Extended Chandrasekhar Recursions

Ali H. Sayed and Thomas Kailath

***Abstract*—We extend the discrete-time Chandrasekhar recursions for least-squares estimation in constant parameter state-space models to a class of structured time-variant state-space models, special cases of which often arise in adaptive filtering. It can be shown that the much studied exponentially weighted recursive least-squares filtering problem can be reformulated as an estimation problem for a state-space model having this special time-variant structure. Other applications arise in the multichannel and multidimensional adaptive filtering context.**

I. INTRODUCTION

The discrete-time Chandrasekhar recursions for linear least-squares estimation in constant-parameter systems were first presented nearly two decades ago [1]–[4]. The point was that the celebrated Kalman filtering algorithm based on the discrete-time Riccati recursion applied equally to time-invariant, i.e., constant parameter, and time-variant state-space models. This is a strength, but on the other hand one might expect some computational reductions when the model is time-invariant. Replacing the Riccati recursion by the Chandrasekhar recursions does allow such a reduction, from $O(n^3)$ to $O(n^2)$ elementary computations per step, where n is the state dimension. The computational reduction can be very significant in applications where n is quite large (see e.g., [5]–[7]).

There have been some efforts over the years to obtain extensions to time-variant state-space models, and progress in this area has come about through a particular application. In the last few years, there has been a great interest (see e.g., [8]–[12]) in fast versions of recursive least-squares (RLS) algorithms for adaptive filtering and control. These fast RLS algorithms are rather complicated to describe and derive, involving a large number (10–20) of variables and subscripts. In independent work, Houacine *et al.* [13], [14], and Slock [15] showed that some of these rather complicated fast RLS algorithms could be described and derived much more compactly and simply by recasting the problem in a form to which the Chandrasekhar recursions could be applied. Some manipulation was required to be able to do this because the "natural" model for the problem is not time-invariant; in adaptive filtering the output system matrix is a function of the data, which of course changes with time.

Motivated by this and related problems, we have shown that the Chandrasekhar recursions can be extended to a certain class of time-variant systems in which the time-variation takes place in a certain structured manner. The extended Chandrasekhar recursions are easy to verify, once they have been discovered. This short note is devoted to describing and establishing these extended recursions. Structured time-variations of this sort arise, as mentioned above, in various adaptive filtering problems (and their dual control versions), and may be encountered in other areas as well.

Manuscript received October 2, 1992; revised January 21, 1993. This work was supported in part by the Air Force Office of Scientific Research, Air Force Systems Command under contract AFOSR91-0060 and by the Army Research Office under contract DAAL03-89-K-0109. The work of Dr. Sayed was also supported by a fellowship from Fundação de Amparo à Pesquisa do Estado de São Paulo, and by Escola Politénica da Universidade de São Paulo, Brazil.

A. H. Sayed is currently with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106.

T. Kailath is with the Information Systems Laboratory, Stanford, CA 94305.

IEEE Log Number 9214007.

We first give a brief review of the Riccati-based Kalman filter. Consider a $p \times 1$ process $\{y_i\}$ with an n-dimensional state-space model

$$\boldsymbol{x}_{i+1} = F_i\boldsymbol{x}_i + G_i\boldsymbol{u}_i$$

$$\boldsymbol{y}_i = H_i\boldsymbol{x}_i + \boldsymbol{v}_i \qquad \text{for } i \geq 0 \tag{1}$$

where $\{F_i, G_i, H_i\}$ are known matrices with dimensions $n \times n$, $n \times m$, and $p \times n$, respectively. We assume that $\boldsymbol{x}_0$, $\boldsymbol{u}_i$, and $\boldsymbol{v}_i$ are stochastic variables that satisfy

$$E\boldsymbol{x}_0 = \overline{\boldsymbol{x}}_0, \quad E(\boldsymbol{x}_0 - \overline{\boldsymbol{x}}_0)(\boldsymbol{x}_0 - \overline{\boldsymbol{x}}_0)^* = \Pi_0,$$

$$E\boldsymbol{u}_i\boldsymbol{x}_0^* = E\boldsymbol{v}_i\boldsymbol{x}_0^* = \boldsymbol{0}, \quad E\boldsymbol{v}_i = E\boldsymbol{u}_i = \boldsymbol{0},$$

$$E\begin{bmatrix}\boldsymbol{u}_i\\ \boldsymbol{v}_i\end{bmatrix}[\boldsymbol{u}_j^* \quad \boldsymbol{v}_j^*] = \begin{bmatrix}Q_i & C_i\\ C_i^* & R_i\end{bmatrix}\delta_{ij}.$$

The symbol δ_{ij} is the Kronecker delta function, * denotes Hermitian conjugation (complex conjugation for scalars), and E denotes expected value. Let $\hat{\boldsymbol{x}}_{i|i-1}$ and $\hat{\boldsymbol{y}}_{i|i-1}$ denote the linear least-squares estimates of $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$ given $\{\boldsymbol{y}_0, \cdots, \boldsymbol{y}_{i-1}\}$, respectively. The Kalman filter [16] computes these quantities via the recursions

$$\hat{\boldsymbol{y}}_{i|i-1} = H_i\hat{\boldsymbol{x}}_{i|i-1}$$

$$\hat{\boldsymbol{x}}_{i+1|i} = F_i\hat{\boldsymbol{x}}_{i|i-1} + K_iR_{\epsilon,i}^{-1}\epsilon_i \tag{2}$$

where $\epsilon_i = \boldsymbol{y}_i - H_i\hat{\boldsymbol{x}}_{i|i-1}$, $R_{\epsilon,i} = E(\epsilon_i\epsilon_i^*)$, and $K_i = E(\boldsymbol{x}_{i+1}\epsilon_i^*)$. Kalman showed that K_i and $R_{\epsilon,i}$ can be computed via the expressions: $K_i = F_iP_{i|i-1}H_i^* + G_iC_i$ and $R_{\epsilon,i} = H_iP_{i|i-1}H_i^* + R_i$, where $P_{i|i-1}$ is the error covariance in the one-step prediction of $\boldsymbol{x}_i$, $P_{i|i-1} \equiv E(\boldsymbol{x}_i - \hat{\boldsymbol{x}}_{i|i-1})(\boldsymbol{x}_i - \hat{\boldsymbol{x}}_{i|i-1})^*$, and satisfies the Riccati difference recursion: $P_{0|-1} = \Pi_0$

$$P_{i+1|i} = F_iP_{i|i-1}F_i^* - \overline{K}_{p,i}\overline{K}_{p,i}^* + G_iQ_iG_i^*, \quad \overline{K}_{p,i} \equiv K_iR_{\epsilon,i}^{-*/2}. \tag{3}$$

We shall define the square root (factor) of a matrix A as a lower triangular matrix, denoted $A^{1/2}$, such that $A = A^{1/2}A^{*/2}$. We shall also denote $(A^{1/2})^* = A^{*/2}$ and $(A^{1/2})^{-1} = A^{-1/2}$ so that $A^{-1} = A^{-*/2}A^{-1/2}$.

We can check that the number of operations, i.e., multiplications and additions, needed in going from index i to index $(i+1)$ in the Riccati recursion (3) is $O(n^3)$, and this is true whether or not the state-space model has constant parameters. However, one expects a computationally more efficient procedure in the case of time-invariant (also called constant-parameter) systems $\{F, G, H, Q, R, C\}$. Indeed, it has been shown [1]–[4], [17], [18] that in the constant-parameter case the complexity can be reduced to $O(n^2\alpha)$ per iteration, where the so-called displacement rank α is given by

$$\begin{aligned}\alpha &= \text{rank}\,(F\Pi_0F^* + GQG^* - \overline{K}_{p,0}\overline{K}_{p,0}^* - \Pi_0)\\ &= \text{rank}\,(P_{1|0} - P_{0|-1}).\end{aligned}$$

This is achieved by using the so-called Chandrasekhar recursions to compute $\{K_i, R_{\epsilon,i}\}$ for use in the formulas (2). There are many forms for the Chandrasekhar recursions [1]–[3], but we shall give here perhaps the simplest (so-called square-root) version [4].

0018-9375/94$04.00 © 1994 IEEE

II. The Square-Root Chandrasekhar Filter

Let $\delta P_i = P_{i+1|i} - P_{i|i-1}$. It turns out that for constant-parameter systems, the quantity δP_i often has low rank (examples are given later in this section), much less than n, and this fact can be exploited to find a lower complexity algorithm. Observe that δP_i is a Hermitian matrix, so that it has only real eigenvalues. We can factor it (nonuniquely) as

$$\delta P_i = P_{i+1|i} - P_{i|i-1} = L_i S_i L_i^* \tag{4}$$

where S_i is an $\alpha \times \alpha$ signature matrix *viz.* a diagonal matrix with as many $\pm 1's$ on the diagonal as δP_i has positive and negative eigenvalues. In fact, it turns out, as shown ahead, that S_i is the same for all i, that is $S_i = S$, $\forall i$. We now form the prearray

$$A_i = \begin{bmatrix} R_{\epsilon,i}^{1/2} & HL_i \\ \overline{K}_{p,i} & FL_i \end{bmatrix}. \tag{5}$$

Let Θ_i be any $J = (I \oplus S_i)$—unitary matrix ($\Theta_i J \Theta_i^* = J$) that triangularizes A_i. That is,

$$A_i \Theta_i = \begin{bmatrix} X & \mathbf{0} \\ Y & Z \end{bmatrix}.$$

Comparing the entries on both sides of the equality $A_i J A_i^* = A_i \Theta_i J \Theta_i^* A_i^*$ we get

$$\begin{aligned} XX^* &= R_{\epsilon,i} + HL_iS_iL_i^*H^* \\ &\times =^{(4)} R_{\epsilon,i} + H(P_{i+1|i} - P_{i|i-1})H^* \\ &= R_{\epsilon,i} + R_{\epsilon,i+1} - R_{\epsilon,i} = R_{\epsilon,i+1}. \end{aligned}$$

So we can choose $X = R_{\epsilon,i+1}^{1/2}$. Moreover,

$$\begin{aligned} YX^* &= K_i + FL_iS_iL_i^*H^* \\ &= K_i + F(P_{i+1|i} - P_{i|i-1})H^* = K_{i+1} \end{aligned}$$

and hence, we can identify $Y = \overline{K}_{p,i+1}$. Finally,

$$\begin{aligned} YY^* + ZS_iZ^* &= K_iR_{\epsilon,i}^{-1}K_i^* + FL_iS_iL_i^*F^* \\ &= K_iR_{\epsilon,i}^{-1}K_i^* + F(P_{i+1|i} - P_{i|i-1})F^*. \end{aligned}$$

Therefore,

$$ZS_iZ^* = P_{i+2|i+1} - P_{i+1|i} = L_{i+1}S_{i+1}L_{i+1}^*$$

where we have used definition (4). Hence, we can choose $S_{i+1} = S_i = S$ =signature matrix and identify Z as L_{i+1}. So we are led to the following so-called square-root Chandrasekhar recursions

$$\begin{bmatrix} R_{\epsilon,i}^{1/2} & HL_i \\ \overline{K}_{p,i} & FL_i \end{bmatrix} \Theta_i = \begin{bmatrix} R_{\epsilon,i+1}^{1/2} & \mathbf{0} \\ \overline{K}_{p,i+1} & L_{i+1} \end{bmatrix} \tag{6}$$

where Θ_i is any $J = (I \oplus S)$—unitary matrix that produces the block zero entry on the right-hand side of (6). We can verify that each iteration takes only $O(n^2\alpha)$ computations when $n > p$, as is often the case.

This particular approach to the Chandrasekhar recursions is of course not the way they were originally derived. For more motivation, not necessary here, as to the particular choice of the prearray (5), see [3] and also [19], [22]. Let us consider two special cases [2]:

- $\Pi_0 = 0$: In this case, $P_{1|0} = GQG^*$ (assuming $C_i = 0$) and we can choose $L_0 = GQ^{1/2}$ and $S = I$. Moreover, Θ_i is any usual unitary matrix.
- $\Pi_0 = \overline{\Pi}$, that is Π_0 is the unique nonnegative-definite solution of (F assumed stable) $\overline{\Pi} = F\overline{\Pi}F^* + GQG^*$. In this case, we get $P_{1|0} - P_{0|-1} = -\overline{K}_{p,0}\overline{K}_{p,0}^*$. So we can choose $L_0 = \overline{K}_{p,0}$ and $S = -I$. Now the matrix Θ_i is a $J = (I \oplus -I)$—unitary matrix.

III. Structured Time-Variant Models

The derivation of the Chandrasekhar recursions (6) is based on the fact that δP_i has low rank for constant-parameter systems, as expressed in (4). We now show that these recursions can be extended to a class of time-variant state-space models that exhibit a certain structure in their time-variation.

The computational advantage of the Chandrasekhar recursions stems from the fact that they propagate the low rank factor L_i instead of $P_{i+1|i}$, where L_i is defined via relation (4). A direct generalization would be to consider differences of the form $P_{i+1|i} - \Psi_i P_{i|i-1} \Psi_i^*$, where the Ψ_i are convenient time-variant matrices that result in a low rank difference, say of rank α. That is

$$P_{i+1|i} - \Psi_i P_{i|i-1} \Psi_i^* \equiv L_i S_i L_i^* \tag{7}$$

for some $n \times \alpha$ matrix L_i (we shall also show that for the special time-variant models to be introduced here we shall have $S_i = S$, $\forall i$).

We consider again the state-space model given by (1), and we shall say that it is a structured time-variant model if there exist $n \times n$ matrices Ψ_i such that F_i, G_i, and H_i vary according to the following rules:

$$H_i = H_{i+1}\Psi_i, \quad F_{i+1}\Psi_i = \Psi_{i+1}F_i, \quad G_{i+1} = \Psi_{i+1}G_i. \tag{8}$$

It is clear that constant-parameter systems satisfy (8) with $\Psi_i = I$. Other special cases of (8) also arise in adaptive filtering as noticed in [20]–[22] and in Section V. We first assume that the covariance matrices R_i, Q_i, and C_i are time-invariant whereas F_i, H_i, and G_i vary in time according to (8). We shall verify in the next section that these restrictions can be relaxed in order to allow for time-variant R_i, Q_i and C_i.

The reason for imposing the conditions specified in (8) will become clear as soon as we give a simple algebraic verification of the proposed recursion (they can also be justified by noting that under these constraints the covariance matrix of the output process *viz.* $\mathcal{R} = [\text{cov}(\boldsymbol{y}_i, \boldsymbol{y}_j)]_{i,j=0}^{\infty}$, possesses a time-invariant displacement structure as detailed in [19]–[22]).

IV. Extended Chandrasekhar Recursions

We derive here the extended Chandrasekhar recursions associated with time-variant models as above, in both the normalized and unnormalized (square-root or array) forms.

A. Square-Root Form

We form the prearray (which should be compared with (5))

$$A_i = \begin{bmatrix} R_{\epsilon,i}^{1/2} & H_{i+1}L_i \\ \Psi_{i+1}\overline{K}_{p,i} & F_{i+1}L_i \end{bmatrix}$$

and let Θ_i be any $J = (I \oplus S_i)$—unitary matrix that triangularizes A_i. That is

$$A_i\Theta_i = \begin{bmatrix} X & \mathbf{0} \\ Y & Z \end{bmatrix}.$$

Comparing the entries on both sides of the equality $A_iJA_i^* = A_i\Theta_iJ\Theta_i^*A_i^*$, we get (here we use the condition $H_{i+1}\Psi_i = H_i$)

$$\begin{aligned} XX^* &= R_{\epsilon,\,i} + H_{i+1}L_iS_iL_i^*H_{i+1}^* \\ &= R_{\epsilon,\,i} + H_{i+1}(P_{i+1|i} - \Psi_iP_{i|i-1}\Psi_i^*)H_{i+1}^* \\ &= R_{\epsilon,\,i} + H_{i+1}P_{i+1|i}H_{i+1}^* - H_iP_{i|i-1}H_i^* = R_{\epsilon,\,i+1}. \end{aligned}$$

So we can choose $X = R_{\epsilon,\,i+1}^{1/2}$. Moreover (we now use the conditions on F_i and G_i),

$$\begin{aligned} YX^* &= \Psi_{i+1}K_i + F_{i+1}L_iS_iL_i^*H_{i+1}^* \\ &= \Psi_{i+1}K_i + F_{i+1}(P_{i+1|i} - \Psi_iP_{i|i-1}\Psi_i^*)H_{i+1}^* \\ &= \Psi_{i+1}K_i + F_{i+1}P_{i+1|i}H_{i+1}^* - \Psi_{i+1}F_iP_{i|i-1}H_i^* \\ &= \Psi_{i+1}K_i + K_{i+1} - G_{i+1}C - \Psi_{i+1}K_i + \Psi_{i+1}G_iC = K_{i+1} \end{aligned}$$

and hence $Y = \overline{K}_{p,\,i+1}$. Finally (we now use the condition on G_i),

$$\begin{aligned} YY^* + ZS_iZ^* &= \Psi_{i+1}K_iR_{\epsilon,\,i}^{-1}K_i^*\Psi_{i+1}^* + F_{i+1}L_iS_iL_i^*F_{i+1}^* \\ &= \Psi_{i+1}K_iR_{\epsilon,\,i}^{-1}K_i^*\Psi_{i+1}^* \\ &\quad + F_{i+1}(P_{i+1|i} - \Psi_iP_{i|i-1}\Psi_i^*)F_{i+1}^* \\ &= \Psi_{i+1}K_iR_{\epsilon,\,i}^{-1}K_i^*\Psi_{i+1}^* + F_{i+1}P_{i+1|i}F_{i+1}^* \\ &\quad - \Psi_{i+1}F_iP_{i|i-1}F_i^*\Psi_{i+1}^*. \end{aligned}$$

Therefore,

$$ZS_iZ^* = P_{i+2|i+1} - \Psi_{i+1}P_{i+1|i}\Psi_{i+1}^* \equiv L_{i+1}S_{i+1}L_{i+1}^*$$

and we see that we can choose $S_{i+1} = S_i$ and identify Z as L_{i+1}. Therefore, we are led to the following (square-root) extended Chandrasekhar recursions

$$\begin{bmatrix} R_{\epsilon,\,i}^{1/2} & H_{i+1}L_i \\ \Psi_{i+1}\overline{K}_{p,\,i} & F_{i+1}L_i \end{bmatrix}\Theta_i = \begin{bmatrix} R_{\epsilon,\,i+1}^{1/2} & \mathbf{0} \\ \overline{K}_{p,\,i+1} & L_{i+1} \end{bmatrix} \tag{9}$$

where Θ_i is any $J = (I \oplus S)$—unitary matrix that produces the block zero entry on the right-hand side of the last expression. These equations should be compared with the Chandrasekhar recursions derived in Section II. The differences are that F_{i+1}, H_{i+1}, and Ψ_{i+1} appear on the left-hand side of the above expression instead of F, H, and I, respectively.

B. Unnormalized Form

It is sometimes convenient to express the extended Chandrasekhar recursions (9) in an unnormalized form. For this, we consider the following alternative factorization (compare with (7))

$$P_{i+1|i} - \Psi_iP_{i|i-1}\Psi_i^* = -L_i^{(u)}R_{r,\,i}^{-1}L_i^{*(u)} \tag{10}$$

where $R_{r,\,i}$ is an $\alpha \times \alpha$ matrix that is not necessarily a signature matrix. However, comparing with (7) we see that if we factor $R_{r,\,i}$ as $-R_{r,\,i} = R_{r,\,i}^{1/2}S_iR_{r,\,i}^{*/2}$, where S_i is a signature matrix, then $L_i = L_i^{(u)}R_{r,\,i}^{-*/2}$ and we are back to the same form given in (7). If we instead continue with (10), then we can check that the following array transformation holds

$$\begin{bmatrix} R_{\epsilon,\,i} & H_{i+1}L_i^{(u)} \\ \Psi_{i+1}K_i & F_{i+1}L_i^{(u)} \\ L_i^{*(u)}H_{i+1}^* & R_{r,\,i} \end{bmatrix}\Sigma_i = \begin{bmatrix} R_{\epsilon,\,i+1} & \mathbf{0} \\ K_{i+1} & L_{i+1}^{(u)} \\ \mathbf{0} & R_{r,\,i+1} \end{bmatrix} \tag{11}$$

where a particular form for Σ_i is

$$\Sigma_i = \begin{bmatrix} I_p & -R_{\epsilon,\,i}^{-1}H_{i+1}L_i^{(u)} \\ -R_{r,\,i}^{-1}L_i^{*(u)}H_{i+1}^* & I_\alpha \end{bmatrix}. \tag{12}$$

It can be verified that the above Σ_i satisfies the generalized unitarity relation

$$\Sigma_i\begin{bmatrix} R_{\epsilon,\,i+1}^{-1} & \mathbf{0} \\ \mathbf{0} & -R_{r,\,i+1}^{-1} \end{bmatrix}\Sigma_i^* = \begin{bmatrix} R_{\epsilon,\,i}^{-1} & \mathbf{0} \\ \mathbf{0} & -R_{r,\,i}^{-1} \end{bmatrix}. \tag{13}$$

Combining (11) and (12) will give us the Chandrasekhar recursions in equation form rather than as an array transformation—see [4] for related discussions, where the array form was obtained starting with (11), (12).

C. Time-Variant C_i, Q_i, R_i

We now drop some of the earlier restrictions and consider time-variant models that obey the following rule (which is a relaxation of (8) and includes the cross term C_i—see the model (1)):

$$H_i = H_{i+1}\Psi_i, \quad F_{i+1}\Psi_i = \Psi_{i+1}F_i, \quad G_{i+1}C_{i+1} = \Psi_{i+1}G_iC_i. \tag{14}$$

If we define $T_i = G_iQ_iG_i^*$, then time-variant R_i and Q_i can be handled in much the same manner as before by introducing the differences (as in [3]):

$$\delta R_i = R_{i+1} - R_i, \quad \delta_{\Psi_{i+1}}T_i = T_{i+1} - \Psi_{i+1}T_i\Psi_{i+1}^*$$

and factoring them as $\delta R_i = V_iM_iV_i^*$, $\delta_{\Psi_{i+1}}T_i = X_iE_iX_i^*$, where M_i and E_i are signature matrices. If we now define the (time-variant) signature matrix $J_i = (I \oplus S_i \oplus M_i \oplus E_i)$, then following the same reasoning as in Section IV-A, we get the following recursive array

$$\begin{bmatrix} R_{\epsilon,\,i}^{1/2} & H_{i+1}L_i & V_i & \mathbf{0} \\ \Psi_{i+1}\overline{K}_{p,\,i} & F_{i+1}L_i & \mathbf{0} & X_i \end{bmatrix}\Theta_i = \begin{bmatrix} R_{\epsilon,\,i+1}^{1/2} & \mathbf{0} & \mathbf{0} \\ \overline{K}_{p,\,i+1} & L_{i+1} & \mathbf{0} \end{bmatrix}$$

where Θ_i is any J_i—unitary matrix that produces the block zero entries on the right-hand side of the last expression. Unlike the case of constant R_i, Q_i, and C_i, the rank of L_{i+1} can increase or decrease with i [3] (that is, the signature matrix S_i now varies with i).

$$P_{i+1|i} - \Psi_iP_{i|i-1}\Psi_i^* = L_iS_iL_i^*, \qquad S_i \text{ is } \alpha_i \times \alpha_i.$$

V. An Application to the RLS Problem

We now illustrate a particular application of the extended Chandrasekhar recursions by considering an important special case of (8) that arises in the recursive least-squares problem in adaptive filtering. The basic problem reads as follows: given pairs of data points $\{\boldsymbol{u}_i, d(i)\}$, $i = 0, 1, \cdots, N$, where $\boldsymbol{u}_i$ is a $1 \times M$ row vector that consists of the values of M input channels at time i,

$\boldsymbol{u}_i = [u_1(i)\ u_2(i)\ \cdots\ u_M(i)]$, ($d(i)$ and $u_j(i)$, $j = 1, \cdots, M$, are assumed scalar for simplicity), we are required to determine the linear least-squares estimate of an $M \times 1$ column vector of unknown tap weights, $\boldsymbol{w} = [w_1\ w_2\ \cdots\ w_M]^T$, so as to minimize the exponentially weighted error sum

$$\mathcal{E} = (\boldsymbol{w} - \overline{\boldsymbol{w}})^* \Pi_0^{-1} (\boldsymbol{w} - \overline{\boldsymbol{w}}) + \sum_{i=0}^{N} \lambda^{N-i} |d(i) - \boldsymbol{u}_i \boldsymbol{w}|^2 \tag{15}$$

where $\overline{\boldsymbol{w}} = E\boldsymbol{w}$, $E(\boldsymbol{w} - \overline{\boldsymbol{w}})(\boldsymbol{w} - \overline{\boldsymbol{w}})^* = \Pi_0$, and the parameter λ is often called the forgetting factor, since past inputs are (exponentially) weighted less than the more recent values. In several applications, the input channels exhibit the shift structure: $u_j(i) = u_{j-1}(i-1)$. That is, if we denote the value of the first channel at time i by $u(i)$, then this corresponds to having an input row vector $\boldsymbol{u}_i$ of the form $\boldsymbol{u}_i = [u(i)\ u(i-1)\ \cdots\ u(i-M+1)]$. It can be shown (see eg. [20]–[22] for details—see also [13]–[15] for an alternative and related discussion) that this is equivalent to a state-space estimation problem by considering the following $(N+1)$—dimensional state-space model

$$\boldsymbol{x}_{i+1} = \lambda^{-1/2} \boldsymbol{x}_i, \quad \boldsymbol{x}_0 = [\boldsymbol{w}^T 0]^T$$

$$y(i) = \boldsymbol{h}_i \boldsymbol{x}_i + v(i), \quad Ev(i)v^*(j) = \delta_{ij} \tag{16}$$

where $\boldsymbol{h}_i = [u(i)\ u(i-1)\ \cdots\ u(0)\ \boldsymbol{0}_{N-i}]$ is a $1 \times (N+1)$ row vector, $y(i) = d(i)/(\sqrt{\lambda})^i$, and $\boldsymbol{x}_i$ is an $(N+1) \times 1$ state-vector with trailing zeros, $(\sqrt{\lambda})^i \boldsymbol{x}_i = [\boldsymbol{w}^T\ 0]^T$. An initial state covariance matrix (with trailing zeros) is assumed *viz.* $E(\boldsymbol{x}_0 - \overline{\boldsymbol{x}}_0)(\boldsymbol{x}_0 - \overline{\boldsymbol{x}}_0)^* = \Pi_0 \oplus \boldsymbol{0}$, where Π_0 is an $M \times M$ positive-definite matrix. The corresponding Kalman equations can now be written as

$$\hat{x}_{i+1|i} = \lambda^{-1/2} \hat{\boldsymbol{x}}_{i|i-1} + \boldsymbol{k}_i r_{\epsilon,i}^{-1} [y(i) - \boldsymbol{h}_i \hat{\boldsymbol{x}}_{i|i-1}]$$

$$r_{\epsilon,i} = 1 + \boldsymbol{h}_i P_{i|i-1} \boldsymbol{h}_i^*, \quad \boldsymbol{k}_i = \lambda^{-1/2} P_{i|i-1} \boldsymbol{h}_i^*$$

$$P_{i+1|i} = \lambda^{-1} [P_{i|i-1} - P_{i|i-1} \boldsymbol{h}_i^* r_{\epsilon,i}^{-1} \boldsymbol{h}_i P_{i|i-1}] \tag{17}$$

with $P_{0|-1} = \Pi_0 \oplus \boldsymbol{0}$. The gain vector $\overline{\boldsymbol{k}}_{p,i} = \boldsymbol{k}_i r_{\epsilon,i}^{-*/2}$ also has trailing zeros *viz.* $\overline{\boldsymbol{k}}_{p,i} \equiv [\boldsymbol{c}_i^T\ 0]^T$. However, though time-variant, the special structure of $\boldsymbol{h}_i$ *viz.* $\boldsymbol{h}_i = \boldsymbol{h}_{i+1} Z$, where Z is the lower triangular shift matrix, can be further exploited to reduce the operation count to $O(M)$. Observe that the above relation shows (along with $F_{i+1}Z = ZF_i$, since $F_i = \lambda^{-1/2} I$) that the state-space model (16) is a special structured time-variant model. The reduction in operation count can now be achieved by using a special case of the extended Chandrasekhar recursions (9) with $\Psi_i = Z$, $F_i = \lambda^{-1/2} I$. To apply these recursions, we first introduce the (nonunique) factorization $L_0 S L_0^* = P_{1|0} - Z P_{0|-1} Z^*$, where L_0 and S are $(N+1) \times \alpha$ and $\alpha \times \alpha$ matrices, respectively. The factor L_0 is clearly of the form $L_0 = [\tilde{L}_0^T\ \ 0]^T$, where $\tilde{L}_0$ is $(M+1) \times \alpha$. Let $\tilde{\boldsymbol{h}}_i$ be the row vector of the first $M+1$ coefficients of $\boldsymbol{h}_i$. Writing down the extended Chandrasekhar recursions (9), we obtain

$$\begin{bmatrix} r_{\epsilon,i}^{1/2} & \tilde{\boldsymbol{h}}_{i+1} \tilde{L}_i \\ \begin{bmatrix} 0 \\ \boldsymbol{c}_i \end{bmatrix} & \lambda^{-1/2} \tilde{L}_i \end{bmatrix} \Theta_i = \begin{bmatrix} r_{\epsilon,i+1}^{1/2} & 0 \\ \begin{bmatrix} \boldsymbol{c}_{i+1} \\ 0 \end{bmatrix} & \tilde{L}_{i+1} \end{bmatrix} \tag{18}$$

where Θ_i is any $J = (1 \oplus S)$—unitary matrix that produces the zero entry on the right hand-side of the above expression. The computational complexity of each step is $O(\alpha M)$ where the value of α depends on the choice of Π_0. This recursion is a square-root version of fast RLS algorithms discussed in the literature [11], [12].

We should remark that the connection between the Chandrasekhar recursions and fast RLS algorithms has been pointed out earlier by Houacine *et al.* [13], [14] by constructing a time-invariant state-space model and by using the time-invariant Chandrasekhar recursions (Section II). Slock [15] also discussed, in greater details, the connection between the (unnormalized) Chandrasekhar recursions and the fast transversal filter (FTF) algorithm of Cioffi and Kailath [12] by using an infinite dimensional time-invariant state-space model and showing the relation of the involved quantities to those in the FTF algorithm.

We addressed here the same connection within the framework of structured time-variant models, which includes as a special case, the particular time-invariant models of Houacine and Slock. We then showed that the square-root version of the fast RLS algorithm followed by choosing a convenient Ψ_i in the extended Chandrasekhar recursions (9) (which happens to be $\Psi_i = Z$). However, the extended formulation in (9) allows us to readily consider more general cases. The point is that though we assumed that the channel inputs obey a shift structure *viz.* $u_j(i) = u_{j-1}(i-1)$, our derivation makes it clear that we can also obtain fast algorithms for other cases where the input channels exhibit a generalized shift structure. For example, if the input vector $\boldsymbol{u}_i$ satisfies a relation of the form $\boldsymbol{u}_i = \boldsymbol{u}_{i+1} \Psi$, for some constant matrix Ψ, then the associated state-space estimation problem reduces to that of a time-variant structured model, and we can write down the corresponding extended Chandrasekhar recursions. For matrices Ψ that are relatively sparse, in the sense that $\Psi \overline{\boldsymbol{k}}_{p,i}$ requires $O(M)$ operations, and for appropriate choices of Π_0, we are also led to a fast RLS algorithm. Furthermore, the state-space model for the (scalar) RLS problem has a row H matrix, whereas our derivation (see state-space model (1)) allows for models with more general matrices H. Such models arise for example, in the multichannel, multidimensional, and/or nonlinear adaptive problems [23], where in many instances, choice of Ψ with a block shift structure is convenient, such as: $\Psi = Z \oplus Z \oplus \cdots \oplus Z$. Finally, the above framework also allows us to derive the so-called QR and lattice adaptive algorithms as detailed in [21], [22].

VI. Conclusion

We have extended the Chandrasekhar recursions to a class of structured time-variant models and we have derived the corresponding square-root (or array) forms in both the normalized and unnormalized forms. An application to the much studied exponentially weighted recursive least-squares filtering problem has been briefly discussed. Further applications to multichannel and multidimensional adaptive filtering, and extensions to alternative windowing schemes will be discussed elsewhere.

References

[1] T. Kailath, M. Morf, and G. S. Sidhu, "Some new algorithms for recursive estimation in constant linear discrete-time systems," in *Proc. Seventh Princeton Conf. Inform. Sci. Syst.*, Princeton, NJ, pp. 344–352, 1973.

[2] M. Morf, G. S. Sidhu, and T. Kailath, "Some new algorithms for recursive estimation in constant, linear, discrete-time systems," *IEEE Trans. Automat. Contr.*, vol. AC-19, no. 4, pp.315–323, Aug. 1974.

[3] M. Morf and T. Kailath, "Square root algorithms for least squares estimation," *IEEE Trans. Automat. Contr.*, vol. AC-20, no. 4, pp. 487–497, Aug. 1975.

[4] T. Kailath, A. C. Vieira, and M. Morf, "Orthogonal transformation (square-root) implementations of the generalized Chandrasekhar and generalized Levinson algorithms," in *Lecture Notes in Control and Information Sciences*, A. Bensoussan and J. Lions, eds. New York: Springer-Verlag, vol. 32, pp. 81–91, 1978.

[5] A. A. Desalu, L. A. Gould, and F. C. Schweppe, "Dynamic estimation of air pollution," *IEEE Trans. Automat. Contr.*, vol. AC-19, no. 6, pp. 904–910, Dec. 1974.

[6] D. Saint, X. C. Du, and G. Demoment, "Image restoration using a noncausal state-space model and a fast 2-D Kalmar filter," in *Proc. Conf. Mathematics in Signal Processing*, Bath, UK, 1985.

[7] A. K. Mahalanabis and K. Xue, "An efficient two-dimensional Chandrasekhar filter for restoration of images degraded by spatial blur and noise," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-35, no. 11, pp. 1603–1610, Nov. 1987.
[8] S. Haykin, *Adaptive Filter Theory*, 2nd ed. Englewood Cliffs, NJ: Prentice Hall, 1991.
[9] J. G. Proakis, C. M. Rader, F. Ling, and C. L. Nikias, *Advanced Digital Signal Processing.* New York: Macmillan, 1992.
[10] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification.* Cambridge, Ma: MIT Press, 1983.
[11] G. Carayannis, D. Manolakis, and N. Kalouptsidis, "A fast sequential algorithm for least squares filtering and prediction," *IEEE Trans. Acoustics, Speech, Signal Process.*, vol. ASSP-31, no. 6, pp. 1394–1402, Dec. 1983.
[12] J. Cioffi and T. Kailath, "Fast recursive-least-squares transversal filters for adaptive filtering," *IEEE Trans. Acoustics, Speech, Signal Process.*, vol. ASSP-32, pp. 304–337, Apr. 1984.
[13] A. Houacine and G. Demoment, "Chandrasekhar adaptive regularizer for adaptive filtering," in *Proc. IEEE ICASSP*, Tokyo, Japan, 1986, pp. 2967–2970.
[14] A. Houacine, "Regularized fast recursive least squares algorithms for adaptive filtering," *IEEE Trans. Signal Process.*, vol. SP-39, no. 4, pp. 860–870, Apr. 1991.
[15] D. T. M. Slock, "Fast algorithms for fixed-order recursive least-squares parameter estimation," PhD dissertation, Stanford Univ., Stanford, Ca, 1989.
[16] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME J. Basic. Eng.*, vol. 82, pp. 34–45, Mar. 1960.
[17] G. S. Sidhu, "A shift invariance approach to fast estimation algorithms," PhD dissertation, Stanford Univ., Stanford, Ca, 1975.
[18] M. Morf, "Fast algorithms for multivariable systems," PhD dissertation, Stanford Univ., Stanford, Ca, 1974.
[19] A. H. Sayed, H. Lev-Ari, and T. Kailath, "Lattice filter interpretations of the Chandrasekhar recursions for estimation and spectral factorization," in *Proc. IEEE ICASSP*, San Francisco, Ca, vol. 4, pp. 549–552, Mar. 1992.
[20] A. H. Sayed and T. Kailath, "Structured matrices and fast RLS adaptive filtering," in *Proc. 2nd IFAC Workshop on Algorithms and Architectures for Real-Time Control*, P. J. Fleming and W. H. Kwon, eds. Seoul, Korea: Pergamon, pp. 211–216, Sept. 1992.
[21] ——, "A state-space approach to adaptive filtering," in *Proc. IEEE ICASSP*, Minneapolis, Mn, vol. 3, pp. 559–562, Apr. 1993.
[22] A. H. Sayed, "Displacement structure in signal processing and mathematics," PhD dissertation, Stanford Univ., Stanford, Ca, Aug. 1992.
[23] B. H. Khalaj, A. H. Sayed, and T. Kailath, "A unified approach to multichannel least-squares algorithms," in *Proc. IEEE ICASSP*, Minneapolis, Mn, vol. 5, pp. 523–526, Apr. 1993.

On the Computation of Upper Covariance Bounds for Perturbed Linear Systems

P. Bolzern, P. Colaneri, and G. De Nicolao

***Abstract*—Motivated by previously published results, the computation of upper covariance bounds for perturbed linear systems is considered. It is shown that, for a wide choice of cost functions, the bound optimization problem is convex with respect to a scalar parameter. The analysis hinges on the properties of a H_∞-type Riccati equation.**

I. Introduction

The present note is motivated by the paper [1] where the computation of upper covariance bounds for perturbed linear systems is addressed. Among other things, in [1] it was observed, without any further consideration, that the bound optimization problem might not be convex. The main contribution of the present note is to show that convexity is actually guaranteed for a class of cost functions, including the one considered in [1].

Consider the time-invariant continuous-time linear system

$$\dot{x}(t) = A_0 x(t) + D_0 w(t)$$

where A_0 is stable and $w(t)$ is a white noise signal of unit intensity. Then, the asymptotic state covariance $X_0 = X_0' \geq 0$ is the unique solution of the algebraic Lyapunov equation

$$A_0 X_0 + X_0 A_0' + W_0 = 0$$

where $W_0 = D_0 D_0'$. As pointed out in [1], perturbations in the system matrix A_0 are inevitable in practice, so that the real system matrix is $A_0 + \Delta A$, where ΔA keeps into account model uncertainties. Correspondingly, as long as $A_0 + \Delta A$ remains stable, the state covariance $X = X'$ of the perturbed system is the unique solution of

$$(A_0 + \Delta A)X + X(A_0 + \Delta A)' + W_0 = 0.$$

In [2] and [1], the problem of obtaining an upper bound for the perturbed state covariance X was dealt with. In particular, it was shown that upper covariance bounds are provided by the solutions of a suitable H_∞-type Riccati equation.

Theorem 1 [1]: Let the uncertainty set Ω be defined as $\Omega \triangleq \{\Delta A\colon\ \Delta A\,\Delta A' \leq \overline{A}\}$, where $\overline{A}$ is a given nonnegative matrix. Suppose A_0 is stable and $(A_0 + \Delta A,\, W_0)$ is stabilizable $\forall \Delta A \in \Omega$. If there exist a real $\beta > 0$ and $\overline{X} \geq 0$ satisfying the Riccati equation

$$A_0 \overline{X} + \overline{X} A_0' + \overline{X}\,\overline{X}/\beta + \beta \overline{A} + W_0 = 0 \qquad (1)$$

then $A_0 + \Delta A$ is asymptotically stable and $X \leq \overline{X}$, $\forall \Delta A \in \Omega$. ∎

Manuscript received October 15, 1992; revised January 5, 1993. This work was supported by CNR Special Project "Algorithms and Architectures for Identification and Robust and Adaptive Control," MURST Project "Model Identification, Systems Control, Signal Processing," and Centro di Teoria dei Sistemi of the CNR, Milano, Italy.

P. Bolzern and P. Colaneri are with the Dipartimento di Elettronica e Informazione, Politecnico di Milano, 20133 Milano, Italy.
G. De Nicolao is with the Dipartimento di Informatica e Sistemistica, Università di Pavia, 27100 Pavia, Italy.

IEEE Log Number 9214006.