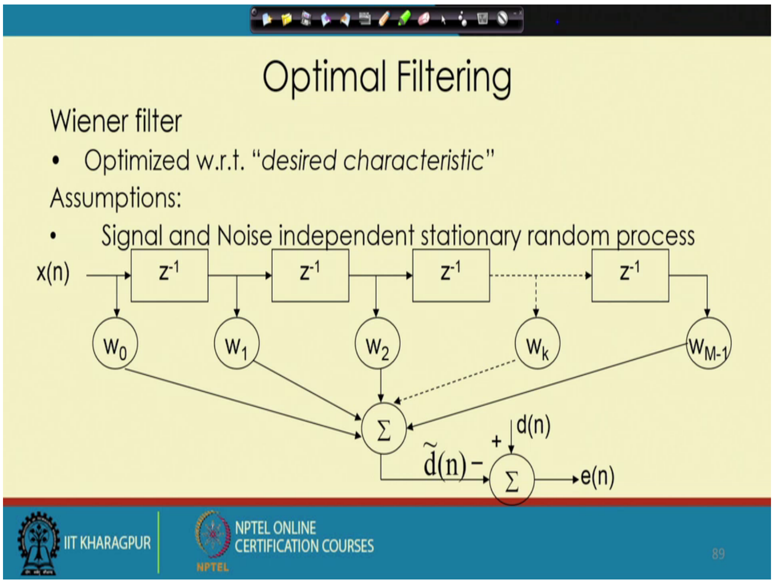
**Biomedical Signal Processing**
**Prof. Sudipta Mukhopadhyay**
**Department of Electrical and Electronics Communication Engineering**
**Indian Institute of Technology, Kharagpur**

**Lecture - 11**
**Artifact Removal (Contd.)**

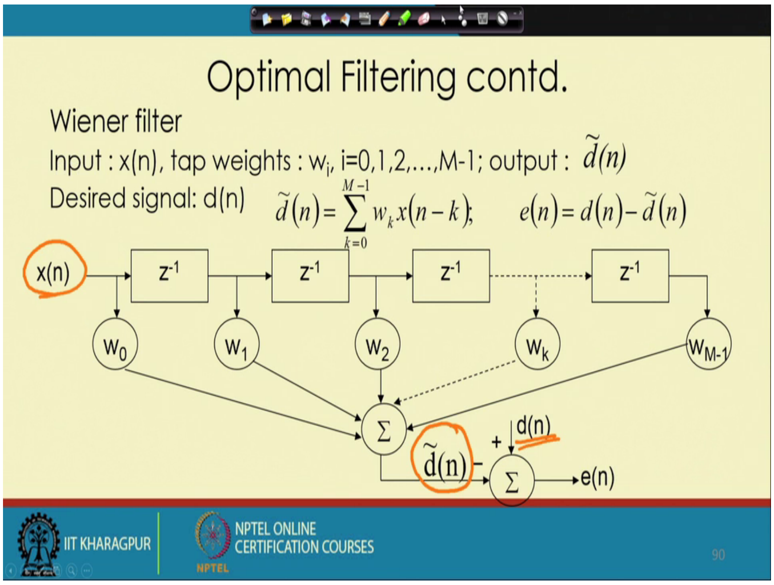
Good morning. So, today we will start with the optimal filtering. And first what we need to realize that we talk about the optimal filtering the technique we will use that Wiener filter, the first the point of it is optimized with respect to a desired characteristics, so first we should try to understand that what we mean by that the optimal characteristics.

(Refer Slide Time: 00:40)



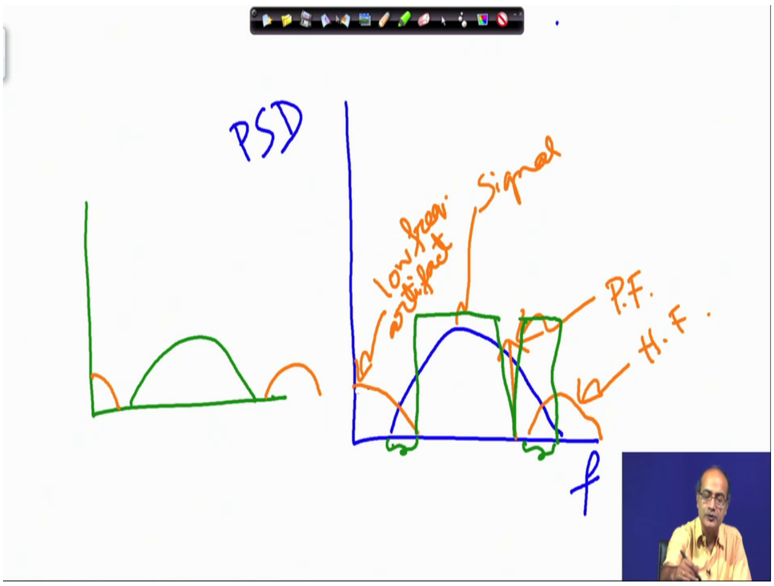

So, let us take an example.

(Refer Slide Time: 00:45)




So, because take a signal we had looking at the frequency domain plot and here we are drawing the PSD and for that here is the say signal spectra some arbitrary spectra we have taken, and let us take that there should be some noise spectra also.

(Refer Slide Time: 01:05)




This is the low frequency one I think you should use a different color for that. So, and here the high frequency side, so this is for the say signal this is for the low frequency artifacts, this is high frequency artifacts and we can have one contribution of the say power frequency this is power frequency hum. Now we have shown only the positive part of the axis; however, if we actually take a real signal as in this case we should have a reflection of the whole thing in the left hand side also.

Now, if we try to have a filter for it say we can draw a filter like this that we can take a characteristics uniform then to avoid that we can have a notch and then it can we again uniform and commit here. So, this could be our filter. Now here what we see there are number of choices, the first point that here what we have trying that we try to have it like a notch filter to avoid the power frequency hum, but this power frequency hum when we are trying to reduce then we are actually losing some part of the signal also. Now should we really make it 0 or will have a simple deep here so that some part of the signal we get, but at the time that along with that we need to then accept some part of the power frequency signal also.

Let us look at the, that left boundary of the band pass signal. Here the way we have taken that we wanted to avoid the low frequency artifact completely and by this choice that we are actually eliminating that, but what is happening as it is all that in this zone in this band that there is some actually signal energy also we are losing that. Now that is not certainly desirable, we should try to keep all the signal energy and to reject all that the noise energy.

In the right hand side boundary if you look at here then by choosing it here we are making sure that we are taking all the signal energy, but then for these band we are accepting a good part of the high frequency noise also. So, here lies the dichotomy that, so long there is no overlap for example, if you had the scenario like this that we have here is a signal and say that here is a the noise spectra, then we could easily separate them with a filter. But when we have the overlapping spectra between the signal and the noise then we have some actually problem in choosing the boundary and in that case what best we can do that is the goal of that this optimal filtering. Because here we cannot say simply that if I select it here that we can avoid all the noise actually energy while without sacrificing the signal energy. Our goal should be then that do you try to get the signal the as much as possible of the signal energy, but as low as possible the noise energy. So, that is what is actually part we can talk about that optimality.

And when we talk about the word optimal actually another word should follow optimal in what sense if we do not tell that then actually description is incomplete. So, slowly we will come to that point. Here just a sake of brevity we have not mentioned there we follow that. Now, to go for these optimal filtering or here the Wiener filter which has gone in the name of the inventor, that we look at the assumptions we need to have on the signal and the noise. The signal here that it should be that and the both the noise it should be stationary. Stationary means it should be weak sense stationary because that is the one what we can get and this two should be independent if you want to make it little loose then you can tell that they need to be uncorrelated at least.

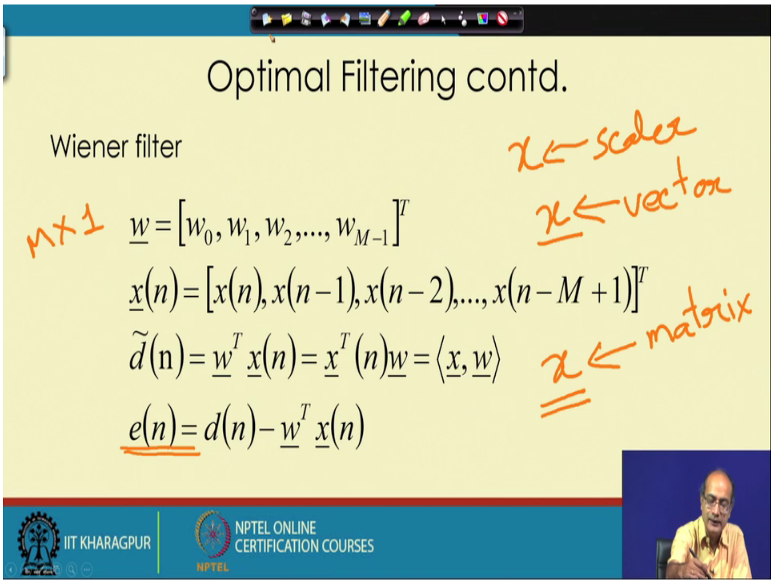
So, this much is required. So, they are two random processes which is stationary and independent that is the minimum condition it is needed for the Wiener filtering design or to show that we have a design where we get a optimal filter. Now let us look at the signal our signal is here that x n that is the observed signal and here it consist of that the real signal plus noise. So, we pass it through a tab delay filter or some registers we can say and with the help of this delay line we get the present value of the signal as well as few passed value of the signal they are multiplied with respective weights and we add them up and at the end of it what we get we tell that this is the desired signal; the d tilde n.

And to get optimal filter we also need something called that that what is the reference or what is the desired signal and with respect to that if we subtract from the desired signal we get some error. So, we will try to minimize that. So, when we told that we are talking about an optimal filter and it should have some sense of optimality we talk about that this error should be minimized in mean square error terms. So, that is the sense we use. So, let us proceed with it and we move forward, we describe that what is that the input is x the tap weights are w i, i equal to 0 to M minus 1. That means the equal number of actually that input signals what we have that many number of weights should be required. And output is d tilde n and desired signal is d n and we can actually write it in this form that in the set of as equation that is first equation is this, that it is nothing but that some actually weighted you can say average of the that our the present and the past inputs and error is the difference between the desired signal and the prediction for the desired signal that is d tilde n.

And here one point I would just like to actually point out that here that w k this index is suggesting that with which lag of the input it is multiplied and it does not have any actually that index with n. What it suggest, thus that this weights they are not varying with time whereas, x n is changing with time, signal is changing, that the desired signal is changing with time and the prediction of the desired signal also changing with time. And for the time being I would request you to assume that some d n is there. Now there could be a practical question that you get d n is there then why we are doing all such thing we can take d n and we can proceed with that. So, let us shelf that idea for that time being that question we just keep it for the future, we will take it up later.

Let us go through the deduction of this and learn about this Wiener filter and then we will get back to it that how we get d n for the practical use. So, let us proceed then, for it that.

(Refer Slide Time: 14:20)



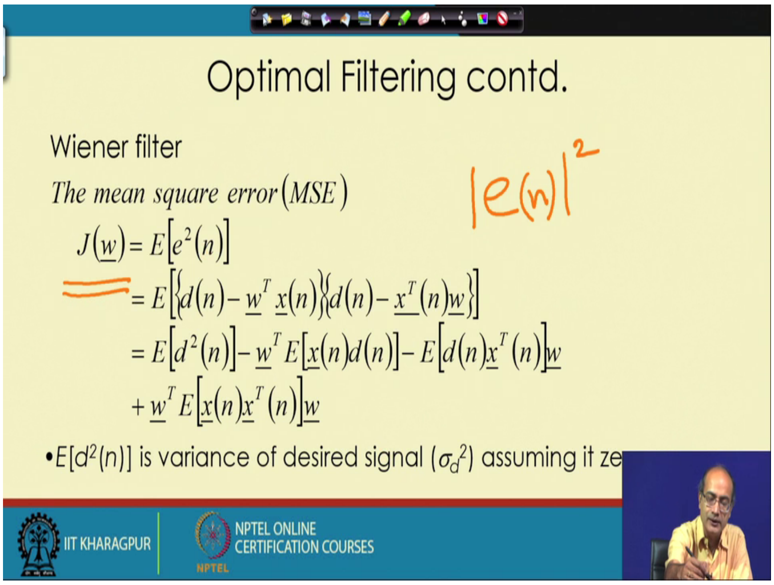

So, for the Wiener filter that here we are now writing down the different that the constituents first is a weight and here I think we should first make some of these things clear that when we write say x this is a scalar. See when we use when we read the notations in the book then we have a choice of number of fonts, but when we write by pen and paper, then we need to be very careful that we are not that accurate to reproduce this fonts. So, we are following some simple notation. Say if x is a scalar then we would tell that x one bar below it would be a vector and we will have another notation will use it a little later that if we write two bar then it will be a matrix. So, we keep it simple we follow this notation so that we can actually right it with pen and paper easily and it becomes unambiguous. Otherwise, sometimes fonts look very close to each other and we get confused that whether it is a scalar or vector or whatever it could be.

So, with that notation what we get that first we have written that the weights as w with a bar below it that is the weight and it is a column vector. And what kind of column vector it is? M by 1 dimensional one column vector. Same way the input signal x n it has been written in a vector form and it is again a column vector because we have written a row and then transpose and with that that again we are having a M by 1 column vector.

Next, we get that what is the output at actually that instant d n before d n. Let us go back again x n and note that this vector also is changing over the time. So, we have an index n and that the configuration of it is the first point or the top most point in this column is the present value and then we moving forward in the past that we are going back that x n minus 1 x n minus 2 in that way, in that way that we have M minus 1 lags of x n.

And when you talk about d n it is a scalar d tilde n. So, say d tilde n how we get that we get by the dot product of these actually w and x n vectors. So, we can write it that w transport x n or x transposed w in both the ways. And then we have the error e n again e n is a scalar it is nothing but the difference between the that desired signal d n minus that d tilde n here we have directly put that value of d tilde n. So, this is a error we compute. So, this is what we would be trying to actually minimize.

(Refer Slide Time: 16:14)




So, here we define our objective function as J w; that means, these objective function this is a function of the weights. See only these weights are in our hand to control this all the other things that what is the input signal that is what we observe, what is a kind of actually that the signal we have chosen, what is the noise and what is the desired signal nothing is actually in our control we are given with them or we will get them. Only thing what we can actually make a change that is with the weights or w, so this objective function that mean square error or in short MSE that is the function of the weights we have written it as J w.

And when we talk about optimal filter then what we mean that it should be optimal in the mean square error sense. The meaning of it is that we would like to have as minimum error as possible; that means, we please refer to that that figure we have drawn earlier that we wanted to show that whenever we are choosing a filter and the situation is so that there is overlap in spectra of the signal and the noise either we pick up some part of the signal or we sacrifice some part of the (Refer Time: 18:00) where actually part of the desired signal we sacrifice or we pick up some part of the noise.

So, here we are having a balance between that, so that the error in the desired signal is minimum. If we include mode of noise then we are moving away from the goal if we lose some part of the desired signal then also we are moving away from the court, so we are just trying to have a balancing act between these two. And for that we have taken the expected value of the squared error why you have taken that that here we have taken e n that because we are dealing with the real signal we are happy with actually that e square n, otherwise we have to take actually the mod of e n square if the e n is a complex signal we may have to take this one, but here the things a simple that we are dealing with real signal, so we are happy to take the just e n square.

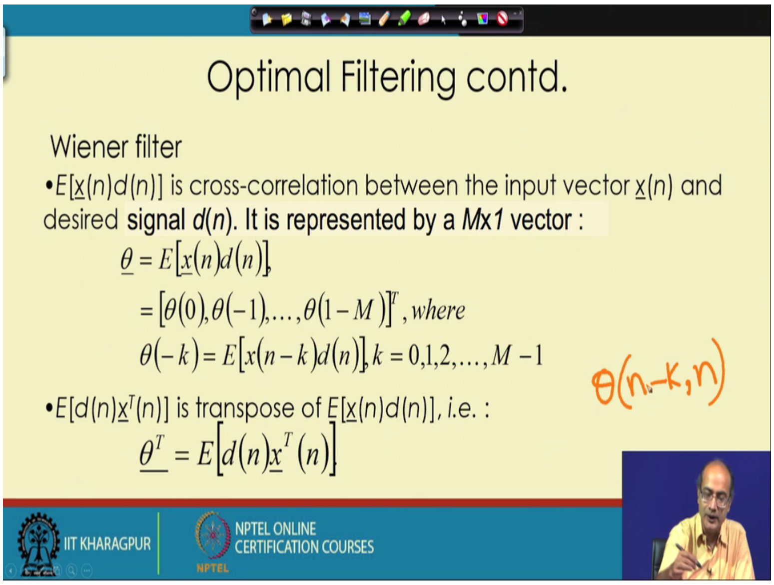
And we are taking the expected value of it the reason is that as a signal x n is a random signal our desired signal d n is a random signal and why it is so, our desired signal is a random signal noise is also random, their constituents are random and even error is also random. So, when we talk about minimization of it we should take actually the expected value of it or we should look the look at that error and on an average basis and try to minimize that quantity. So, we have taken the expectation of it.

Now, in the next step what we do, that we have expanded this e n we have actually replaced it with the that the values what we have actually shown in the previous page that the x where that expression of e n we actually make use of that and we put that the two terms for e n into e n and when we multiply that we get a term that with d n square two terms which are cross between that vector x n and d n that is a desired signal and the vector of observe signal. Here is the second term for that and of course, that we have that the weights and the last term is actually the cross product of the two vectors the that they are representing the observe signal or their value at the that particular instant n.

Now, this quantity will give us a matrix, but the thing to note that we knew we should check the dimensionality that we will get that overall each of this terms has to be a scalar because the error is a scalar quantity. So, in out of these the first thing we could recognize that expected value of d square n it is a variance of the desired signal and assuming it is zero mean. We told you earlier that whenever we are working with signal processing the first thing we assume or sometimes even without stating, we take that signal is zero mean unless it has some special meaning in it. So, we have taken the desired signal is zero mean and in that case this quantity that expected value of d square n it gives us the variance of the desired signal. So, that is the first observation.

Now, let us look at what are the other consequence here.

(Refer Slide Time: 22:36)

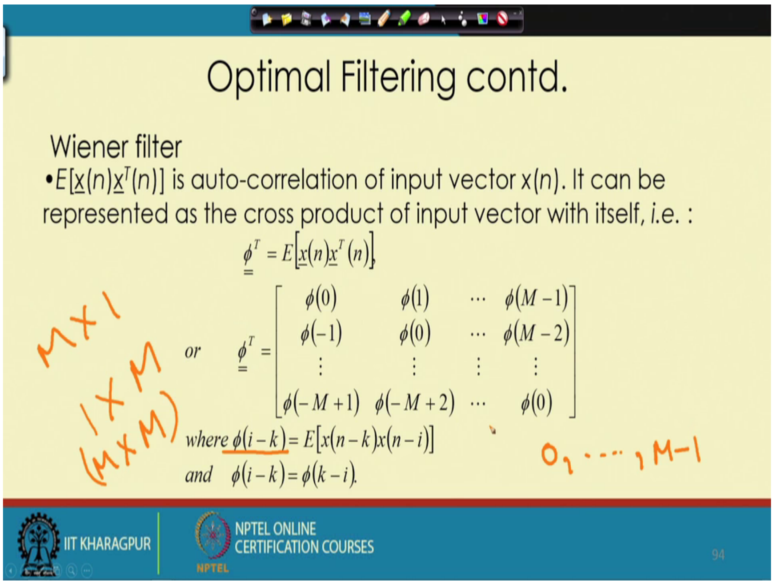



Second, that the part what will look at that we get that x n vector and d this is a cross product between the input vector x n and the desired actually signal d n. Now these thing we get a actually the cross correlation between the two signals and we represent that by theta, we represent that by this term theta. Here we should note one thing that we have not given any index with the time. The theta is not a function of the time.

Now why it is so, because we have assumed both that signals they are stationary. So, the cross correlation of that that kind of signal they would not change with time. So, we have not used any index for the theta, but theta is a vector and for that vector theta that we have that the different components that the first term is theta 0, next term is theta 1 and going in that way up to theta 1 minus n and the form of it that that each of this term we can write that theta minus k equal to x expected value of x n minus k into d n. Had it not been a stationary signal one of them then we should have actually we would have to write this one that instead of theta k we need to write it as theta n minus k comma n in that way.

We could not simplify it and represented by the lag between the two instances and simplify this one and then we get the transposed of it also that we get a transpose of this term. So, theta T also we get, so that is what the third second and the third term we get.

(Refer Slide Time: 25:36)




So, and then we look at that the whole thing that we have another term that is the fourth term what we have received here that is the cross between the two values that is x n and x n this cross of that. And here if you check the dimensionality the x n has a dimensionality say dimensionality is M by 1 and x transposed it has 1 by M.

So, at the end what we are getting a auto multiplication M by M 1 matrix we are getting here. So, that is the, as we are taking the cross between the same signal that is a observe signal x we call it is a autocorrelation and we are getting a autocorrelation matrix of the dimension M by n and we represent that by that the term phi or phi transpose that is as it is given here. And please look at the notation that we have given the two bars below the phi to signify that it is a matrix and the consequence are phi we are marking them as from phi 0 to one side that M minus 1 and it is starting from phi minus M plus 1 to phi M minus 1. And for each of this term phi i minus k the expression is that it is expected value of x n minus k into x n minus i. And again let me remind you that we can put it in this simple form because that of the assumption of stationarity. Had it not been a stationary signal we could not be able to write it in such a simple form.

Another thing we note here that as we are talking about the expected value of the product of these two and product is a actually that cumulative operation, we can change the order of them and if you change the order of them instead of phi i minus k you will get phi k minus i. So, what we get that phi i minus k equal to phi k minus i. So, essentially what we get that these terms are actually that across the point 0 that the phi terms autocorrelation terms they are symmetric for the real signal and for that here to populate these matrix we need not actually the values from that phi minus M plus 1 to phi M minus 1, if we have the values from say starting from 0 to say if we can have the values for M minus 1 that will be sufficient to populate this matrix. So, that is the first initial description about that optimal filter.

Thank you.