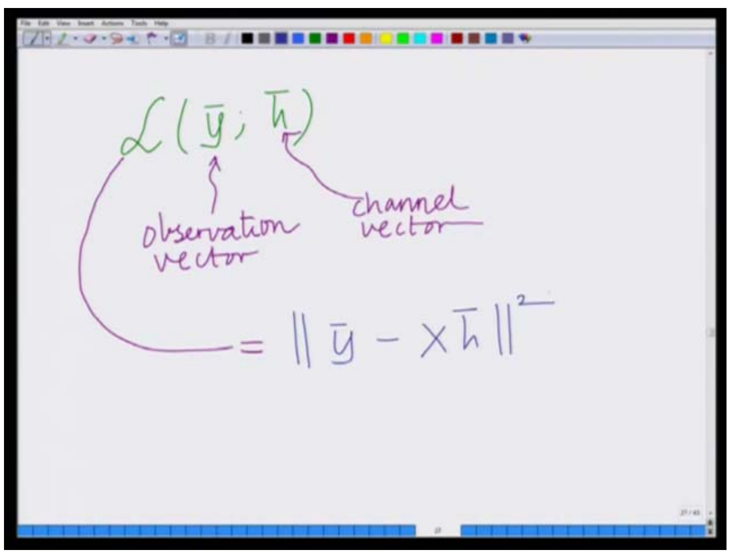
Estimation for Wireless Communications-MIMO/OFDM Cellular and Sensor Networks.
Professor Aditya K Jagannatham.
Department of Electrical Engineering.
Indian Institute of Technology Kanpur.
Lecture -15.
Least-Squares Cost Function for Vector Parameter Estimation Vector Derivative/Gradient.

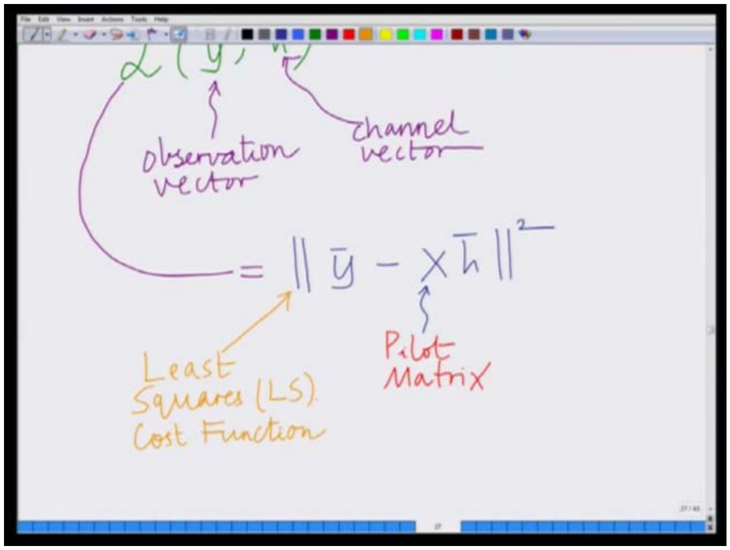
Hello, welcome to another module in this massive open online course on estimation for wireless communication systems. So, we are talking about the estimation of a vector parameter H bar, specifically we are looking at the estimation of this channel vector that arises in a multiple antenna wireless communication system and yesterday we have derived that the log likelihood function corresponding to the estimation of this vector parameter H bar is given by the least requires cost function, which can be described as.

(Refer Slide Time: 0:43)



This is my log likelihood function, that is log likelihood of the observations Y bar parameterised by H bar, this is your observation vector. This is your channel vector, which is the unknown parameter vector, remember we are considering specifically a downlink multiple antenna scenario and this likelihood function or this basically log likelihood function, this is given as a least squares cost function Y bar - X H bar whole square where Y bar is the observation vector, X is remember, we call the pilot matrix.

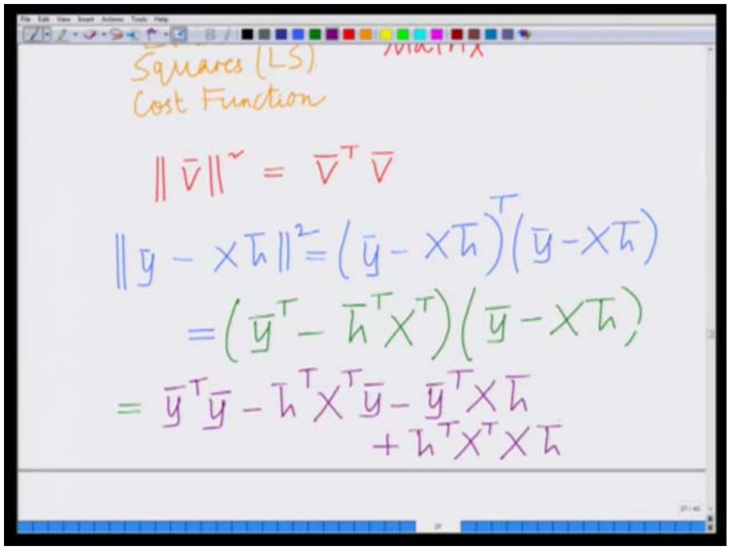
(Refer Slide Time: 1:30)



Yah, so this is just to remind you of these things, this is the pilot matrix and this is basically your least squares cost function.

This is basically your least squares cost function, also abbreviated as the LS cost function. Okay, so we have derived this log likelihood function for the channel vector or multiple antenna channel estimation and we have shown this is given by your least squares cost function and the channel estimate, maximum likelihood channel estimate H hat is the value of, is the parameter vector H bar which minimises or which basically which minimises this least squares cost function. So, towards that, towards finding this maximum likelihood estimate, let us 1st simplify this least squares cost function a little bit.

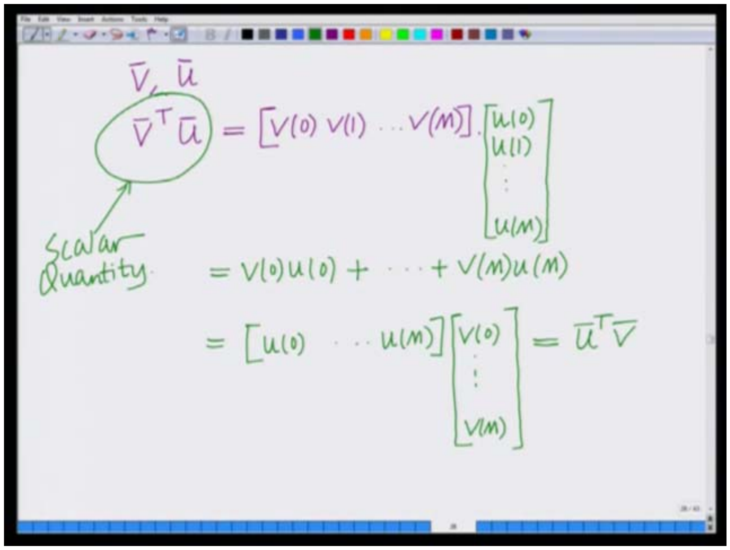
(Refer Slide Time: 2:35)



Now, we know from our knowledge of matrices and vectors, that for any vector V bar, norm of V bar square, this is what we have seen many times before is V bar transpose V bar. So, norm of V bar square or Euclidean norm square of any vector V bar is V bar transpose V bar, which basically implies that, now I can simplify this quantity that is your Y bar - X H bar norm square is basically Y bar - X H bar, is basically this quantity Y bar - X H bar transpose times Y bar - X H bar and now therefore this can be further simplified as basically.

Y bar - X H bar transpose is Y bar transpose - X H bar transpose is H bar transpose X transpose times Y bar - X H bar. Now, multiplying out the terms Y bar transpose Y bar - H bar transpose X transpose Y bar - Y bar transpose or - Y bar transpose X H bar + H bar transpose X transpose X H bar. This is the cost function that we get, all right. So, this is what we have used, we have used the property that norm V bar square or norm V bar square of any vector V bar is V bar transpose V bar, to simplify this least squares cost function norm of Y bar - HX bar square.

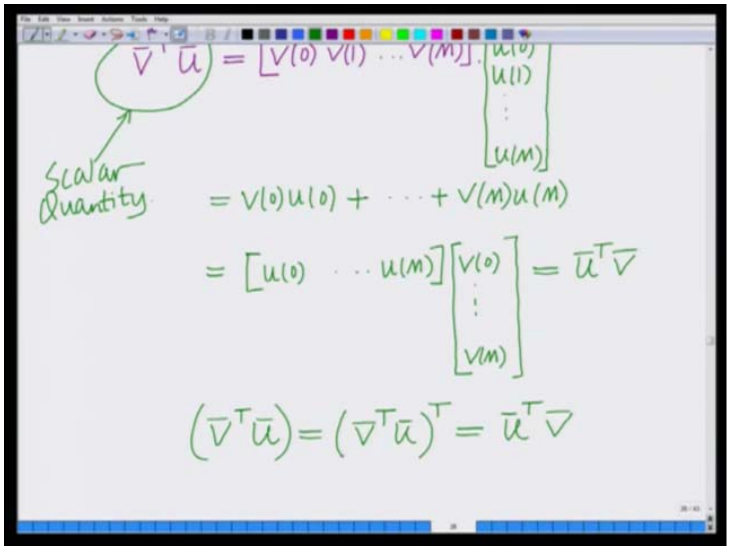
(Refer Slide Time: 4:44)



Now let us use another property, for any 2 vectors V bar and U bar, this is also something we have seen for any 2 vectors V bar and U bar, V bar transpose U bar which is a scaler quantity, remember for this vectors, V bar and V bar which are basically let us say you are M dimensional vectors or n-dimensional vectors, whatever these might be, any 2 vectors, V bar transpose U bar is V0, V1,… VM to U0, U1… Up to UM, which is basically equal to the summation that is V0 U0 + so on up to VM UM, which is also you can see U bar transpose V bar, that is U0 up to UM times V0 up to VM because this quantity, this is also you equal to U bar transpose V bar because this quantity, V bar transpose U bar, you can see this is clearly a number, there is also basically, it is not a vector, it is a scaler quantity.

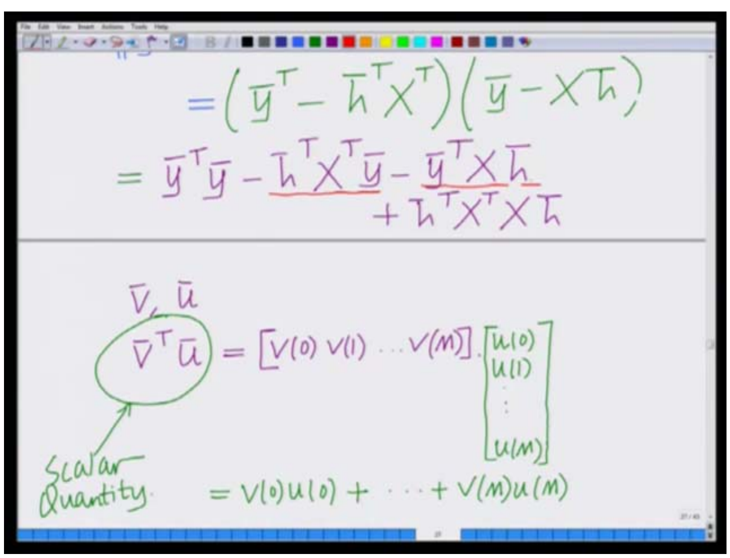
It is not a vector, it is a scaler quantity, that is basically, it is simply a number. Therefore if you take the transpose of a number, if you take the transpose of a number, basically you get the number itself. Therefore the transpose of this V bar transpose U bar is basically to itself.

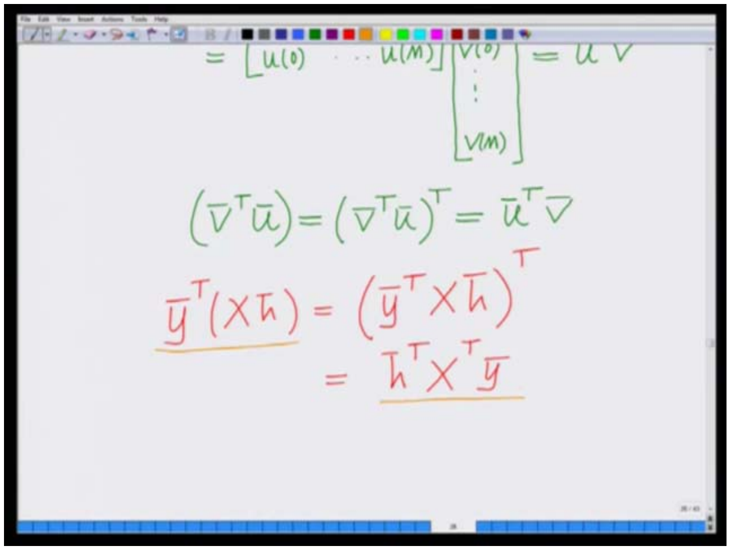
(Refer Slide Time: 6:45)



Therefore what we have, is we have V bar transpose U bar equal to V bar transpose U bar transpose equal to basically U bar transpose V bar. So, for any 2 vectors V bar and U bar, we have V bar transpose U bar is equal to U bar transpose V bar. Now, we are going to use that property to simplify this cost function that we have developed for the least squares cost function. That is the simplification, further simplify this least squares cost function. Now, if you look at your least squares cost function, you can notice these 2 terms, one is your H bar X transpose Y bar, Y bar transpose X H bar and these quantities are nothing but the transpose of each other.

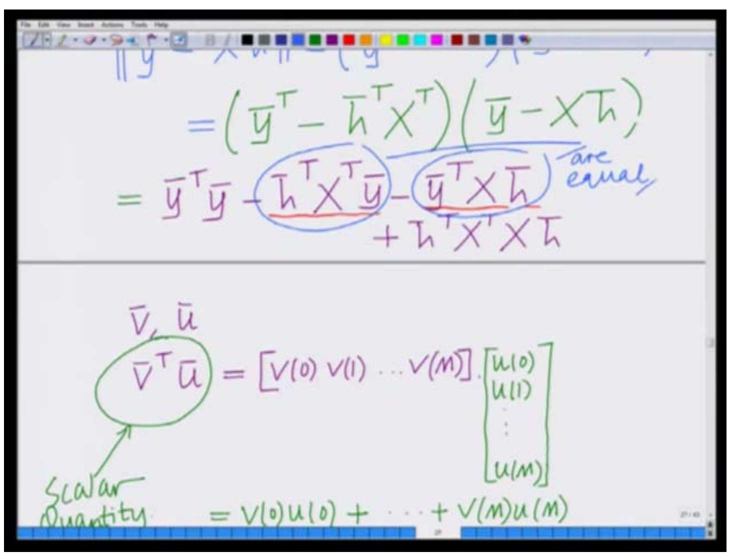
(Refer Slide Time: 7:22)





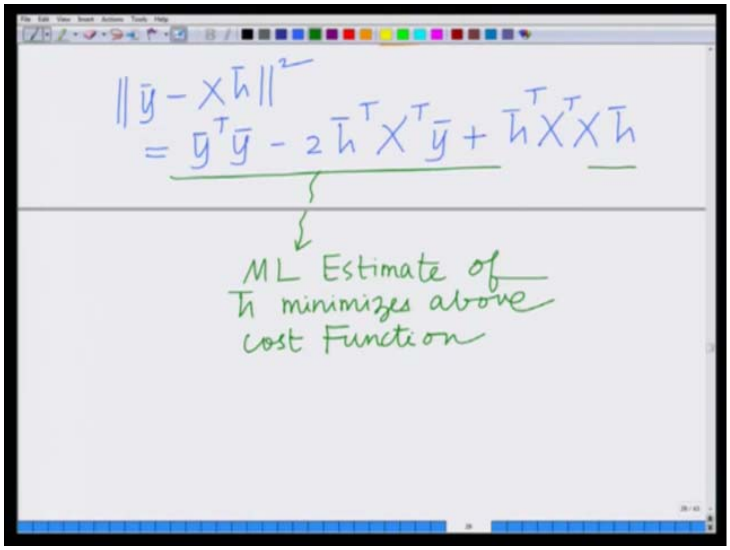
For instance, let us look at this quantity Y bar into Y bar transpose X H bar, so this quantity is basically equal to the, this is a number, so this is basically equal to the transpose of itself Y bar transpose X H bar transpose which is basically equal to H bar transpose X transpose Y bar, so these 2 quantities, Y bar transpose X H bar and H bar transpose X transpose Y bar are basically equal, which means basically you have,

(Refer Slide Time: 8:16)



what you have is basically these 2 quantities here, these 2 quantities are equal. These 2 quantities are equal, therefore your least squares cost function can be further simplified as

(Refer Slide Time: 8:34)

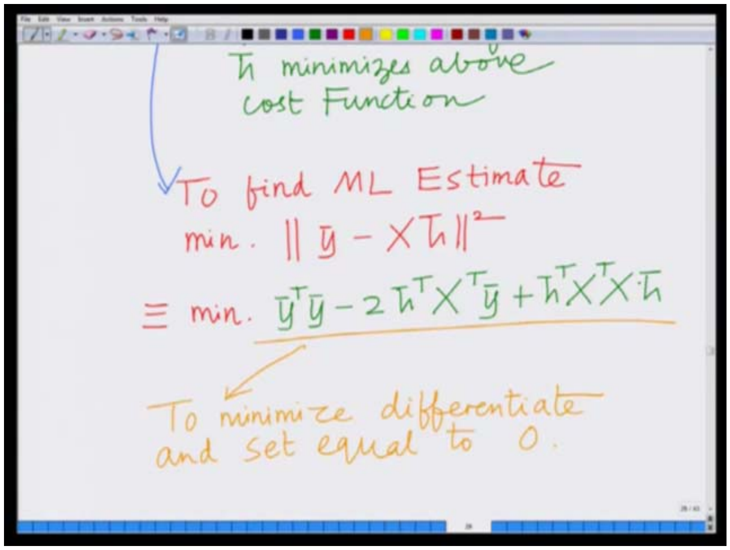


norm Y bar - X H bar square equals Y bar transpose Y bar - 2 H bar transpose X transpose Y bar + H bar transpose X transpose X times H bar.

And now what we can see is basically the maximum likelihood estimate of the channel vector H is the one which minimises this cost function. So, the ML estimate basically can be found as the minimum of this cost function. ML estimate, the ML estimate of H bar minimises the above cost function. That is known as, that gave already seen, that is known as the least squares estimate of the maximum likelihood estimate in this case of the channel vector. Now, how do you find the H bar which minimises the cost function? And for that basically as you all know for any function to find the minimum, that is if it is differentiable, I can basically differentiate it and set it equal to 0 to find the point where the minima is.

So, basically I have to differentiate this cost function with respect to the channel vector H bar and set it equal to 0 to find the, to find the H bar for with this cost function is minimum. So, basically to find the maximum likelihood estimate, to find your ML estimate,

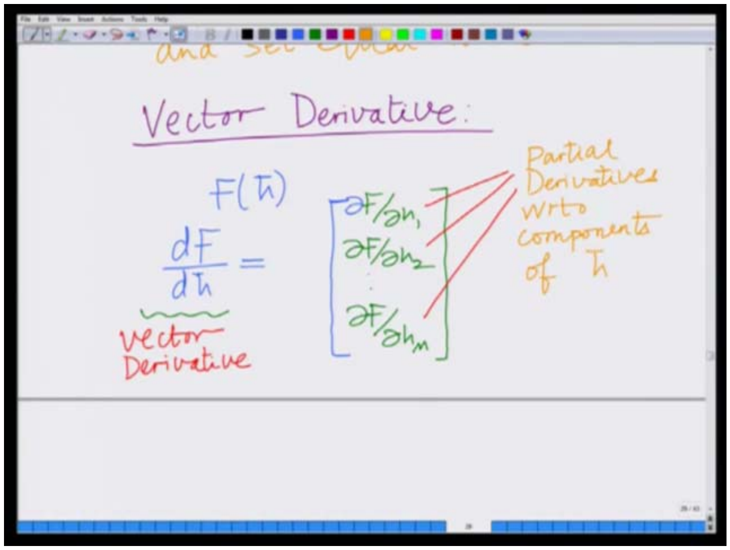
(Refer Slide Time: 10:29)



I have to minimise, which is basically in fact equivalent to minimising, minimising the simplified cost function Y bar transpose Y bar - 2 H bar transpose X transpose Y bar + X bar transpose X transpose X H bar. And to minimise this, we have to differentiate and set equal to 0. That is we have to differentiate with respect to the parameter vector H bar and set it equal to 0 and the point at which it is 0, basically that corresponds to the minimum and that is basically, that H bar is basically the maximum likelihood estimate.

However, differentiate it with respect to H bar, we recall that H bar is the vector, so we have to basically define this notion of a vector derivative which is basically similar to the gradient with respect to a vector.

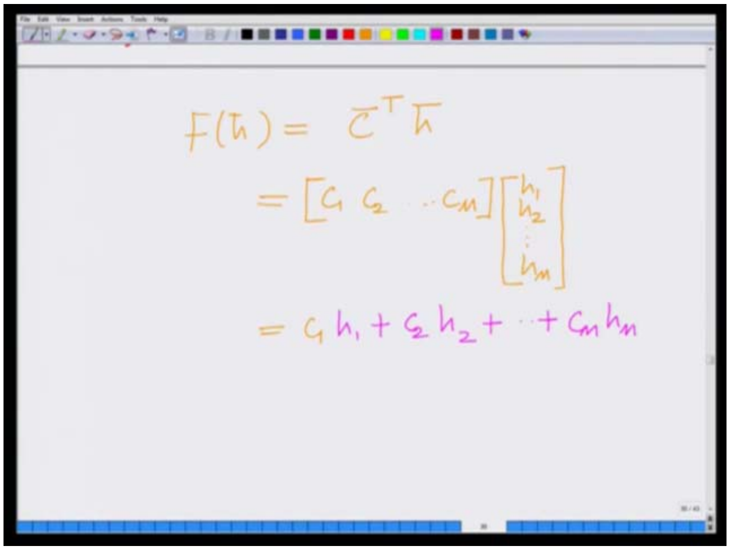
(Refer Slide Time: 12:18)



So, let us define the vector, so for any vector derivative, so let us again, although some of you might already be familiar with this notion of a vector derivative, or gradient for the sake of completeness, let us now define this vector derivative, that is for a function F of any vector H, the derivative, that is dF, I can write this as dF by dH, that is the derivative with respect to the vector H is nothing but the vector of the partial derivatives with respect to the components of H.

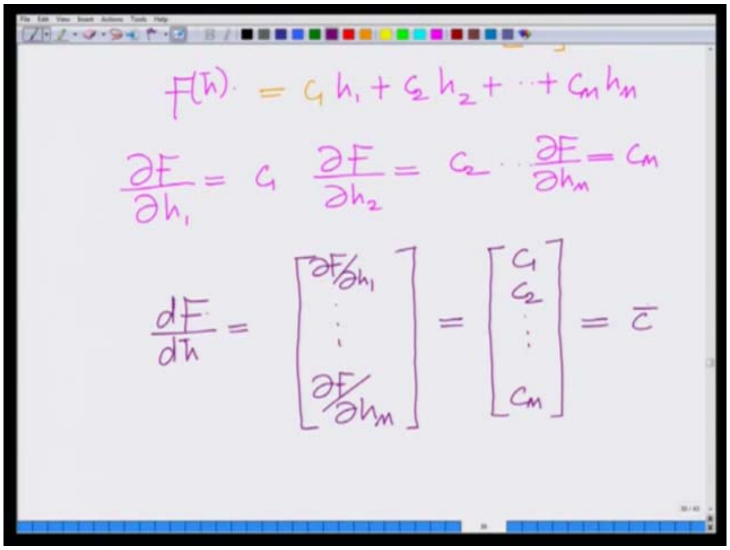
That is I have to differentiate with respect to each component of H. Dow F by dow H1, dow F by dow H2, so on up to dow F over dow HM, that is with respect to, that is simply differentiate, and this is a natural definition, that is simply differentiate the function vector H component of the vector H, that is nothing but the gradient or the vector derivative. So, this is basically your vector derivative, this is your vector derivative and these are your partial derivatives with respect to components of H. Partial derivative with components, with respect to components of the vector H bar.

(Refer Slide Time: 14:06)



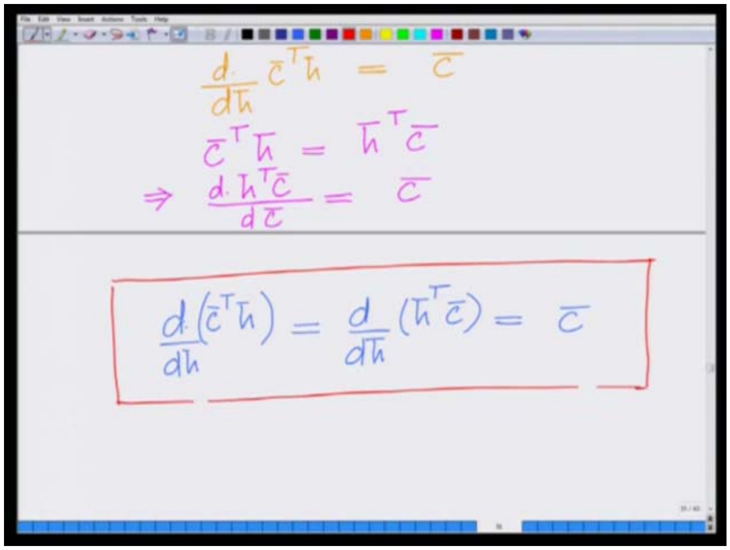
For instance, let us take a look at a simple example, let us consider the simplest of functions, for instance if my function F of H bar is basically some vector C bar transpose times H bar which is basically your C0, C1 or C1, C2, CM up to C2, CM up to H1, H2 up to HM. Which is basically nothing but C1 H1 + C2 H2 so on up to CM HM. So, this is basically your C1 H1 + C2 H2 + so on up to CM HM, now you can clearly see that the partial derivative,

(Refer Slide Time: 15:09)



this is your ass of H bar. Now you can clearly see that the partial derivative with respect to H1 is C1, partial derivative with respect to H2 is C2 and so on and so forth partial derivative with respect to the Mth component HM equals CM, so no surprise there. And what you can see is basically now, if you summarise this thing,

(Refer Slide Time: 16:44)



therefore the derivative with respect to H bar is basically your partial derivative with respect to C1 up to your partial I am sorry, partial derivative with respect to the individual components H1 up to HM which is nothing but your C1, C2 up to CM which is basically the vector C bar.

So, okay, so if C bar is a constant vector and my function of the vector H bar is C bar transpose H bar, then the derivative of that with respect to H bar is naturally C bar. And you can see this as an analog to the scaler derivative, that is we take a constant K multiply it with X then the derivative with respect to X is simply K. So, this is simply a natural extension of that to the vector scenario, that is instead of K times X, you are looking at C bar transpose times X bar. So the derivative of, basically let me summarise this, the derivative of this this quantity, derivative with respect to H bar of this function C bar transpose a bar is basically C bar. And now also realise something very straightforward, that is we have C bar, as we have seen again, several times before, C bar transpose H bar is H bar transpose C bar which implies the derivative of H bar transpose C bar with respect to C bar is also C bar.

And therefore now I summarise this set of relations as the derivative of C bar transpose H bar with respect to H bar equals the derivative with respect to H bar or H bar transpose C bar and that is basically your and that is basically nothing but your C bar. So, this is the, this is basically the vector derivative. So, this is basically the relation for your vector derivative or simple relation for the vector derivative. Okay. And so basically what we have seen in this module so far is basically we have seen your lease square cost function, all right. And we have simplified this we square cost function using several properties, we have simplified this lease square cost function, we have defined this concept of vector derivative and now we are trying to explore the properties of this vector derivative.

Remember towards differentiating that least squares cost function, basically setting it to 0 so we can find the value of the vector parameter H bar for which that lease squares cost function will minimise. And that will basically reverse the maximum likelihood estimate of the channel vector H bar. So, let us stop this module over here and we will continue with other aspects of this vector derivative, that is other properties of vector derivative and applying this vector derivative to the lease squares cost function itself in the next module. Alright, so let us stop here.