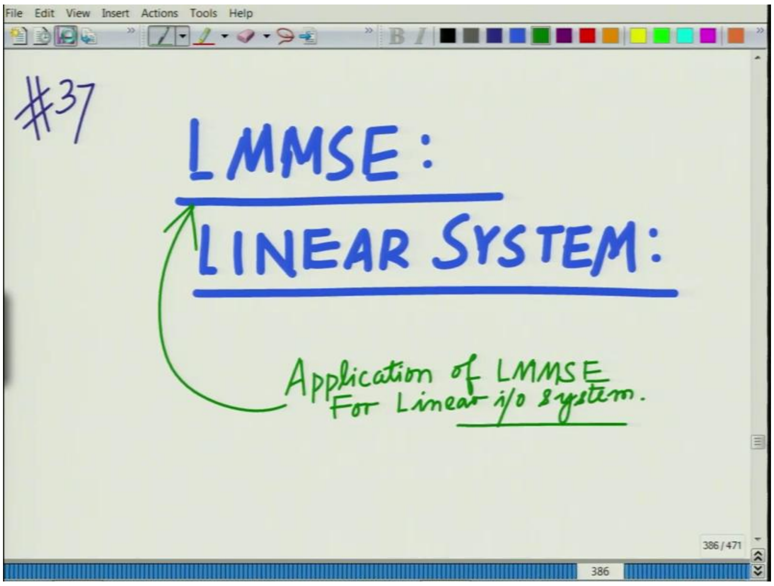
Applied Linear Algebra for Signal Processing, Data Analytics and Machine Learning
Professor Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur
Lecture – 37
LMMSE Estimation in Linear Systems

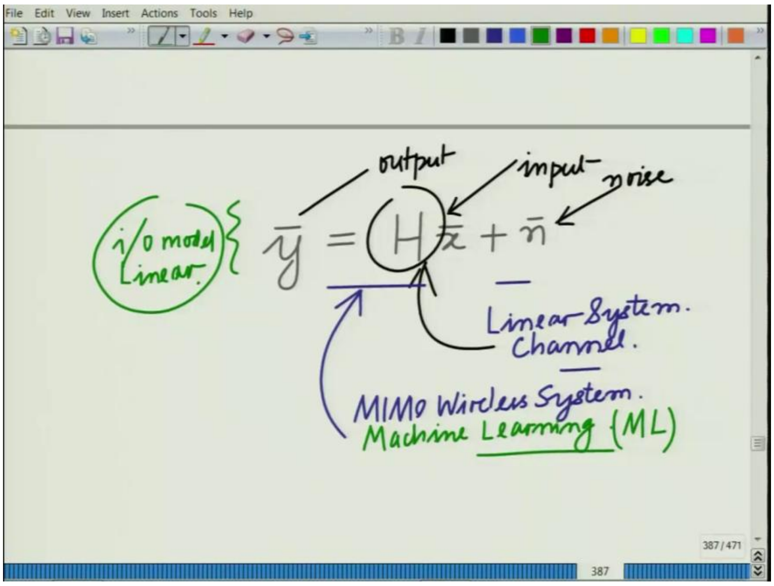
Hello, welcome to another module in this massive open online course. So, we are looking at the LMMSE or the Linear Minimum Mean Square Error Estimation principle. Now, let us look at an application of the LMMSE for a linear system.

(Refer Slide Time: 00:35)



So, we want to look specifically at the application of the LMMSE, we want to look at an application of that specifically in the context of now a linear system that is the linear input output system. So, what we mean by this is what happens in a that is we want to look at the application of LMMSE for a linear input output system, for linear input output system.

(Refer Slide Time: 1:48)



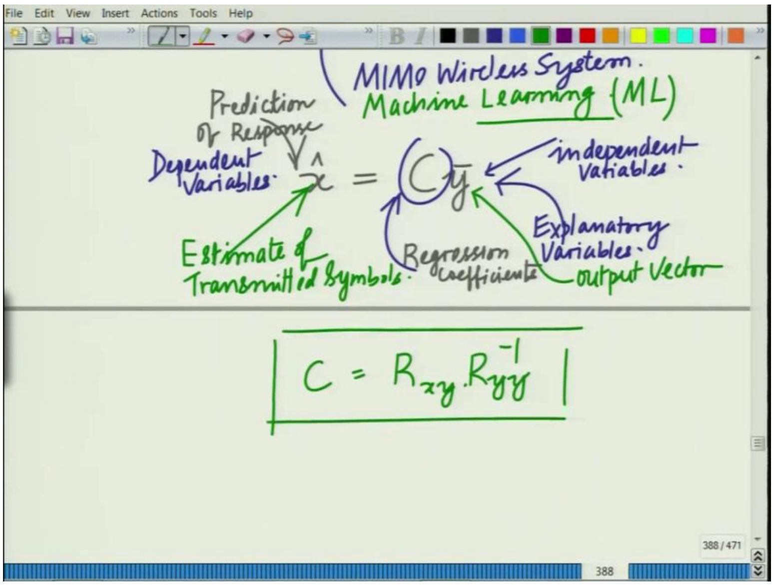
So, 1 can model such a system that is a linear input output system as y bar equals H times x bar plus n bar y bar is you can think of this as the output x bar, the input n bar is the noise, H is the system or basically the linear system or basically for instance, if you consider a wireless system this will be the channel.

So, for instance this would be the channel in a wireless system, for instance this can be the model as we have seen before this can be a MIMO wireless system that is multiple input multiple output, that is a multiple input multiple output wireless system or this can also be a machine learning.

We have also machine learning where you have observations and based on these observations now, you want to build a regressor, and the regressor model is linear. So, essentially so this can also be used in M L or this can be used as I have said in wireless communication so you can use it in machine learning, this is an input output model. So, this is essentially your the IO model which is a linear model.

Now, 1 thing you have to observe is only you have to remember as I have already mentioned that the LMMSE is always linear irrespective of the input output models, LMMSE is the best linear minimum mean squared error estimate. So that is always linear irrespective of the fact that, irrespective of the fact that irrespective of whether the input output model is linear or nonlinear. Now for the specific case when the input output model is linear, what happens to the LMMSE estimate that is essentially what we want to look at.

(Refer Slide Time: 4:21)

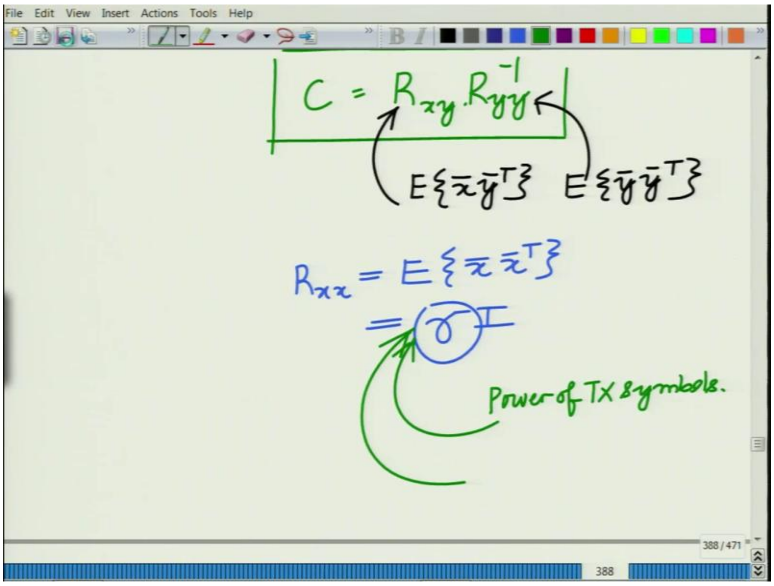


So, now, we want let us say we want to now again build x hat equals C y bar where your y bar is as we have said in the example of regression is your y bar is essentially your independent variables or essentially your explanatory variables. So, these are you can think of this as your explanatory and see these are basically your regression coefficients.

You want to determine C and x hat, this is essentially your prediction of the of the, this is the prediction of the response or essentially your dependent variables ? Or in the wireless channel, this would be estimate of the channel estimate of the transmitted symbols, this would be estimate of the transmitted symbols and y bar would be essentially your output symbols.

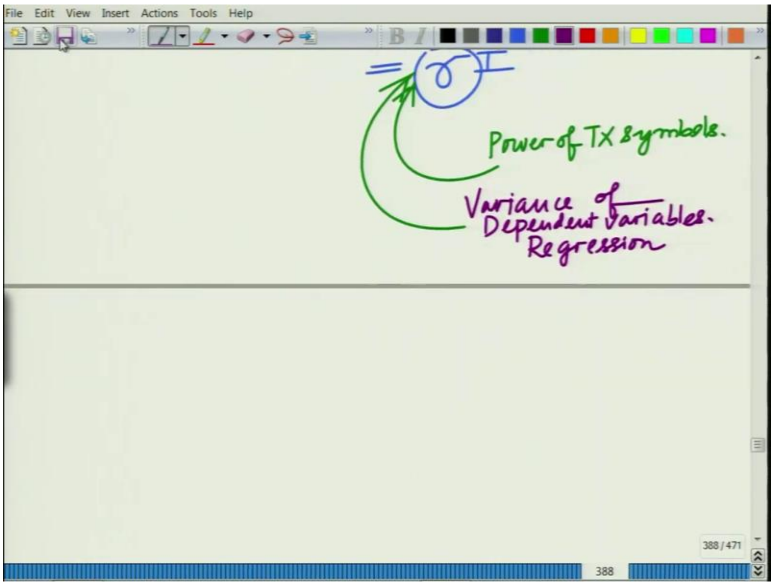
So, depending on which context you are looking at it, that is either you are looking at the machine learning or wireless. So, this would be your output vector al, depending on the context in which you are looking at. And as we have seen for LMMSE that is this regression coefficients, this is given as R x y into R y y inverse. So C equals R x y into R y y inverse, this is essentially the beautiful result that we have for the LMMSE.

(Refer Slide Time: 6:52)



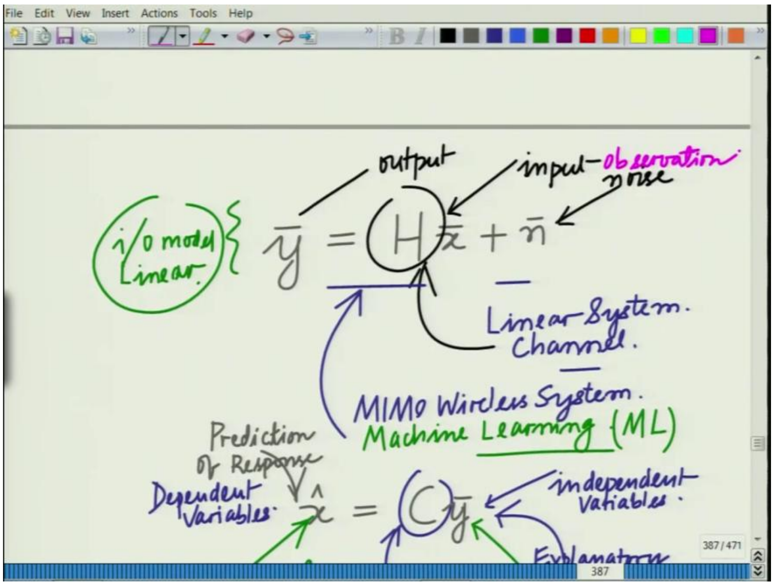
Now R x x we know this is essentially expected value x bar y bar transpose. R y y this is essentially expected value of y bar y bar transpose. Now, let us form R now R x x let us form this as x, to do this, we will need R x x which is expected value of x bar x bar transpose, let us set this as gamma times identity, where you can think of this vector gamma as either the power of the transmitted symbols or this is essentially also your variants of the dependent variables.

(Refer Slide Time: 7:55)



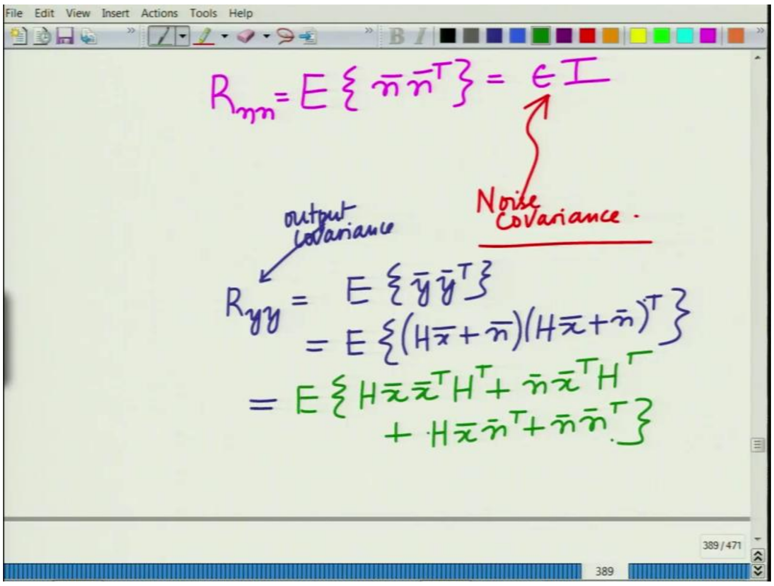
So, variants of the, variants of the dependent variables for your regression problem. And we can set R n n that is the observation your n bar this you can think of this as this is your observation noise.

(Refer Slide Time: 8:26)



Observation or essentially you can also think of this as the measurement noise, you are making the measurements y bar. So, this is the noise in the measurements and the corresponding covariance is basically the measurement noise covariance we can set that as epsilon times identity.

(Refer Slide Time: 8:49)

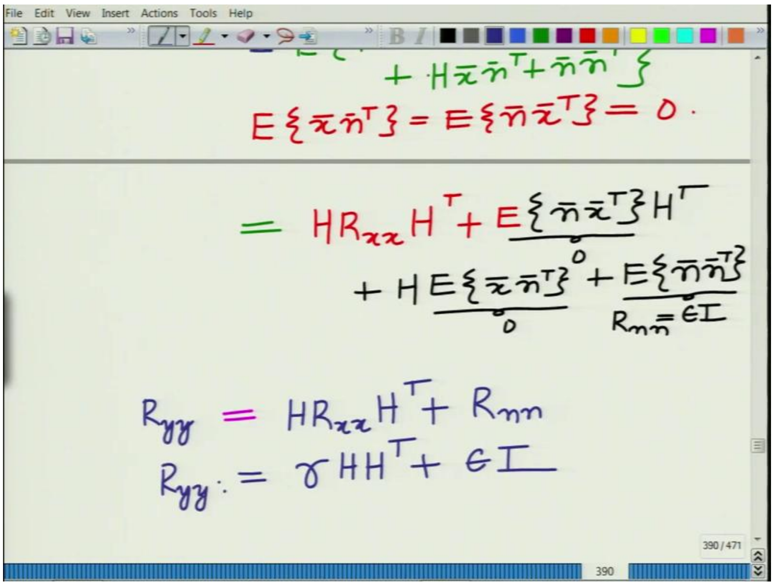


So, we can set expected n bar, this is your R x x you can call this as R n n expected x bar n bar n bar transpose, this is you consider this as epsilon times identity, this is your essentially this is your noise covariance, this is your noise covariance. And therefore, if we for now we want to form remember first we want to find the R y y which is output covariance or essentially you can think of this as the covariance of the explanatory variables R y y. So, this is the output covariance.

This is the expected value of y bar y bar transpose which is expected value of H x bar plus n bar times H x bar plus n bar transpose which is essentially expected value of H x bar x bar transpose H transpose plus n bar x bar transpose H transpose plus H x bar n bar transpose plus n bar n bar transpose which now, if you move the expected expectation operator inside.

So, you will have terms such as expected value of x bar x bar transpose which is nothing but R xx, expected value of x bar n bar transpose and expected value of n bar x bar transpose, these are the cross covariance between the symbols and the noise or this is a cross covariance between essentially your, it is a cross covariance between the response and the noise so this we can set as 0.

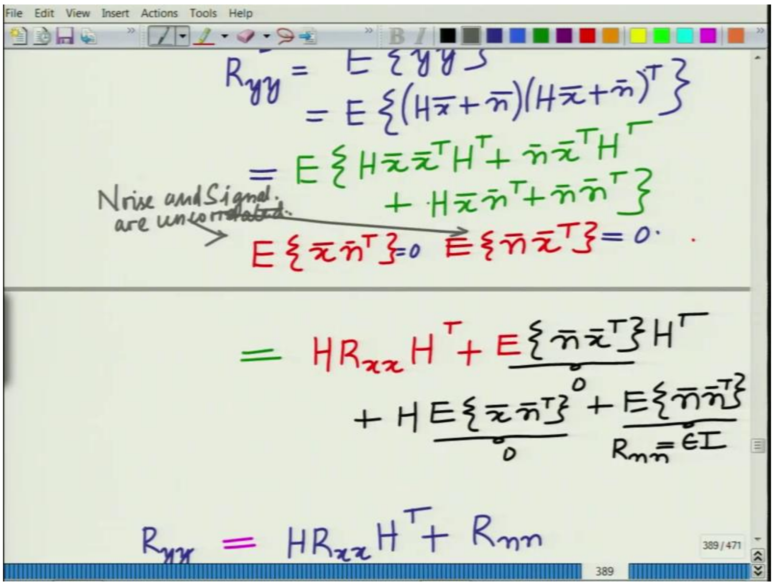
(Refer Slide Time: 11:35)



This is the cross covariance between the symbols and the noise which we are going to set equal to 0 and then you have the noise covariance. So, the interesting part here is I can set the cross covariance between the noise and the transmit symbols, this is equal to 0. So, this reduces to H, expected value of x bar x bar transpose this is R x x H transpose plus this is 0, you have expected value of n bar x bar transpose times H transpose so this is 0 plus H times expected value of x bar n bar transpose this is 0 plus expected value of n bar n bar transpose which is this is your R n n which is epsilon times identity.

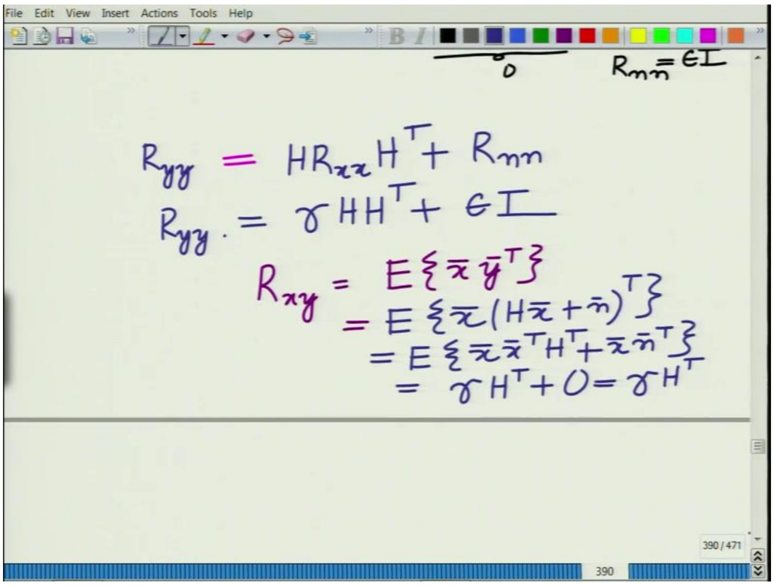
So, finally, simplifying this you will have H R x x H transpose plus R n n, this is your R y y which for this case is essentially your gamma H H transpose plus epsilon times identity. So, this is your output covariance, this is your R y y. So, the property that we have used here is that expected value of x bar n bar transpose is equal to expected value of x bar transpose is 0.

(Refer Slide Time: 13:49)



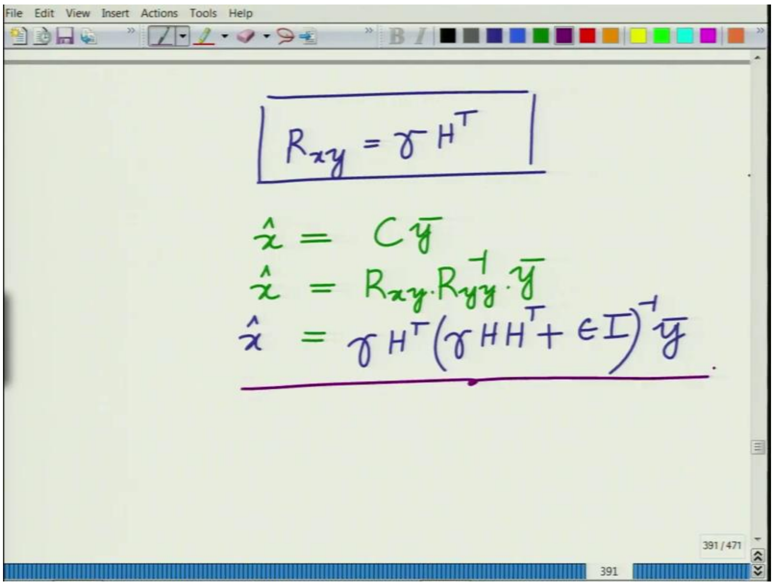
So this is 0 and this is essentially that the property, this is arises from the property that the noise and signal are uncorrelated, noise and signals are uncorrelated.

(Refer Slide Time: 13:57)



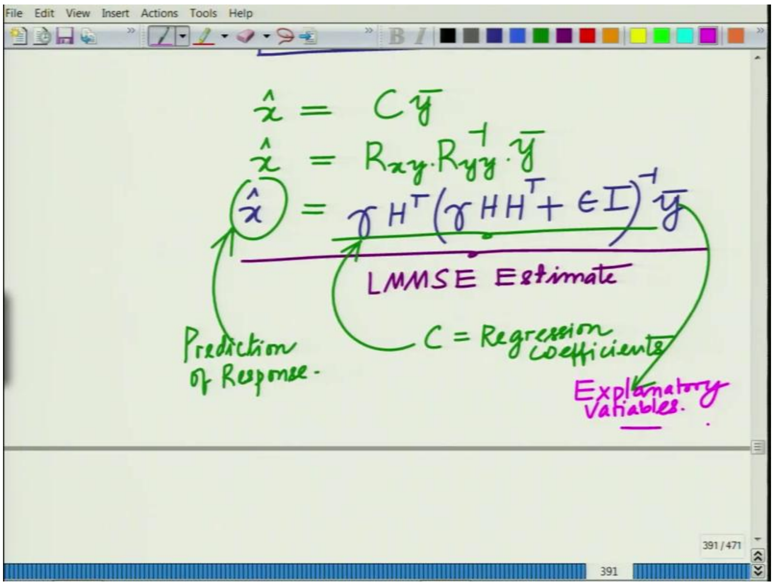
And we other quantity we have to determine is basically R x y which is again the cross covariance between X and Y, this is expected value of x bar y bar transpose. So, this is essentially the expected value of x bar y bar transpose which is if you simplify this substitute for y bar, this is expected value of x bar H x bar plus n bar transpose which is again expected value or x bar x bar transpose H transpose plus x bar n bar transpose which is essentially you will get once again gamma times H transpose plus expected value of x bar n bar transpose is 0 so gamma H transpose. So, that is essentially what this is.

(Refer Slide Time: 14:55)



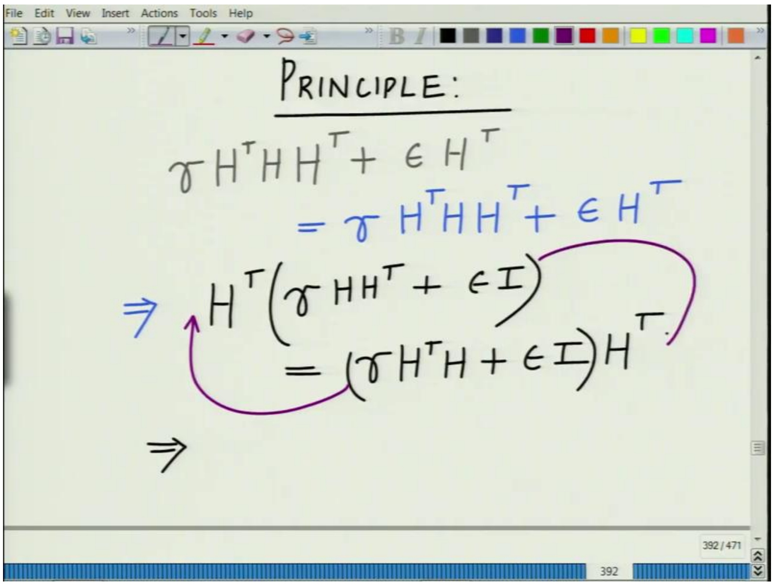
So, R x y equals gamma H transpose fairly straightforward. And therefore, now x hat that is the prediction of the response, the prediction can be determined as x hat equal to C times y bar, where C is R x y to R y y inverse so this will be equal to C times y bar, which is essentially R x y times R y y inverse times y bar which is essentially equal to gamma H hat times gamma H H gamma H transpose plus gamma H H transpose times epsilon I inverse times y bar, this is your LMMSE estimate. This is your LMMSE estimate, you can think of this matrix. So, this matrix these are essentially your regression coefficients. So, this is your C, these are your regression coefficients.

(Refer Slide Time: 16:42)



This is your essentially your response or the prediction of the response. These are your explanatory, these are your explanatory or independent variables, essentially that is what you have. So, essentially you have obtained the very interesting expression that is x hat that is estimate of you can also think of this as the estimate of the transmit symbol vector equals gamma H transpose time, which is R x y times gamma H H transpose plus epsilon I inverse which is R y y inverse times y. Now, let us further simplify this. So, if we simplify this what we get is let us use the following tool.

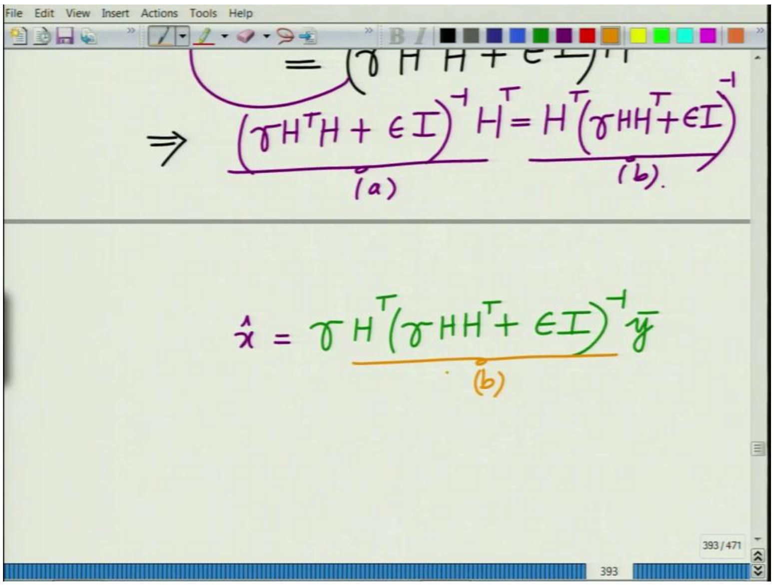
(Refer Slide Time: 18:02)



To simplify we will discover or we will use the following principle, let us discover this principle let us start on both sides with gamma H transpose H H transpose plus epsilon H transpose is equal to gamma H transpose H H transpose plus epsilon H transpose, on both sides we have this quantity. So now left and is both gamma H transpose H H transpose plus epsilon H transpose. So, we have the same quantity on left and. Now here on the left hand side take gamma H transpose common on the left.

So that gives gamma H trans or that gives H transpose times gamma H H transpose plus epsilon i. On the hand side take H transpose outside on the. So, that gives gamma H transpose H plus epsilon times identity times H transpose. Now, bring this over here. So, bring this over here and we bring this over here after inverting.

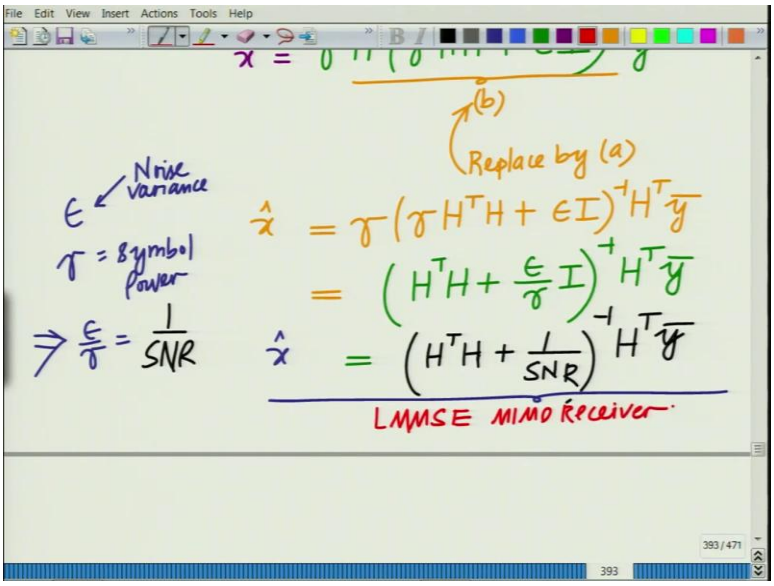
(Refer Slide Time: 19:48)



So, that essentially means that gamma H transpose H plus epsilon I inverse H transpose equals H transpose gamma H H transpose plus epsilon I inverse. Let us call this a, let us call this b, we have a equal to b. Now, if you look at this we have x hat equals gamma times H transpose gamma H H transpose plus epsilon i inverse into y bar.

Now, if you look at this quantity that is this underlined quantity H transpose gamma H transpose epsilon I inverse, this is essentially your quantity b. So, I can replace it this by a so replace so this is essentially your quantity b.

(Refer Slide Time: 21:09)

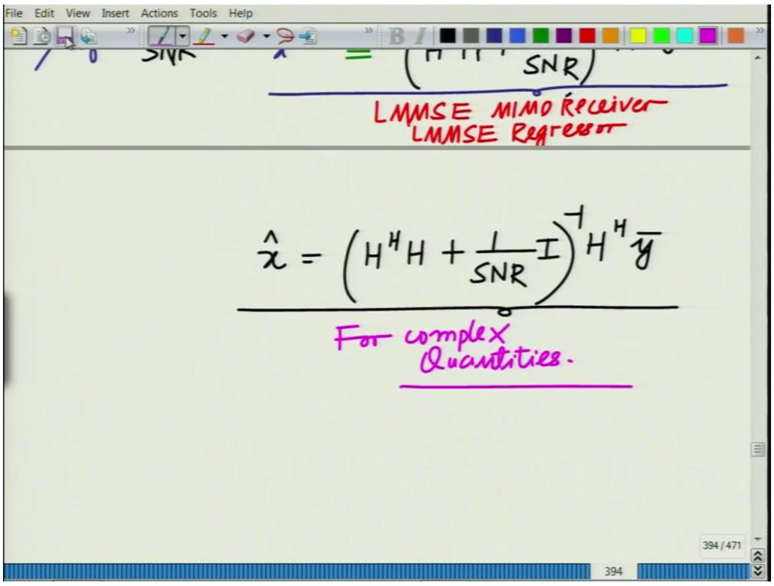


So replace by a which becomes gamma, a is nothing but gamma H transpose H plus epsilon I inverse H transpose y bar. And now you take the gamma inside that will be your, if you take it inside it becomes 1 over gamma because there is an inverse so that will be H transpose H plus epsilon over gamma I inverse H transpose y bar which you can interestingly now write as x hat, remember epsilon is nothing but noise variance and gamma is the symbol power or signal variance, because it is expected value of x bar x bar transpose. So this implies epsilon over gamma is noise variance by symbol power so this is essentially your what we call in communication as 1 over SNR.

So, this becomes a very interesting formula H transpose H plus 1 over SNR inverse H transpose y bar, this is essentially what we call simply in communication as the LMMSE receiver or the MMSE receiver. So, for the MIMO this essentially becomes your LMMSE MIMO receiver or you can also think of this as the LMMSE regressor, you can also think of this as the LMMSE linear linear minimum mean squared error regressor al.

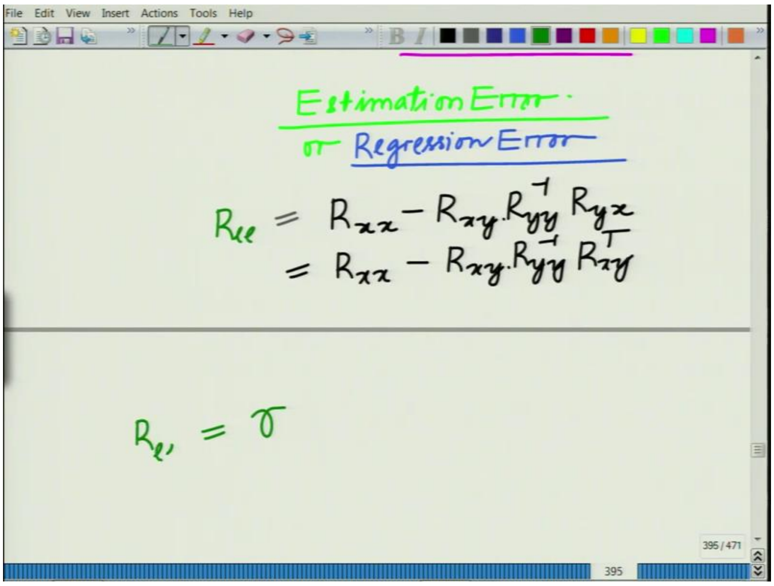
So, that is the expression that is x hat equals H transpose H plus 1 over SNR times H transpose y bar. And in fact for complex channel matrix because in communication typically the matrices and quantities are common, you can simply replace the transpose by the Hermitian.

(Refer Slide Time: 23:40)



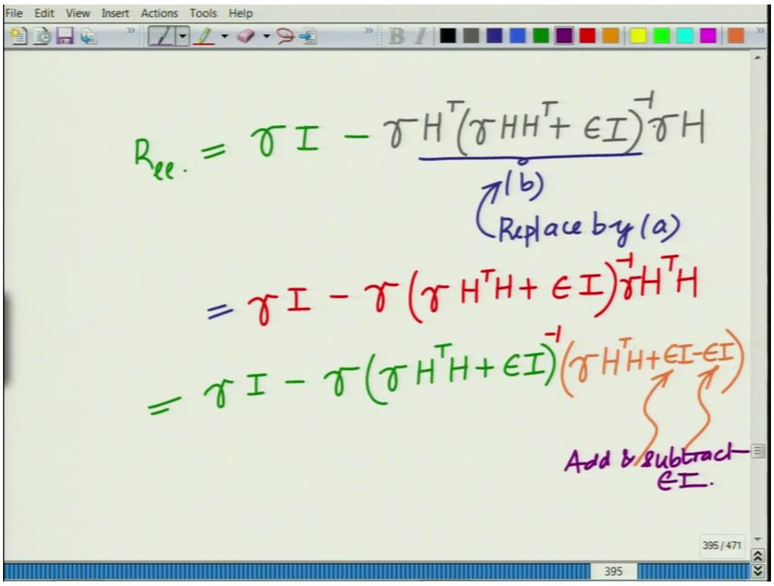
So, for complex you can write this as x hat, for complex quantities this simply becomes x hat equals H hermitian H plus 1 over SNR identity inverse times H hermitian y, this is essentially for complex quantities, this is essentially what we have for complex quantities. And now, let us find the error variance that is the what is the regression error and what is the variance of this regression error or this prediction error or we can also think of it for a communication scenario as the estimation error covariance. So we ask the question, what is the, we want to find the estimation error covariance. So, how do we find the estimation error?

(Refer Slide Time: 25:02)



Or basically you can think of this as the regression and this we have seen is basically this is equal to R x x minus R x y R y y inverse into R y x, which I can write as R x x minus R x y we have calculated, R y y we have calculated R y y inverse, R y x is essentially R x y transpose. So, substituting these quantities, this becomes, I am just going to substitute these quantities so this is you can call it this as R e e regression error covariance.

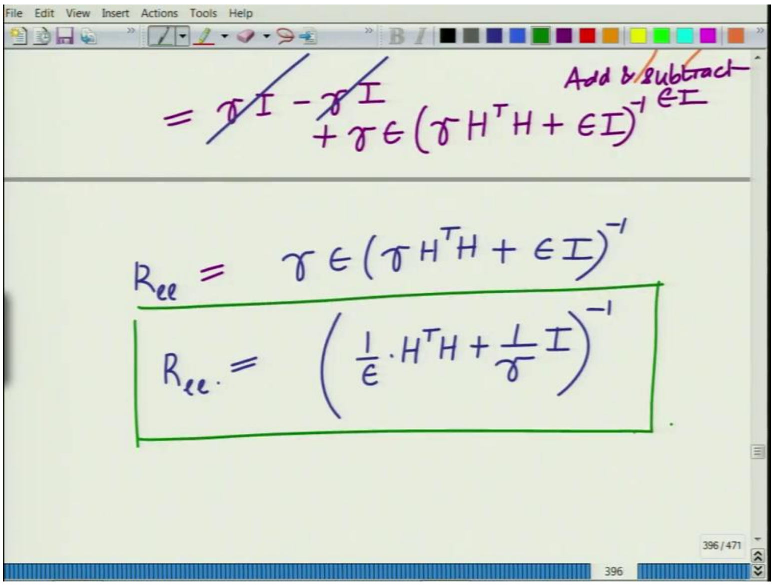
(Refer Slide Time: 26:02)



So, this becomes gamma times identity minus R x y is gamma H transpose times gamma H H transpose plus epsilon identity inverse times R x y transpose which is gamma H. Now, if you look at this quantity, let us look at this H transpose gamma H H transpose plus epsilon, this is nothing but this quantity is nothing but this quantity is basically your B, what we call above as this is b, so once again replace by a, so this will become your gamma I minus. So, this will become gamma I minus gamma times H transpose H, this will become gamma H transpose H plus epsilon identity inverse into H transpose H. Of course, there is going to be another gamma over here.

And now, I am going to add and subtract epsilon identity and you will see something very interesting happening. So, this is gamma I minus gamma gamma H transpose H plus epsilon identity times this inverse of this times gamma H transpose H plus epsilon identity minus epsilon identity. So, we are adding and subtracting, for simplification add and subtract epsilon I.

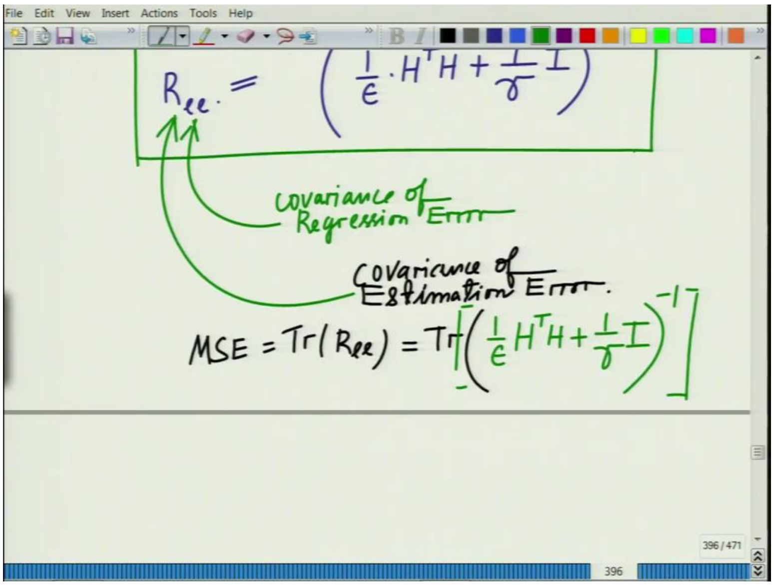
(Refer Slide Time: 28:15)



Now if you look at this, this is essentially this quantity, if you look at this quantity gamma H transpose x plus epsilon I, this is essentially gamma H transpose H plus epsilon I. So, this inverse into this, this becomes identity. So, we will have gamma I minus gamma identity plus gamma epsilon times gamma H transpose H plus epsilon identity inverse.

So, this will be now these things cancel and therefore this will be your error covariance will be gamma epsilon times H transpose H plus epsilon identity which is essentially if you take gamma and epsilon inside this will be 1 over epsilon that is 1 over the noise power times H transpose H plus 1 over gamma 1 over a signal power times the identity and of course, the inverse of that. So, that is essentially what this is going to be. So, this is essentially your this is essentially your covariance of the regression error.

(Refer Slide Time: 29:49)



Or you can also think of this as the covariance of estimation error for the LMMSE receiver. And of course, your MSE that is a minimum MSE is going to be the trace of this which is nothing but the trace of 1 over epsilon H transpose H plus 1 over gamma identity inverse. So trace of this is essentially your, this is your MSE that is the mean squared error that is you take the sum of the diagonal elements of the covariance matrix because those correspond to the terms expected value of magnitude e 1 square, expected value of magnitude e 2 square and so on.

So, you take the sum of those elements on the diagonal you get the mean squared error that is essential in this case, either your mean square estimation error or your mean Square regression error. And as you have already seen before, the higher your cross covariance, the better is your estimate, the lower is your mean squared error.

And that thing also comes out from this that is if you look at here, 1 over epsilon and gamma, you can also deduce very easily from this what happens to this regression error or this estimation error as a function of this epsilon and gamma, you should be able to easily deduce that.

I leave that as an exercise to you but essentially these are the basics of the LMMSE principle which has again widespread applications, regression, wireless communication, so on and so forth. So, let us stop here and let us continue this discussion in the subsequent modules. Thank you very much.