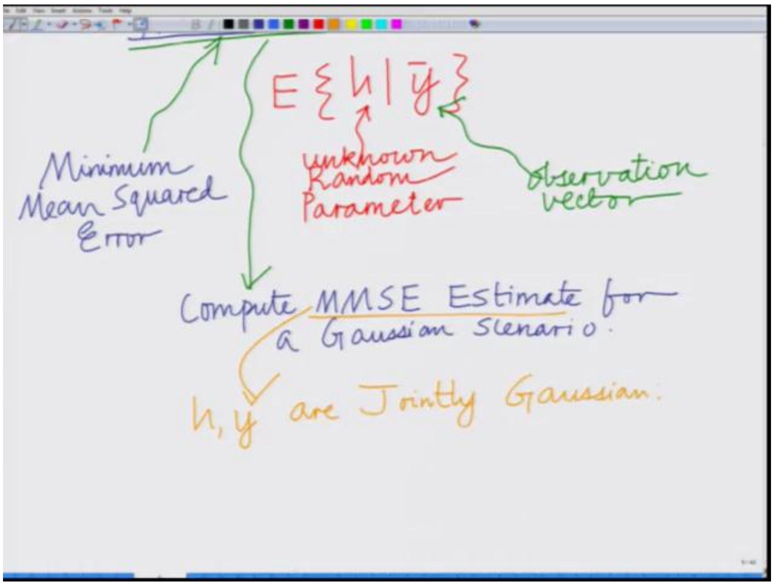
Bayesian/MMSE Estimation for MIMO/OFDM Wireless Communications
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture - 03
Derivation of Minimum Mean Squared Error (MMSE) Estimate for Gaussian Parameter – Part1

Hello. Welcome to another module in this massive open online course on Bayesian, MMSE Estimation for Wireless Networks. Previously we have seen that the MMSE or the minimum mean squared error estimator of an unknown parameter h is given as the expected value of h given y bar where y bar is the observation vector.

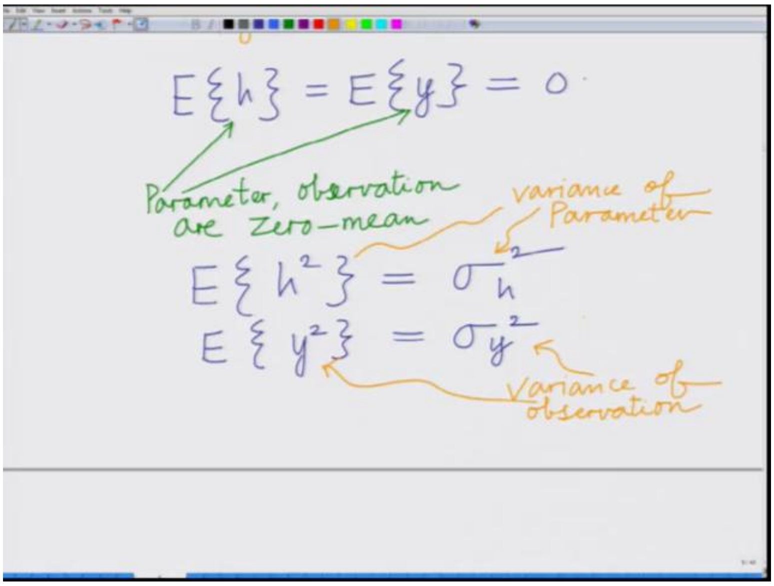
(Refer Slide Time: 00:40)



So, we have seen at the MMSE estimate of this parameter h is basically expected value of h given y bar right; h is a unknown random parameter, y bar is the observation vector and MMSE of course this stands for the minimum mean squared error. And this is what we have seen so far; MMSE stands for the minimum mean squared error. This is an important aspect of the estimation.

So, let us now illustrate how to compute this MMSE estimate for a Gaussian scenario. So, what we want to do is today’s module is that we want to compute this MMSE estimate specifically for a Gaussian scenario or basically when the parameter h and the observation y are jointly Gaussian. What we want to do is we want to compute MMSE estimate for a; let us see on this MMSE estimate looks like when the parameter h comma observation y, we are not considering observation we are considering a single observation y in the parameter h and the observation y are jointly; and this is important when they are jointly Gaussian that is random parameter h and it is corresponding observation y are jointly Gaussian random variables. We want to compute the MMSE estimate for this scenario which is nothing but the expected value of h given.

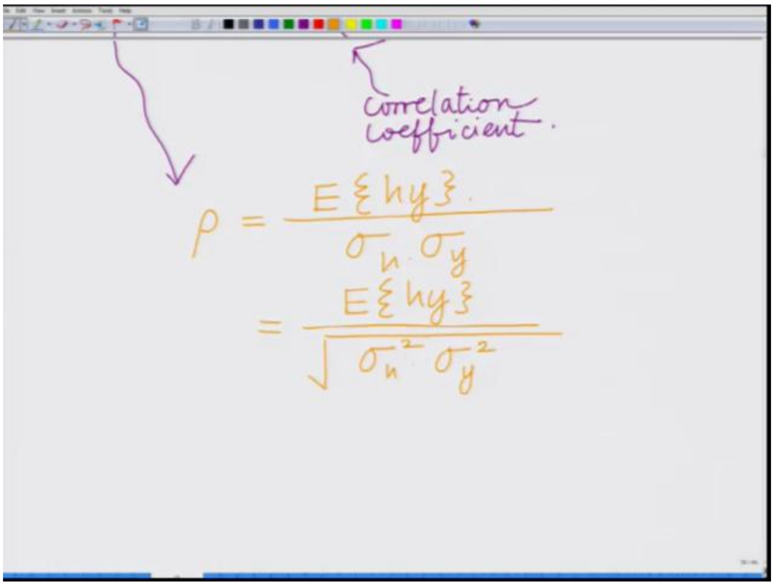
(Refer Slide Time: 03:26)



So, to start with let us consider zero-mean random variables that is let us say that expected value of h; to make thing simple let us consider expected value of h equal to expected value of y for 0. This basically implies that the parameter h and the observation are zero-mean random variables, they have zero-mean. Also let the variances be given as expected value of h square which is the variance of h since it is the zero-mean random variable equal sigma h square, that is the variance of the parameter is sigma h square expected value of y square equals sigma y square.

This is the variance of the observation so let us write this down; this is the variance of the parameter, this is the variance of your observation y or rather this is the variance of the parameter and this is the variance of the observation.

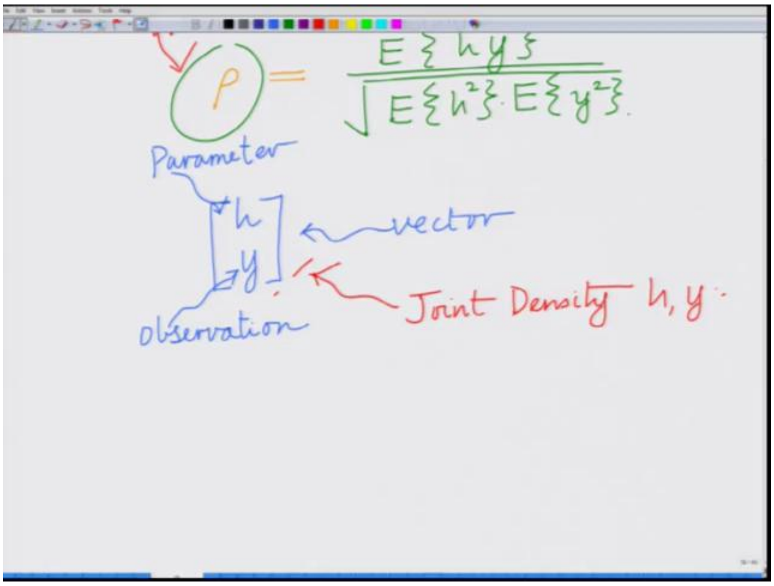
(Refer Slide Time: 05:08)



And also we need the covariance that is we need the expected value of h times y we need the covariance or that is expected value of h or rather the cross covariance expected value of h y which will be denote by rho sigma h into sigma y. With this rho, yes we met already familiar with this rho is the correlation coefficient of the parameter h and the observation.

This is the correlation of coefficient of the parameter h and the observation y that is this correlation coefficient is defined as rho, it can be seen that the correlation coefficient is given as rho equals the cross covariance expected value of h y divided by divided by sigma h into sigma y that is the product of the standard deviation sigma h square is the variance, sigma h is the standard deviation of h, sigma y square is the variance of y, sigma y is the standard deviation of y which can be written as expected value of h into y divided by under root sigma x square sigma y square.

(Refer Slide Time: 06:43)

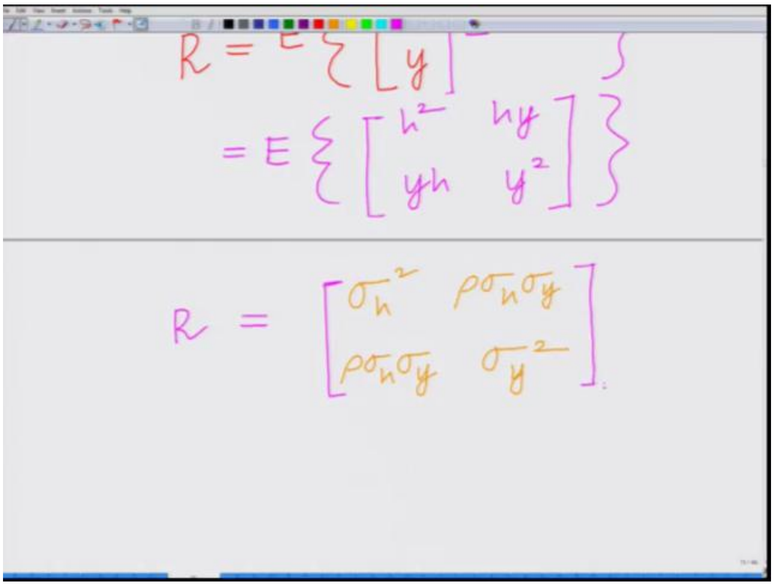


Which is nothing but basically that is rho equals expected value of h and y sigma h square expected value of h square and sigma y square is expected value of y square. So, this is the definition of your correlation coefficient. Remember this is a rho which is the correlation coefficient. One of the most important and interesting properties of the correlation coefficient is that the magnitude of correlation coefficient is always less than 1. It lies between minus 1 and 1 it indicates the extent of the correlation between h and y. And of course for Gaussian random variable said h and y this is also something that more many of we are familiar with Gaussian random variables h and y, if the correlation coefficient is 0 for any two random variable h and y are uncorrelated.

However, for Gaussian random variables; that is jointly Gaussian random variables h and y if the correlation coefficient is 0 then h and y are independent as per this property holds only for Gaussian random variable h and y. See these are some of the properties of the correlation coefficient.

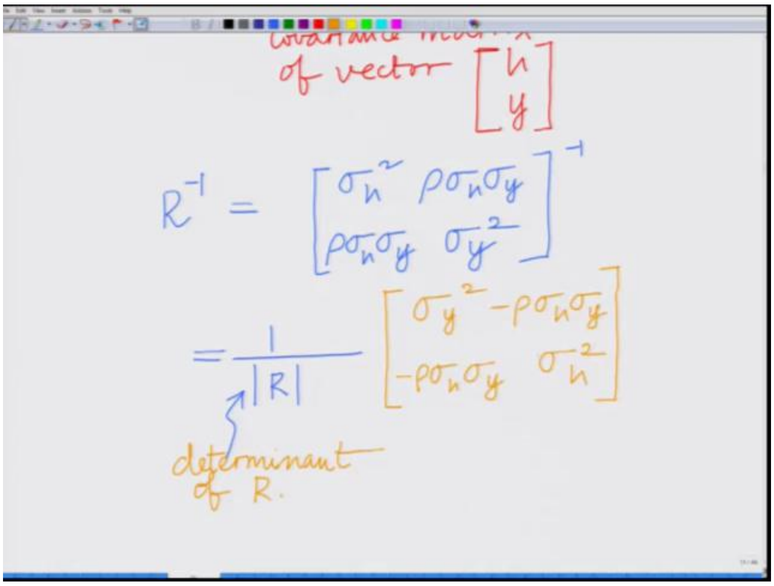
Now remember we want to come up with the MMSE estimate of h given y. So, towards this end first we want to construct the joint distribution of h and y. So, let us look at this vector let us define this vector; let us construct this vector this is our vector which we has basically the parameter and the observation, and what you want to come up with is we want to come up with the joint density or the joint property density function of h the parameter h comma the observation y.

(Refer Slide Time: 09:14)



Towards this end let us define the covariance matrix of this vector as expected value of this vector times it is transpose. We want to define the covariance matrix R as expected value of vector h y times it is transpose which is the expected value of this matrix this matrix which is now we can see h square hy yh y square. And expected value of this is now let us look at it lets take the expected value of each term expected value of h square is of course this is sigma h square expected value of y square is sigma y square expected value of hy is rho sigma h sigma y expected value of yh is rho sigma h sigma y, and this is the convenience matrix of the random vector h comma y.

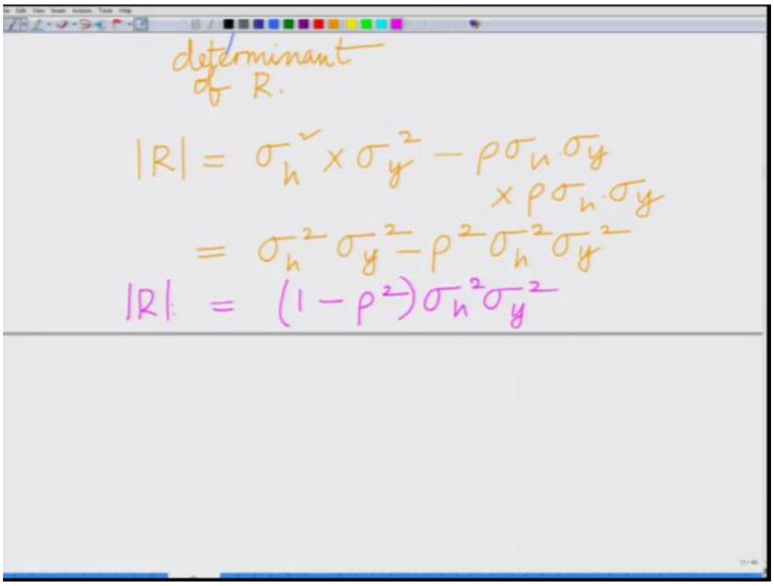
(Refer Slide Time: 10:37)



This is your covariance matrix of the vector h comma y where h denotes your observation and yh denotes the parameter and y denotes the observation. So, we have defined or we have rather derived the covariance matrix of the vector hy where h denotes the parameter y denotes the observation. In terms of the variances of hy and also the cross covariance that the cross correlation between the h and y.

And of course we also need the R inverse that is the inverse of this covariance matrix and that is given as follows that is the inverse of this matrix sigma h square, sigma y square, rho sigma h sigma y, rho sigma h sigma y inverse and that is given 1 over the determinant of this matrix R this stands for the determinant times. The inverse of this 2 cross 2 matrix basically we are interchange the diagonal element sigma y square and the negative of the off diagonal elements right minus rho is sigma h sigma y; this is the inverse of the matrix of course we can also compute what is the determinant of this matrix R.

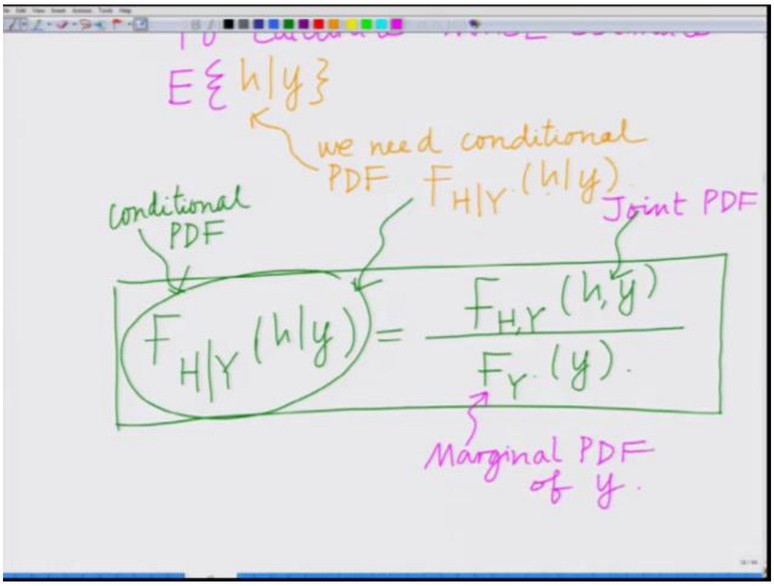
(Refer Slide Time: 12:37)



The determinant of this matrix R you can see is basically sigma h square and sigma y square minus rho sigma h sigma y times rho sigma h sigma y which is basically your sigma h square sigma y square minus rho square times, sigma h square sigma y square and this is equal to 1 minus rho square sigma h square sigma y square; this is the (Refer Time: 13:28) determinant of covariance matrix.

Now remember the LMSE estimate of the MMSE estimate is given as a expected value of h given y expected value of the parameter h given the observation factor y for that we need the conditional probability density function of h given y. So we need to find the conditional probability density function of the parameter h given observation y.

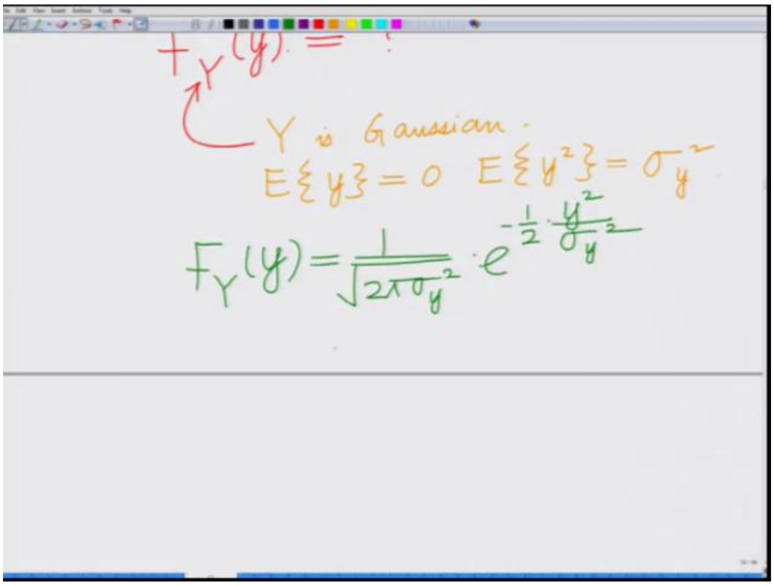
(Refer Slide Time: 14:01)



That is we need to calculate the MMSE estimate which is basically expected value of h given y we need the conditional PDF F of H we need the conditional probability density function of the parameter given the parameter h given the observation. Now, the conditional probability density functions. We know that this conditional probability density of function F of H given Y and this is important the conditional probability density of function of h given y this is basically given as the joint probability density of function of H comma Y divided by the marginal probability density function of h.

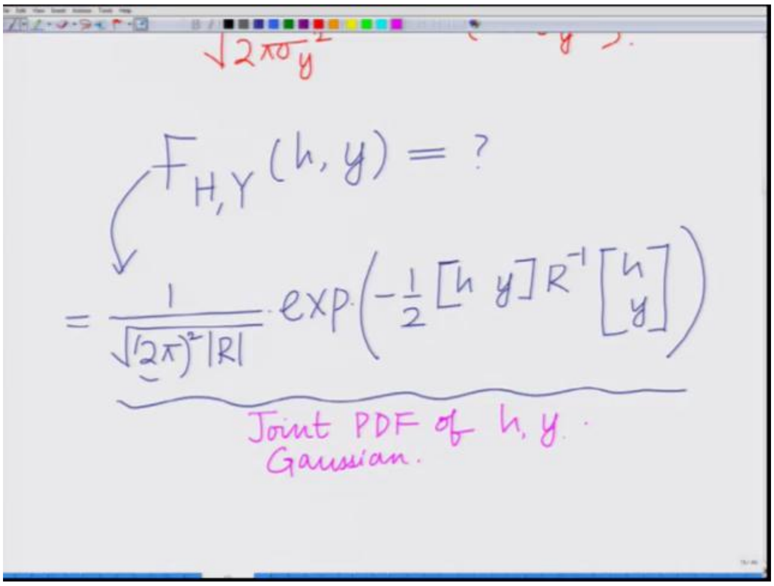
So, this the important relation keep in mind and this is the relation we are now going to use to compute this conditional. Remember this is the conditional PDF, this is the joint PDF of h comma y and this is the marginal PDF of your observation. So, we can compute the conditional PDF of the observation h of the parameter h given the observation y. From the joint probability density function of h comma y divided by the marginal probability density function of y. We need to compute each these components.

(Refer Slide Time: 16:29)



Now, first we know the marginal probability density of function of y. This is easy to compute what is this we know that Y is Gaussian random variable expected value of y equal to 0 and expected value of y square equal sigma square or equal sigma y square rather. So, Y is a Gaussian random variable with mean 0 variance sigma y square which implies F of Y of y equals 1 over square root of 2 phi sigma square e raise to minus 1 over twice y square divided by sigma y square.

(Refer Slide Time: 17:29)

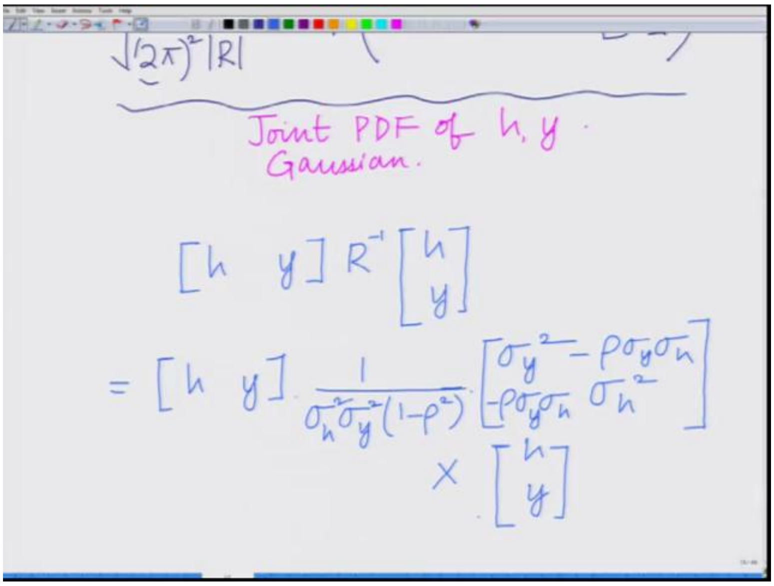


I can also just write denoting this e as x I can also write this as 1 over square root of 2 phi sigma y square x minus y square divided by 2 sigma y square, where e and x are basically same thing where e raise to x is basically x of x. So, the marginal probability density function of this Gaussian random variable y with mean 0 and variance sigma y square is 1 over square root of 2 phi sigma y square times x of minus y square divided by 2 sigma y square. So, that is the easy part.

Now, we want to come up with the joint probability density functions of h comma y. So, what my question is now or I have to rather find the joint probability density of function of h comma y. This is slightly complicated expression, this is given as basically let we write it down and then we can see what this is given as 1 over square root 2 phi square determinant of R x for e to the power of minus half h comma y or the rho vector h y times R inverse times the column vector h y; this is basically the joint probability density of function, this is the expression for the joint PDF of h comma y which is in fact Gaussian; h comma y are jointly Gaussian random variable.

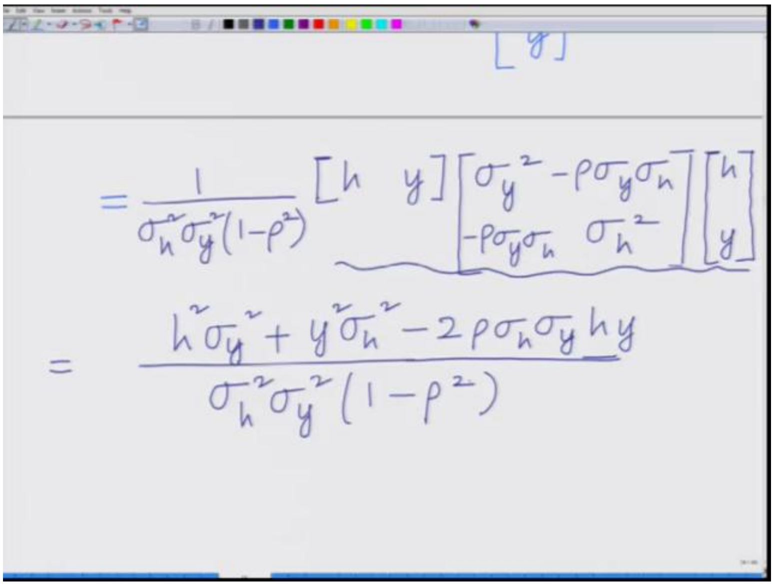
So, this is an expression or the joint probability density of function of the jointly Gaussian random variables h comma y. Where, h is the unknown parameter, y is the corresponding observation. Now, let us evaluate this joint probability density of function of h comma y. And this to evaluate this first we need h R inverse we already calculated R inverse now let us evaluate what is h y what is rho vector I want to evaluate this part that is the rho vector hy times R inverse times the column vector hy that is the rho vector hy.

(Refer Slide Time: 20:23)



We know that R inverse is 1 over the determinant of R which is sigma h square sigma y square into 1 minus rho square times sigma y square sigma h square minus rho sigma y sigma h minus rho sigma y sigma h times. The column vector h comma y.

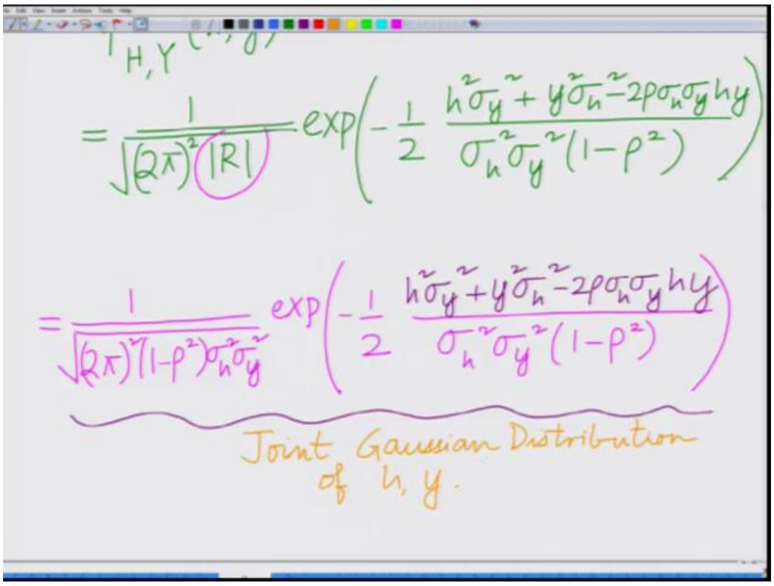
(Refer Slide Time: 21:08)



And this can be simplified as now this is a scalar quantity so I will bring it affront so this is 1 over sigma h square sigma y square into the rho vector h y times this matrix which is sigma y square sigma h square minus rho sigma y sigma h minus rho sigma y sigma h times, well let me write it over here times the column vector h y. And now I can simplify this as 1 over; let me simply this I can simplify this as 1 over well I can simplify write this expression in the denominator I have sigma h square sigma y square times 1 over rho square in the numerator I have h square sigma y square plus y square sigma h square that is I am evaluating this part over here this is the part which corresponds to the numerator which is the h square sigma y square plus y square sigma h square minus 2 rho sigma h sigma y times h into y.

This is basically the term which corresponds to the rho vector h y times R inverse where R is the covariance matrix corresponding to the vector h y times the column vector h y.

(Refer Slide Time: 23:27)



And therefore, the joint probability density of function is given as, remember we go back to our earlier relation for this joint probability density of functions where joint probability density of function that can be obtained from remember this relation that we have over here that gives the joint probability density of function. And therefore that is now F of H comma Y of h comma y which is equal to now 1 over 2 phi square determinant of R times e to the power of minus half, the denominator I have sigma h square sigma y square into 1 minus rho square, in the numerator I have h square sigma y square plus y square sigma h square minus 2 rho sigma h sigma y h comma y; write this clearly minus 2 rho sigma h sigma y and times h y.

Now I can substitute this value corresponding to this determinant of R we also know what is the determinant of the covariance matrix R that is 1 over. Let me again write this clearly over here that is 1 over 2 phi whole square times the determinant of R this is basically 1 minus rho square sigma h square sigma y square times e to the power of minus half basically sigma h square sigma y square into 1 minus rho square and on the numerator we have h square sigma y square plus y square minus h square minus 2 rho sigma h sigma y into h comma y and this is therefore the joint distribution what is this, this is my joint Gaussian distribution of the parameter h and the observation y.

Now, we have the joint distribution of h and y so we have two aspects; we have the joint distribution; this is your joint distribution, we have the marginal probability density of function; so this is basically your marginal probability density of function of random variable y. Now we substitute this joint distribution and the marginal distribution joint probability density of function of h comma y in the numerator and the marginal probability density of function of the y in the denominator and this expression to get the conditional probability density of function of the random variable or the parameter h and y.

Once we get the conditional probability density of function of h given y we can evaluate the mean of this conditional probability density of function to derive the MMSE estimate of the parameter h given the observation y. And that is something that we are going to do in the subsequent module. So, we will stop here.

Thank you very much.