MULTI-CHANNEL PARAMETRIC ESTIMATOR FAST BLOCK MATRIX INVERSES

S. Lawrence Marple Jr.*
School of Electrical Engineering and Computer Science
Oregon State University
Corvallis, OR 97331
Marple@eeecs.oregonstate.edu

Phillip M. Corbell, Muralidhar Rangaswamy†
Air Force Research Laboratory
Sensors Directorate
Hanscom AFB, MA 01731
Phillip.Corbell@hanscom.af.mil,
Muralidhar.Rangaswamy@hanscom.af.mil

ABSTRACT
The optimal (adaptive) linear combiner (beamformer) weights for a sensor array are expressed in terms of the inverse of the multi-channel (MC) covariance matrix. Rather than form an estimate of the covariance matrix directly from the available data and inverting it, an alternative direct estimate of the inverse may be obtained by forming parametric MC linear prediction estimates and then expressing the inverse in terms of these parametric MC estimates. The resulting parametric estimate of the inverse is typically more accurate than inverting the estimate of the covariance matrix. This paper reveals, for the first time, the structure of the inverse of the covariance matrix for the MC version of the covariance least squares linear prediction algorithm. The inverse structure involves products of triangular block MC Toeplitz matrices, which leads to fast computational solutions for the optimal weights.

Index Terms— Adaptive Arrays, Matrix Decomposition, Matrix Inversion, Radar Signal Processing, Multichannel

1. INTRODUCTION
The output $y = w^T x$ of a linear combiner (space-time filter) operating on an arbitrary geometry sensor array is expressed in terms of an inner product of a multi-channel (MC) weight vector $w$ of dimension $M \times 1$ and a comparable dimension MC data vector from the array. For a stationary array (position does not change with time), the MC data vector is a function of just spatial dimension $M$. For a non-stationary array (position changes with time), the MC data vector is a function of both spatial dimension $M$ and temporal dimension $N$. An example that motivates this paper is the radar space-time adaptive processing (STAP) problem, as summarized in [1]. It is well known that the optimal (adaptive) weights that minimize the variance of the output $y$ while passing a signal from a preferred steering vector direction represented by the vector $e$ is

$$ w = R^{-1} e \quad (1) $$

for which the MC covariance matrix $R$ is typically estimated as

$$ \hat{R} = \frac{1}{R} \sum_{r=1}^{R} x[r] x^H[r] \quad (2) $$

over $R$ measured realizations of the MC data vector $x$. However, we have found that improved estimates of the inverse MC covariance matrix may be obtained using MC parametric estimation algorithms because the inverse covariance matrix can be expressed directly in terms of the estimated MC parameters. This paper summarizes the relationship of the inverse to the MC parameters. Details of the actual performance of the adaptive array using the MC parametric approach will be found in companion papers presented at the Asilomar 2006 and DASP 2006 conferences. For reference, we first illustrate the inversion formula in terms of the MC parameters in the known covariance case, and show that the inverse can be expressed in terms of products of block triangular Toeplitz structures. Next, we reveal for the first time the inverse obtained from MC data using the MC parametric approach will be expressed in terms of products of block triangular Toeplitz structures, which leads to fast computational solutions for the optimal weight vectors (covered in the other conference papers).

2. MULTI-CHANNEL LINEAR PREDICTION
PARAMETRIC MODEL AND INVERSE FOR KNOWN COVARIANCE
The MC forward linear prediction error (LPE) of model order $p$ for the MC signal $x[n; r]$ of $M$ channels at each realization $r$ is defined as

$$ e_p[n; r] = x[n; r] + \sum_{k=1}^{p} A_p[k] x[n - k; r] \quad (3) $$
Multi-Channel Parametric Estimator Fast Block Matrix Inverses

School of Electrical Engineering, and Computer Science, Oregon State University, Corvallis, OR, 97331

Approved for public release; distribution unlimited

with dimension $M \times 1$, in which $A_p[k]$ are the MC forward linear prediction parameter matrices of dimension $M \times M$. This may be expressed in vector inner product form as

$$ e_p^n[n; r] = (I \; a_p) \; x_p^n[n; r] $$

for which $a_p = (A_p[1] \ldots A_p[p])$ is a $M \times M p$ block row vector of the MC forward LP parameters and

$$ x[n; r] = \begin{pmatrix} x_1[n; r] \\ \vdots \\ x_M[n; r] \end{pmatrix} \; ; \; x_p^n[n; r] = \begin{pmatrix} x_1[n; r] \\ \vdots \\ x[n; r] \end{pmatrix} $$

are $M \times 1$ and $M(p + 1) \times 1$ column vectors, respectively, of data samples. Similarly, the MC backward linear prediction error (LPE) is defined as

$$ e_p^b[n; r] = x[n - p; r] + \sum_{k=1}^{p} B_p[k] x[n - p + k; r] $$

for which $b_p = (B_p[p] \ldots B_p[1])$ is a $1 \times p$ block row vector (scalar dimension: $M \times M p$) of the MC backward LP parameters (note that the backward LP parameter vector has reversed time indexing relative to the forward LP parameter vector). The $M \times M$ MC forward LPE variance is both defined and expressed as

$$ P_a^p = \mathcal{E} \{ e_p^n[n](e_p^n[n])^H \} = (I \; a_p) \; R_p \; (I \; a_p)^H $$

in which $R_p$ is the $M \times M$ MC block Toeplitz autocovariance matrix [it is also hermitian symmetric with a scalar dimension of $M(p + 1) \times M(p + 1)$]


composed of the stationary $M \times M$ scalar-dimensioned MC autocovariance matrix elements

$$ R[m] = \begin{pmatrix} r_{11}[m] & \cdots & r_{1M}[m] \\ \vdots & \ddots & \vdots \\ r_{M1}[m] & \cdots & r_{MM}[m] \end{pmatrix} $$

in which $r_{jk}[m]$ is the scalar cross-covariance between channels $j$ and $k$ at time lag index $m$. In a similar manner, the $M \times M$ MC backward LPE variance is given by

$$ P_b^p = \mathcal{E} \{ e_p^b[n](e_p^b[n])^H \} = (b_p \; I) \; R_b \; (b_p^H \; I) $$

The choices for the forward and backward LP parameters $a_p$ and $b_p$ that minimize the variances $P_a^p$ and $P_b^p$ can be shown [3] to satisfy the pair of MC Yule-Walker normal equations

$$ (I \; a_p) \; R_p = (P_a^p \; 0_p) $$

$$ (b_p \; I) \; R_p = (0_p \; P_b^p) $$

in which $0_p$ is a $M \times M p$-array of all-zeros. Define the $M \times M$ MC partial correlation coefficient

$$ \Delta_p = \mathcal{E} \{ e_p^{a-1}[n](e_p^{b-1}[n-1])^H \} $$

and the $M \times M$ MC lattice filter parameters $\Gamma_p^a = A_p[p]$ and $\Gamma_p^b = B_p[p]$; note that $A_p[p] \neq B_p[p]$, in contrast to the one-dimensional case where they were identical. The fast MC Levinson algorithm [3] can solve the $M(p + 1) \times M(p + 1)$ Eqs. 11 and 12 in a number of computational steps proportional to $(M p)^2$, rather than proportional to the usual $(M p)^3$, while avoiding the explicit formation of the block matrix $R_p$. Assuming the initialization $P_0 = P_0 = R[0]$, the four steps of the MC Levinson-like recursion, that solves for the MC LP parameters $a_k$ and $b_k$ and the MC LPE variances $P_k^a$ and $P_k^b$ over orders $k = 1, \ldots, p$, are

$$ \Delta_k = (I \; a_{k-1}) \begin{pmatrix} R[k] \\ \vdots \\ R[1] \end{pmatrix} $$

$$ \Gamma_k^a = -\Delta_k (P_{k-1}^a)^{-1} $$

$$ \Gamma_k^b = -\Delta_k (P_{k-1}^b)^{-1} $$

$$ (I \; a_k) = (I \; a_{k-1}) \; 0 + \Gamma_k^a (0 \; b_{k-1} \; I) $$

$$ (I \; b_k) = (0 \; b_{k-1} \; I) + \Gamma_k^b (I \; a_{k-1} \; 0) $$

$$ P_k^a = (I - \Gamma_k^a P_k^a) P_{k-1}^a $$

$$ P_k^b = (I - \Gamma_k^b P_k^b) P_{k-1}^b $$

This algorithm reduces to 1-D Levinson recursion when $M = 1$ (one channel).

The preferred form of the block Toeplitz inverse [3] [2] which is useful to this project, and which also depends only on the last parameter order computed, is

$$ R_p^{-1} = A_p^a R_p^a A_p - B_p^b R_p^b B_p $$

in which the block triangular Toeplitz matrices are given as

$$ A_p = \begin{pmatrix} I & A_p[1] & \cdots & A_p[p] \\ 0 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & A_p[p] \\ 0 & 0 & \cdots & I \end{pmatrix} $$

II - 1138
3. MULTI-CHANNEL COVARIANCE LEAST SQUARES LINEAR PREDICTION PARAMETRIC MODEL AND INVERSE

A full least squares minimization against all the MC linear prediction parameter matrices \( \mathbf{a}_p \) and \( \mathbf{b}_p \) simultaneously is the basis of the MC covariance method of least squares linear prediction. To set up the problem for a least squares solution, the MC forward LPE and MC backward LPE of order \( p \) can be expressed as \( M \times (N-p)R \) block row vectors when there are a finite number of \( N \) samples and \( R \) realizations for \( M \) channels

\[
\mathbf{e}_p = \left( \mathbf{e}_p^a[p+1] ; \ldots ; \mathbf{e}_p^a[N; R] \right) = (\mathbf{I} \ \mathbf{a}_p) \ \mathbf{X}_p \ \mathbf{e}_p^b = \left( \mathbf{e}_p^b[N; 1] ; \ldots ; \mathbf{e}_p^b[p+1; R] \right) = (\mathbf{b}_p \ \mathbf{I}) \ \mathbf{X}_p
\]

in which the \( M(p+1) \times (N-p)R \) block rectangular Toeplitz data matrix \( \mathbf{X}_p \) and \( M \times R \) block data vector \( \mathbf{X}[k] \) are defined as

\[
\mathbf{X}_p = \begin{pmatrix}
\mathbf{X}[p+1] & \cdots & \mathbf{X}[N-p] & \cdots & \mathbf{X}[N]
\end{pmatrix}
\]

\[
\mathbf{X}[k] = \begin{pmatrix}
\mathbf{x}[k; 1] & \cdots & \mathbf{x}[k; R]
\end{pmatrix}
\]  

(26) \hspace{1cm} (27)

The magnitude squared error sums (or variance estimates, if divided by \( (N-p)R \)) can then be expressed as

\[
\mathbf{P}_p^a = (\mathbf{e}_p^a)^* (\mathbf{e}_p^a)^H
\]

\[
\mathbf{P}_p^b = (\mathbf{e}_p^b)^* (\mathbf{e}_p^b)^H
\]  

(30) \hspace{1cm} (31)

Minimizing the trace of each of these estimated variances \([3]\) then leads to the following \((p+1)\)-block-dimensioned normal equations

\[
(\mathbf{I} \ \mathbf{a}_p) \ (\mathbf{X}_p \mathbf{X}_p^H) = (\mathbf{P}_p^a \ 0_p)
\]

\[
(\mathbf{b}_p \ \mathbf{I}) \ (\mathbf{X}_p \mathbf{X}_p^H) = (0_p \ \mathbf{P}_p^b)
\]  

(32) \hspace{1cm} (33)

In general, the MC forward and backward LP parameters in the least squares case are not complex conjugates of the other, unlike the situation in the one-channel known autocovariance case. Comparing Eqs. 32 and 33 to the recursions in the known covariance case of the previous section, one can see that the block non-Toeplitz \((p+1) \times (p+1)\) least-squares-based product matrix \( \mathbf{X}_p \mathbf{X}_p^H \) has replaced the role of the block Toeplitz \((p+1) \times (p+1)\) autocovariance matrix \( \mathbf{R}_p \). When the MC covariance least squares linear prediction algorithm is used, the inverse \((\mathbf{X}_p \mathbf{X}_p^H)^{-1}\) will be used in lieu of \( \mathbf{R}_p^{-1} \) in all the detection statistics discussed in \([5]\). Despite the fact that least-squares normal Eqs. 32 and 33 do not have a block Toeplitz matrix, the non-Toeplitz \( \mathbf{X}_p \mathbf{X}_p^H \) matrix is the product of two rectangular-shaped block Toeplitz data matrices \( \mathbf{X}_r \) and \( \mathbf{X}_c^H \). This is sufficient to generate an order-recursive fast computational algorithm that requires a number of computational operations proportional to \( p^3 \), rather than the usual solution proportional to \( p^5 \), to solve simultaneously for both forward \( \mathbf{a}_p \) and backward \( \mathbf{b}_p \) MC LP parameters. The MC least-squares-based fast algorithm requires the introduction of two additional gain adjustment vectors, defined as

\[
\mathbf{c}_p = (\mathbf{X}_p \mathbf{X}_p^H)^{-1} \mathbf{X}_p^H[p+1] = (\mathbf{x}_{p+1}^H[N] \ldots \mathbf{x}_{p+1}^H[N-p])
\]

\[
\mathbf{d}_p = (\mathbf{X}_p \mathbf{X}_p^H)^{-1} \mathbf{x}_{p+1}^H[N] = (\mathbf{x}_{p+1}^H[p+1] \ldots \mathbf{x}_{p+1}^H[1])
\]  

(34) \hspace{1cm} (35)

in which the \( R \times (p+1) \) gain vectors are \( \mathbf{c}_p = (\mathbf{c}_p[0] \ldots \mathbf{c}_p[p]) \) and \( \mathbf{d}_p = (\mathbf{d}_p[0] \ldots \mathbf{d}_p[p]) \). We will also need to define the \( M \times M \) gain vector variances

\[
\mathbf{P}_p^c = \mathbf{I} - \mathbf{X}_p[p+1] (\mathbf{X}_p \mathbf{X}_p^H)^{-1} \mathbf{X}_p^H[p+1] = (\mathbf{I} \ \mathbf{a}_p)^* \mathbf{X}_p \mathbf{X}_p^H (\mathbf{I} \ \mathbf{a}_p)^H
\]

\[
\mathbf{P}_p^d = \mathbf{I} - \mathbf{x}_{p+1}^H[N] (\mathbf{X}_p \mathbf{X}_p^H)^{-1} \mathbf{x}_{p+1}^H[N] = (\mathbf{I} \ \mathbf{b}_p)^* \mathbf{X}_p \mathbf{X}_p^H (\mathbf{I} \ \mathbf{b}_p)^H
\]  

(36) \hspace{1cm} (37)

Note that trace \( \{\mathbf{P}_p^c\} \geq 0 \) and trace \( \{\mathbf{P}_p^d\} \geq 0 \). The key MC Levinson-like order-update recursions in the MC covariance least squares linear prediction case are

\[
(\mathbf{I} \ \mathbf{a}_p) = (\mathbf{I} \ \mathbf{a}_{p-1} \ 0_p) + \mathbf{P}_p^c (\mathbf{0}_p \ \mathbf{b}_{p-adj}^c \ \mathbf{I})
\]

\[
(\mathbf{I} \ \mathbf{b}_p) = (\mathbf{0}_p \ \mathbf{b}_{p-adj}^b \ \mathbf{I}) + \mathbf{P}_p^c (\mathbf{I} \ \mathbf{a}_{p-1} \ 0_p)
\]  

(38) \hspace{1cm} (39)

and the time-adjusted LP parameter \( \mathbf{b}_{p-adj} \) time-update recursion has the form

\[
\mathbf{b}_{p-adj} = \mathbf{b}_p + C_1 \mathbf{c}_p - 1 + C_2 \mathbf{d}_p - 1
\]  

(40)

for unspecified \( M \times M \) matrix constants \( C_1 \) and \( C_2 \). Similar order and time updates for \( \mathbf{c}_p \), \( \mathbf{d}_p \), \( \mathbf{c}_p^b[n] \), and \( \mathbf{c}_p^a[n] \) are a part of the fast MC algorithm, but not shown here.

The inverse matrix \((\mathbf{X}_p \mathbf{X}_p^H)^{-1}\) can be explicitly expressed as

\[
(\mathbf{X}_p \mathbf{X}_p^H)^{-1} = \mathbf{A}_p \mathbf{P}_p^a \mathbf{A}_p^H - \mathbf{B}_p \mathbf{P}_p^b \mathbf{B}_p^H
\]

\[
+ \mathbf{C}_p - 1 \mathbf{P}_p^c \mathbf{C}_p^b - 1 - \mathbf{D}_p - 1 \mathbf{P}_p^b \mathbf{D}_p - 1
\]  

(41)
in which the block triangular Toeplitz matrices are

\[
A_p = \begin{pmatrix}
I & 0 & \cdots & 0 & 0 \\
A_p[1] & I & \cdots & 0 & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
A_p[p-1] & A_p[p-2] & \cdots & I & 0 \\
\end{pmatrix}
\]

\[
B_p = \begin{pmatrix}
0 & 0 & \cdots & 0 & 0 \\
0 & 0 & \cdots & 0 & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\end{pmatrix}
\]

\[
C_{p-1} = \begin{pmatrix}
c_{p-1}[0] & 0 & \cdots & 0 & 0 \\
0 & 0 & \cdots & 0 & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
c_{p-1}[p-2] & c_{p-1}[p-3] & \cdots & 0 & 0 \\
c_{p-1}[p-1] & c_{p-1}[p-2] & \cdots & c_{p-1}[0] & 0
\end{pmatrix}
\]

\[
D_{p-1} = \begin{pmatrix}
d_{p-1}[0] & 0 & \cdots & 0 & 0 \\
0 & 0 & \cdots & 0 & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
d_{p-1}[p-2] & d_{p-1}[p-3] & \cdots & 0 & 0 \\
d_{p-1}[p-1] & d_{p-1}[p-2] & \cdots & d_{p-1}[0] & 0
\end{pmatrix}
\]

\[
P_p = \begin{pmatrix}
(P_p)^{-1} & 0 & \cdots & 0 & 0 \\
0 & (P_p)^{-1} & \cdots & 0 & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & (P_p)^{-1} & 0 \\
0 & 0 & \cdots & 0 & (P_p)^{-1}
\end{pmatrix}
\]

are formed from the forward linear prediction parameters \( A_p[m] \), forward linear prediction squared error variance \( P_p^a \), backward linear prediction parameters \( B_p[m] \), backward linear prediction squared error variance matrix \( P_p^b \), gain adjustment block vector parameters \( c_p[m] \) and \( d_p[m] \), and matrix gain adjustment variances \( P_p^c \) and \( P_p^d \). These are defined as matrix equations associated with the covariance linear prediction case. The structures of \( P_p^b, P_p^c \), and \( P_p^d \) are similar to that shown for \( P_p^a \).

4. REFERENCES


