Optimal Filtering with Multirate Observations

Ryan J. Kuchler and Charles W. Therrien
Department of Electrical and Computer Engineering
Code EC/Ti
Naval Postgraduate School
Monterey, CA 93940
Tel: 831.656.3347  FAX: 831.656.2760
e-mail: therrien@nps.navy.mil

Abstract—This paper considers the problem of optimal filtering for several channels of observations occurring at different sampling rates. The problem is formulated in general terms with arbitrary sampling rates for the observations. The set of observation sequences (channels) are jointly cyclostationary and the resulting optimal filters are linear and periodically time-varying (LPTV). A formula for the period of the system is given and the explicit forms of the Wiener-Hopf equations and the estimation error variance are derived. Simulation results are presented for estimation of a signal with known correlation function in noise using two observation sequences with sampling rates differing from that of the underlying signal to be estimated.

I. THE MULTIRATE SYSTEM

In this paper we consider the problem of optimal filtering where the observation sequences are sampled at different rates. This problem occurs, for example, when two or more sensors with different resolutions are observing the same underlying phenomena. A typical model for $M$ observation sequences in this case is shown in Fig. 1 where $d[n]$ is the signal of interest, $\gamma_i[n]$ represents the impulse response of linear time-invariant filters for output $i$, $\eta_i[n]$ represents the independent noise sequences and $x_i[n]$ represents the observation sequences reduced in rate by decimation with factors $K_i$.

As will be shown, $h^{(i)}[n]$ represents the coefficients of $M$ linear periodically time-varying (LPTV) filters. Their outputs are added to produce the estimate $\hat{d}[n]$ at the full rate.

A simplified version of this problem has been addressed in [1] and [2] where two sequences are observed, one at the full rate and the other decimated by a factor of 2. A least squares formulation of the problem is reported in [3], where the decimation factor of the second observation sequence is arbitrary. In this paper we consider the general case where the $M$ sampling rates each may be different from that of the quantity to be estimated and different from each other by arbitrary factors.

This problem is closely related to the problem of estimation for cyclostationary random processes considered in [4]. However, we avoid the matrix spectral factorization problem cited there by considering the use of FIR filters. Such FIR filters are often preferred in practice since they are simple to implement and avoid possible problems of instability associated with recursive filters.

Fig. 3 shows an example of a 2-signal multirate system where the decimation factors for the $x_1[m_1]$ and $x_2[m_2]$ sequences are 2 and 3, respectively, and the filter orders are $P_1 = 4$ and $P_2 = 2$. This figure will be referred to throughout the paper in discussing the estimation equations. By inspection, the periodicity of the signals and LPTV filters in this particular system shown is six.
II. FUNDAMENTAL RATE AND SYSTEM PERIODICITY

In order to adequately describe the relationships between all the signals in the system, a sampling rate needs to be defined by which all signals can be referred against. This fundamental rate $F$ is the least common multiple of all sampling rates in the system, or $\text{lcm}(F_i, \ldots, F_M)$, where $\text{lcm}(\cdot)$ denotes the least common multiple and $F_i$ is the sampling rate of the $i^{th}$ signal. For example, if $x_1[m_1]$ is sampled at 12 kHz and $x_2[m_2]$ is sampled at 8 kHz, then the fundamental rate would be 24 kHz.

All the signals in the multirate system are decimated versions of a signal at the fundamental rate. Therefore we can define a decimation factor for the $i^{th}$ signal as $K_i = \frac{F}{F_i}$.

These decimation factors are useful in determining a system periodicity. The set of observations considered in this problem are jointly cyclostationary with periodicity $K$. The period of the LPTV filters is equal to the cyclostationary period $K$ and can be computed as follows. Assume there are $M$ observed signals with associated decimation factors, $K_i$, $i = 1, 2, \ldots, M$, then the cyclostationary period $K$ is equal to the least common multiple of the observed signal decimation factors, or $K = \text{lcm}(K_1, \ldots, K_M)$.

For example, if $K_1 = 2$, $K_2 = 4$, $K_3 = 6$, the period of the system is 12. In order to form the optimal estimation of $\hat{d}[n]$, a set of twelve distinct system equations must be developed.

III. SYSTEM EQUATIONS

To derive the system equations, let us consider $M$ sequences, $x_i[m_i]$ through $x_M[m_M]$, ordered such that $K_i < K_{i+1} < \cdots < K_M$. The estimation of the desired sequence can be approached as follows. Let us write the desired signal at time $n$ as

$$d[n] = d[K \cdot l + k]$$

where $l = \lfloor n/K \rfloor$ and $k = n \mod K$. The variable $k$ is the periodicity index with $0 \leq k \leq K - 1$.

Let the order of the $i^{th}$ filter be denoted by $P_i$. Then for any time $n$ and corresponding value $k$, the observation sequence in $x$ can be expressed as

$$\tilde{x}_i^{(k)}[n] = \tilde{D}_i^{(k)} \tilde{\tilde{x}}[n]$$

where $\tilde{\tilde{x}}[n]$ is given by

$$\tilde{\tilde{x}}[n] = \begin{bmatrix} x_1[n] & x_1[n-1] & \ldots & x_1[n-P_i \cdot K_i + 1] \end{bmatrix}^T$$

Note that $\tilde{\tilde{x}}[n]$ consists of all possible values of $x_i[n]$ (observed or unobserved) and $\tilde{x}_i^{(k)}[n]$ represents just the observed values used by the filter. The decimation matrix $\tilde{D}_i^{(k)}$ is a matrix of 1’s and 0’s used to extract elements of $\tilde{\tilde{x}}[n]$ as observations to be placed in $\tilde{x}_i^{(k)}[n]$. For example, the first two decimation matrices for $\tilde{\tilde{x}}_1[n]$ in our example are

$$\tilde{D}_1^{(0)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

and

$$\tilde{D}_1^{(1)} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

These matrices can be presented in a more compact form with the use of the Kronecker product, specifically.

$$\tilde{D}_1^{(0)} = I \otimes [1 \ 0]$$

and

$$\tilde{D}_1^{(1)} = I \otimes [0 \ 1]$$
where \( I \) is the identity matrix of size equal to the filter order \( P_i \) of the associated observation signal.

By virtue of (1), the autocorrelation matrix \( \tilde{R}^{(i)} \) of \( \tilde{x}_i[n] \) is given by

\[
\tilde{R}^{(i)} = D_i^{(i)} R_i D_i^{(i)T}.
\]

(2)

The cross-correlation matrices, \( \tilde{R}_{y}^{(i)} \) and \( \tilde{r}_{di}^{(i)} \) maintain a dependency on the associated lag times and are given by

\[
\tilde{R}_{y}^{(i)} = D_i^{(i)} R_y D_i^{(i)T}
\]

(3)

\[
\tilde{r}_{di}^{(i)} = D_i^{(i)} r_d D_i^{(i)T}
\]

(4)

It should be noted that the dependency on the observation lag times is completely determined by the decimation matrices.

The optimal estimate is formed by summing the output of linear periodically time-varying (LPTV) filters, one for each observed sequence. (The filtering of the data for the high rate filter of order \( P = 4 \) and the low rate filter of order \( P_2 = 2 \) is depicted in Fig. 3.) Let \( h_i^{(i)} \) represent the vector of filter coefficients for signal \( i \) at periodicity index \( k = n \mod K \); then at any particular time \( n \) the estimate is given by

\[
\hat{d}_i[n] = \sum_{j=1}^{M} \tilde{x}_i^{(i)T}[n] h_i^{(i)} = \sum_{j=1}^{M} \tilde{r}_{di}^{(i)T} \tilde{x}_i^{(i)}[n].
\]

(5)

The notation \( \hat{d}_i[n] \) is actually redundant since \( k = n \mod K \); however, we use this notation to emphasize that the statistical properties of the estimate and the error are periodically time varying.

IV. MULTIRATE WIENER-HOPF EQUATIONS

By defining the error of the estimate as

\[
\varepsilon_i[n] = d[n] - \hat{d}_i[n]
\]

and applying the orthogonality principal, we can write

\[
E\left\{ \tilde{x}_i^{(i)}[n] \varepsilon_i[n] \right\} = E\left\{ \tilde{x}_i^{(i)}[n] \left[ d[n] - \sum_{j=1}^{M} \tilde{x}_j^{(i)T}[n] h_j^{(i)} \right] \right\}
\]

\[
= \tilde{r}_{di}^{(i)*} - \sum_{j=1}^{M} \tilde{R}_{y}^{(i)} h_j^{(i)*} = 0
\]

or

\[
\tilde{r}_{di}^{(i)*} = \sum_{j=1}^{M} \tilde{R}_{y}^{(i)} h_j^{(i)*}
\]

(6)

By representing (6) in Matrix form, and using the hermitian symmetry property \( \tilde{R}_{y}^{(i)} = \tilde{R}_{y}^{(i)T} \), we obtain

\[
\begin{bmatrix}
\tilde{R}_{y}^{(i)} & \tilde{R}_{y}^{(i)} & \cdots & \tilde{R}_{y}^{(i)} \\
\tilde{R}_{y}^{(i)T} & \tilde{R}_{y}^{(i)} & \cdots & \tilde{R}_{y}^{(i)} \\
\vdots & \vdots & \ddots & \vdots \\
\tilde{R}_{y}^{(i)T} & \tilde{R}_{y}^{(i)T} & \cdots & \tilde{R}_{y}^{(i)}
\end{bmatrix}
\begin{bmatrix}
h_1^{(i)} \\
h_2^{(i)} \\
\vdots \\
h_M^{(i)}
\end{bmatrix}
= \begin{bmatrix}
\tilde{r}_{di}^{(i)*} \\
\tilde{r}_{di}^{(i)*} \\
\vdots \\
\tilde{r}_{di}^{(i)*}
\end{bmatrix}
\]

(7)

where the correlation and cross-correlation elements are defined by (2) through (4).

The associated error variance, \( \sigma_e^2 = E_1 \left[ d[n] \varepsilon_i[n] \right] \), is

\[
\sigma_e^2 = \begin{bmatrix}
\tilde{r}_{di}^{(i)*} & \cdots & \tilde{r}_{di}^{(i)*}
\end{bmatrix}
\begin{bmatrix}
\tilde{R}_{y}^{(i)} & \cdots & \tilde{R}_{y}^{(i)} \\
\tilde{R}_{y}^{(i)T} & \cdots & \tilde{R}_{y}^{(i)} \\
\vdots & \ddots & \vdots \\
\tilde{R}_{y}^{(i)T} & \tilde{R}_{y}^{(i)T} & \cdots & \tilde{R}_{y}^{(i)}
\end{bmatrix}
\begin{bmatrix}
h_1^{(i)} \\
h_2^{(i)} \\
\vdots \\
h_M^{(i)}
\end{bmatrix}
\]

(8)

Observe that the mean-square error \( \sigma_e^2 \) is periodically time varying. In order to establish a single figure of merit for the system, we define the average error as

\[
\sigma_a^2 = \left( \prod_{k=0}^{K-1} \sigma_e^2 \right)^{1/K}
\]

This geometrical average is chosen, since the error, when expressed in dB becomes the arithmetic average of the \( \sigma_e^2 \) values in dB.

V. FILTER COEFFICIENTS

Solving the multirate Wiener-Hopf equations, the filter coefficients for the direct form optimal filters of Fig. 2 can be written recursively as:

\[
h_i^{(i)} = E_i^{-1} \left[ \tilde{x}_i^{(i)} - G_i^{(i)} \right]^{-1} \left[ \tilde{x}_i^{(i)} - G_i^{(i)} \right]^{-1}
\]

(9)

where the terms \( G_i^{(i)} \) and \( E_i^{(i)} \) are defined as

\[
G_i^{(i)} = \begin{bmatrix}
\tilde{R}_{y}^{(i)} & \cdots & \tilde{R}_{y}^{(i)} \\
\tilde{R}_{y}^{(i)T} & \cdots & \tilde{R}_{y}^{(i)} \\
\vdots & \ddots & \vdots \\
\tilde{R}_{y}^{(i)T} & \tilde{R}_{y}^{(i)T} & \cdots & \tilde{R}_{y}^{(i)}
\end{bmatrix}
\]

and

\[
E_i^{(i)} = \begin{bmatrix}
0 & \cdots & 0 \\
\tilde{R}_{y}^{(i)} & \cdots & \tilde{R}_{y}^{(i)} \\
\vdots & \ddots & \vdots \\
\tilde{R}_{y}^{(i)T} & \tilde{R}_{y}^{(i)T} & \cdots & \tilde{R}_{y}^{(i)}
\end{bmatrix}
\]

where \( 1 \leq i \leq M \).
The error variance for the direct form solution is as shown in (8).

Further manipulation of the filter coefficients allows the optimal filter to be represented in the form shown in Fig. 4. We refer to this as the innovations representation.

\[
E^{(k)} = \begin{bmatrix} \hat{R}_i^{(k)} - \tilde{R}_i^{(k)} \\ \vdots \\ \hat{R}_{(i-1)i}^{(k)} \end{bmatrix}, \quad i = 1 \leq i \leq M
\]

The iterative error variance equation is

\[
\sigma_j^2(k) = \sigma_{j-1}^2(k) - \left( \begin{bmatrix} \hat{f}_d^{(k)} - \tilde{f}_d^{(k)} \\ \vdots \\ \hat{f}_{d(i-1)}^{(k)} \end{bmatrix} \right)^T E_j^{(k)-1} \left( \begin{bmatrix} \hat{f}_d^{(k)} - \tilde{f}_d^{(k)} \\ \vdots \\ \hat{f}_{d(i-1)}^{(k)} \end{bmatrix} \right)
\]

This form is useful because each signal is modified to remove any information from it that is already contained in the higher rate signals of the system. If the highest rate signal is sampled at the fundamental sampling rate, then the filter for the first signal is the LTI filter derived from the Wiener-Hopf equations for one signal. If the highest rate signal is not at the fundamental sampling rate, then the associated filter still will be an LPTV filter.

VI. PERFORMANCE RESULTS

As an example of multirate optimal filtering, let us consider a system with two observation sequences \(x_1[n]\) and \(x_2[n]\), with decimation factors of \(K_1 = 2\) and \(K_2 = 5\). Two different types of signals were analyzed for comparison: a 2nd order AR process and a signal with two sinusoids. The signal equations are shown in Table 1; the driving noise in the AR process is WGN with mean zero and variance one.

<table>
<thead>
<tr>
<th>Signal Type</th>
<th>Signal Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd Order AR</td>
<td>(s[n] = 0.9s[n-1] - 0.8s[n-2] + w[n])</td>
</tr>
<tr>
<td>2 Sinusoids</td>
<td>(s[n] = 4\cos(0.2\pi n) + 2\cos(0.02\pi n))</td>
</tr>
</tbody>
</table>

The associated correlation functions for each of the signals are provided in Table 2.

<table>
<thead>
<tr>
<th>Signal Type</th>
<th>Autocorrelation Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd Order AR</td>
<td>(R_1(l) = \begin{cases} 3.7114 \cos(0.8944) &amp; l \geq 0 \ -3.7114 \cos(1.180) &amp; l &lt; 0 \end{cases})</td>
</tr>
<tr>
<td>2 Sinusoids</td>
<td>(R_2(l) = 8\cos(0.2\pi l) + 2\cos(0.02\pi l), \quad \forall l)</td>
</tr>
</tbody>
</table>

The two observation sequences are defined as

\(x_1[n] = s[n] + \eta_1[n]\) and \(x_2[n] = s[n] + \eta_2[n]\), where \(\eta_1[n]\) and \(\eta_2[n]\) are independent white noise sequences with variances chosen such that the Signal to Noise Ratio (SNR) varied over the range, \(-6\) dB, \(-3\) dB, \(-1.7\) dB, \(0\) dB, \(1.7\) dB, \(3\) dB.
Simulations were conducted with filter orders of 5 for both $P_1$ and $P_2$. Plots showing the theoretical error variance for the high rate signal in a 0 dB environment and the low rate signal varying over the range from -6 to 6 dB are shown in Fig. 5 and Fig. 6.

As these two figures show, even though the low rate signal performs worse than the high rate signal over the range of SNR for both signal types, by processing the two signals together using the multirate technique significant improvements can be obtained. For the case of the 2nd Order AR process, improvements up to 2.5 dB were obtained, and for the sinusoidal signal improvements of up to 6 dB were seen.

Table 3. ERROR VARIANCES FOR 2 SINUSOIDS (ALL VALUES IN dB)

<table>
<thead>
<tr>
<th>Filter Orders</th>
<th>$\sigma^2_{x_1}$</th>
<th>$\sigma^2_{x_2}$</th>
<th>$\sigma^2_{x_i}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P = 5, Q = 5$</td>
<td>$-0.43$</td>
<td>$3.27$</td>
<td>$13.98$</td>
</tr>
<tr>
<td>$P = 30, Q = 5$</td>
<td>$-13.54$</td>
<td>$-12.26$</td>
<td>$13.98$</td>
</tr>
<tr>
<td>$P = 5, Q = 30$</td>
<td>$-6.54$</td>
<td>$3.27$</td>
<td>$9.39$</td>
</tr>
<tr>
<td>$P = 30, Q = 30$</td>
<td>$-17.32$</td>
<td>$-12.26$</td>
<td>$9.39$</td>
</tr>
</tbody>
</table>

From Table 3, one can see that the use of both sets of observations results in significant improvement (3 dB to 5 dB) over using either $x_1[n]$ or $x_2[n]$ separately, even though the error associated with filtering using only $x_1[n]$ is large compared to that of $x_2[n]$.

VII. CONCLUSIONS

In this paper, we addressed the problem of optimal filtering of multiple channels observed at different sampling rates. We developed the explicit form of the Wiener-Hopf equations using linear periodically time-varying filters. The error variance of the optimal filter is periodically time-varying. Using the geometric average of the periodic error variances as a performance measure, we demonstrated that optimal filtering of multiple channels can provide improved performance over optimally filtering one channel, even if the secondary channel has high error variances. The largest performance gains occurred when large filter orders were used on the secondary channel of a sinusoidal signal. Small improvements were found in the AR process. However, the low order nature of the AR process used can account for the small improvements, since a filter order of 5 already over estimates the order of the process.

In addition, simulations were conducted with filter orders for $P_1$ and $P_2$ ranging from 5 to 30 for an environment with a Signal to Noise Ratio of 6 dB. Selected results are tabulated below in Table 3 for the signal with two sinusoids showing the differences in filtering with both signals and filtering using only one signal. Results are not provided for the AR process. Because of the low order of the AR process, varying the filter orders between 5 and 30 has no appreciable effect on the error variance.

REFERENCES