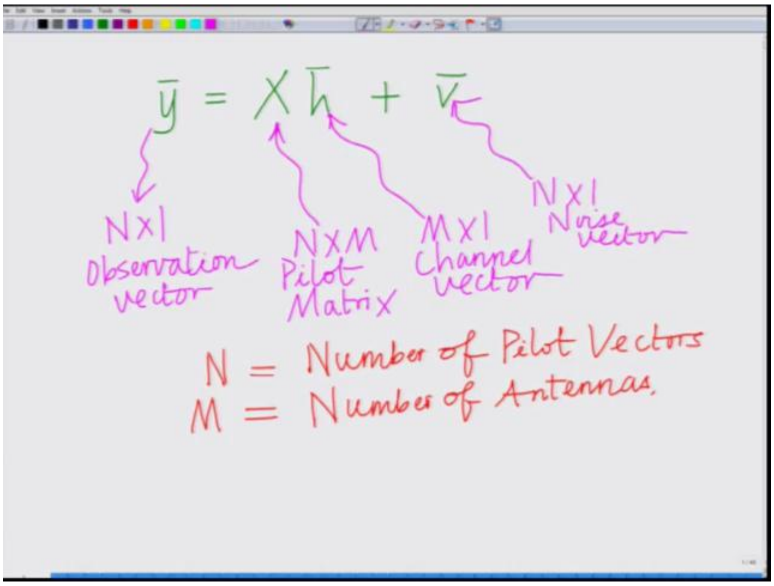
Bayesian/MMSE Estimation for MIMO/OFDM Wireless Communications
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture – 19
Linear Minimum Mean Squared Error (LMMSE) Estimate for Multi-Antenna Downlink Wireless Channel Estimation - Part II

Hello. Welcome to another module in this massive open online course on Bayesian MMSE estimation for wireless communication. So, we are looking at MMSE estimation for a vector parameter, and as a particular example, or as a particular instance of its application, we are considering the estimation of a downlink wireless, of a downlink multi antenna wireless channel; that is, we are considering the base station which has multiple antennas, m antennas to be more specific, and a user who has a single antenna; and, we are considering transmission from the base station to the user; that is, we are considering the downlink, and we have illustrated how to perform channel estimation for this downlink.

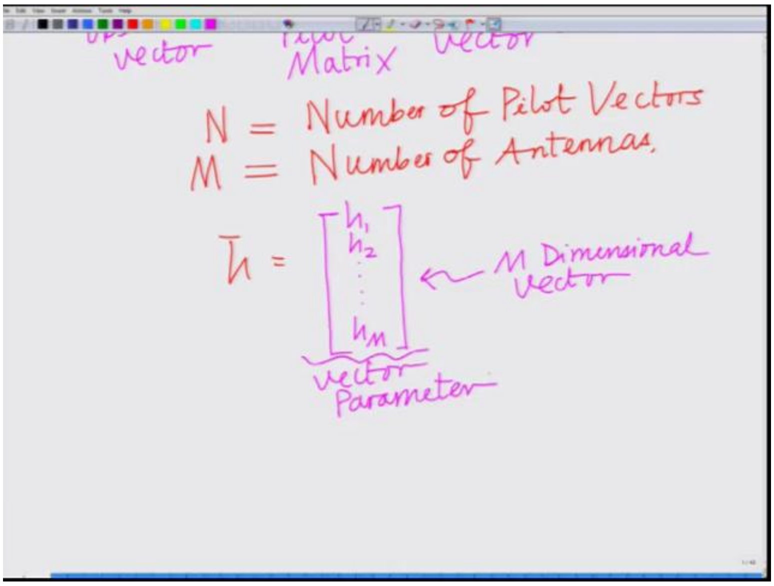
(Refer slide Time: 00:55)



So, the system model that we have considered is your y bar equals X h bar, plus v bar. This y bar is your N cross 1 observation vector; X, this is the N cross M pilot matrix. Recall that, your N equals number of observation, or basically, number of pilot vectors;

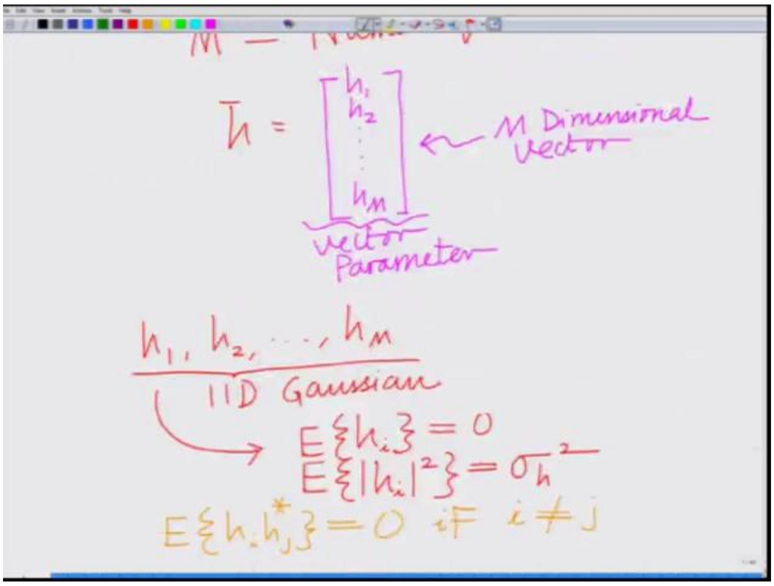
that is, basically your number of. And, M equals to number of antennas; h bar is the M cross 1 channel vector; and, v bar is the N cross 1 noise vector. h bar is, naturally, it is an M cross 1 channel vector. It contains the m channel coefficients, one channel coefficient corresponding to each transmit antenna, and the single receive antenna, alright.

(Refer slide Time: 02:53)



So, for instance, if we consider ith transmit antenna, the channel coefficient between the ith transmit antenna at the base station, and the single antenna of the user, this is denoted by the coefficient h i. So, h bar is this vector parameter; that is what we are talking about. h bar is equal to, if you recall, this is basically h 1, h 2, h m; this is your m dimensional vector, and therefore, we have vector parameter. And further, we are considering the Bayesian vector parameter. Recall that, we are considering a Bayesian vector parameter which, in which, we are assuming each coefficient h i to be Gaussian in nature, with zero mean and variance sigma h square, alright. So, this is a Bayesian, MMSE estimation scenario, alright. So, there is some prior information about these channel coefficients, or this channel vector to be estimated.

(Refer slide Time: 03:48)

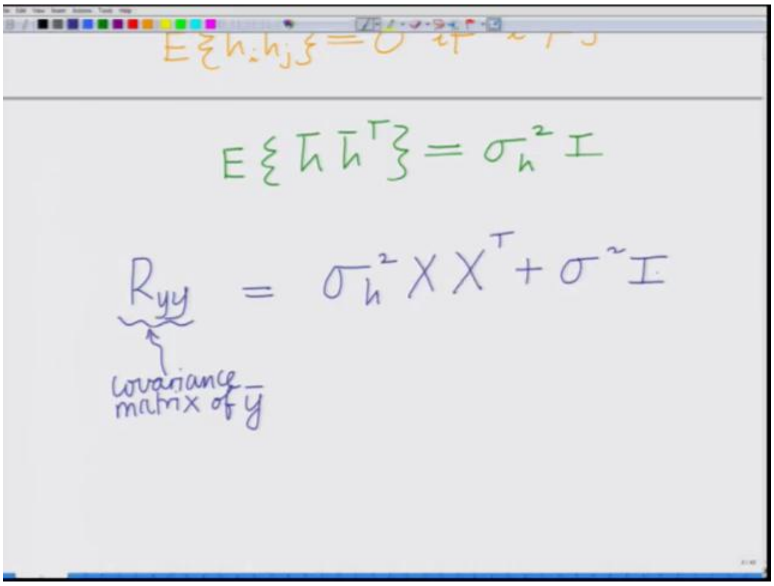


So, we are considering h 1, h 2, h m to be more specific. We are considering h 1, h 2, h m, these are IID Gaussian, which means, basically, independent, identically distributed Gaussian, with expected, each h i equal to zero, expected value; that is, the mean of each coefficient h i equal to zero; the variance, expected magnitude h i square equals sigma h square. Further, they are independent, which means, your expected value of h i into h j, if they are complex, h j conjugate equals zero, if i is not equal to j.

Now, currently, we are considering only real symbols, alright. So, I can simply call this as expected value of h i times h j. But, if I have complex coefficients, and we can use this MMSE estimation framework also when the channel coefficients are complex, alright. I will shortly show you how to extend this to a complex vector parameter estimation scenario; and, it is fairly simple; it is a relatively straight forward extension from the estimation of the real parameter to estimation of a complex parameter, alright. So, expected value of h i times h j equals zero, if i is not equal to j, and the parameters, h 1, or, the coefficients h 1, h 2, up to h m are real parameters.

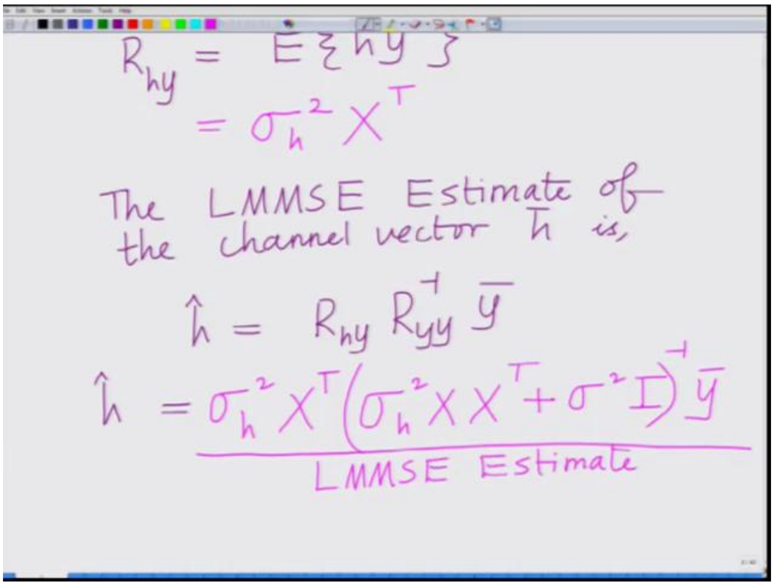
We are considering a real estimation scenario so far. And therefore, the covariance, if you look at the covariance, we have expected value of h bar into h bar transpose, once again, considering real parameters, that will be, sigma h square times, sigma h square times identity. And, we also derived for this model that, we have given over here, that is, we have your y bar equals X h bar plus v bar. We have derived the various quantities; we have derived the covariance matrix of your, the covariance matrix of the observation vector y; that is, R y y, this is the covariance of observation, and this, we have said is sigma h square X X transpose plus sigma square times identity.

(Refer slide Time: 06:12)



We have also derived the cross covariance, or the cross covariance between R h y. This is nothing, but expected value of h bar y bar transpose; this covariance matrix, this is nothing, but expected value of y bar y bar transpose.

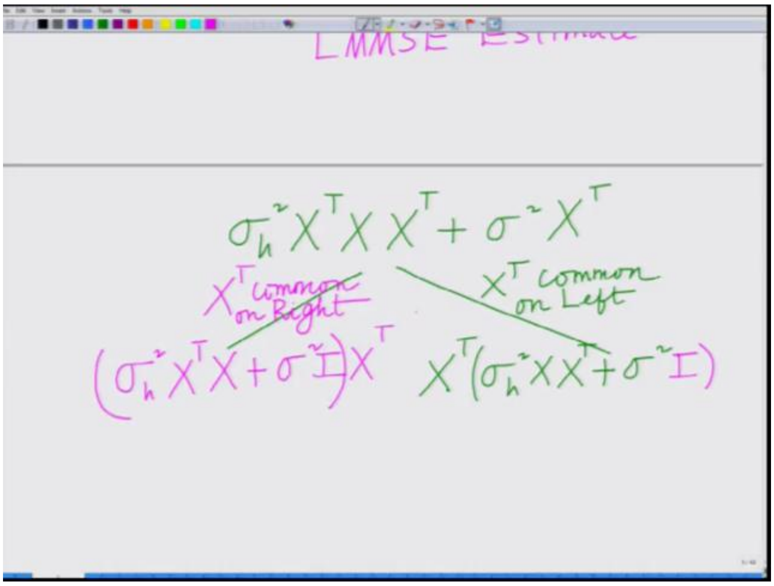
(Refer slide Time: 06:45)



So, this is expected value of y bar y bar transpose; this is expected value of h bar y bar transpose, which we have derived as sigma h square times X transpose. Therefore, the LMMSE estimate, remember, even when it is Gaussian, when the channel coefficient are non– Gaussian, we can talk about the LMMSE estimate. Of course, if they are Gaussian, then, the LMMSE estimate itself is the MMSE estimate; that is, the minimum mean square error of the estimate. Anyway, considering a general scenario, probably not necessarily Gaussian channel coefficients h, the LMMSE estimate, the LMMSE estimate, the L LMMSE estimate is h hat equals, this is the estimate; that is, R h y into R y y inverse into y bar, which is h hat equals.

Now, I have to substitute R h h; this is sigma h square X transpose, times sigma h square, X X transpose, plus sigma square times identity inverse into y bar. So, this is the expression for the LMMSE estimate. This is the expression for the LMMSE estimate. Now, of course, the LMMSE estimate itself is the MMSE estimate, if the channel vector h h bar is Gaussian, alright.

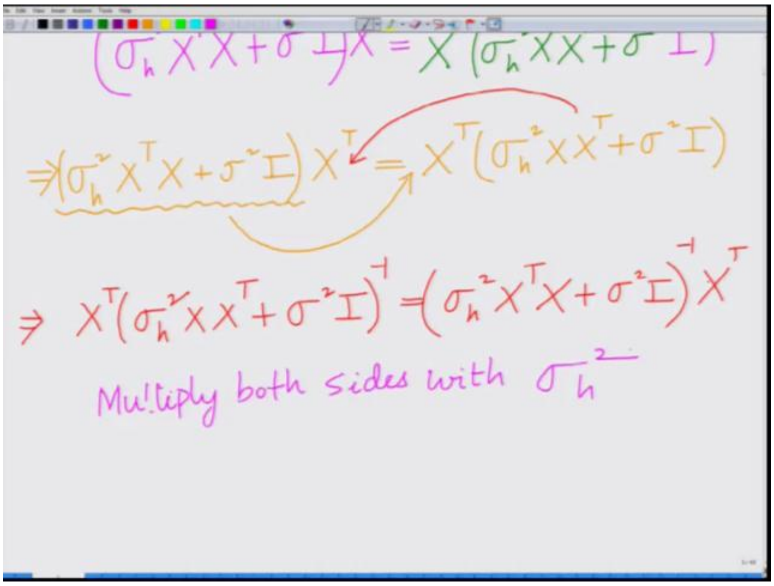
(Refer slide Time: 09:19)



Now, what we are going to do is, let me simplify this LMMSE estimate in a form, that is much more convenient, and I am going to illustrate that. So, now, first realize, and again, similar to what we have done several times before, realize, I will start with this quantity, sigma h square X transpose X X transpose plus sigma square times identity; and, I can expand this in two ways. I can either, I can take X transpose common on the left. So, taking X transpose common on the left that will give me X transpose sigma h square X X transpose plus sigma square. I have changed this identity; this should be X transpose; that will be sigma square times identity.

Now, I can also take a look at this; I can also take X transpose common on the right. So, X transpose, taking it outside on the right, what I am going to have, I am going to have sigma h square X transpose X plus sigma square identity times X transpose. And naturally, since both these quantities are derived from the same quantity, therefore, I have these two quantities are equals, since both are derived from same quantities, which implies basically, what I have, let me just write this clearly; I have sigma h square.

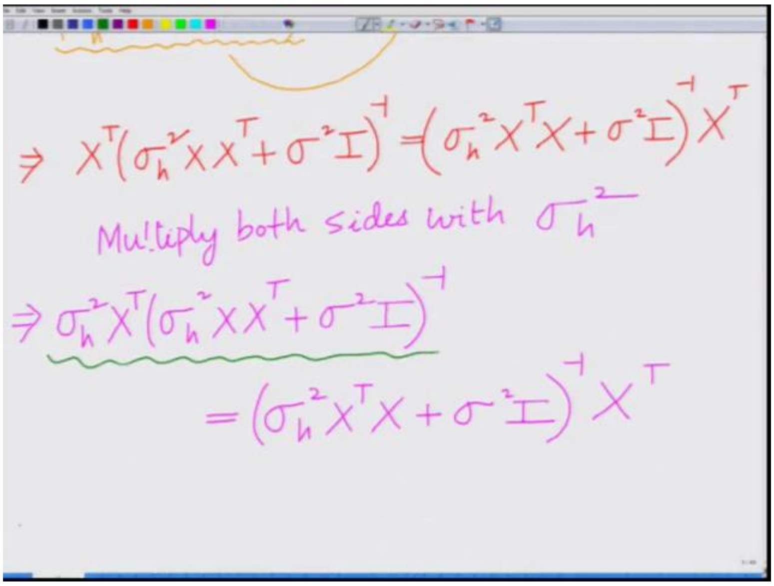
(Refer slide Time: 11:02)



Let me just write it; sigma h square X transpose X plus sigma square identity into X transpose equals X transpose sigma h square X X transpose, plus sigma square identity. And now, since both these quantities are equal now, I can multiply with this inverse of sigma h square X transpose X plus sigma square identity on the left. So, this will come over here; this will become the inverse, and come over here, and sigma h square X X transpose sigma square identity, I can take its inverse, and I can bring it on the right over here. And therefore, now, what you will observe, you can observe that, what i have is, well, X transpose times sigma h square X X transpose plus sigma square identity inverse; this is equal to sigma h square X transpose X plus sigma square identity inverse times X transpose. And, this is the result that we have.

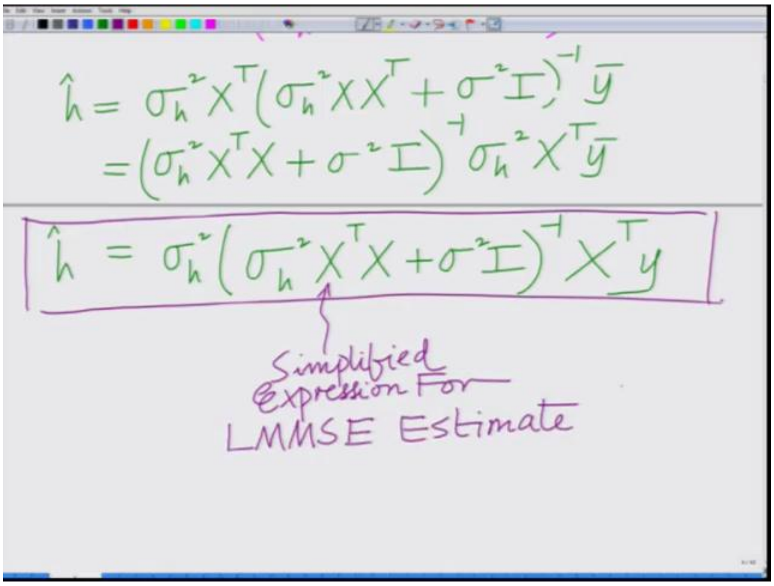
Now, I can multiply both sides with sigma h square; that is the last step. Multiply both sides with, multiply both sides with sigma h square, and that implies, that is very simple of course, sigma h square, this is the scalar.

(Refer slide Time: 13:00)



So, I can simply multiply both sides by sigma h square. So, I have sigma h square X transpose sigma h square X X transpose plus sigma square identity inverse; this is equal to sigma h square X transpose X plus sigma square identity inverse times X transpose. And now, if I look at this, now, look at this quantity. This quantity is sigma h square X transpose times sigma h square X X transpose plus sigma square i inverse. Now, if you look at this quantity, look at that, that quantity is exactly equal to this quantity. If I call this quantity as your star, this quantity is exactly equal to this quantity, is exactly your star.

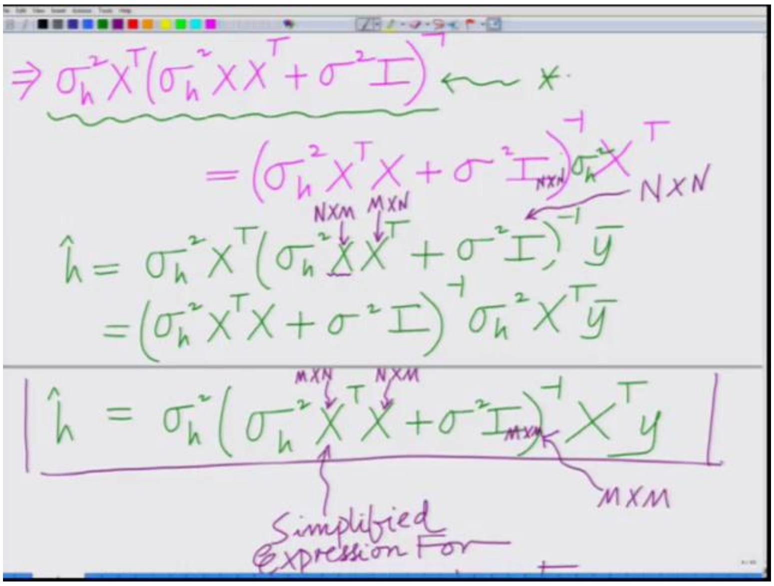
(Refer slide Time: 14:18)



So, this quantity is equal to this quantity, which means, I, can replace this star by sigma h square X transpose X plus sigma square times identity inverse into X transpose. And therefore, I will have h hat equals, previously it was sigma square h square X transpose sigma h square X X transpose plus sigma square identity inverse into y bar. I can equivalently write this as, now, replacing star by sigma h square X transpose X plus sigma square identity inverse times. Of course, I am missing a sigma h square here, because I am multiplying both sides by sigma h square. So, times sigma h square is a scalar quantity. So, I can multiply it anywhere. So, sigma h square X transpose y bar.

In fact, I can simply bring the sigma h square all the way to the left; I can write this sigma h square sigma h square X transpose X plus sigma square identity inverse into X transpose y bar; that is your, that is your h hat, and this is the simplified expression. This is what we are saying as your simplified expression for the MMSE estimate; or rather, LMMSE estimate. This is the simplified expression for the LMMSE estimate. So, h hat equals sigma h square times sigma h square h X transpose X plus sigma square identity inverse into X transpose y. Now, why we are saying this is a simplified estimate, because, look at this. If you look at the first expression, this expression, X into X transpose; look at this. So, if you have X into X transpose, this is an N cross M; X is N cross M; X transpose is M cross N; as a result, this whole matrix will be N cross N.

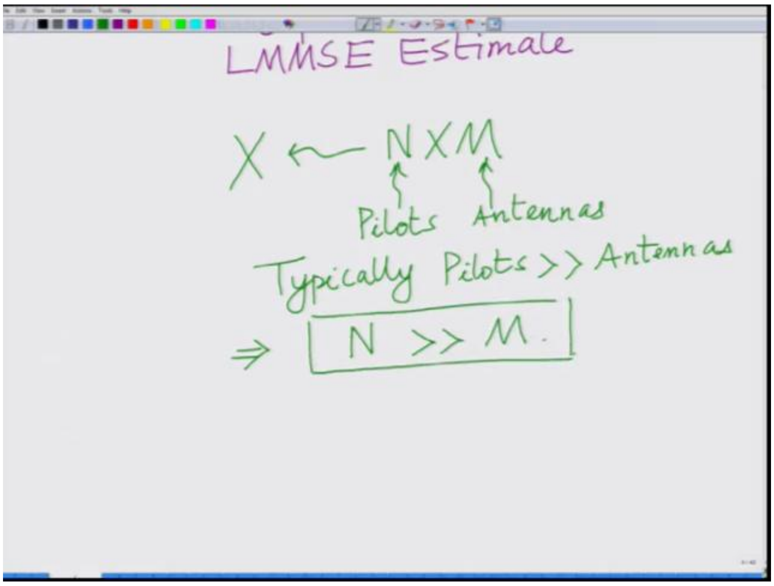
(Refer slide Time: 16:32)



So, this X X sigma h square, X X transpose plus sigma square identity is an N cross N. So, this is the inverse; we have to compute the inverse. In the first technique, we have to compute the inverse of an N cross N matrix. But, if you look at this quantity, X transpose; X transpose is basically, X transpose. This is basically M cross N; X is basically N cross M. So, this is basically your sigma h square X transpose X plus sigma square identity, this is an M cross M. In fact, if you look at this identity, the identity here is M cross M; the identity here is N cross N. So, in the first case, you are inverting an N cross N matrix, where N is the number of pilot vectors.

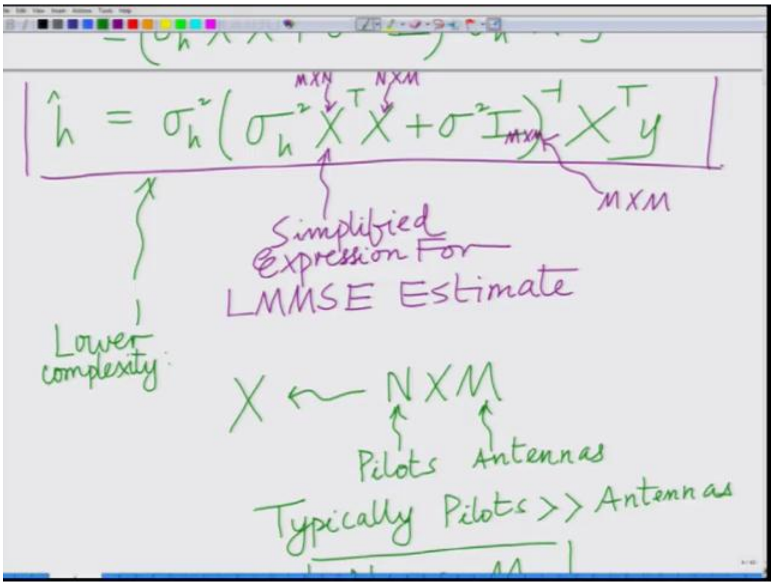
In the second case, you are inverting an M cross M matrix, where M is the number of antennas. And typically, the number of pilot symbols, pilot symbol vectors transmitted, is always much more than the number of antennas; that is, if we look at N cross M, that is, if we look at X; remember, if you look at X, X is N cross M, N is number of pilots; M is antennas.

(Refer slide Time: 18:01)



Typically, number of pilots is much greater than the number of antennas. Number of pilots is much greater than the number of antennas, implies, your N is typically much greater than M. Therefore, therefore, it is much easier to compute the inverse of an M (Refer Time: 18:50), because N is much larger. So, N, first matrix is N cross N; which means, it is much larger matrix. So, it is much more, much more to difficult to inverse, alright. The second matrix is an M cross M matrix, since M is much smaller. So, it is much easier, it is easier to compute the inverse of this M cross M matrix, rather than N cross N matrix, the second one that is, the second implementation, has a much lower complexity.

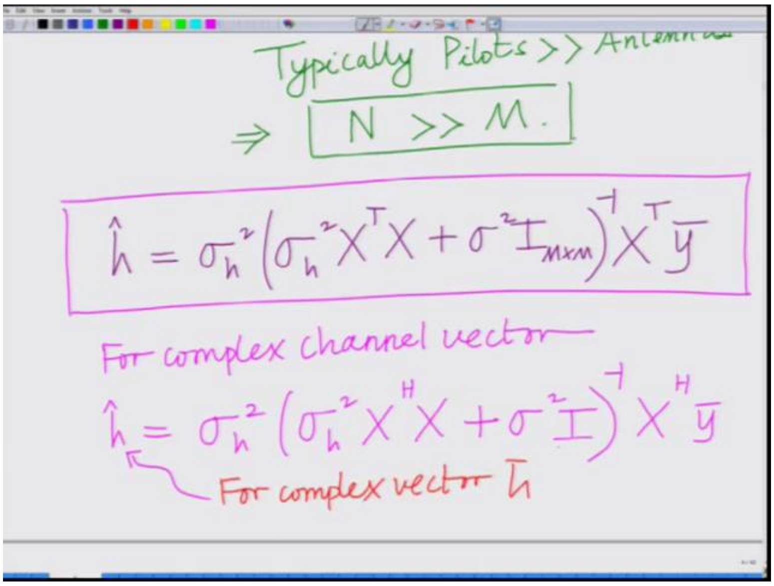
(Refer slide Time: 19:17)



So, this one has a lower complexity. And, in fact, we had seen that, earlier, we had seen a special case of this, where M is equal to 1, when we have a single channel coefficient, then, M cross M becomes simply 1. So, the inverse is simply a scalar, correct. We have, might have N pilots symbols, that is a 100 pilot symbols, but if M equal to 1, M cross M is simply 1 cross 1, which means, it is a scalar quantity. So, inverse is, of a scalar is simply the reciprocal, alright. And, in fact, that is the property that we had used earlier also to simplify the, simplify the LMMSE, or the MMSE estimate for a channel estimation, for the estimation of a fading channel coefficient, with a single antenna, single transmit and single receive antenna.

And now, of course, we are extending to the estimation of a channel vector corresponding to M antennas. So, what we are saying is, we can do this with complexity of inversion of an M cross M matrix, where M is the number of antennas, rather than N cross N matrix, where N is the number of pilots symbols.

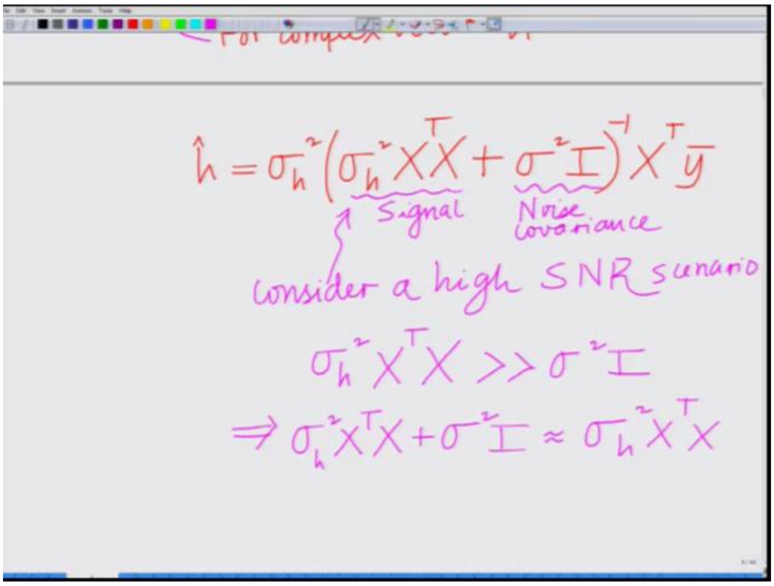
(Refer slide Time: 20:21)



So, let me again write that down, summarize that, the channel, final channel estimate is h hat equals sigma h square times sigma h square X transpose X plus sigma square identity. This is M cross, this is an M cross M identity matrix inverse X transpose into into y bar. So, this is the LMMSE estimate. And now, the additional thing that we can look at here, is that, for complex, I can simply replace that transpose by Hermitian. For complex channel vector, let me just try to write down for a complex channel vector, this is the trick that we use fairly off, and you simply replace that transpose by the Hermitian; that is all you have to do. sigma h square sigma h square X Hermitian X plus sigma square identity inverse x bar Hermitian y bar, this is for a complex vector h bar.

Remember, we said in the beginning that, this can be expanded, extended also, to complex parameters. So, when the channel vector h bar is complex, in relatively straight forward fashion, the simple trick is to be replace that transpose by the Hermitian. So, it will become sigma h square times sigma h square into X X Hermitian X plus sigma square identity inverse times X Hermitian y bar; that is the simple extension to the, when the channel vector h bar is complex; that is, the coefficients are complex, the transmitted pilots symbol are complex; the noise is complex; hence the received observations at the receiver are also complex. For the complex estimation scenario, this is the straight forward exchange.

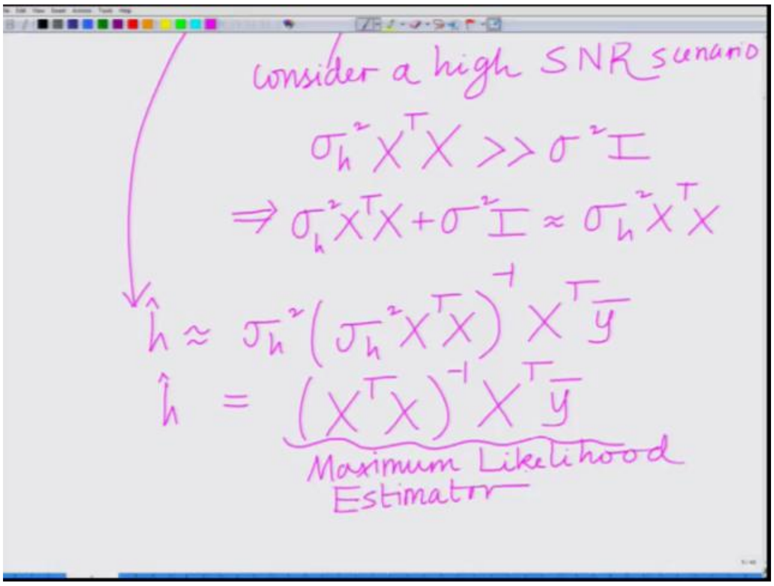
(Refer slide Time: 22:50)



And, one small point to observe here, which is very interesting, that is, we look at this estimate h hat equals sigma h square sigma h square X transpose; going again back to the real scenario; sigma X times X transpose y. Now, consider the scenario where high signal to noise power ratio scenario; consider, consider a high S N R scenario. In a high S N R scenario, remember, what will you have? The signal power is much greater than the noise power. So, look at this, this is the signal component. This is your noise covariance. In high signal to noise power ratio, the signal power, or signal covariance, is much stronger, much stronger, than the noise covariance.

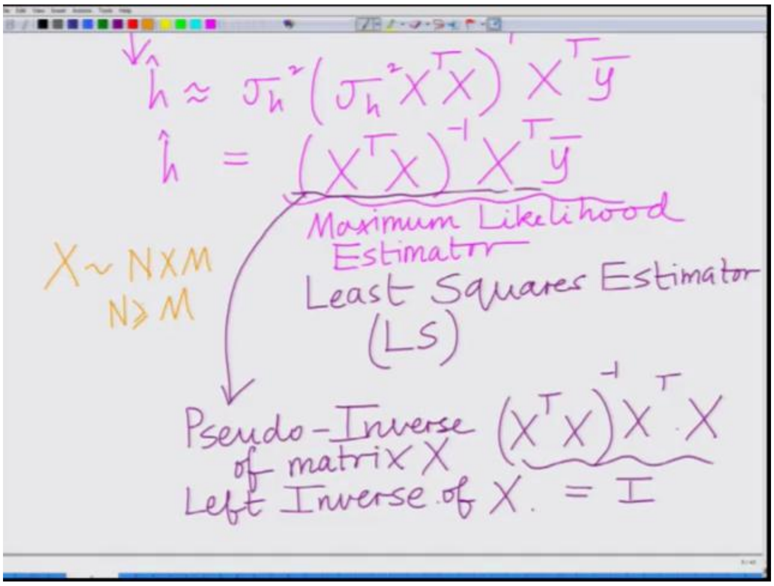
Therefore, what will we have is, this sigma h square will be much larger; sigma h square X transpose X will be much larger than sigma square identity, which implies that, I can neglect, that is, sigma h square, which implies basically, sigma h square X transpose X plus sigma square identity is approximately equal to sigma h square X transpose X; because, the noise power, or noise covariance, is insignificant compared to signal covariance. I can approximate this quantity, sigma h square X transpose X plus sigma square identity simply by sigma h square X transpose X.

(Refer slide Time: 24:46)



And now, you can see something very interesting. Therefore, h will be approximately equal to, or h hat is approximately equal to, sigma h square sigma h square X transpose X inverse times X transpose y bar, which is equal to, now, if I bring the sigma h square outside, this is a scalar quantity. This is equal to simply X transpose X inverse X transpose y bar. And, this is simply, you can recall, this is the maximum likelihood estimator.

(Refer slide Time: 25:44)

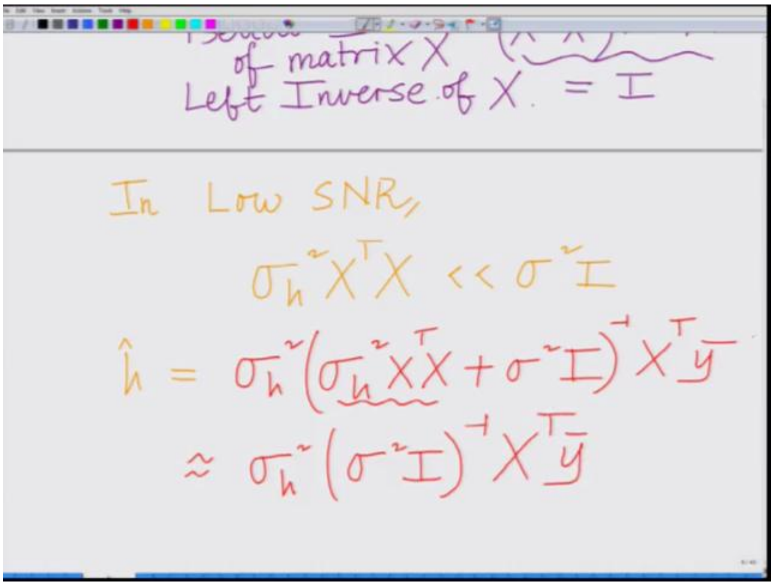


This is the maximum likelihood estimator, and if you look at this, this quantity. In fact, this is also known as the maximum likelihood, or this is known as the least squares estimator, or the L S. Least square is basically short; it can be shortened as L S. And, if you can look at this quantity, X transpose X inverse X transpose, this is known as the pseudo inverse of the matrix X. The reason being, if you look at X transpose X inverse X transpose, and multiply that by X, now, you have X transpose X inverse into X transpose X. This is equal to identity. So, X transpose X inverse X transpose is pseudo inverse, or, this also equal to the left inverse of X of X, pseudo inverse of matrix X. So, X transpose X inverse X transpose is the pseudo inverse of X. It is also the left inverse of X, because if you multiply it on the left of X, you get the identity matrix.

Even though X might not be an invertible matrix, because, N is greater than or equal to M. Because, remember, again, this is very interesting; X is N cross M; and, N we are saying, is greater than, can be greater than M; N can be greater, number of pilot symbols can be greater than the number of antennas. So, if N is equal to M, of course, it is a square matrix, and can be inverted, alright. If inverse exists; but, if N is greater than M, then, no inverse exists for X. Therefore, this matrix which we will multiply by on the left by X gives the identity; this is known as pseudo inverse, or basically, this is acting as an inverse of X. This is the pseudo, or sort of a fictitious inverse of this matrix X. And, what is interesting is, at the high signal, in a high signal to noise power ratio, this MMSE estimator again reduces to the least squares estimator, that is the maximum likelihood estimator alright.

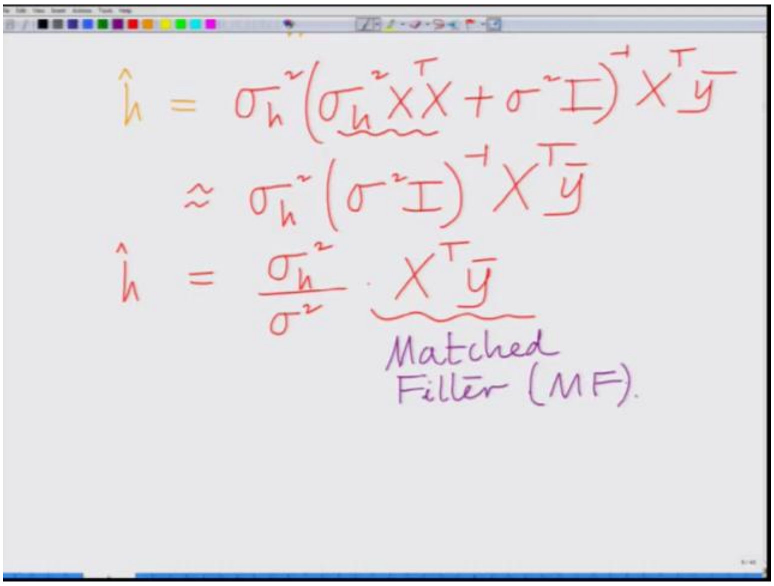
What is more interesting, what is also interesting is, what happens in low signal to noise power ratio, in low S N R, in this case, the signal covariance will be very small compared to the noise covariance; that is, sigma h square X transpose X is much smaller than sigma square times identity.

(Refer slide Time: 28:04)



Therefore, we have h, we observe closely, h hat is sigma h square sigma h square X transpose X plus sigma square identity inverse X transpose y. We are saying this is very small; this is much smaller compared to sigma square identity. Therefore, this can be approximated as sigma h square times sigma square identity inverse X transpose y bar. And, I have to write y bar over here, and what is this? Now, if you look at this sigma square, it is a scalar.

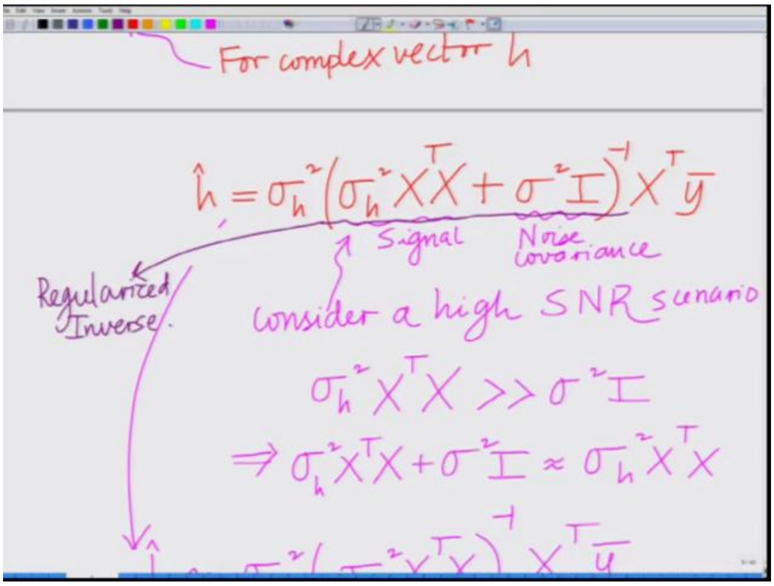
(Refer slide Time: 29:08)



So, this is simply sigma h square divided by sigma square times X transpose y bar. This is simply X transpose y bar; this is simply a matched filter, or your M F, because your signal is X, right; your y bar equals X h bar plus v bar; you are simply multiplying by X transpose into y bar. And, of course, it is a scaling factor, right; sigma h square divided by sigma square; this is simply your matched filter. So, it is very interesting; at high S N R, it becomes the least square estimator, the LMMSE estimator; and, at low S N R, it becomes the matched filter. So, that is the, that is a very interesting aspect.

And, the other interesting aspect of this is that, this matrix sigma h square X X transpose X into sigma square times identity, this inverse of this, always exists even when X transpose X is not invertible; even when X transpose X is not invertible. So, this is also known as a regularized inverse.

(Refer slide Time: 30:15)



So, this is known as a, the sigma square identity is known as a regularized, regularization term. So, this known as, what it says is, if X transpose X is not invertible, that is, you cannot compute the least squares estimate, all we can, what we can do is, we can add a simple sigma square times identity term to that, compute the inverse of it. So, the inverse becomes stable; this is known as the regularized inverse. So, this LMMSE estimate is beneficial, because the inverse is always stable; even when X transpose X is ill-behaved, it cannot be inverted, sigma h square X transpose X plus sigma square times identity is always invertible, alright. So, the inverse is always stable. So, the LMMSE estimate is stable, alright, which means, your solution can be computed in a very controlled fashion, right, without computing the inverse of, without, unlike least squares when X transpose X is not invertible, then, the solution will become out of bounds, alright. It is similar to taking inverse of zero.

So, the LMMSE estimate will always be a stable estimate, because of the regularization nature of it, where you are adding sigma square identity to X transpose X, and then taking the inverse. So, there are a lot of interesting interpretations of this LMMSE estimate, particularly for this vector scenario, alright. So, with this, basically we have simplified the LMMSE estimate for the computation, for the channel, channel vector h bar, and we also demonstrated several interesting simplification, several interesting observations, what happens in the low S N R regime, what happens in higher S N R regime, etcetera, alright. Then, in the subsequent modules we will do a simple example, to further illustrate this concept.

Thank you very much.