

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

**Lecture – 13**  
**Artifact Removal (Contd.)**

So, for the optimal filter this is the equation we get here, that we have in that equation that for the optimal weight with the help of the autocorrelation function and the cross correlation function.



(Refer Slide Time: 00:19)

Optimal Filtering contd.

Wiener filter  
or,  $\sum_{i=0}^{M-1} w_{oi} \phi(k-i) = \theta(k), k = 0, 1, 2, \dots, M-1.$  ✓

considering that the stationarity of signal i.e.  
 $\phi(i-k) = \phi(k-i)$  and  $\theta(-k) = \theta(k).$  X.Y  
FT(X ⊗ Y)  
= FT(X).FT(Y)

The Wiener-Hopf equation can be written in the form of convolution relationship  
 $w_{0k} * \phi(k) = \theta(k)$

 IIT KHARAGPUR
 NPTEL ONLINE CERTIFICATION COURSES
100

And again we recall that the condition we have taken for the stationarity of the signal. Here the signal means not only the stationarity of the original signal actually we are talking about the stationarity of the desired signal and the noise also and feature independent with each other. That, using that that and for the real signal we can actually write that phi i minus k equal to phi k minus i and same way theta minus k equal to theta k. Now, using these facts and from the Wiener equation instead of actually the product form or dot product we can say we can write it in terms of convolution also.

So, if you change phi k minus i by phi i minus k then actually we can say that it becomes a convolution term between the weight and that autocorrelation. Now, what is the benefit in writing it in that way? We need to recall that some of the facts of the Fourier transform that if we have two signals say x and y, now if we take the Fourier transform of these two

that is Fourier transform of  $x$  dot  $x$  convolution say  $y$  then after taking the Fourier transform it becomes the product of the individual term Fourier transform of  $x$  and Fourier transform of  $y$ . So, we would be able to actually make use of that fact and for that actually we have written this equation in the convolution form.

(Refer Slide Time: 02:55)

Optimal Filtering contd.

Applying Fourier transform to the equation, we get:

$$W(\omega) \cdot S_{xx}(\omega) = S_{xd}(\omega)$$

or,  $W(\omega) = \frac{S_{xd}(\omega)}{S_{xx}(\omega)}$ ,

where,  $S_{xx}(\omega)$  and  $S_{xd}(\omega)$  are PSD of input signal and cross-spectral density (CSD) between input signal and desired signal.

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES | 101

So, after taking the Fourier transform in the left hand side what we get that we have this term  $W$  omega.  $W$  omega is the Fourier transform of the tap weight then if we take the Fourier transform of the autocorrelation what we get we get the PSD of the signal. So, the second term is PSD of the that observe signal  $x$  and to differentiate it with actually that the cross PSD of the signal and the desired signal observed signal and the desired signal we use this indexes we write here  $S_{xx}$  for the PSD of the observe signal and  $S_{xd}$  for the cross PSD or cross for spectral density that in between the input signal and the desired signal.

So, using these equations now we can write that the spectrum of the PSD of the signal it is nothing, but that the ratio of the cross PSD between that observe signal and the desired signal divided by the PSD of that observe signal. So, from here itself we get an idea if you look at the spectrum of that the tap weights then where it to be high all those places where the signal and the desired actually that see that observe signal and the desired signal they have high cross correlation.

If the cross correlation is very high then only that  $\omega$  would be high for that value of  $\omega$ . If it happened the cross correlation between the desired signal and the input signal is low and there is some finite energy in the observe signal this ratio would be low. So, that gives us some idea about that how we can how what would be the characteristics of actually  $\omega$ .

Now, will proceed further to understand it in a more unambiguous way.

(Refer Slide Time: 05:55)

Optimal Filtering contd.

Let us consider:  $x(n) = d(n) + m(n)$   
or,  $\underline{x}(n) = \underline{d}(n) + \underline{m}(n)$  n x 1

where, ' $\underline{\cdot}$ ' represents the vector notation.

$$\begin{aligned} \underline{\Phi} &= E[\underline{x}(n)\underline{x}^T(n)] = E[\{\underline{d}(n) + \underline{m}(n)\}\{\underline{d}(n) + \underline{m}(n)\}^T] \\ &= E[\underline{d}(n)\underline{d}^T(n)] + E[\underline{d}(n)\underline{m}^T(n)] + E[\underline{m}(n)\underline{d}^T(n)] + E[\underline{m}(n)\underline{m}^T(n)] \\ &= E[\underline{d}(n)\underline{d}^T(n)] + E[\underline{m}(n)\underline{m}^T(n)] \\ &= \underline{\Phi}_d + \underline{\Phi}_m \end{aligned}$$

So, first let us go back that what we have taken we have started with the observe signal that is  $x(n)$  which consists of the desired signal and the noise  $m(n)$ . And these two both are independent and they are stationary signals, thereby  $x(n)$  is also stationary signal.

Now, we can write the same that  $x(n)$  as well as  $d(n)$  and  $m(n)$  in vector form of the dimension that  $M$  by  $1$ . Now, in that way that what we can do we can now rewrite the autocorrelation matrix  $\Phi$ . So, which can be written as expected value of  $x(n)$  and  $x(n)$  times transposed  $n$ . And now, actually replacing  $x(n)$  the vector  $x(n)$  with its corresponding constituent that is  $d(n)$  plus  $m(n)$ , so we can put it in that way. And we if we take the multiplication we get four terms, first term is with just containing that only the desired signal second and the third term they are having the cross between the desired signal and the noise and the fourth one is having the only the noise term.

Now, using that what we get that these two cross terms they should go to 0 because what we have assumed that the desired signal and the noise they are actually independent with each other. So, using that condition of independence, we can get that these two that the value should go to 0 or even if it is uncorrelated we tell that if it is uncorrelated if we take the expectation of the product that value would go to 0.

So, these two terms goes to 0 leading to the simplification of this term  $\phi$  and we are left with the two terms which gives the first gives the autocorrelation matrix of the observe signal sorry the desired signal here and the second term is the autocorrelation matrix of the noise. So, if we have some idea about the desired signal and some idea about the noise that then we can actually compute them and we have already shown that how that can be done in practice. So, we can get rough idea about the desired signal we can compute these terms that and we can get the autocorrelation matrix of the noise as well as for the signal.

So, first let us go back that what we have taken we have started with the observe signal that is  $x \times n$  which consists of the desired signal and the noise  $m \times n$ . And these two both are independent and they are stationary and signals thereby  $x \times n$  is also stationary signal. Now, we can write the same that  $x \times n$  as well as  $d \times n$  and  $m \times n$  in vector form of the dimension that  $M$  by 1. Now, in that way that what we can do we can? Now rewrite the autocorrelation matrix  $\phi$ , which can be written as expected value of  $x \times n$  and  $x$  times transposed  $n$  and now, actually replacing  $x \times n$  the vector  $x \times n$  with its corresponding constituent that is  $d \times n$  plus  $m \times n$ . So, we can put it in that way. And we if we take the multiplication we get four terms first term is just containing that only the desired signal, second and the third term they are having the cross between the desired signal and the noise and the fourth one is having the only the noise term.

Now, using that that what we get that these two cross terms they should go to 0 because what we have assumed that the desired signal and the noise they are actually independent with each other. So, using that condition of independence, we can get that these two that the value should go to 0 or even if it is uncorrelated we tell that if it is uncorrelated if you take the expectation of the product that value would go to 0. So, these two terms goes to 0 leading to the simplification of these term  $\phi$  and we are left with the two terms which gives the first term gives the auto correlation matrix of the observe signal sorry, the desired signal here and the second term is a autocorrelation matrix of the noise.

So, if we have some idea about the desired signal and some idea about the noise that then we can actually compute them and we have already shown that how that can be done in practice. So, with that rough idea about the desired signal we can compute this terms that and we can get the autocorrelation matrix of the noise as well as for the signal.

(Refer Slide Time: 13:36)




Optimal Filtering contd.

Furthermore:

$$\begin{aligned} \underline{\theta} &= E[x(n)d(n)] = E[\{d(n) + m(n)\}d(n)] \\ &= E[d(n)d(n)] + E[m(n)d(n)] \\ &= E[d(n)d(n)] \\ &= \underline{\Phi}_{1d}, M \times 1 \text{ autocorrelation vector of desired signal.} \end{aligned}$$

The optimal Wiener filter is then given by:

$$\underline{w}_0 = (\underline{\Phi}_d + \underline{\Phi}_m)^{-1} \underline{\Phi}_{1d}$$

Now, let us get that what would be the term theta which is the cross correlation between the observe signal and the death desired signal  $d(n)$ , the vector  $x(n)$  and the signal  $d(n)$ . Now, here we can actually for the purpose of simplification we can rewrite the vector  $x(n)$  in terms of the  $d(n)$  and  $m(n)$ . And while do that we get two terms one involving only  $d(n)$  [noise], another the cross between the noise vector and the desired signal  $d(n)$ .

Now, again here we can make use of the fact that because these two are independent the expected value of the product would go to 0. So, as these term is going to 0 we are left with the first term only which is nothing, but actually that a vector of the auto correlation matrix that it is the same actually signal that is the desired signal we are dealing with and we have the fast actually that column of that. So, that is why we have written here that this one with  $\Phi_{1d}$ . So, it is the first column of the autocorrelation matrix or we can say it is the autocorrelation vector of the desired signal.

So, in this way we can rewrite the optimal weight. Now, it is given by that the inverse of  $\Phi_d$  plus  $\Phi_m$  that some of it the inverse of that into  $\Phi_{1d}$ . So, this is the way we can actually right about that the optimal weight  $W_0$ . So, that is one of the thing that how we

can in practice we would drive the  $W$  to 0 when we do not actually directly get  $d$ . We did not have to get the exact values rather we can actually have those autocorrelation function at the cross correlation actually autocorrelation function of the desired signal and autocorrelation function of the noise. So, if we can get that that is sufficient to get the optimal weight. So, that is the way actually it is computed in practice.

(Refer Slide Time: 16:47)

Optimal Filtering contd.

Frequency response of Wiener filter is given by:

$$S_{xx}(\omega) = S_d(\omega) + S_m(\omega)$$

and  $S_{xd}(\omega) = S_d(\omega)$

which leads to

$$W(\omega) = \frac{S_d(\omega)}{S_d(\omega) + S_m(\omega)} = \frac{1}{1 + \frac{S_m(\omega)}{S_d(\omega)}}$$

where,  $S_d$  and  $S_m$  represent the PSDs of desired signal and noise respectively.

$\phi_{xx} = \phi_d + \phi_m$

Now, we come to the next step of it again we go back to the frequency domain. If you go back to the frequency domain we have one actually benefit here that we can have a better understanding about the process, that we can write the PSD of actually  $x$  equal to the PSD of that the desired signal plus actually that of the that noise. It is coming from the fact that we are found that  $\phi_{xx}$  equal to  $\phi_{dd}$  which for simplification that we have use a notation  $\phi_d$  and then we have that  $\phi_m$  we could write from  $\phi_{mm}$  also and taking the Fourier transform we are getting this one.

And with that that; and we can rewrite in the spectral domain that the spectrum of the tap weight as that  $S_d$  divided by  $S_d$  plus  $S_n$  or by dividing both the numerator and the denominator by  $S_d$  we can write it in this form. So, what we get out of it that when say the noise is very high then the denominator will have high value, so at that particular frequency the spectrum of the tap weight should be very low because the denominator is high and thereby it will help in rejecting the noise. Now, if it happens that the signal is

also strong at that point then again the numerator sorry denominator will get actually shrink and the value of the depth that spectrum of the tap weight will increase.

Now, let us look at the situation where the noise spectrum is absent or 0 in that case irrespective of the value of  $d$ . So, long it is not 0 we get this term as 0. So, then the tap weight will reach to the value 1 or you can become much high. So, where thus noise is not present it will give access to the depth our signal spectrum in undistorted way it will accept all the signal.

When the noise is very strong and our desired signal is very low it will this denominator will increase and it will reject actually it will it should reject all that that signals that is because that is coming from the noise so the tap weight will become near 0. And when intermediate situation occurs; that means, both the signal and the noise both are present in that case the spectrum will be in some intermediate value. So, that gives us a much better idea that how we can actually get the optimal tap weights or how it works.

And another important information we get out of this equation. To get the optimal tap weight what is the minimum thing, we need we need the spectral characteristics or the PSD of the desired signal and spectral characteristics of the noise or to be more specific the additive noise these two are sufficient to actually define that tap weights in the spectral domain and from there we can get what will be the tap weights in the time domain.

So, with that we complete this discussion on the optimal filtering.

Thank you.