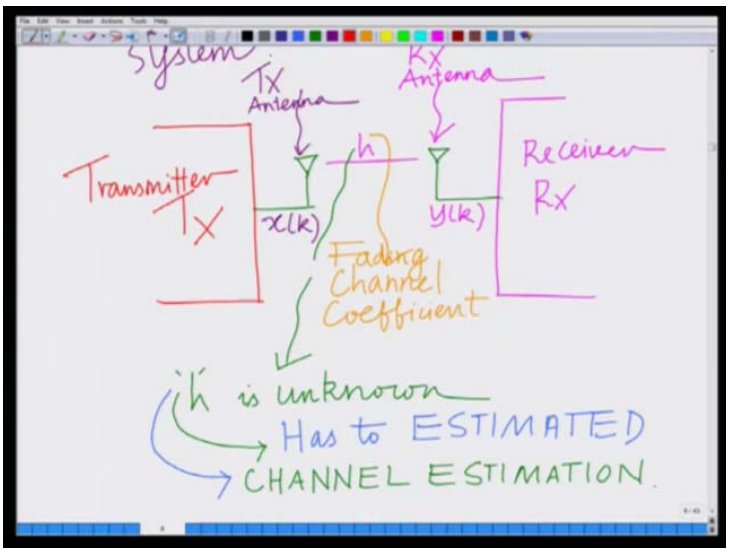
Estimation for Wireless Communications-MIMO/OFDM Cellular and Sensor Networks.
Professor Aditya K Jagannatham.
Department of Electrical Engineering.
Indian Institute of Technology Kanpur.
Lecture -08.
Wireless Fading Channel Estimation - Pilot/Training Based Maximum Likelihood (ML) Estimate.

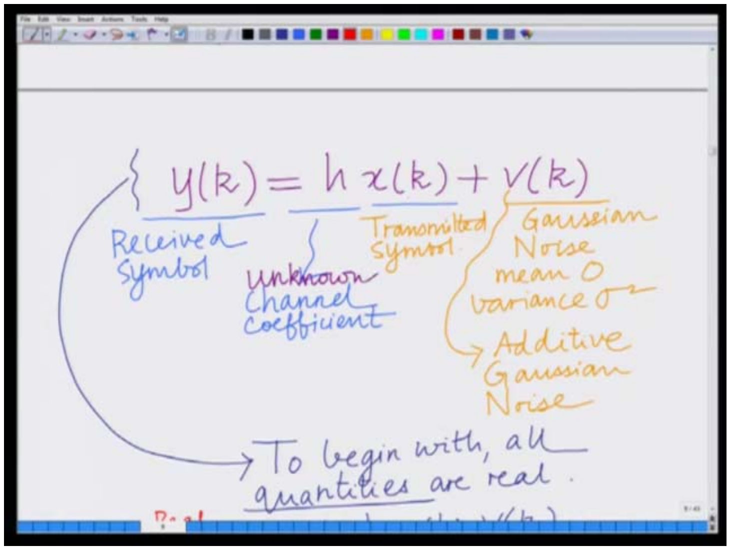
Hello, welcome to another module in this massive open online course on estimation for wireless communications where we are looking at the development of a scheme for the estimation of the unknown channel coefficient h,

(Refer Slide Time: 0:31)



so we said that in a typical wireless communication system, what we have is basically we have the transmitted antenna, receive antenna, the channel coefficient h, XK is the transmitted symbol, YK is the received symbol and

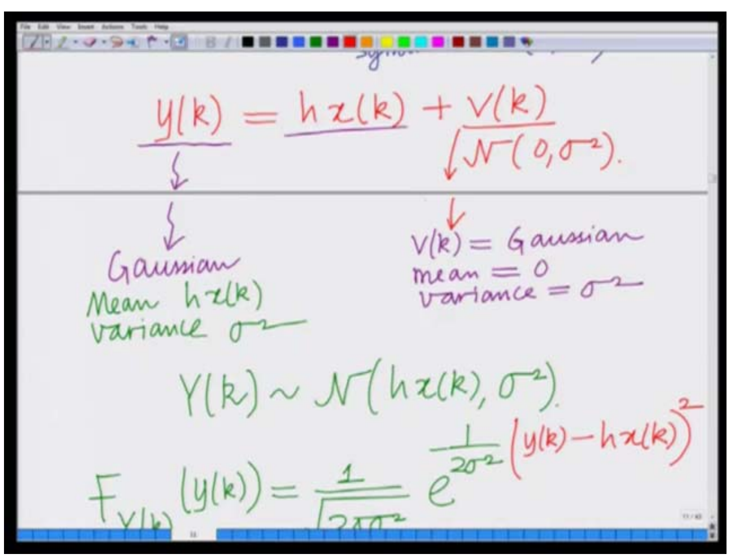
(Refer Slide Time: 0:43)



we have said that this can be represented as YK equals h times XK + VK where VK is the additive Gaussian noise and we also said that these XK, the symbols XK are, transmitted symbols for the, for the purpose of for the purpose of channel estimation, we transmit a sequence of known pilot symbols that is X1, X2,… XN which are also, which are also known as the pilot symbols or the training symbols.

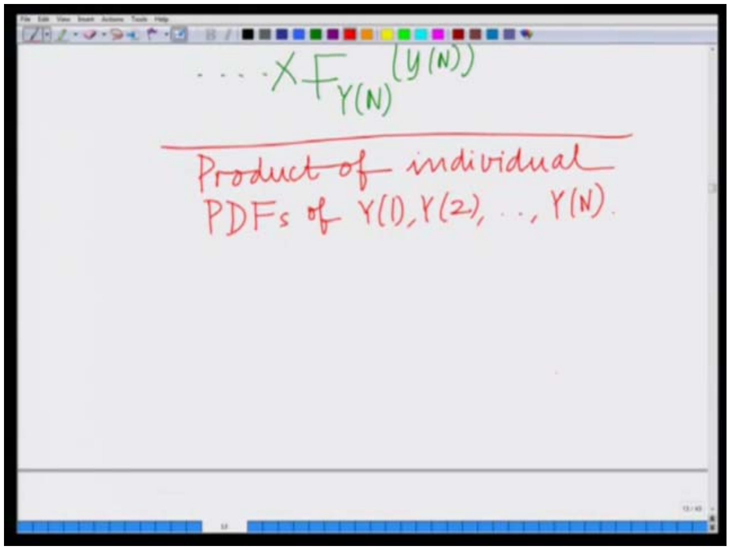
Yet, from the training symbols and the observed corresponding observed training outputs Y1, Y2… YN, we estimate the wireless channel.

(Refer Slide Time: 1:19)



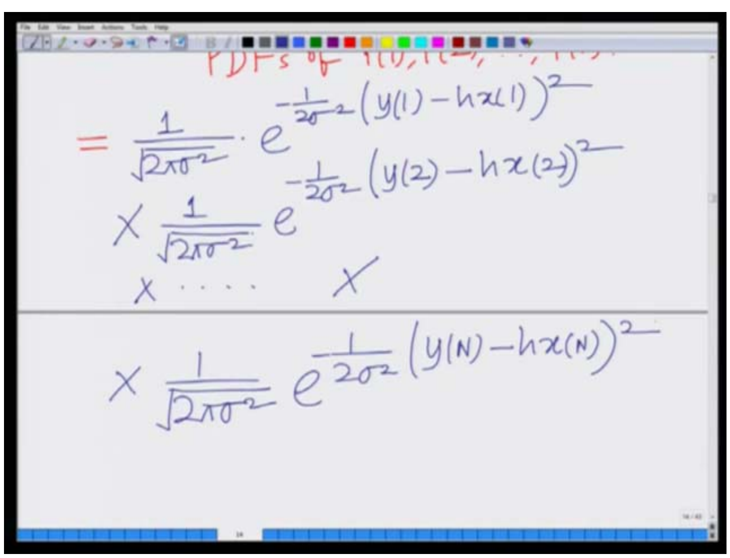
We have described that for each Gaussian noise, that is when noise V1 is Gaussian, that is VK is Gaussian with mean 0 and variance Sigma square, the output YK is Gaussian with mean h times XK and variance Sigma square

(Refer Slide Time: 1:41)



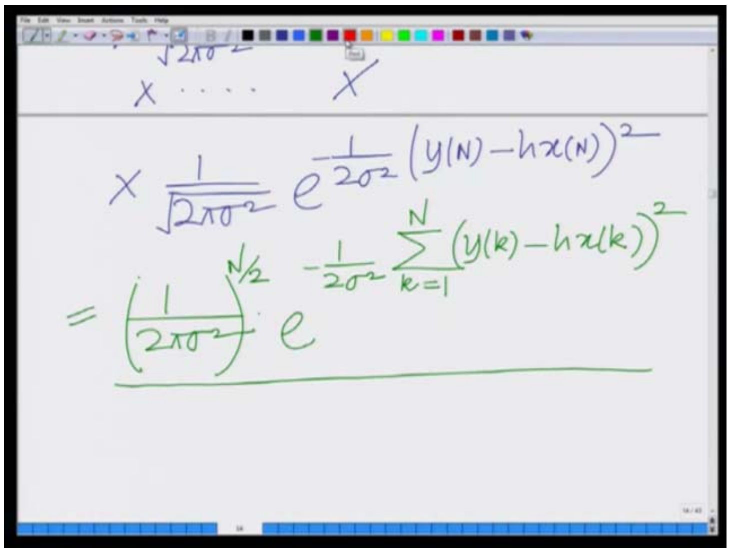
and therefore we have developed the expression for the PDF of the observation, probability density function of the observation YK and we also said that if the noise samples V1, V2… VK are independent, then the observations Y1, Y2… YN are independent and therefore the joint PDF of the observations is given as the product of the individual PDFs or probability density functions which can be written as,

(Refer Slide Time: 1:58)



now we know an expression for the probability density function of Y1 or each YK, therefore I can write this as 1 over 2 pie Sigma square E raised to -1 over 2 Sigma square YK or Y1, 1st we have the PDF of Y1, Y1 - h X1 square times 1 over under root 2 pie Sigma square E raised to -1 over 2 Sigma square Y2 - h X2 whole square, so on and so forth until the product 1 over 2 pie Sigma square E raised to -1 over 2 Sigma square E raised to -1 over 2 Sigma square YN - h times XN square which is basically equal to…

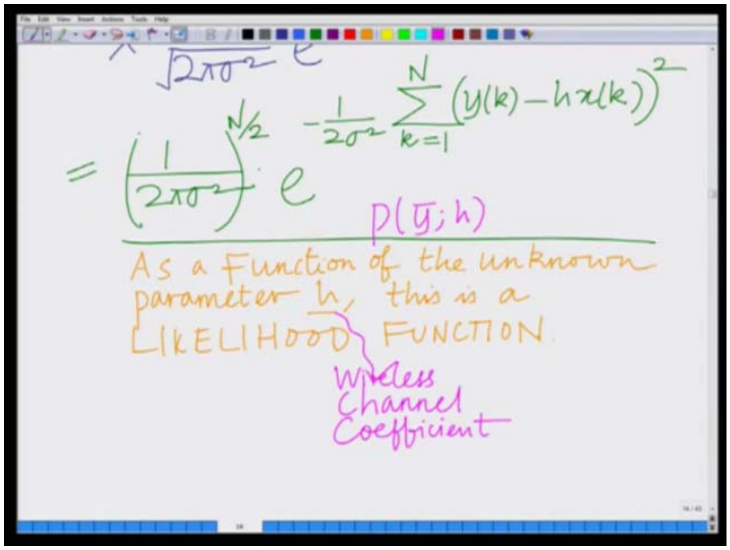
(Refer Slide Time: 3:30)



Now this is basically equal to 1 over, this can be now, combining all the terms and multiplying all the terms, we have 1 over 2 pie Sigma square to the power of N over 2, E raised to -1 over 2 Sigma square summation, all the terms in the exponent will add up. K equals 1 to N, YK - h XK whole square and therefore this is the probability joint PDF of the observations Y1, Y2… YN, this is the joint PDF of the observations Y1, Y2… YN.

Also remember in the context of the sensor network, we said if we view it as a function of unknown parameter h, this is a likelihood function of the unknown parameter h.

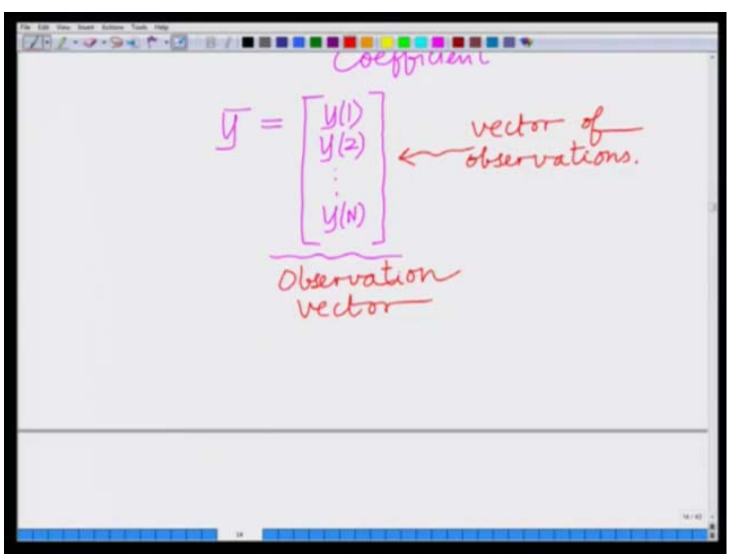
(Refer Slide Time: 4:34)



So, viewed as a function of unknown parameter h, so viewed as a function of unknown parameter h, so as now, as a function, as a function of the unknown parameter h, this is a, what is this, this is your likelihood function.

The likelihood function, likelihood function of the unknown parameter h which is basically or channel coefficient. This h is nothing but the unknown parameter, which is the wireless, in fact this is your wireless, this is your wireless channel coefficient. I can represent this likelihood function as P of Y bar parameterised by h by Y bar is nothing but your vector of observations,

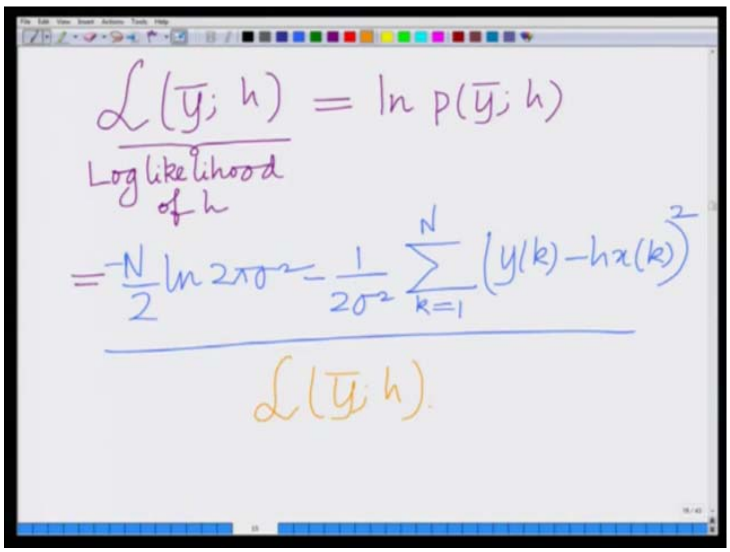
(Refer Slide Time: 5:48)



that is Y bar equals Y1, Y2… Up to YN. So, what is this Y bar Y1, Y2… YN, this is nothing but the vector of observations which is also termed as the observation, this is also termed as your observation vector or the vector of… This is basically the vector of observations.

N dimensional vector of observations that this is the joint PDF of the observations Y1, Y2… YK Y1, Y2… YN, which is viewed as a function of the unknown parameter h, this is termed as a likelihood function. And now I take, now the next step, similar to the sensor network scenario, the next step that we have to do is to take the Logarithm of this likelihood function which is termed as the log likelihood.

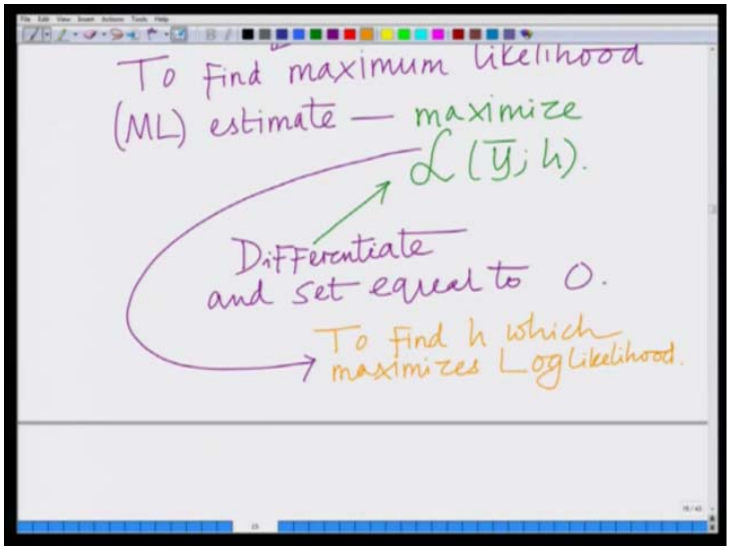
(Refer Slide Time: 7:01)



So, let us go through that step, then it takes the logarithm, it is similar to what we have already seen in the context of a sensor network, the log likelihood, this is the log likelihood function of the unknown parameter h which is the natural logarithm of the likelihood function parameterised by h, which of course now when you should take the logarithm, you can simplify this, you can write this as N over 2 or - N over 2, the natural logarithm of 2 pie Sigma square -1 over 2 Sigma square summation K equals 1 to N YK - hK whole square.

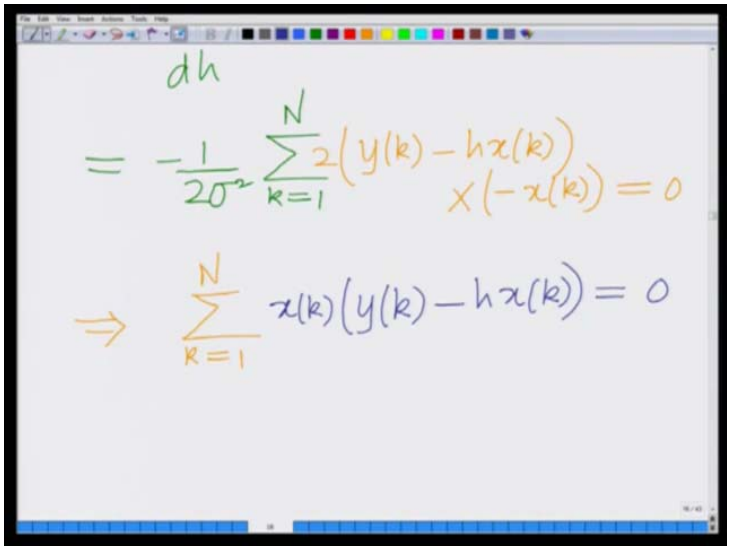
So this basically is your log likelihood function, this is easy to see, this is your log likelihood function. Again following a similar procedure to that we have implied for the sensor network. Find the joint PDF of the observations as a function of the unknown parameter h, it is the likelihood function, take the logarithm of that, you get the log likelihood function. Now, one has to find the maximum of log likelihood function, that is find the value of h for which this log likelihood is maximised, that gives us the maximum likelihood estimate of the unknown parameter h.

(Refer Slide Time: 8:33)



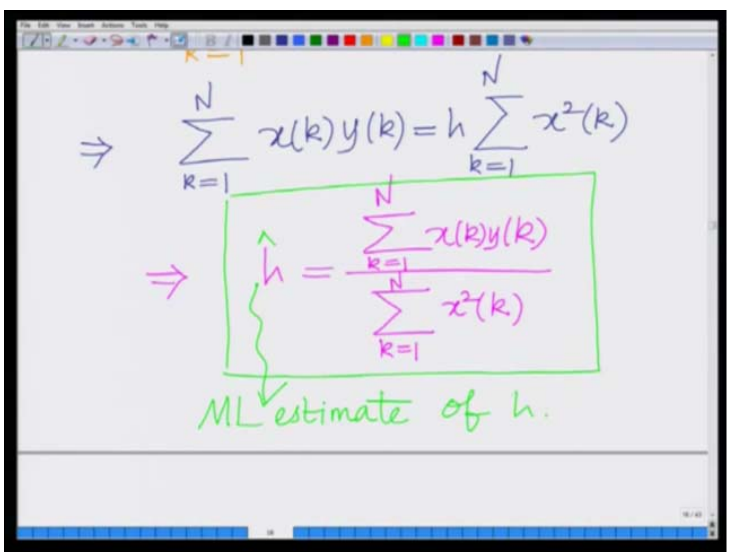
That is to find the ML estimate, to find maximum likelihood or basically ML estimate, one has to basically maximise the log likelihood function, that is Y bar h. Which means basically differentiate this, differentiate and set equal to 0. To find the value of parameter h which maximises the value of log likelihood function, that is to find h which maximises your, to find the value of X which maximises the log likelihood function. And now if we differentiate the log likelihood function, if I differentiate this log likelihood function and you can check with respect to h what we have is when we differentiate this log likelihood function, this - N over 2 log 2 pie Sigma square, this is a constant, so the derivative of this is 0,

(Refer Slide Time: 10:19)



what remains is basically your -1 over 2 Sigma square summation K equals to equal to 1 to N, YK - h, twice YK - h XK times multiplied by - XK and that has to be equal to 0 and this is 0, implies one over Sigma square goes away, that is a constant implies K equals 1 to N XK times YK - h XK equals 0.

(Refer Slide Time: 11:19)

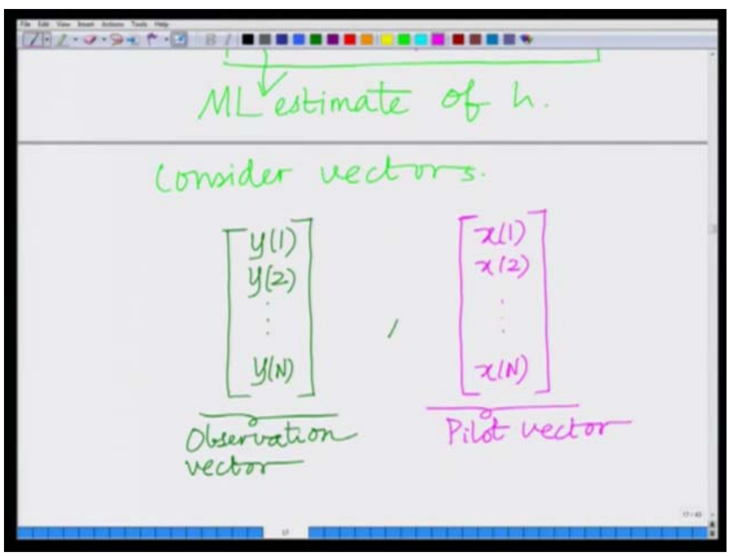


And this implies, basically now, you can simplify this, this implies basically that summation K equals 1 to N XK times YK equals h times summation K equals to 1, XK into XK is X square of K implies the value of h for which this is maximised, is summation K equals to 1 to N XK YK divided by summation K equals to 1 to NX square K and this value of h, where this is maximised, this is the maximum likelihood estimate and this is denoted by h hat which is basically your ML estimate or the maximum likelihood estimate, this is the maximum likelihood estimate of the channel coefficient h.

So, this h hat, with a summation K equals to 1 to N XK YK divided by summation K equals to 1 to NX square of K, this is the value of h for which the log likelihood function of the parameter h that is maximised, we obtained this by differentiating the log likelihood function and setting equal to 0, this value of h, that is when the log likelihood is maximised, it is denoted by h hat and this is known as the maximum likelihood estimate of the channel coefficient.

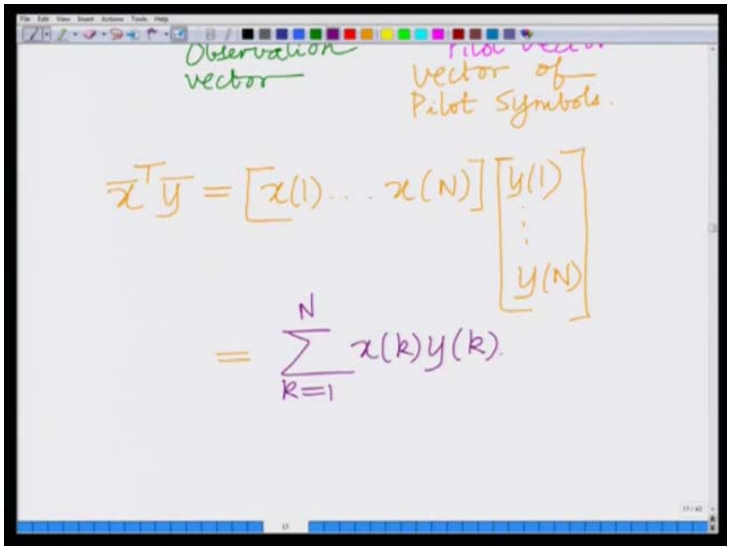
And also one can now although we derived this expression, , one can now be present this in a much more comprehensive manner or in a much more compact fashion by using vector notation.

(Refer Slide Time: 13:20)



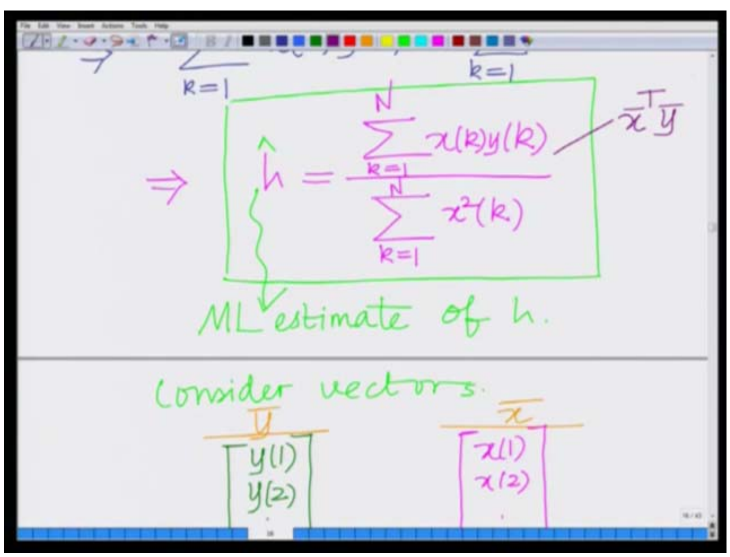
Consider, one can represent it in much more compact fashion using vector notation, so consider vectors, we already find this Y1, Y2… Up to YN, this is the observation vector, remember, in the previous module we have already defined as, this is the observation vector and now similarly let us define a pilot vector.

(Refer Slide Time: 14:15)



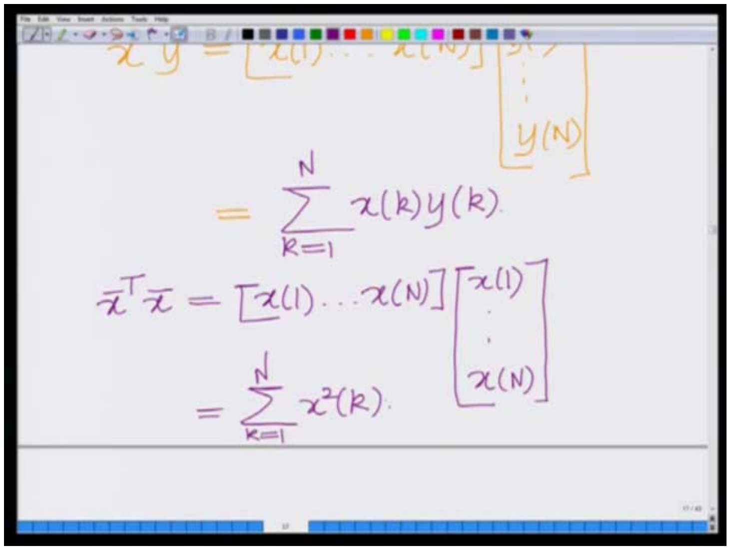
This is a vector of pilot symbols, this is your pilot symbol pilot vector or this is basically your vector of pilot symbols, this is the vector of pilot symbols, if I look at, I am going to denote this observation vector already we have represented this observation vector by Y bar, let us denote this pilot vector by X bar that is X1, X2,… XN. Then, now you can look at it, if I look at X bar transpose Y bar, that is basically equal to your row vector X1 up to XN times your column vector Y1 up to YN which is basically nothing but this.

(Refer Slide Time: 15:16)



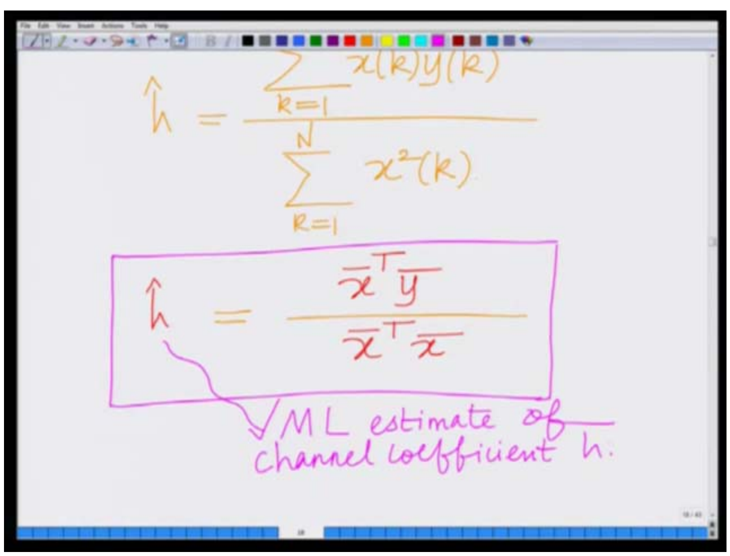
You can clearly see, this can be simplified as summation K equals to 1 to N XK times YK. Therefore the numerator of this expression, that is if I look at the numerator of this expression, that is X bar transpose Y bar.

(Refer Slide Time: 15:27)



Similarly if I look at the denominator, consider now X bar transpose X bar, that is equal to the row vector X1 up to XN times the column vector Y1 sorry times the column vector X1 up to XN and this is nothing but summation K equals to 1 to N X square K and you can see this is basically the denominator, that is X bar transpose X bar,

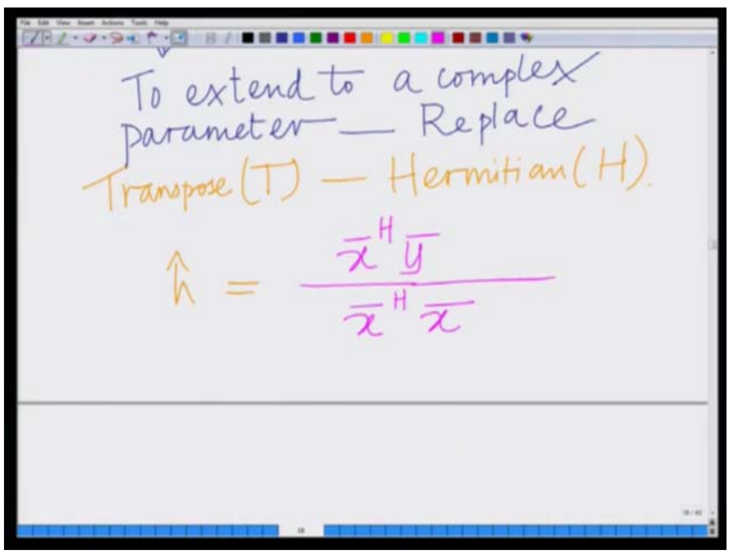
(Refer Slide Time: 16:19)



therefore the channel estimate h hat can also be written as h hat equals summation K equals to 1 to N XK YK divided by summation K equals to 1 to N X square of K which is basically as we have seen can now be simplified as X bar transpose Y bar divided by X bar transpose X bar and this is another interesting or another compact way to represent the maximum likelihood.

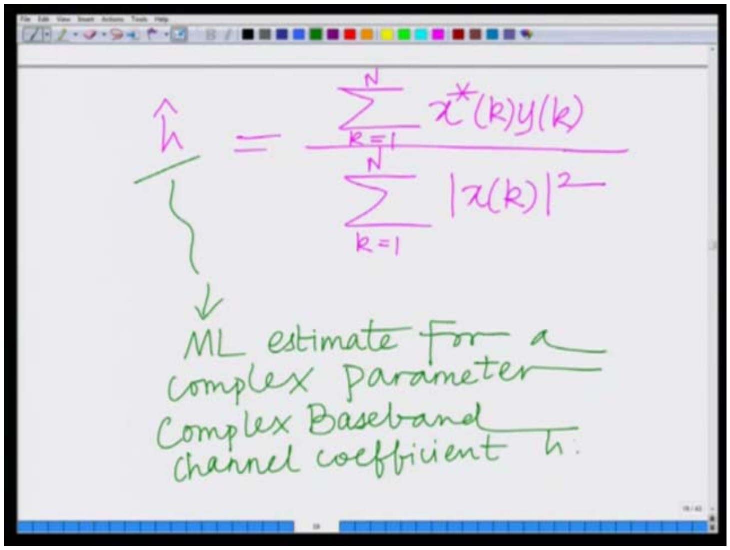
Let us remind ourselves again, this is the ML estimate of channel coefficient, this is the ML estimate of the channel coefficient h, alright. Now, one can again simplify this extend this naturally to the complex parameter scenario. For the complex parameter, a simple trick is to replace the transposed by the hermitian operator and we will also keep using this trick often. To extend this to a complex parameter h, simply replace the transpose by hermitian.

(Refer Slide Time: 17:49)



So, to extend this to a complex parameter, to extend to a complex parameter, replace your transpose by the hermitian, replace the transpose by the hermitian which is basically your conjugate transpose operator. So, therefore for a complex parameter or a complex channel coefficient, this is the maximum likelihood estimate that is X bar transpose becomes X bar hermitian Y bar divided by X bar hermitian X bar

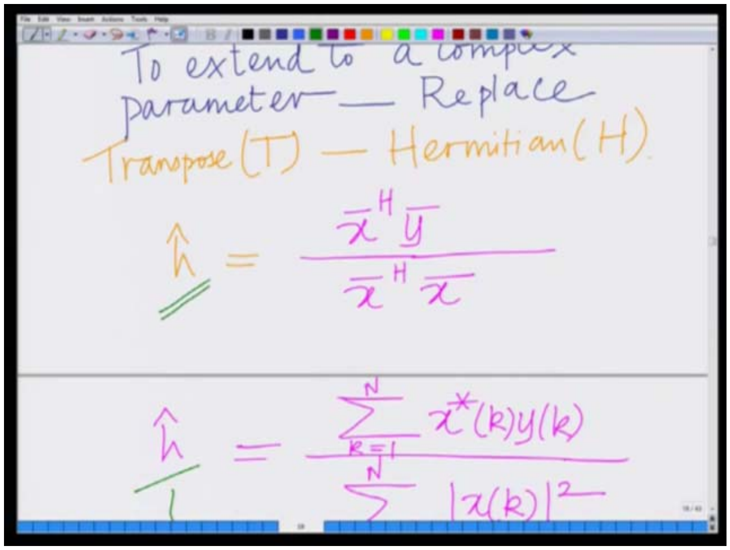
(Refer Slide Time: 18:49)



which can now naturally be simplified as summation K equals to 1 to N X conjugate K times, that is a conjugate or the transmitted pilot symbol divided by K equals to 1 to N X conjugate K times XK which is magnitude XK square, that is the magnitude of the complex symbol. This is the estimate of a complex parameter, maximum likelihood estimate for a, this is the ML estimate for a complex parameter or complex channel coefficient or a complex parameter or your complex baseband or basically your complex baseband channel coefficient h.

That is X bar hermitian Y bar divided by X bar hermitian h bar where as we have seen Y bar is the vector of observation symbols and X bar is the vector of ah the transmitted pilot symbols and this is the maximum likelihood estimate, remember,

(Refer Slide Time: 20:13)



this expression is the maximum likelihood estimate for your complex unknown complex baseband channel coefficient h.

So, what we have seen in this module is basically we have formulated, we have extended this maximum likelihood estimation framework for a wireless scenario or a wireless communications system scenario between the channel coefficient h between the transmitter and receiver is unknown, so we have explored all we have seen a framework where the transmitter transmits a sequence of pilot symbols and you have corresponding to those you have the received pilot outputs that is Y1 up to YN from this knowledge of the transmitted pilot symbols and the observed pilot outputs Y1, Y2… YN, we have basically derived the maximum likelihood estimate of the unknown channel coefficient h and this is known as the pilot based channel estimation scheme in the context of a wireless communication system or this is also known as the training symbol based channel estimation scheme which is very important in the context of a wireless communication system because the channel coefficient h is unknown and has to be estimated at the receiver prior to the beginning of the communication so that the information symbols can be decoded at the receiver.

So, this is the maximum likelihood estimation, the procedure for maximum likelihood estimation of the unknown channel coefficient. We will stop this module here and look at other aspects and properties of this maximum likelihood estimate of the channel coefficient in the subsequent modules. Thank you very much.