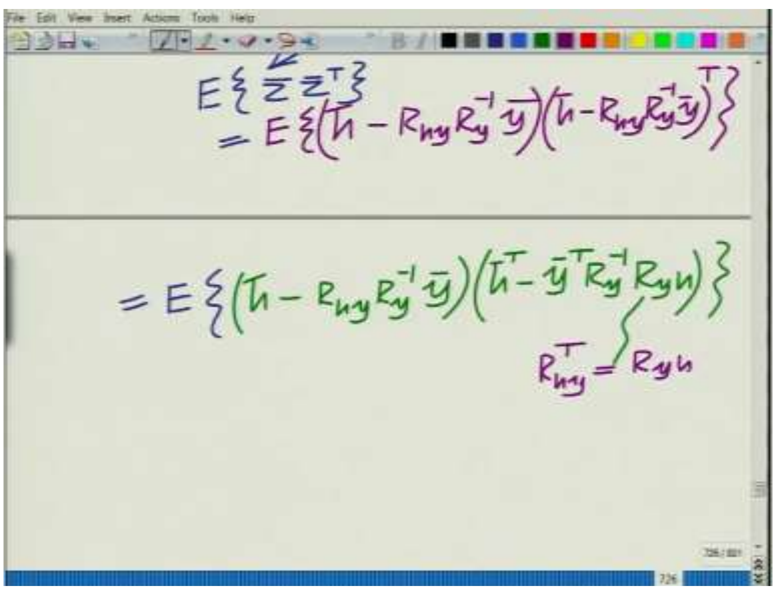
Now, to do that let us start first once again with the covariance of this quantity z bar given y bar, where you remember z bar is something we had defined as h bar minus Rhy into Ry inverse into y bar. So, we had constructed this vector z bar, remember to aid us in this derivation, so z bar equals h bar minus Rhy Ry inverse into y bar.

Now, we ask the question what is the covariance of this, now the other thing that we had seen is that expected value of z bar y bar transpose this equal to 0, which implies these are uncorrelated, implies z bar comma y bar these are uncorrelated and this also implies z bar comma y bar are independent, these are independent, z bar y bar are independent you can also say these are independent since they are Gaussian.

And therefore, now what does that mean and therefore, now this now since z bar and y bar these are independent, this essentially implies covariance of z bar given y bar this is simply equal to the covariance of z bar, because z bar and y bar are independent. So, probability density function of z bar given y bar is simply the probability density function of z bar. Therefore, the covariance of z bar given y bar simply the covariance of z bar, because z bar and y bar are essentially independent.

(Refer Slide Time: 5:36)

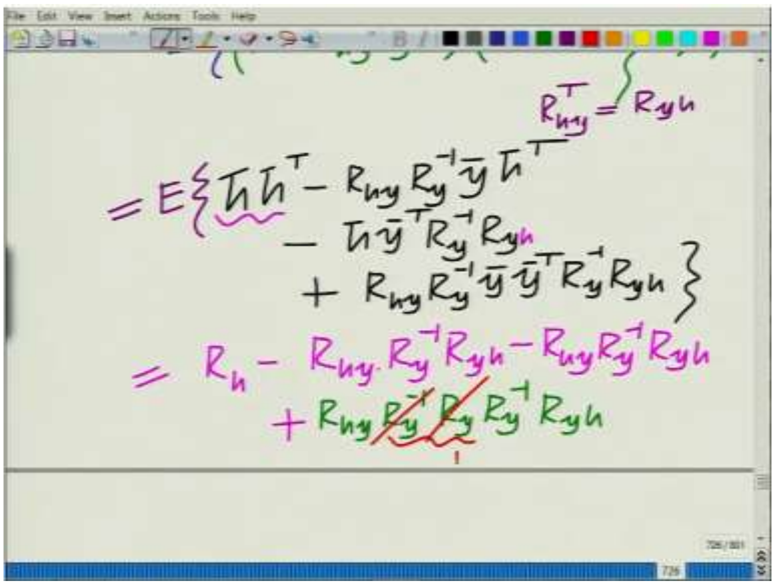




Now, we simplify this covariance of z bar, this you can derive as covariance, remember z bar is a 0 mean random variable. So, covariance of z bar is expected value of z bar z bar transpose which if you simplify it use the expression for z bar this expected value of, you have your h bar minus Rhy, h bar minus Rhy Ry inverse y bar into h bar minus Rhy Ry inverse y bar transpose which is equal to expected value of, let me simplify this, h bar minus Rhy Ry inverse y bar into h bar transpose minus Rhy or taking the transpose of this quantity.

This becomes y bar transpose Ry inverse transpose is Ry inverse, because Ry is a symmetric matrix and Rhy transpose is Ryh. So, we use the property here that Rhy transpose is nothing but Ryh, this is something that we had established before, this is easy to see because Rhy is nothing but expected of hy transpose and Ryh is expected of y bar h bar transpose and these are transpose of each other.

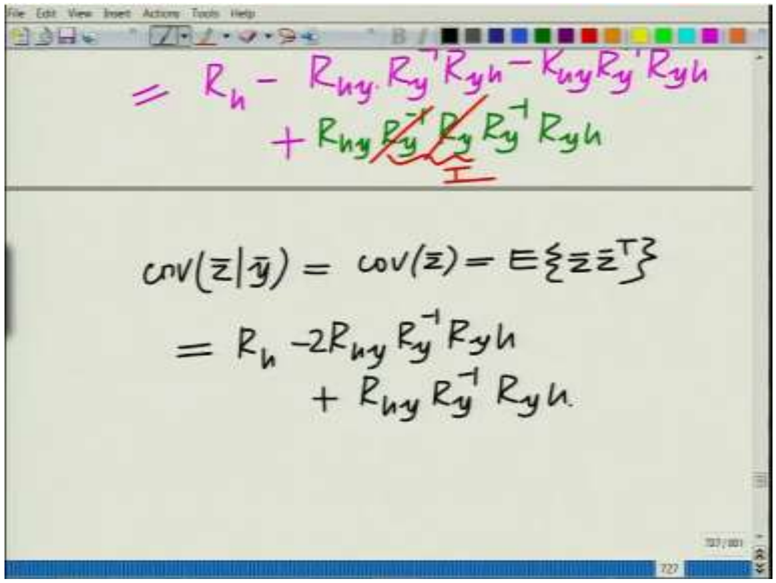
(Refer Slide Time: 7:28)



Therefore, now if we simplify this quantity you will have, this is the expected value of, well writing a term by term h bar h bar transpose minus Rhy Ry inverse y bar h bar transpose minus h bar y bar transpose times Ry inverse into Ryh into Ryh plus Rhy into Ry inverse into y bar into y bar transpose Ry inverse Ryh and this is equal to, take the expected value of h bar h bar transpose this is Rh minus Rhy into Ry inverse into expected value of y bar x bar transpose.

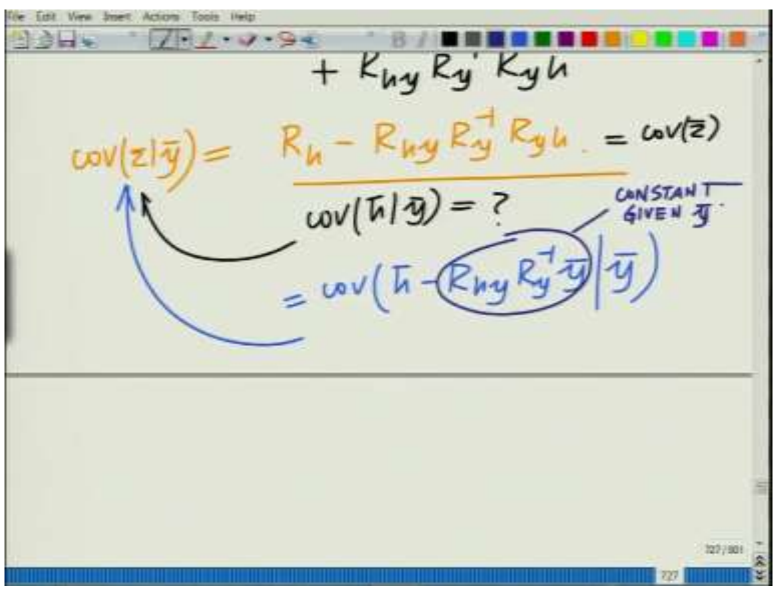
This is Ryh minus expected of h bar y bar transpose, this is Rhy into Ry inverse, I am sorry, this has to be Ryh subscript h, Ryh plus you have Rhy Ry inverse expected value of y bar y bar transpose this is Ry into Ry inverse into Ryh and needless to say this Ry inverse into Ry this is identity, so this becomes your, these things cancel so this becomes identity.

(Refer Slide Time: 9:29)



And therefore, you are left with the result that the covariance matrix of z bar given y bar equals the covariance of z bar which is essentially since this z bar is 0 mean, this is z bar z bar transpose, this is Rh minus, you can see Ry inverse Ryh minus twice this plus again the same quantity Rhy Ry inverse Ryh.

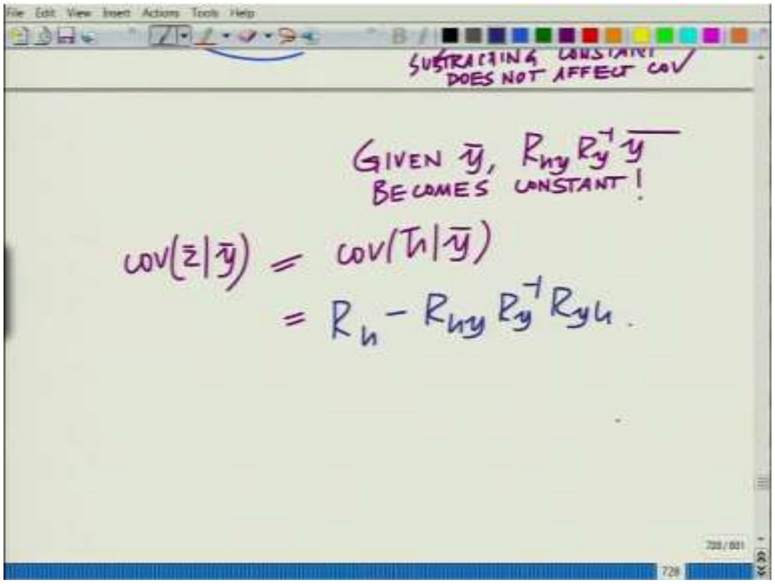
(Refer Slide Time: 10:14)



And now if you simplify this you get this is essentially Rh minus Rhy Ry inverse Ryh, so this is the covariance of z bar given y bar, which is very interesting or which is essentially also the covariance of z bar. Now, remember ultimately we want to ask the question what is the covariance of h bar given y bar, this is the thing that we set out to derive and if you look at this we substitute z bar, we substitute for the vector z bar here which is z bar equals.

Remember, h bar minus Rhy into Ry inverse into y bar so this is equal to the covariance of h bar minus Rhy Ry inverse y bar given y bar, but now look at this, the covariance of now this quantity given y bar, this quantity is a constant, this quantity is a constant given y bar.

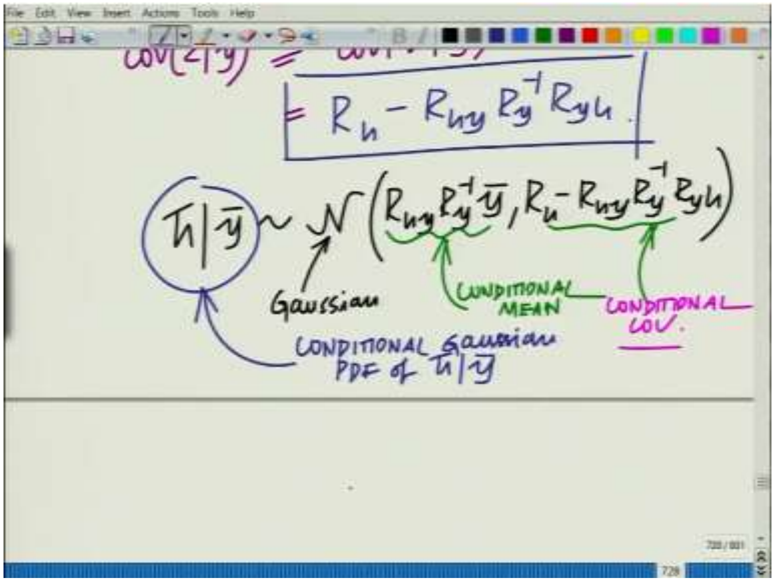
(Refer Slide Time: 11:44)



So, given y bar, given y bar you have Rhy Ry inverse y bar this quantity becomes constant and therefore, this is equal to simply now you are subtracting a constant from h bar. Now, subtracting a constant does not affect the covariance or variance therefore, this is also equal to the covariance of h bar given y bar.

So, therefore, since subtracting a constant does not affect the covariance, so subtracting constant does not affect constant, so this is equal to the covariance of h bar given y bar. So, what we are saying prude is the covariance of z bar given y bar equals covariance of h bar given y bar which now you can write as Rh minus Rhy into Ry inverse Ryh this is a essential.

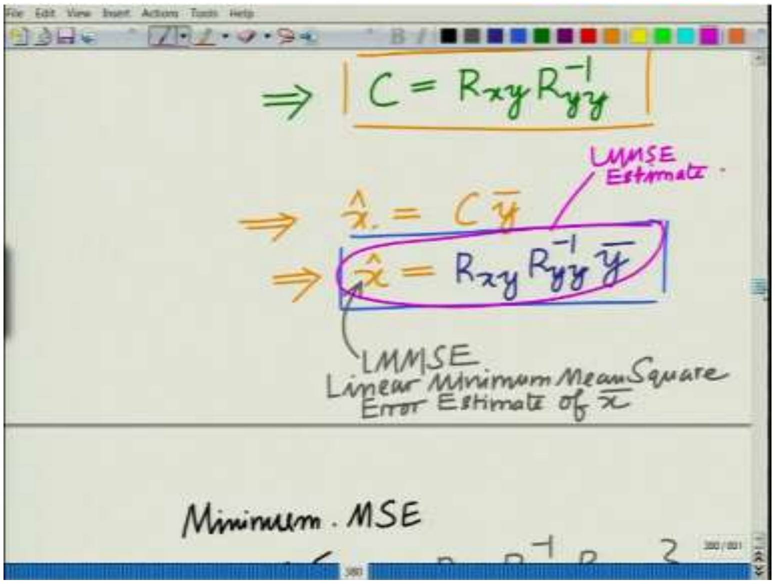
(Refer Slide Time: 13:25)



And therefore, now you can write h bar given y bar this is Gaussian, remember this shows that the PDF is Gaussian or multivariate Gaussian, the mean is Rhy Ry inverse y bar, this is the conditional mean and the conditional covariance is Rh minus Rhy into Ry inverse into Ry. This is the conditional Gaussian PDF, so this is the conditional Gaussian PDF of h bar given y bar, this is the conditional Gaussian PDF of h bar given y bar, this is the conditional mean, this is the conditional covariance, this is the conditional mean, this is a conditional covariance.

Now, if you look at it, now if you go if you think about it there is something that is very interesting, there is something that we have already seen, my claim is there is something that we already seen, if you go all the way back and take a look at what is known as this principle of this LMMSE estimation, the Linear Minimum Mean Square Error Estimate, if you go all the way back.

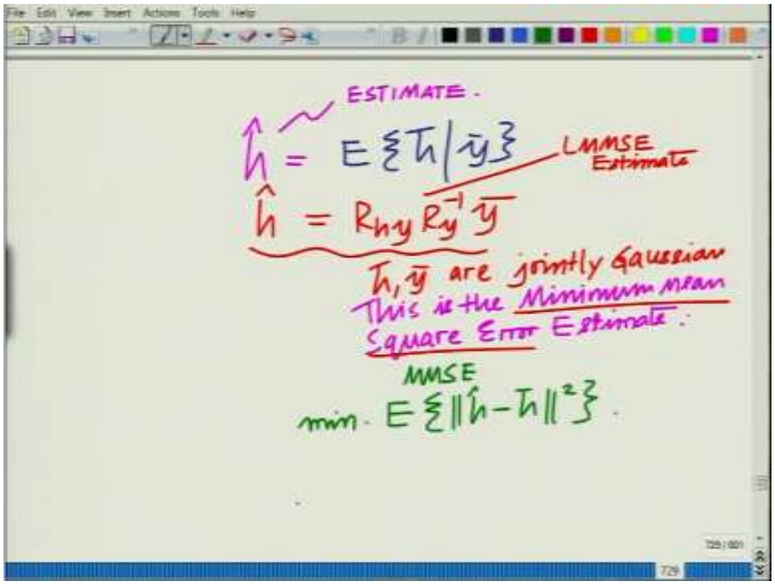
(Refer Slide Time: 15:16)

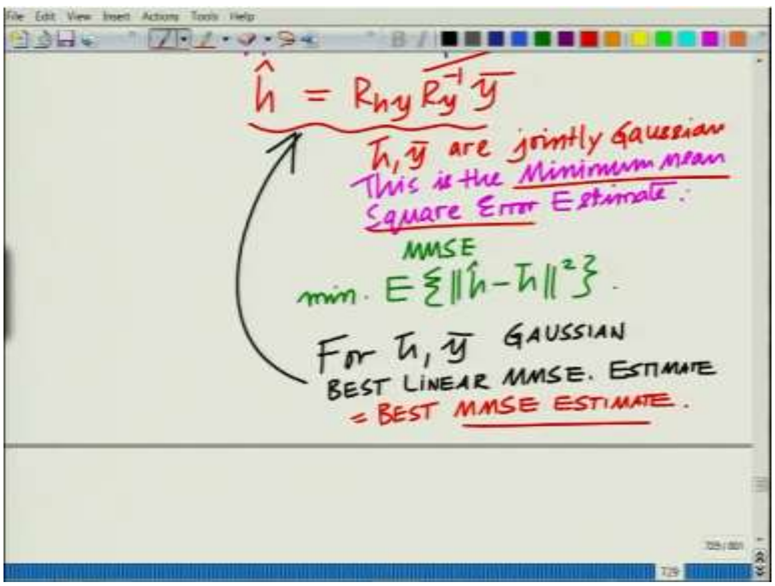


Let me just refresh your memory a little bit, you go all the way back and you look at this quantity, if you look at this quantity this is something that you should have already seen which is essentially that ah this thing that is if you look at the LMMSE estimate, this is of x bar given y bar, that is essentially same thing we have replaced x by h, this essentially is your LMMSE estimate.

And now if you look at this you will exactly get the same kind of expression what you are observing here, is that if you denote this by what is known as the MMSE, this quantity is what is known as the MMSE estimate, that is the Minimum Mean Square Error estimate.

(Refer Slide Time: 16:12)





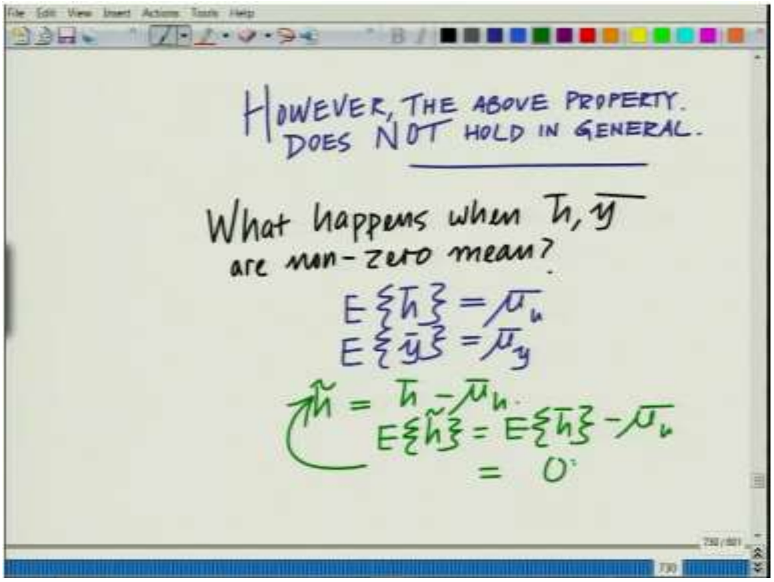
That is it can be shown that if you set the estimate, remember h hat is unknown, if you set the estimate of h hat, remember h bar is an unknown quantity, we are trying to estimate h bar after observing y bar and therefore, one way to derive this estimate of h bar is to set it as the conditional mean of h bar given y bar, that is set it as expected value of h bar given y bar which essentially is the Rhy Ry inverse y bar which is Rhy into Ry inverse into y bar and which is exactly as the LMMSE estimate.

So, which is exactly as the LMMSE estimate and this quantity here when h and y bar are jointly Gaussian this becomes the LMMSE estimate, so when h bar, this is the MMSE estimate, that is the minimum mean, MMSE by MMSE mean this is the minimum mean squared error estimate. So, this is becomes the MMSE estimate, that is it minimizes, that is if you ask the quantity what is the estimate h hat such that h hat minus x bar norm square, the mean of that, the mean square error is minimum that is this MMSe estimate.

So, what you see is that the MMSE estimate is equal to the LMMSE estimate when the h bar and y bar are jointly Gaussian, which is what we already seen, which the LMMSE estimate is something that we already seen, what we are seeing now is something new that is the MMSE estimate and much harder to derive, but for the Gaussian the MMSE estimate is equal to the LMMSE estimate, that is the best linear estimate, that is the best linear estimate itself is the best estimate however, this does not hold for other PDF.

So, something important to note here is for Gaussian, that is for Gaussian for h bar y bar Gaussian the best that is LMMSE, that is the best linear estimate, best linear MMSE estimate equals the best MMSE estimate. However, this is an important property of the Gaussians, this is a very important property for Gaussian random vectors however, this does not hold for others.

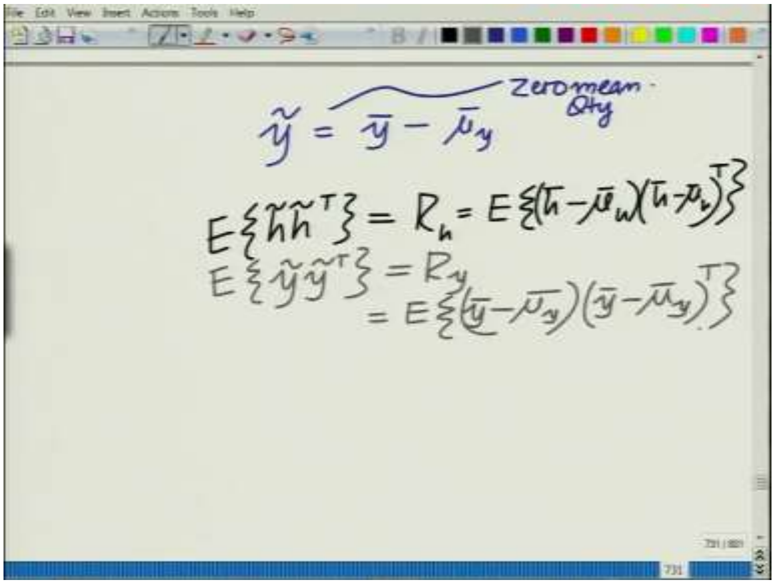
(Refer Slide Time: 20:01)

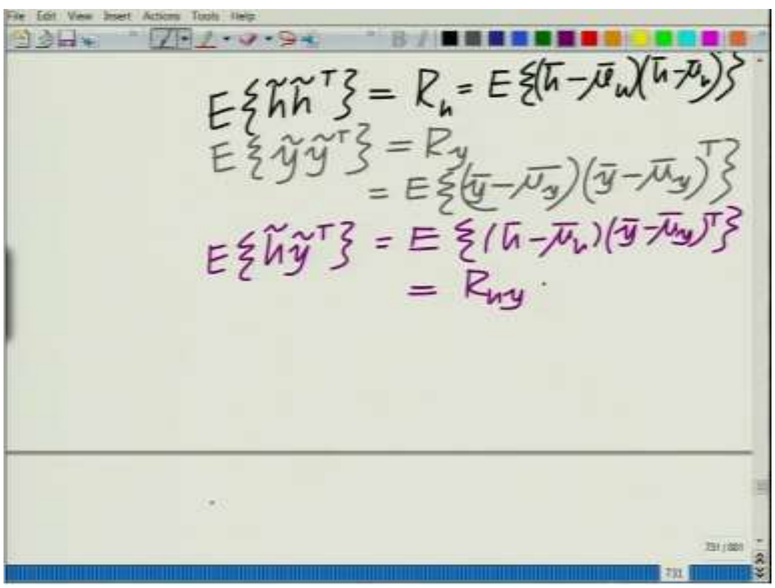


In general however, this is a very interesting property however, the above property does not hold in general, this something where is very important, that is the best LMMSE estimate need not be the best MMSE estimate, this only holds when it is a Gaussian.

Now, let us consider what happens when you have 0 mean, I think we have already seen this but we can also consider this once more, what happens when h bar comma y bar are non 0 mean? Well, very simple, if they are non 0 mean we define these new quantities, let us say expected value of h bar equal to h mu h bar and expected value of y bar equal to mu y bar then we define these new quantities, we define h tilde equals h bar minus mu bar h and now observe that expected value of h tilde, this tilde quantity equals expected value of h bar minus mu h bar which is equal to 0.

(Refer Slide Time: 22:03)

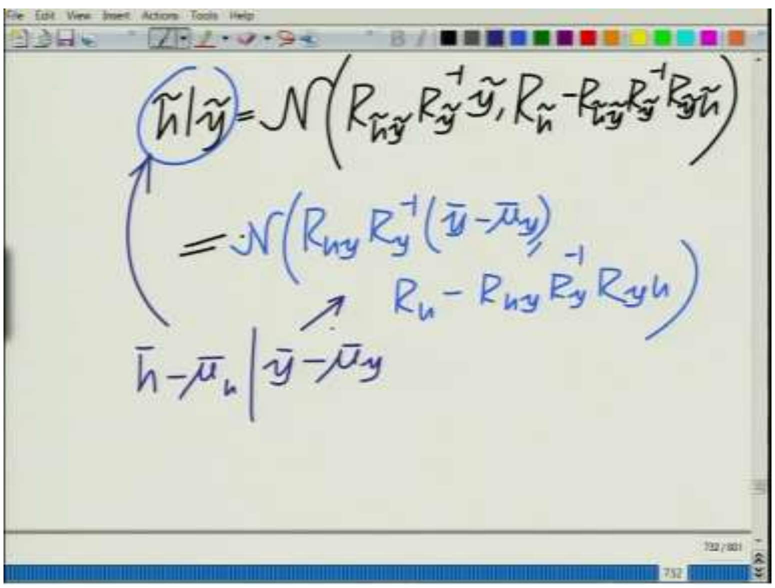


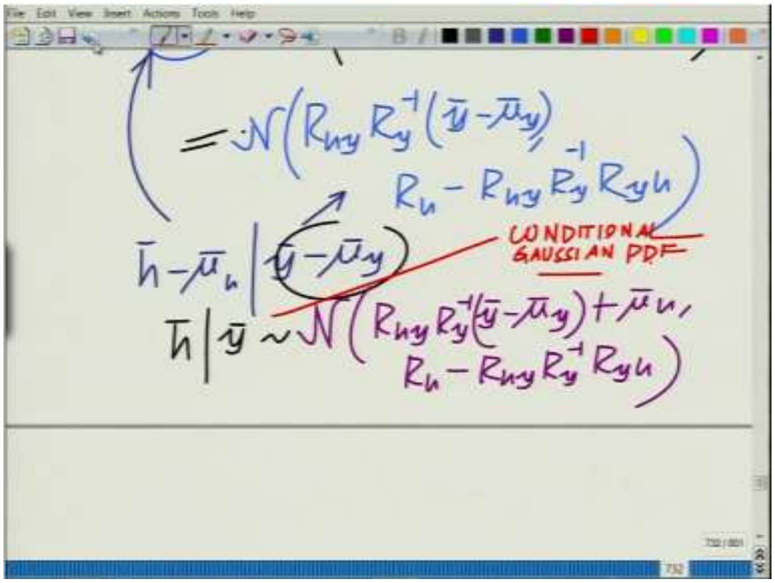


Similarly, we define these tilde quantity y tilde equals y bar minus the expected value of y bar that is mu y bar and this is also another zero mean quantity. Now, we can estimate in terms of this tilde quantities now we say expected value of h tilde h tilde transpose let us say this is equal to its still denoted by Rh, which is Rh which is nothing but expected value of because this is nothing but expected value of Rh is now nothing but expected value of mu bar h bar minus mu bar h into h bar minus mu bar h transpose, because the shifting by the mean that does not affect the covariance. So, covariance of h bar and h tilde will still be the same.

Similarly, the covariance of y bar and y tilde that is expected value of, y tilde y tilde transpose equals Ry equals expected value of y bar minus mu bar y transpose y bar minus mu, I am sorry, y bar minus mu y into y bar minus mu y by transpose. And similarly, you have your Rhy that will become expected value of are h tilde y tilde transpose equals are h bar minus mu bar h into y bar minus mu bar y transpose this is equal to your Rhy.

(Refer Slide Time: 24:00)





And now you put these things together now you will have the conditional PDF of h tilde given y tilde this is basically a Gaussian with mean you will have Rh tilde y tilde into Ry tilde inverse into y tilde into the conditional covariance that is Rh tilde minus Rh tilde y tilde into Ry tilde inverse into Ry tilde h tilde, which again if you replace these quantities by their logical equivalents it is not very difficult to see that this is kind Gaussian.

Because Rh tilde y tilde is Rhy this is Ry inverse and this one becomes y bar minus mu bar y and this quantity the other quantity that is the covariance becomes Rh minus Rhy becomes Ry inverse into Ryh and this is h tilde given y tilde but h tilde, remember what is h tilde? h tilde is simply your h bar minus mu bar h conditioned on y bar minus mu bar y this has this PDF.

So, you are simply subtracting which means h bar, to get the PDF of h bar all you have to do or h bar given y bar minus mu y bar that is y bar minus mu y bar or h bar given simply y bar, because y bar and y bar minus mu r that is PDF of h bar given y bar is the same as the PDF of h bar given y bar minus mu y bar because these are y bar and y bar minus mu y bar simply differ by the shift.

So, this becomes now all you have to do is you have to add the mean to this. So, h bar minus mu h bar as this mean, so h bar will have the mean Ry inverse y bar minus mu bar y plus mu bar h and the covariance will not, the covariance does not remain unchanged, does not change because remember, simply shifting by the mean or shifting by a mean does not affect and this is the final expression for the conditional Gaussian PDF and which is very useful.

And as I said arises, this is the expression for the conditional Gaussian PDF which is very useful and arises very frequently for instance arises in machine learning, data analytics, signal processing, communication is there everywhere in the principle and this remember I already told you relates to the principle of MMSE estimation, the minimum mean squared error estimation. So, this has a lot of use in that sense.

So, that is wonderful, so I think we have covered this, we have, I hope all of you have been able to appreciate and completely understand this concept of the conditional probability, Gaussian probability density function that is when h bar and y bar are jointly Gaussian random vectors given y bar what is the conditional PDF of h bar and that then determines what is the estimate MMSE estimate of h bar given y bar which is nothing but the conditional expectation of h bar given y bar. So, let us stop here and we will continue with equivalent modules, thank you very much.