**Abstract**

We consider the problem of estimating the arrival times of overlapping ocean-acoustic signals from a received signal which consists of an unknown deterministic signal along with scaled and delayed versions due to multipath propagation plus additive noise. Our objective is to simultaneously determine the transmitted waveform and the arrival times. The proposed algorithm obtains approximately maximum likelihood estimates of the arrival times and the parameters which characterize the unknown signal. Our assumptions are that the number of the different paths is known and that the signal must belong to a parametric class of signals. We demonstrate the algorithm on a class of signals consisting of gated sinusoids.
Transmit Signal Extraction in a Multipath Environment

R.J. Vaccaro and E. Maragakis
Department of Electrical Engineering
University of Rhode Island

R.L. Field
NOARL, Code 244
Stennis Space Center

Abstract

We consider the problem of estimating the arrival times of overlapping ocean-acoustic signals from a received signal which consists of an unknown deterministic signal along with scaled and delayed versions due to multipath propagation plus additive noise. Our objective is to simultaneously determine the transmitted waveform and the arrival times. The proposed algorithm obtains approximately maximum likelihood estimates of the arrival times and the parameters which characterize the unknown signal. Our assumptions are that the number of different paths is known and that the signal must belong to a parametric class of signals. We demonstrate the algorithm on a class of signals consisting of gated sinusoids.

1 Introduction

Time delay estimation is a well known problem occurring frequently in the fields of sonar, radar and geophysics. In this problem the received waveform consists of delayed and scaled replicas of the transmitted signal. This is the result of different reflections and attenuation of the signal in the channel.

The received waveform \( r(t) \) can be described mathematically as

\[
r(t) = \sum_{k=1}^{M} a_k s(t - t_k) + n(t), \quad 0 \leq t \leq T \tag{1}
\]

where \( s(t) \) is the transmitted signal, \( a_k \) the attenuation factor for path \( k \), \( t_k \) the time delay for path \( k \), \( M \) the number of different paths and \( n(t) \) a noise component. In our development we assume that the noise is white Gaussian.

The classical method for estimating the times of arrival is correlating the received waveform with the transmitted waveform. The peaks in the correlator output give the estimates of the arrival times. It can be shown that if the signals are separated in time by more than the duration of the signal autocorrelation function, the correlator is equivalent to the MLE [1]. Other approaches are given in [2], [3], and [4].

\footnote{This work was supported in part by Grant NO0014-89-K-6003 from the Naval Ocean and Atmospheric Research Laboratory (NOARL)}

A completely different approach has recently been proposed by Kirsteinz [5, 6]. The basic idea of this approach is to look at the problem in the frequency domain. Since a delay in the time domain is equivalent to multiplication by an exponential in the frequency domain, the frequency domain problem is one of fitting weighted complex exponentials to the spectrum of the received signal. Utilizing an iterative method of fitting complex exponentials as in [7] and [8], this approach provides a way of estimating the times of arrival. In this algorithm the number of different paths must be known and the spectrum of the source signal must be nonzero. The requirement that the number of paths must be known is not too restrictive since in many cases the number of different paths can be determined from the geometry of the channel.

All the above methods require the source signal to be known. In our problem we assume that the source signal is not known. Our objective is to simultaneously obtain good estimates for both the delays and the source signal with a minimal amount of computations. In our formulation we assume that the source signal belongs to a parametric class of signals which means that it can be completely determined by a vector of parameters. A rectangular pulse for example can be completely characterized by its duration, its amplitude and its starting point. By assuming that the source signal belongs to a certain class of signals we have to estimate a much smaller number of parameters and the problem comes much better defined. In our development we will assume that the number of paths is also known. Using those two assumptions we will develop a method of obtaining approximate maximum likelihood estimates of the time delays and the source signal parameters.

2 The maximum likelihood estimator

Making the assumption that the signal belongs to a parametric class of signals, we can rewrite equation (1) as

\[
r(t) = \sum_{k=1}^{M} a_k a(t - t_k; \theta) + n(t), \quad 0 \leq t \leq T \tag{2}
\]

where \( \theta \) is the vector of parameters which characterize the source signal. In the case that \( n(t) \) is white Gaussian noise the least squares estimator is also the maximum likelihood estimator.
estimator (Helstrom, pp 199 [9]). Therefore the MLE is given by
\[
\min_{\lambda, A} \int_{-\infty}^{\infty} |R(\omega) - S(\omega; \theta) \sum_{k=1}^{M} a_k e^{j\lambda_k \omega}|^2 d\omega
\]  

where \( R(\omega) \) and \( S(\omega; \theta) \) are the Fourier transforms of \( r(t) \) and \( s(t; \theta) \), respectively. It should be noted that this expression for the MLE is only valid for white Gaussian noise. However even if the noise is not Gaussian the estimator may still be useful as the least squares estimator.

The problem now is to approximate \( R(\omega) \) by a weighted sum of complex exponentials. If we sample the frequency functions with spacing \( \Delta \omega \), we have
\[
e^{-j\lambda \omega} - e^{-j\lambda_k \omega} = e^{j\lambda_k \omega}
\]

where \( \lambda_k = -\tau_k \Delta \omega \).

After sampling the integrals are approximated by sums and the MLE is given approximately by
\[
\min_{\lambda, A} \sum_{n=0}^{N-1} |R(n + q) - w(n + q) \sum_{k=1}^{M} a_k e^{j\lambda_k (n-q)}|^2
\]  

where \( w(n; \theta) = S(\Delta \omega n; \theta) \), and \( L \) is the total number of points in the discrete Fourier transform applied to the sampled data.

Remarks

Before giving an algorithm for estimating parameters from frequency domain data, we first give some features of the frequency domain formulation of the MLE.

1. Note that the delay parameters are estimated as real numbers even when using sampled data. There is no need for interpolation as there would be in a time-domain formulation.

2. The frequency-domain formulation is equivalent to modeling the spectrum of the received signal as a weighted sum of complex exponentials with real-valued coefficients. The complex exponentials do not occur with conjugate symmetry, in general. The fitting of unweighted complex exponentials with complex amplitudes is a well-known problem, and accounting for the weights is simple. However, constraining the amplitudes to be real when the data is complex has apparently not been considered before. We show in the next section how to include this constraint.

3 The known signal algorithm

Using the notation developed in the previous section, we now consider the case when the source signal is known and is narrowband. We show in the next section how the known-signal algorithm can be used iteratively in the case when the signal is not known.

When the signal is narrowband, most of the energy of the signal is concentrated in the passband. For example in a gated sinusoid most of the energy of the signal is concentrated in the main lobe around the center frequency. In this case, we do not have to include all frequency points in the minimization; we only have to include those which contain some signal energy. If we take \( N \) points starting at \( q \) corresponding to positive frequencies where the spectrum of the transmitted signal is nonzero, we can define the following error function
\[
E_1(\lambda, A) = \sum_{n=0}^{N-1} |R(n + q) - w(n + q) \sum_{k=1}^{M} a_k e^{j\lambda_k (n-q)}|^2
\]  

Note that the above equation is not equivalent to the MLE expression shown in (5) because it only corresponds to the positive frequency portion of the spectrum. The conjugate symmetric portion of the spectrum corresponding to negative frequencies must also be included to obtain the MLE error expression. We first define some notation.

Let
\[
r = [R(q) \, R(q+1) \, \cdots \, R(q+N-1)]^T
\]
\[
a = [a_1 \, a_2 \, \cdots \, a_M]^T
\]
\[
W = \text{diag} \{w(q) \, w(q+1) \, \cdots \, w(q+N-1)\}
\]
\[
A(\lambda) = \begin{bmatrix}
e^{j\lambda (e+1)} & e^{j\lambda (e+1)} & \cdots & e^{j\lambda (e+1)} \\
e^{j\lambda (e+1)} & e^{j\lambda (e+1)} & \cdots & e^{j\lambda (e+1)} \\
\vdots & \vdots & \ddots & \vdots \\
e^{j\lambda (e+N-1)} & e^{j\lambda (e+N-1)} & \cdots & e^{j\lambda (e+N-1)}
\end{bmatrix}
\]
\[
P(\lambda) = WA(\lambda)
\]

Then
\[
E_1(a, \lambda) = \|r - P(\lambda)a\|^2.
\]  

This is the error expression considered in [6]. However, the MLE error expression must also include the conjugate symmetric portion of the spectrum as shown below
\[
E(a, \lambda) = \|r - P(\lambda)a\|^2 + \|r^* - P^*(\lambda)a\|^2,
\]  

where the superscript * refers to complex conjugation. Note that the vector \( a \) is not conjugated in the second term of the above equation since it is assumed to be a real number. Thus the formulation in (8) is equivalent to the constraint that the amplitudes be real valued.
For any fixed \( \lambda \), the coefficients \( a_1 \) which minimize \( E \) are given by

\[
  a_1 = P^*(\lambda)r = (P^H P)^{-1}P^H r
\]

Note that the resulting vector \( a_1 \) will be complex in general substituting (9) into (7) yields

\[
  E_1(\lambda) = \| (I - P (P^H P)^{-1}P^H) r \|^2 = \| P^*(\lambda)r \|^2
\]

Similar expressions can be written for the true MLE error expression as follows

\[
  E_0(a, \lambda) = \| r - P^*(\lambda) \|^2
\]

for any fixed \( \lambda \), the coefficients \( a \) which minimize \( E \) are given by

\[
  a = \{ P^T P \}^{-1} P^T r
\]

where the superscript \( H \) stands for complex conjugate transpose, and \( \text{Real}(\cdot) \) stands for the real part of a complex number. Note that the above expression always results in real values for the amplitudes.

If we define

\[
  G = \begin{bmatrix} P(\lambda) \\ P^*(\lambda) \end{bmatrix}
\]

and substitute (13) and (14) into (11) we get

\[
  E(\lambda) = \| (I - G(G^H G)^{-1}) r \|^2 = \| G^2(\lambda) r \|^2
\]

The error is now only a function of the unknown time delays. Note that \( G^2(\lambda) \) is the matrix which projects onto the orthogonal complement of the column-space of \( G \).

The expressions for \( E_1 \) (which uses half the spectrum) and \( E \) (which uses the complete spectrum) are of the same amount of noise added. In Fig. 2, we show the error surfaces for \( E_1 \) and \( E \) for a one-path example to gain some insight into the relationship between these two error expressions.

### 3.1 A One-Path Example to Compare \( E \) and \( E_1 \)

The errors \( E_1 \) and \( E \) can be written as functions of the unknown time delays only as shown in (10) and (15). In this section, we give an example of a signal containing a single time delay to compare \( E_1 \) and \( E \). The transmitted signal in this example consists of a 244 Hz gated sinusoid whose duration is 40 ms. The received signal was obtained by delaying the transmitted signal by 50 ms. Thus the error surfaces should have minima at t=0.05 secs.

In Fig. 1, we show the received signal with a moderate amount of noise added. In Fig. 2, we show the error surfaces for \( E \) and \( E_1 \) corresponding to the received data in Fig. 3. Note that both \( E \) and \( E_1 \) have minima at \( t = 0.05 \), but \( E_1 \) is a smooth unimodal function, while \( E \) is a modulated sinusoid. Our attempts to minimize \( E \) converged to some local minimum unless the algorithm was initialized very close to the global minimum. On the other hand, our algorithm converged to the global minimum of \( E_1 \) for a wide range of initializations.

By looking only at Fig. 2, we might conclude that minimizing \( E_1 \) always gives the same results as minimizing the true MLE expression \( E \). However, this is not the case. Fig. 3 shows the same received signal with a large amount of additive noise, and Fig. 4 shows the corresponding error surfaces \( E_1 \) and \( E \). The error surface \( E \) still has a minimum at \( t = 0.05 \) but the minimum of \( E_1 \) occurs at \( t = 0.048 \). Thus it seems that minimizing \( E_1 \) will result in biased estimates of the time delays at low signal-to-noise ratio. Nevertheless, we choose to work with \( E_1 \) instead of \( E \) because it is easier to find the global minimum of \( E_1 \) than it is for \( E \).
As shown in the above example, \( E_i \) will work well provided the signal-to-noise ratio is large enough. The development of an efficient algorithm to find the global minimum of \( E \) for a wide range of initial estimates is an important open problem.

### 3.2 A Perturbation Expansion Approach to Minimizing \( E_1 \)

An iterative algorithm for minimizing \( E_1 \) has been derived in [6]-[8],[10]. However, the algorithm described in these references failed to converge for the example given in the next section of this paper. Here we briefly describe a new approach to minimizing \( E_1 \).

We begin by considering the value of the error function at an increment \( \Delta x \) away from a nominal value \( x \) of the coefficients

\[
E(x + \Delta x) = \| f(x + \Delta x) \|^2.
\]

After some calculation, a first-order perturbation expansion for \( E(x + \Delta x) \) can be obtained. The result is

\[
E(x + \Delta x) \approx \langle f \Delta \xi + \cdots \rangle + \langle \Delta f \Delta \xi \rangle + \cdots
\]

where \( \Delta \xi \) represents the error function to first-order in \( \Delta x \).

Substituting the above expression into (20) with

\[
E(z + \Delta z) = \| f + \Delta f \|^2
\]

where \( z = P_B y + Q_B \Delta B^H P_B^H y + P_B^H \Delta B Q_B y \|^2
\]

we have

\[
E(z + \Delta z) \approx \| f + \Delta f \|^2
\]

where \( z = P_B y + \Delta f \).

Using the estimated time delays \( \tau_s \), calculate new values \( \tau_{s+1}, \tau_{s+1} \) for the frequency and duration.

4. Check for convergence; return to 2.

The questions that we now have to address are how to estimate \( \tau \) and \( \tau \) and how to obtain the first estimates of \( \tau \) and \( \tau \).

Consider the error function for a given set of time delays:

\[
E(f, d) = \sum_{n=0}^{N-1} |R(n+q) - \sum_{k=1}^{m} a_k e^{iA_k(n+q)}|^2
\]

where \( f \) and \( d \) are the frequency and the duration respectively. This expression is not easy to minimize. It is much more complicated than a fitting of exponentials since \( w \) is also nonlinear in \( f \) and \( d \). A multidimensional search is practically impossible because of its computational intensity.

Our algorithm reduces the computations by breaking the problem down and solving for the parameters and the delays as follows:

1. Obtain initial estimates \( f^0, d^0 \) of the unknown frequency and duration.
2. Use \( f^0 \) and \( d^0 \) in the known-signal algorithm to estimate the time delays \( \tau_s^0 \).
3. Using the estimated time delays \( \tau_s^0 \), calculate new values \( f^{s+1}, d^{s+1} \) for the frequency and duration.
4. Check for convergence; return to 2.

The questions that we now have to address are how to estimate \( f^0 \) and \( d^0 \) and how to obtain the first estimates of \( f^0 \) and \( d^0 \).
2. For each value of \( d \) find the best value of \( f \) using a gradient based technique.

3. Repeat 2 until a minimum is found.

Note that during the minimization \( E(f, d) \) will have to be evaluated several times. Each evaluation is made faster by the fact that \( w(f, d) \) can be computed by an FFT.

The last question is how to obtain the initial estimates \( f_0 \) and \( d_0 \). Those estimates do not have to be very accurate, in fact they can be quite crude since they are only used to initialize the algorithm.

For our experimental arrangement the geometry of the channel gives us that the first two paths will be the least attenuated and will probably be overlapping. After filtering the reflections from the first two paths will lie over the noise. Taking advantage of this we can get estimates of \( f \) and \( d \) as follows.

- Obtain the envelope of the signal as follows: for each point take the maximum amplitude over the next 20 points (generally choose a number of points sufficient to cover a period of the sinusoid).
- Take the initial estimate of the duration to be \((Q - 20)/2\).
- The initial estimate of the frequency is found using a standard frequency estimation algorithm on the data points between \( L + 20 \) and \( Q \).

This completes the description of our algorithm. In the next section we will demonstrate the algorithm on both real and simulated data.

5 Example with Experimental Data

In our experiments the geometry of the channel gives us four different paths with the first two less attenuated than the other two. This situation arose in ocean acoustic signals reflected on the surface and the bottom of the ocean. When the data was collected, a hydrophone near the transmitter recorded the actual source signal. The transmitted signal was a gated sinusoid of frequency 244 Hz and duration 40 ms.

A record of received data is shown in Fig. 5. Using the known-signal algorithm presented in section 3 of this paper, we estimated the time delays and amplitudes of the four paths. We then constructed an estimate of the received signal using our estimates of the delays and amplitudes as well as the known source signal. The actual received signal and our reconstructed signal are shown together in Fig. 6, which shows that the estimates provide a good fit to the data.

Next, we applied the unknown signal algorithm presented in the previous section to try to estimate the parameters of the transmitted signal as well as the delays and amplitudes in the received signal. The estimates of both the known signal and unknown signal algorithms are shown in the Table 1.

We believe that the estimates obtained by both algorithms would be improved by minimizing error surface \( E_1 \) instead of \( E \). We also remark that the estimates of the amplitudes are extremely sensitive to the values of the estimated delays, however the amplitudes are always calculated to give a least-squares fit to the data for a given set of time delays.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>True parameter</th>
<th>K.S. estimate</th>
<th>U.S. estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration</td>
<td>244.0</td>
<td>-</td>
<td>239.3</td>
</tr>
<tr>
<td>( r_1 )</td>
<td>0.0400</td>
<td>-</td>
<td>0.0365</td>
</tr>
<tr>
<td>( r_2 )</td>
<td>-</td>
<td>0.1550</td>
<td>0.1550</td>
</tr>
<tr>
<td>( r_3 )</td>
<td>-</td>
<td>0.1640</td>
<td>0.1657</td>
</tr>
<tr>
<td>( r_4 )</td>
<td>-</td>
<td>0.1920</td>
<td>0.1902</td>
</tr>
<tr>
<td>( a_1 )</td>
<td>-</td>
<td>0.6448</td>
<td>-0.2657</td>
</tr>
<tr>
<td>( a_2 )</td>
<td>-</td>
<td>-0.8463</td>
<td>-0.7162</td>
</tr>
<tr>
<td>( a_3 )</td>
<td>-</td>
<td>-0.4237</td>
<td>0.0304</td>
</tr>
<tr>
<td>( a_4 )</td>
<td>-</td>
<td>-0.0207</td>
<td>0.4871</td>
</tr>
</tbody>
</table>

Table 1: Estimates of delays and amplitudes using the known signal (K.S.) and unknown signal (U.S.) algorithms.

References

Figure 1. A received signal consisting of a single path with a moderate amount of additive noise.

Figure 2. The error surfaces $E$ and $E_1$ corresponding to the received signal in Fig. 1. The smooth curve is $E_1$ and the oscillating curve is $E$.

Figure 3. A received signal consisting of a single path with a large amount of additive noise.

Figure 4. The error surfaces $E$ and $E_1$ corresponding to the received signal in Fig. 2. The smooth curve is $E_1$ and the oscillating curve is $E$.

Figure 5. A record of experimental data with four overlapping paths moderate amount of additive noise.

Figure 6. The received signal of Fig. 5 together with the reconstructed signal obtained by the known signal algorithm.