ARRAY SIGNAL PROCESSING IN THE KNOWN WAVEFORM AND STEERING VECTOR CASE

Yi Jiang   Jian Li

Department of ECE,
University of Florida,
Gainesville, FL 32611, USA

Petre Stoica

Dept. of Systems and Control,
Uppsala University,
PO Box 337, SE-751 05, Uppsala, Sweden

ABSTRACT

The amplitude estimation of a signal whose waveform is known (up to an unknown scaling factor) in the presence of interference and noise is of interest in several applications including using the emerging Quadrupole Resonance (QR) technology for explosive detection. In such applications a sensor array is often deployed for interference suppression. This paper considers the complex amplitude estimation of a known waveform signal whose array response is also known *a priori*. We study a practical scenario where the interference and noise is both spatially and temporally correlated. We model the interference and noise vector as a multichannel autoregressive (AR) random process. A cyclic iterative ML (IML) method is presented. We show that in most cases the IML method is superior to its simple ML counterpart that ignores the temporal correlation of the interference and noise.

1. INTRODUCTION

Estimating the signal parameters in the presence of interference and noise via array processing is often encountered in practical applications (see, e.g., [1] and the references therein). In several emerging applications, such as using the Quadrupole Resonance (QR) technology for explosive detection [2], the temporal signal waveform is known *a priori* up to an unknown scaling factor and the array response is also given. In QR applications, for example, one of the sensors receives the signal of interest as well as the interference and noise while the remaining sensors receive the interference and noise only. Hence one of the elements of the array steering vector for the signal of interest is one and the remaining elements are zero. It is well known that the temporal information on the signal can be utilized to effectively suppress the interference and noise and hence to significantly improve the estimation accuracy (see, e.g., [3]). However, exploiting both the temporal and spatial information on the signal for interference suppression and signal parameter estimation is a practically important problem that has not been fully investigated before to the best of our knowledge. We presented in [4] a comparative study of Capon and ML that utilize both the temporal and spatial information on the signal for amplitude estimation in the presence of temporally white but spatially colored interference and noise. We showed that the ML estimate is generally superior to Capon. In this paper, we consider a more general scenario where the interference and noise are *both* spatially and temporally correlated. We model the interference and noise vector as a multichannel autoregressive (AR) random process. A cyclic iterative ML (IML) method is presented. We show that in most cases the IML method is superior to its simple ML counterpart that ignores the temporal correlation of the interference and noise.

This work was supported in part by the National Science Foundation Grant CCR-0104887 and the Swedish Foundation for Strategic Research (SSF).

2. SIMPLE MAXIMUM LIKELIHOOD METHOD

We use the following data model in [4],

$$\mathbf{x}_l = \mathbf{a}\beta s_l + \mathbf{e}_l, \quad l = 1, 2, \cdots, L, \tag{1}$$

where $\mathbf{x}_l \in \mathcal{C}^{M\times 1}$, $l = 1, 2, \cdots, L$, denotes the lth array output vector (with M being the number of sensors and L being the number of snapshots), the array steering vector $\mathbf{a} \in \mathcal{C}^{M\times 1}$ of the signal of interest is known, β is the unknown complex amplitude of the signal whose temporal waveform $\{s_l\}_{l=1}^{L}$ is known. We model the interference and noise term $\mathbf{e}_l \in \mathcal{C}^{M\times 1}$ as a zero-mean temporally white but spatially colored circularly symmetric complex Gaussian random process with an unknown and arbitrary spatial covariance matrix $\mathbf{Q}$. The ML method estimates the signal amplitude by maximizing the likelihood function of the random vectors $\{\mathbf{x}_l\}_{l=1}^{L}$. We show in [4] that

$$\hat{\beta}_{ML} = \frac{\mathbf{a}^H\mathbf{T}^{-1}\bar{\mathbf{x}}}{P_s\mathbf{a}^H\mathbf{T}^{-1}\mathbf{a}}, \tag{2}$$

where

$$\mathbf{T} = \frac{1}{L}\sum_{l=1}^{L}\mathbf{x}_l\mathbf{x}_l^H - \frac{\bar{\mathbf{x}}\bar{\mathbf{x}}^H}{P_s}. \tag{3}$$

$$\bar{\mathbf{x}} = \frac{1}{L}\sum_{l=1}^{L}\mathbf{x}_l s_l^*, \tag{4}$$

and

$$P_s = \frac{1}{L}\sum_{l=1}^{L}|s_l|^2 \tag{5}$$

is the average power of the known waveform. In the following sections, we will refer to the ML method presented here that ignores the temporal correlation of the interference and noise as the simple ML (SML) method only for distinction purposes.

3. ITERATIVE MAXIMUM LIKELIHOOD METHOD

The previous study assumed that the interference and noise term in (1) is spatially colored but temporally white. In this section, we model the interference and noise vector as a multichannel autoregressive (AR) random process and propose an iterative ML (IML) method based on the cyclic optimization approach.

3.1. Data Model

Consider the data model:

$$\mathbf{x}_l = \mathbf{a}\beta s_l + \mathbf{v}_l, \quad l = 1, 2, \ldots, L, \tag{6}$$

which is the same as the one in (1) except that the interference and noise term now satisfies the following AR Equation

$$\mathbf{A}(z^{-1})\mathbf{v}_l = \mathbf{e}_l, \tag{7}$$

where z^{-1} is the unit delay operator,

$$\mathbf{A}(z^{-1}) = \mathbf{I} + \mathbf{A}_1 z^{-1} + \mathbf{A}_2 z^{-2} + \ldots + \mathbf{A}_p z^{-p}, \tag{8}$$

and

$$\mathcal{E}[\mathbf{e}_l\mathbf{e}_n^H] = \mathbf{Q}\delta_{ln}, \tag{9}$$

where δ_{ln} denotes the Kronecker delta. Note that if only the interference component in $\mathbf{v}_l$ is a multichannel AR process while the noise component in $\mathbf{v}_l$ is white temporally, then the interference and noise term will be a multichannel autoregressive and moving average random process, which can still be approximated by a multichannel AR process. The SNR for the data model in (6) is defined as

$$\text{SNR} = \frac{MP_s|\beta|^2}{\text{tr}\,(\mathbf{R}_{\mathbf{v}})}, \tag{10}$$

where $\mathbf{R}_{\mathbf{v}}$ is the covariance matrix of $\{\mathbf{v}_l\}$.

3.2. Algorithm

Conditioned on the first p data vectors $\{\mathbf{x}_l\}_{l=1}^{p}$, maximizing the log-likelihood function is equivalent to minimizing

$$C_1 = \left|\sum_{l=p+1}^{L}\left[\mathbf{A}(z^{-1})(\mathbf{x}_l - \mathbf{a}\beta s_l)\right]\left[\mathbf{A}(z^{-1})(\mathbf{x}_l - \mathbf{a}\beta s_l)\right]^H\right| \tag{11}$$

with respect to both β and $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_p]$. Hence the optimization problem becomes very complicated. Here we propose an iterative ML (IML) approach to solve this problem.

To begin with, we obtain an initial estimate $\hat{\beta}^{(0)}$ of β by using the SML method (cf (2)). For a given estimate $\hat{\beta}$, let $\mathbf{z}_l = \mathbf{x}_l - \mathbf{a}\hat{\beta}s_l$. From (11), we get,

$$\hat{\mathbf{A}} = \arg\min_{\mathbf{A}}\left|\sum_{l=p+1}^{L}\left[\mathbf{A}(z^{-1})\mathbf{z}_l\right]\left[\mathbf{A}(z^{-1})\mathbf{z}_l\right]^H\right| \tag{12}$$

where

$$\mathbf{A}(z^{-1})\mathbf{z}_l = \mathbf{z}_l + [\mathbf{A}_1 \ldots \mathbf{A}_p]\begin{bmatrix}\mathbf{z}_{l-1}\\ \vdots \\ \mathbf{z}_{l-p}\end{bmatrix} \triangleq \mathbf{z}_l + \mathbf{A}\boldsymbol{\phi}_l.$$

It's easy to show that the solution to (12) is

$$\hat{\mathbf{A}} = -\hat{\mathbf{R}}_{\mathbf{z}\boldsymbol{\phi}}\hat{\mathbf{R}}_{\boldsymbol{\phi}}^{-1}, \tag{13}$$

which is the multichannel Prony estimate of $\mathbf{A}$. We assume that the order p of the multichannel random process AR(p) is known. If p is unknown, it can be estimated, for instance, by using the Generalized Akaike Information Criterion (GAIC).

For a given $\hat{\mathbf{A}}$, we obtain an improved estimate of β by minimizing the following:

$$C_2 = \left|\sum_{l=p+1}^{L}\left[\hat{\mathbf{A}}(z^{-1})\left(\mathbf{x}_l - \mathbf{a}\beta s_l\right)\right]\left[\hat{\mathbf{A}}(z^{-1})\left(\mathbf{x}_l - \mathbf{a}\beta s_l\right)\right]^H\right|. \tag{14}$$

First, we consider the case of a known damped (or undamped) sinusoidal signal, i.e., $s_l = e^{(-\alpha_s + j\omega_s)l}$ with known frequency ω_s and damping factor α_s.

Let

$$\mathbf{y}_l = \hat{\mathbf{A}}(z^{-1})\mathbf{x}_l, \quad \mathbf{b} = \hat{\mathbf{A}}(z^{-1})|_{z=e^{-\alpha_s + j\omega_s}}\mathbf{a}.$$

Note that the length of the new data sequence $\mathbf{y}_l$, $l = p+1, \ldots, L$, is $L - p$ instead of L. The solution to the above problem is given by the ML estimator proposed in Section 2:

$$\hat{\beta} = \frac{\mathbf{b}^H\mathbf{T}_y^{-1}\bar{\mathbf{y}}}{P_s\mathbf{b}^H\mathbf{T}_y^{-1}\mathbf{b}} \tag{15}$$

where

$$\mathbf{T}_y = \frac{1}{L-p}\sum_{l=p+1}^{L}\mathbf{y}_l\mathbf{y}_l^H - \frac{\bar{\mathbf{y}}\bar{\mathbf{y}}^H}{P_s}, \qquad \bar{\mathbf{y}} = \frac{1}{L-p}\sum_{l=p+1}^{L}\mathbf{y}_l s_l^* \tag{16}$$

and $P_s = \frac{1}{L-p}\sum_{l=1}^{L-p}|s_l|^2$ is the average power of the known waveform $\{s_l\}_{l=p+1}^{L}$. The IML approach maximizes the likelihood function cyclically. We set $\hat{\mathbf{A}}^{(0)} = \mathbf{0}$ and obtain $\hat{\beta}^{(0)} = \hat{\beta}_{SML}$. We then iterate the following two steps untill the solution converges, i.e., when the two consecutive estimates $\hat{\beta}^{(i)}$ and $\hat{\beta}^{(i+1)}$ are sufficiently close:

$$\hat{\mathbf{A}}^{(i+1)} = \arg\max_{\mathbf{A}} f(\mathbf{x}|\mathbf{A}; \hat{\beta}^{(i)}, \{\mathbf{x}_l\}_{l=1}^{p}), \tag{17}$$

which is given by (13) with $\hat{\beta}$ replaced by $\hat{\beta}^{(i)}$, and

$$\hat{\beta}^{(i+1)} = \arg\max_{\beta} f(\mathbf{x}|\beta; \hat{\mathbf{A}}^{(i+1)}, \{\mathbf{x}_l\}_{l=1}^{p}), \tag{18}$$

which is given by (15) with $\hat{\mathbf{A}}$ replaced by $\hat{\mathbf{A}}^{(i+1)}$.

Obviously the likelihood function never decreases in any iteration. In the simulations reported in the next section we found that IML converges in 2 or 3 iterations. Hence the IML estimator is computationally quite efficient.

Next, we consider the case of an arbitrary known waveform signal. Let

$$\mathbf{y}_l = \hat{\mathbf{A}}(z^{-1})\mathbf{x}_l, \quad \mathbf{a}_l = \hat{\mathbf{A}}(z^{-1})\mathbf{a}s_l, \quad l = p+1, \ldots, L$$

$$\mathbf{Y}_{M\times(L-p)} = [\mathbf{y}_{p+1}\mathbf{y}_{p+2}\ldots\mathbf{y}_L]$$

and

$$\mathbf{G}_{M\times(L-p)} = [\mathbf{a}_{p+1}\mathbf{a}_{p+2}\ldots\mathbf{a}_L].$$

Also let $\mathbf{P}$ be an orthogonal projection matrix defined as

$$\mathbf{P} = \mathbf{G}^H(\mathbf{G}^H)^\dagger, \tag{19}$$

where $(\mathbf{G}^H)^\dagger$ is the Moore-Penrose pseudo-inverse of $\mathbf{G}^H$, and let $\mathbf{P}^\perp = \mathbf{I} - \mathbf{P}$.

Then (14) can be written concisely in a matrix form:

$$\begin{aligned}
\hat{\beta} &= \arg\min_{\beta}\left|(\mathbf{Y}-\beta\mathbf{G})(\mathbf{Y}-\beta\mathbf{G})^H\right| \\
&= \arg\min_{\beta}\left|(\mathbf{Y}-\beta\mathbf{G})(\mathbf{P}+\mathbf{P}^\perp)(\mathbf{Y}-\beta\mathbf{G})^H\right| \qquad (20)\\
&= \arg\min_{\beta}\left|(\mathbf{YP}-\beta\mathbf{G})(\mathbf{YP}-\beta\mathbf{G})^H + \mathbf{YP}^\perp\mathbf{Y}^H\right| \\
&\triangleq \arg\min_{\beta}\left|(\mathbf{YP}-\beta\mathbf{G})(\mathbf{YP}-\beta\mathbf{G})^H + \mathbf{T}_y\right| \qquad (21)\\
&= \arg\min_{\beta}\left|(\mathbf{YP}-\beta\mathbf{G})(\mathbf{YP}-\beta\mathbf{G})^H\mathbf{T}_y^{-1} + \mathbf{I}\right|\left|\mathbf{T}_y\right|,
\end{aligned} \tag{22}$$

Note that minimizing the cost function in (22) requires a two-dimensional search over the parameter (since β is complex-valued). To avoid the search, we use Lemma 1 to obtain an approximate estimate of β.

Lemma 1. *For a large data sample number L, minimizing $F_1 = \left|(\mathbf{YP}-\beta\mathbf{G})(\mathbf{YP}-\beta\mathbf{G})^H\mathbf{T}_y^{-1}+\mathbf{I}\right|$, is asymptotically equivalent to minimizing*

$$F_2 = \operatorname{tr}\left[(\mathbf{YP}-\beta\mathbf{G})^H\mathbf{T}_y^{-1}(\mathbf{YP}-\beta\mathbf{G})\right]. \tag{23}$$

Proof. Omitted. □

It follows from (23) that minimizing F_2 with respect to β yields

$$\hat{\beta} = \frac{\operatorname{tr}(\mathbf{G}^H\mathbf{T}_y^{-1}\mathbf{Y})}{\operatorname{tr}(\mathbf{G}^H\mathbf{T}_y^{-1}\mathbf{G})}, \tag{24}$$

where we remind the reader that $\mathbf{T}_y = \mathbf{YP}^\perp\mathbf{Y}^H$. Because (24) is only an approximate solution to (22) in this more general case, the IML method based on (24) is no longer an iterative ML approach (but we still keep the name for convenience) and consequently it is not theoretically guaranteed that IML will yield a more accurate solution than the SML method. However, in our numerical examples, IML outperforms SML in most cases even for modest data sample lengths. To avoid any "convergence problem" in this case in which IML is no longer an iterative minimizer, we simply pre-impose the number of iterations to be 3.

4. NUMERICAL EXAMPLES

Consider the case where the steering vector is given by $\mathbf{a} = [1\ 0\ 0\ 0]^T$ with $(\cdot)^T$ denoting the transpose. First, we assume that $s_l = 1$, $l = 1, 2, \cdots, L$, which is a sinusoid signal with frequency zero. Then we assume s_l to be a known BPSK signal which stands for an arbitary waveform signal. In all the examples, we assume that $\beta = 1$. We generate a multichannel AR(2) random process with the method in [5]. The autocorrelation matrices are given by

$$[\mathbf{R}_\mathbf{v}(l)]_{mn} = \rho\rho_s^{|m-n|}\exp\{-\rho_t l^2 + j(m-n+l)\omega\}, \tag{25}$$

and

$$\mathbf{R}_\mathbf{v}(l) = \mathbf{R}_\mathbf{v}{}^H(-l), \quad l = 0, 1, \ldots, p, \tag{26}$$

where $\rho = \frac{1}{\text{SNR}}$, ρ_s controls the spatial correlation, ρ_t partly decides the temporal correlation, and ω defines the spectral peak location of the colored interference and noise in each channel. The data sample number is $L = 50$. When we use the true autoregressive matrix $\mathbf{A}$ in the IML instead of the estimated one, we refer to the method as the known-AR ML (KML) approach. We include KML for comparison purposes only. We obtain the empirical MSEs of the estimates by using 500 Monte-Carlo trials.

For the constant signal case, our simulations results showed in Figures 1 and 2 suggest as follows:

A: Both IML and SML work better for large ω and/or small ρ_t, B: IML is slightly worse than SML for small ω and/or large ρ_t,

C: IML is significantly better than SML for large ω and/or small ρ_t.

These observations can be explained by examining the signal and the interference and noise terms in the temporal frequency domain. For large ω, the signal is separated from the interference and noise in the temporal frequency domain, which benefits both methods. Similarly, smaller ρ_t means higher correlation in the temporal time domain or more peaky spectra in the temporal frequency domain. Hence both estimators perform better for this case when ω is away from zero. This explains Observation A. Next, we note that a large ρ_t means low correlation in the temporal domain and hence the interference and noise vector is approximately temporally white. KML is approximately SML in such a case. Since IML is inferior to KML, IML is also slightly worse than SML. For small ω, the signal and the interference and noise terms are not well separated in the temporal frequency domain. This explains Observation B. Observation C is as expected.

Finally, we consider the known BPSK signal case. We see from Figure 3 that the IML method significantly outperforms SML (over 10 dB) even for modestly temporally correlated interference and noise ($\rho_t = 0.1$) although it is slightly inferior to SML when the temporal correlation of the interference and noise is weak ($\rho_t = 2$). Our simulations also suggest that a known wideband signal makes suppressing temporally correlated interference and noise easier than a narrowband one.

5. REFERENCES

[1] A. L. Swindlehurst and P. Stoica, "Maximum likelihood methods in radar array signal processing," *Proceedings of the IEEE*, vol. 86, pp. 421–441, February 1998.

[2] M. Rowe and J. Smith, "Mine detection by nuclear quadrupole resonance," *EUREL International Conference*, no. 431, pp. 62–66, 1996.

[3] J. Li, B. Halder, P. Stoica, and M. Viberg, "Computationally efficient angle estimation for signals with known waveforms," *IEEE Transactions on Signal Processing*, vol. 43, pp. 2154–2163, September 1995.

[4] Y. Jiang, J. Li, and P. Stoica, "Array signal processing for QR," *the 36th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2002.

[5] J. H. Michels, P. Varshney, and D. D. Weiner, "Synthesis of correlated multichannel random processes," *IEEE Transactions on Signal Processing*, vol. 42, pp. 367–375, February 1994.

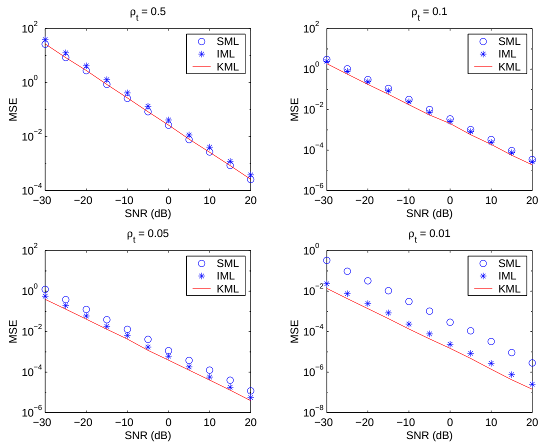

Fig. 1. MSEs of IML, SML and KML estimates for a constant signal vs. SNR when $\omega = 1$ and $\rho_s = 0.6$

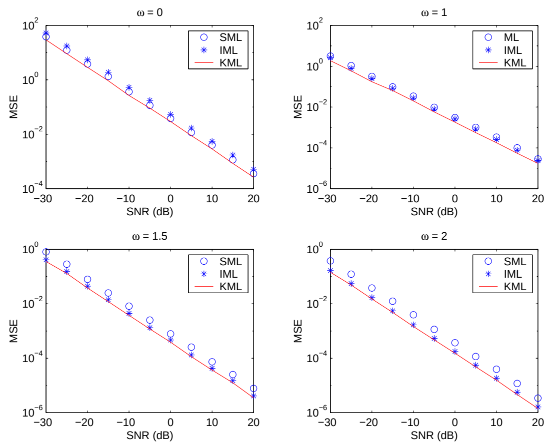

Fig. 2. MSEs of IML, SML and KML estimates for a constant signal vs. SNR when $\rho_t = 0.1$ and $\rho_s = 0.6$.

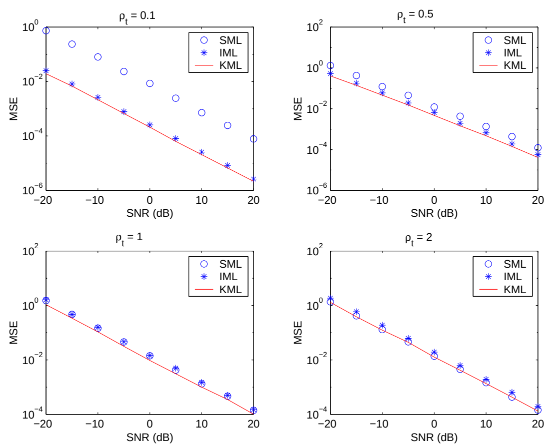

Fig. 3. MSEs of IML, SML and KML estimates for a known BPSK signal vs. SNR when $\omega = 0$ and $\rho_s = 0.6$.