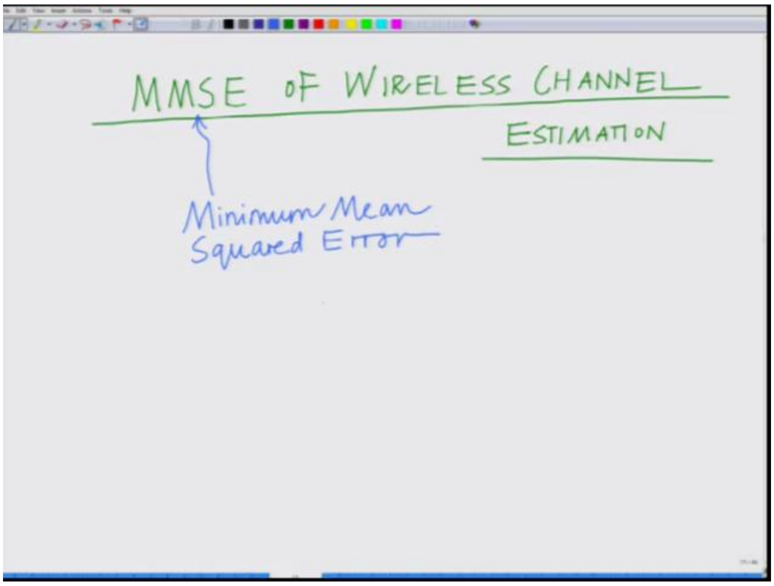
Bayesian/MMSE Estimation for MIMO/OFDM Wireless Communications
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture – 13
Minimum Mean Squared Error (MMSE) for Wireless Fading Channel Estimation-Derivation

Hello, welcome to another module in this massive open online course on Bayesian MMSE Estimation for Wireless Communications. So far we have looked at the MMSE that is the minimum mean squared error of a wireless sensor network. Similarly, let us look at the MMSE that is the minimum mean squared error for the wireless channel estimation problem and the analysis is going to be similar with some slight modifications for the complex channel as coefficient estimation case. So, therefore for the sake of completeness, let us look at the MMSE that is the minimum mean squared error of wireless channel coefficient estimation.

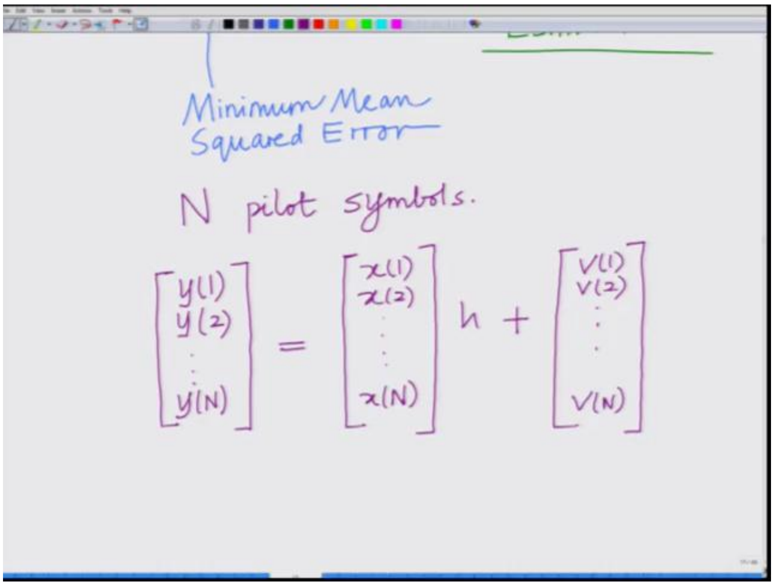
(Refer Slide Time: 00:50)



So, what we are going to look at in this one? Although it is going to be similar to the wireless sensor network MMSE of wireless channel estimation and where we know that this term MMSE stands for the minimum, this is the actual minimum mean squared error.

Let us look at the modal for the estimation and consider the transmission. Of course, we have looked at this modal, but let us just recall this modal.

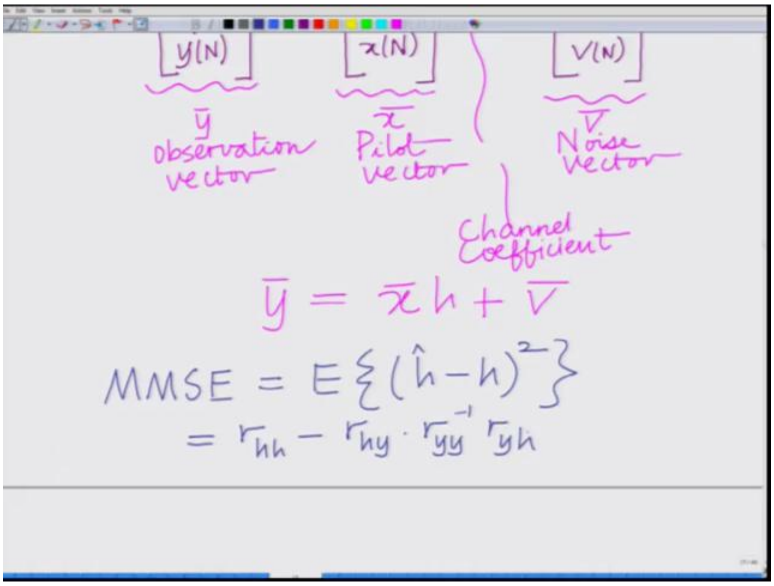
(Refer Slide Time: 02:01)



Let us consider the transmission of N pilot symbols and therefore I can stack the received symbols as a vector that is y 1, y 2 so on up to y N. This is the received or the observed vector which is equal to your pilot vector comprising of the pilot symbols x 1, x 2 so on up to x N times h which is the channel coefficient plus the noise samples of course v 1, v 2 and v N.

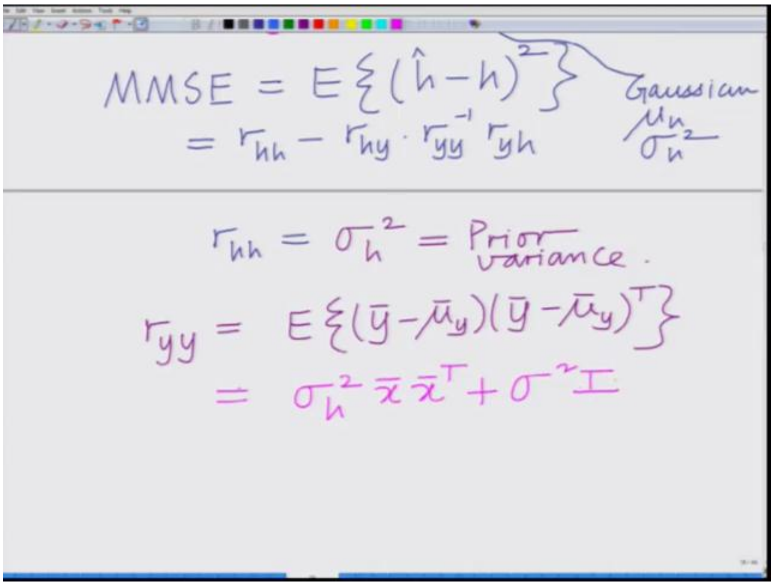
So, we are considering the transmission of N pilots symbols x 1, x 2, x N which were receiving as the pilot vector x bar, the corresponding N received symbols y 1, y 2, y N which basically forms the received vector y bar and the noise vector v 1, v 2, v N which forms the noise vector v bar and of course, the channel coefficient h.

(Refer Slide Time: 03:20)



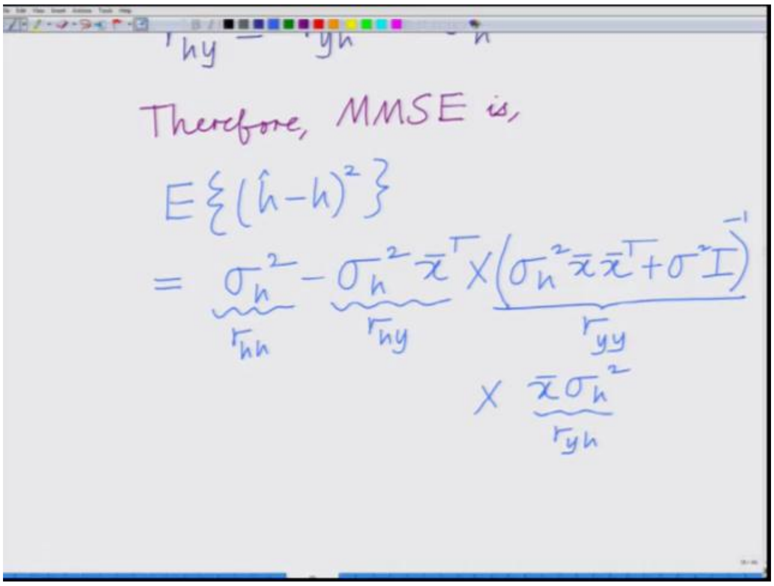
So, the system modal for the estimation is basically your vector y bar, this is the observation vector, this is your pilot vector and this is the v bar which is your noise vector and this is basically your channel coefficient h and estimation of this channel coefficient h is termed as a wireless channel estimation. This is something that we have already seen. Now, what we are going to do is let us first start with the scenario where the channel coefficient h is a real quantity and later we will see how the results can be modified for the scenario where the channel coefficient h is a complex quantity that is known as the complex fading channel coefficient. So, first let us start with the real scenario and therefore in our system modal we have y bar, it was x bar, this is something that we have already seen, x bar h plus v bar and the MMSE that is the minimum mean squared error which is equal to the expected value of magnitude of h hat minus h whole square. Well this is equal to r h h minus r h y into r y y inverse into r y h and remember we are considering a channel coefficient h to be Gaussian, the prior distribution is Gaussian with mean mu h and variance sigma h square.

(Refer Slide Time: 05:44)



So, r h h is equal to your prior variance, this is sigma h square which is basically your prior variance. We already know r h h that is expected value of h square, exhibited value of h minus mu h square equals sigma h square which is the variance of the Gaussian prior for this fading channel coefficient h. Similarly, we have also derived these quantities and you can recall them r y y which is equal to expected value of y minus, in fact, this is expected value of y bar minus mu bar y times y bar minus mu bar y transpose and this quantity is equal to sigma h square times x bar x bar transpose plus sigma square times identity matrix. This is something that we have already derived.

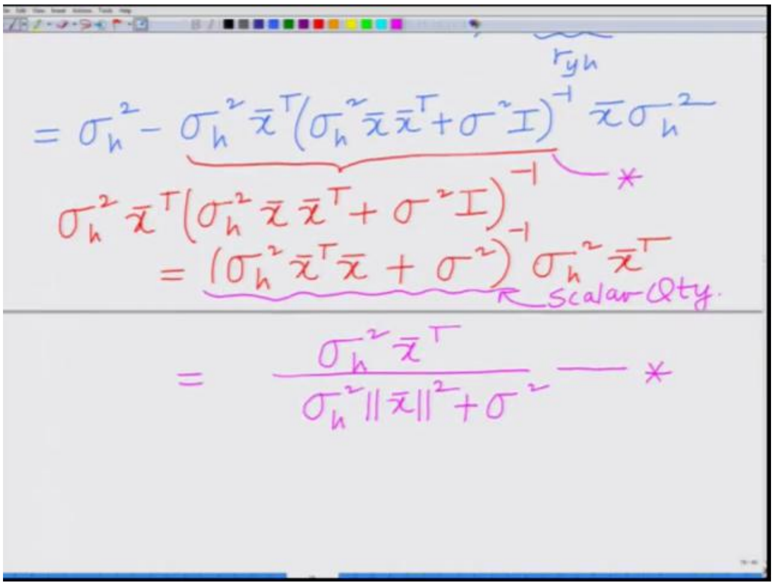
(Refer Slide Time: 07:02)



If you are not clear about this, please check from the lectures of the previous modules and we have the cross co-relation that is r h y which is equal to also r y h transpose. Similarly, remember we are still considering the real scenario. So, r h y equals r y h transpose, if it is complex then it will be the hermitian. We are still considering the real scenario, this will be sigma h square x bar transpose. So, these are the three different quantities r h h, r y y. In fact, four different quantities r h h, r y y, r h y and r y h and r h y and r y h are the transposes of each other. Therefore, the MMSE is you have expected value of h hat minus h square, this is equal to the MMSE that is the minimum mean square error that is equal to now substituting these quantities r h h sigma h square minus r h y that is sigma h square times x bar transpose.

So, this is basically your r h h, this is your r h y times, remember r y y inverse that is sigma h square x bar x bar transpose plus sigma square identity inverse, this is r y y times x bar into sigma h square. So, this is your r y h and basically this is the expression for the MMSE.

(Refer Slide Time: 09:39)

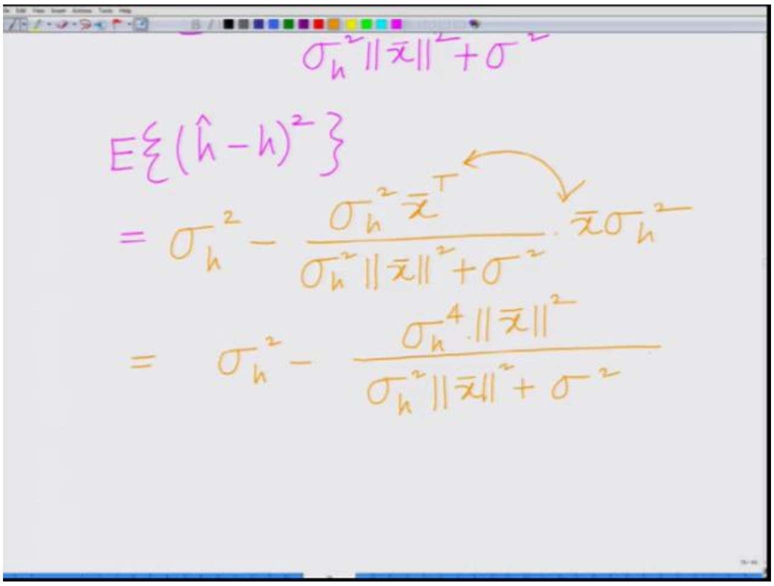


So, let me just write it down once again over here clearly, that is basically sigma h square minus sigma h square into x bar transpose sigma h square x bar x bar transpose plus sigma square identity inverse times x bar into sigma h square. So, this is your MMSE. This quantity is basically your MMSE that is the minimum mean squared error for the estimation of the channel coefficient h. Now, we are going to simplify it and you already know the principle that we use in simplification. We are going to employ a simplification for this and this part can be interestingly simplified as, you can recollect again from the previous modules that sigma h square times x bar transpose. We have demonstrated this also x bar transpose into sigma h square x bar x bar transpose plus sigma square times identity inverse, this is equal to sigma h square x bar transpose x bar plus sigma square inverse times sigma h square x bar transpose.

This is something that we have already shown and now you can see sigma h square x bar transpose x bar is a scalar quantity because x bar, remember we justified this x bar transposes a row vector, x bar is a column vector. So, x bar transpose x bar is a scalar quantity, in fact norm x bar square. So, that is sigma h square norm x bar square plus sigma square which is a scalar quantity and the inverse of this scalar quantity will simply be its reciprocal. In fact, from several principles we approved it. In one of the modules, this also follows from the Matrix Inversion Lemma and you can derive it again the

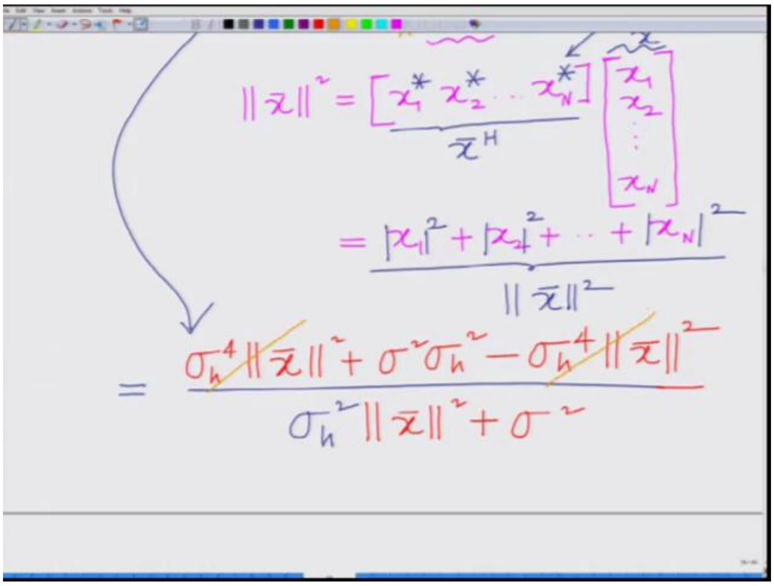
Matrix Woodbury Identity or the matrix inversion lemma and therefore, this is the simplification.

(Refer Slide Time: 12:17)



So, now, we will use this. Let us call this star, this quantity is also star. Now, we will simplify the MMSE by substituting the simplification for this quantity star derived above. This is equal to sigma h square minus sigma h square x bar transpose divided by sigma h square norm x bar square plus sigma square and x bar sigma h square which is equal to sigma h square minus sigma h to the power of 4. Now, you have x bar transpose times x bar that will give a norm x bar square divided by sigma h square norm x bar square plus sigma square and in case you are wondering norm x bar square is this quantity is nothing but just to remind you norm x bar square.

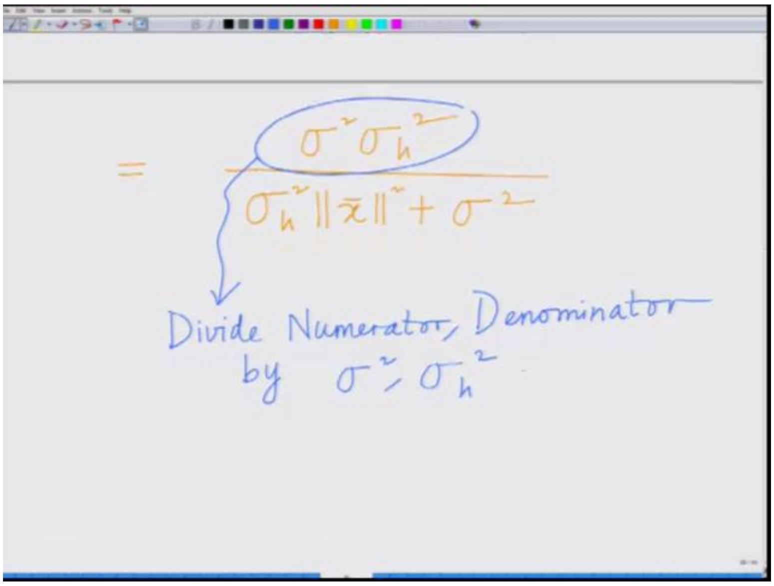
(Refer Slide Time: 14:04)



This quantity is nothing but the row vector x 1 x 2 x N times the column vector x 1 x 2 x N and this will be equal to your x 1 square plus x 2 square plus so on plus x N square. Later, when we consider the complex scenario all we have to do is you have to replace this x 1 x 2 x N by their conjugates. So, this will become x 1 bar or x bar hermitian, this will still be x bar, this is for the complex case remember for the real case of course, complex conjugate will yield the. So, this is for the complex case and this will then be magnitude of x 1 square plus magnitude of x 2 square so on up to magnitude of x N square and this is the quantity we are calling nothing but norm x bar square.

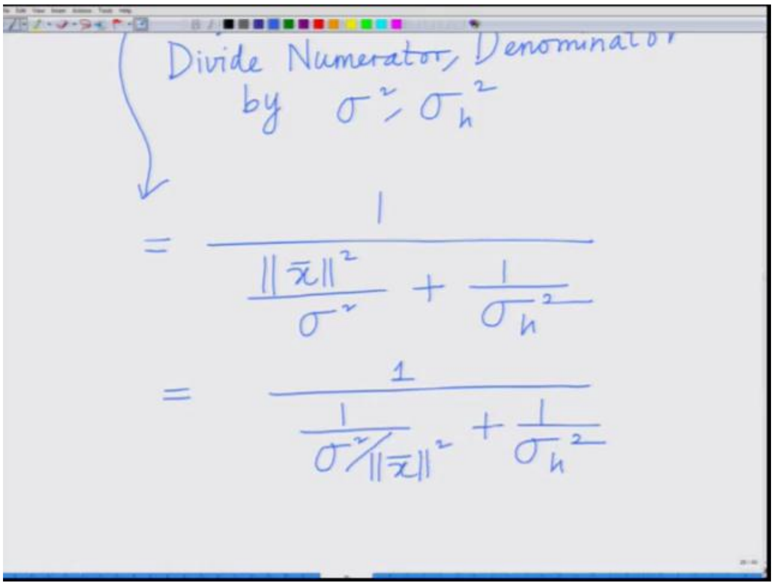
So, norm x bar square for the real vector x bar is x 1 square plus x 2 square so on up to x N square. For a complex vector, that is vector with complex pilot symbols, complex entries, this is simply magnitude x 1 square plus magnitude x 2 square so on up to magnitude x N square, that is the only difference and therefore you can simplify this part. You can simplify it, just carry forward from where we have left off that will be given by you know, now you can take this common sigma h square times norm x bar square plus sigma square sigma h 4 norm x bar square plus sigma square sigma h square minus sigma h to the power of 4 norm x bar square and this sigma h power 4 norm x bar square sigma h power 4 norm x bar square.

(Refer Slide Time: 16:50)



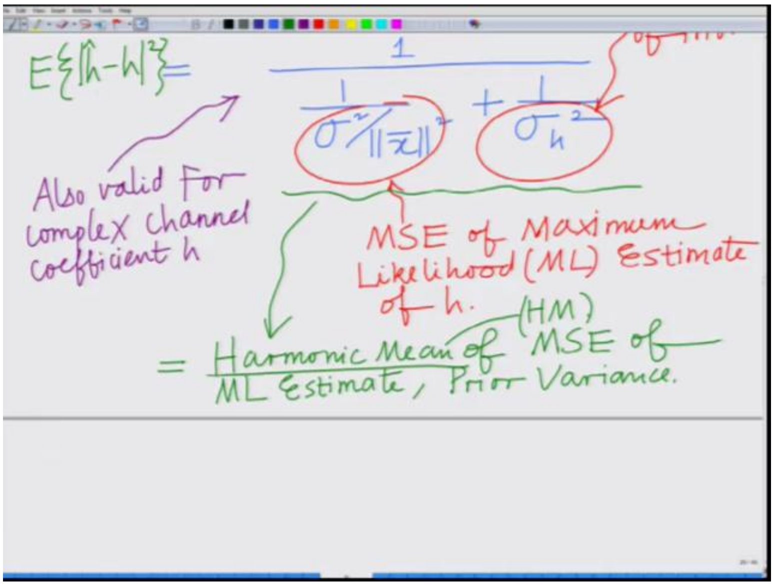
These things cancel and you are left with sigma square sigma h square divided by sigma h square norm x bar square plus sigma square and now we have this expression.

(Refer Slide Time: 17:43)



Now, we divide numerator and denominator by sigma square comma sigma h square and this when you divide numerator and denominator by sigma square sigma h square, what this will give you is of course numerator will become 1, the denominator will become norm x bar square divided by sigma square plus 1 over sigma h square, you can verify this and now what I am going to do? I am just going to rewrite this as follows, I am going to bring the norm sigma square to the denominator and that can be written as 1 over and this can be interestingly written as 1 over sigma square divided by norm x bar square plus 1 over sigma h square. Now, this expression is also valid for the case where the channel coefficient h is complex. All you have to substitute is basically substitute norm x bar square by magnitude x 1 square plus magnitude x 2 square so on up to magnitude x N square.

(Refer Slide Time: 18:52)

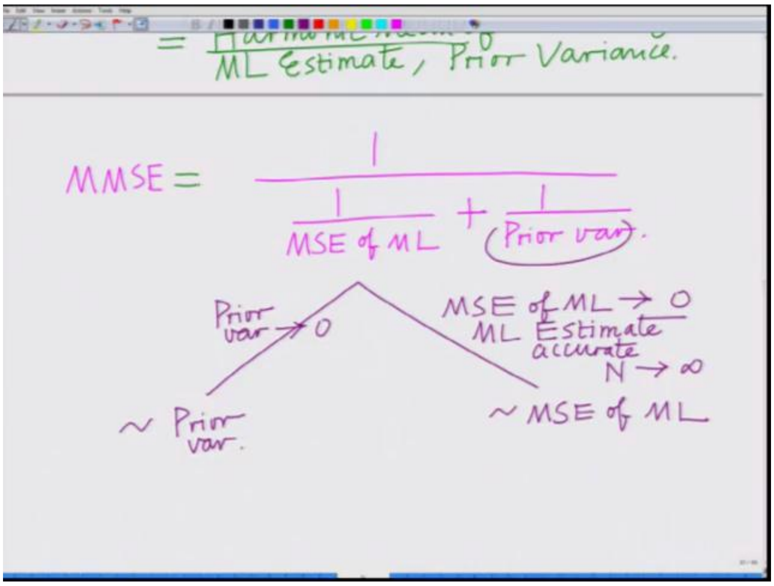


So, let me make a note of that, this is also valid for the complex coefficient h. This is also the MMSE for the estimation of a complex channel coefficient h, it is also valid for a complex and now you can see what is this quantity again, this quantity is the MSE or mean square error of the maximum likelihood or ML estimate and this quantity is the variance of prior that is h is Gaussian channel coefficient with mean mu h and prior variance sigma h square. So, this is basically now you can see the harmonic mean, this is equal to harmonic mean of MSE-Mean Square Error of the ML estimate comma the prior variance.

This is the harmonic mean or basically you can also say the HM right, harmonic mean is nothing but basically just abbreviate this as HM, the harmonic mean of the MSE of the maximum likelihood estimate and the prior variance and remember this is also valid for a complex channel coefficient h except for a complex channel coefficient instead of h hat minus h square, you have to look at magnitude h hat minus h square. For complex of course, this is valid for real coefficient h also, you have to simply look at expected value of magnitude h hat minus h whole square and what you have seen is that the net variance is the harmonic mean of the variances of the ML estimate and of the prior variance and of course all the insides we have drawn for the wireless sensor network case are also valid.

For instance, the harmonic mean of two quantities is less than each of the quantities. So, the harmonic mean of the variance of the ML estimate and the prior variance is less than what the variance of the ML estimate and also less than the prior variance. So, by combining the information from the observations and also basically the prior information, you are getting an estimate which has a lower variance than both the ML estimate which is derived simply from the observations, the prior mean which is derived simply from the prior. So, by combining these two in an MMSE fashion, you are able to get a variance which is lower than both and of course, the other intuition is also valid which is let me again remind this, I mean at the risk of being repetitive, the MMSE equals 1 over 1 over MSE of ML plus 1 over the prior variance.

(Refer Slide Time: 22:13)



If the MSE of the ML estimate is very close to 0 is very small, then the dominating term in the denominator you can observe 1 over MSE of ML will be very large because MSE of ML estimate is very small. So, the denominator will be dominated by 1 over MSE of ML, therefore 1 over 1 over MSE of ML will simply be the MSE of ML. If MSE of ML is equal to zero, you can have again two cases that is, when the MSE of the ML estimate tends to zero that is ML estimate is very accurate and this happens when N tends to infinity, number of observations tends to infinity, this approximately tends to the MSE of ML estimation.

On the other hand, prior variance is very small, it tends to zero. Then, again the denominator is dominated by 1 over the prior variance and therefore, this is approximately equal to simply the prior variance. So, these are the interesting insides again very similar to the wireless sensor network scenario, you can see large number of observations which means the ML estimate is very accurate its variance is very low. So, naturally the final MMSE estimation error will be very close to that of the MSE of the maximum likelihood estimate.

On the other hand, if the prior variance square h square itself is very small which says that you know h hat to h to begin with to a very high degree of accuracy because sigma h square is very small, then the denominator is dominated by 1 over sigma h square and the entire expression will be become approximately sigma x square that is, very similar to the MMSE of estimation is very similar to the prior variance. So, these are the two special cases and for different levels of relative accuracy of the ML estimate versus the prior variance, I mean you get the combination which is in fact the harmonic mean of the MSE of the ML estimate and the prior variance which is lower than each of these quantities. So, that is the interesting aspect and we will start with here.

I will stop here and in the next module we will see briefly what happens when this channel coefficient h is complex, what is the interesting observation there and what is the interesting interpretation specifically for the scenario when the channel coefficient h is complex. So, we will stop this module here.

Thank you.