Proceedings of the
American Control Conference
Seattle, Washington • June 1995

OBLIQUE STATE-SPACE ESTIMATION ALGORITHMS

ALI H. SAYED[1] AND THOMAS KAILATH[2]

[1]Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106.

[2]Department of Electrical Engineering
Stanford University, Palo Alto, CA 94305.

Abstract—We pose and solve oblique estimation problems with and without state-space structure. In particular, we derive an oblique Kalman filter and indicate connections with instrumental-variable (IV) methods and higher-order spectra (HOS) analysis.

I. INTRODUCTION

Oblique projection problems have apparently received little attention in the literature, despite their implicit connections with useful tools in identification and signal processing. In identification problems, for instance, instrumental variable (IV) methods are often employed to guarantee consistent estimators [1, 2, 3]. The connection of these methods to oblique projections is well-known and has been pointed out in [4]. Likewise, in signal processing problems, oblique projections can be used in array processing and communication applications, as well as in higher-order spectra (HOS) analysis [5, 6, 7]. In these applications, the major objective can be interpreted as that of removing undesired interference or noise signals by using oblique operators.

In this paper, we pose two basic oblique estimation problems and establish a relation between their solutions. We then incorporate state-space structure into the statement of the problems and derive an oblique extension of the classical Kalman filter. Connections with the IV and HOS methods are then re-interpreted in terms of the state-space connections, along the same lines of [8]. In particular, a new array algorithm that avoids backsubstitution is suggested.

II. THE WEIGHTED OBLIQUE PROJECTION (WOP) PROBLEM

Let z_1 and z_2 represent two column vectors of n unknown parameters each, and let y_1 and y_2 be two observation vectors that are linearly related to (z_1, z_2):

$$y_1 = A_1 z_1 + v_1 = d_1 + v_1 , \qquad (1)$$
$$y_2 = A_2 z_2 + v_2 = d_2 + v_2 , \qquad (2)$$

where A_1 and A_2 are given matrices, and v_1 and v_2 are noise components. The terms d_1 and d_2 denote the uncorrupted parts $A_1 z_1$ and $A_2 z_2$, respectively.

We may regard the measurements $\{y_1, y_2\}$ as the result of two experiments: the data with subscript $_1$ arises from Experiment I while the data with subscript $_2$ arises from Experiment II.

This work was supported in part by a Research Initiation Award from the National Science Foundation under award no. MIP-9409319, and by the Army Research Office under contract DAAL03-89-K-0109.

The goal is to estimate the unknown vectors (z_1, z_2) and, consequently, the uncorrupted parts (d_1, d_2), from the noisy measurements $\{y_1, y_2\}$. The estimates of (z_1, z_2) will be denoted by $(\hat{z}_1, \hat{z}_2)$. Accordingly, the estimates of (d_1, d_2) will be computed via $\hat{d}_1 = A_1 \hat{z}_1$ and $\hat{d}_2 = A_2 \hat{z}_2$. The corresponding estimation errors will then be given by $\hat{v}_1 = y_1 - A_1 \hat{z}_1$ and $\hat{v}_2 = y_2 - A_2 \hat{z}_2$.

The criterion we shall use for the determination of $(\hat{z}_1, \hat{z}_2)$ will admit an interpretation in terms of oblique projections. For this reason, we shall use the shorthand notation WOP to refer to the resulting so-called *weighted-oblique-projection problem*.

We first state a simplified version of the WOP problem. The estimates $(\hat{z}_1, \hat{z}_2)$ are to be chosen as the *stationary solution* of the two-variable cost function

$$J[z_1, z_2] = [y_1 - A_1 z_1]^* W^{-1} [y_2 - A_2 z_2] , \qquad (3)$$

where W is a given invertible weighting matrix (possibly non-Hermitian). The reason for choosing this criterion is that it leads to a solution that exhibits a *decoupling property* as explained in the sequel.

Indeed, $\hat{z}_1$ ($\hat{z}_2$) is obtained by annihilating the gradient of $J[z_1, z_2]$ with respect to z_2 (z_1^*). This leads to the linear systems of equations (also known as oblique orthogonality conditions)

$$A_2^* W^{-*} [y_1 - A_1 \hat{z}_1] = 0, \quad A_1^* W^{-1} [y_2 - A_2 \hat{z}_2] = 0.$$

Let $\mathcal{A}_1$ and $\mathcal{A}_2$ denote the column spaces of the matrices A_1 and A_2, respectively. It thus follows from the above conditions that $\hat{v}_1$ is W^{-*}–orthogonal to $\mathcal{A}_2$ and $\hat{v}_2$ is W^{-1}–orthogonal to $\mathcal{A}_1$.[1] Assume, for convenience of explanation, that W is the identity matrix. The above facts can then be interpreted as follows: (i) The estimate $\hat{d}_1$ is chosen from the space $\mathcal{A}_1$ (where d_1 lies) in such a way that the resulting error vector, $\hat{v}_1$, is orthogonal to the space $\mathcal{A}_2$. (ii) Likewise, the estimate $\hat{d}_2$ is chosen from the space $\mathcal{A}_2$ (where d_2 lies) in such a way that the resulting error vector, $\hat{v}_2$, is orthogonal to the space $\mathcal{A}_1$.

More generally, we may consider the two-variable cost-function

$$J[z_1, z_2] = z_1^* \Pi^{-1} z_2 + [y_1 - A_1 z_1]^* W^{-1} [y_2 - A_2 z_2] , \qquad (4)$$

where Π is a given invertible weighting matrix (possibly non-Hermitian). In this case, and assuming the invertibility of the matrix $\left[\Pi^{-1} + A_1^* W^{-1} A_2\right]$, the unique stationary solution is given by the expressions

$$\hat{z}_1 = \left[\Pi^{-*} + A_2^* W^{-*} A_1\right]^{-1} A_2^* W^{-*} y_1 , \qquad (5)$$

$$\hat{z}_2 = \left[\Pi^{-1} + A_1^* W^{-1} A_2\right]^{-1} A_1^* W^{-1} y_2 . \qquad (6)$$

[1]Two column vectors p and q are said to be R-orthogonal if $p^* R q = 0$.

It also follows that the value of (4) at its stationary solution is equal to $y_1^* \left[W + A_2 \Pi A_1^*\right]^{-1} y_2$.

A point in favor of the modified criterion (4) is that uniqueness of a solution now requires the invertibility of the matrix $\left[\Pi^{-1} + A_1^* W^{-1} A_2\right]$. This is in contrast to (3), which requires the invertibility of $A_1^* W^{-1} A_2$.

III. An Equivalent Estimation Problem

We now introduce another estimation problem whose (unique) stationary solution will turn out to have the same expression as (5) and (6). We shall refer to it as the *equivalent estimation* (EE) problem. This problem may be formulated in a stochastic setting, where the variables of interest are assumed random. For generality, however, we shall allow the variables to lie in an abstract linear vector space, say $\mathcal{M}$.

The space $\mathcal{M}$ is required to satisfy two simple requirements: (i) $\mathcal{M}$ is linear over $\mathcal{C}$, the set of complex numbers, and (ii) $\mathcal{M}$ possesses a bilinear form, $< .,. >_{\mathcal{M}}$, such that for any $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathcal{M}$, and for any $\alpha, \beta \in \mathcal{C}$, we have $< \alpha\mathbf{a} + \beta\mathbf{b}, \mathbf{c} >_{\mathcal{M}} = \alpha < \mathbf{a}, \mathbf{c} >_{\mathcal{M}} + \beta < \mathbf{b}, \mathbf{c} >_{\mathcal{M}}$ and $< \mathbf{b}, \mathbf{a} >_{\mathcal{M}} = < \mathbf{a}, \mathbf{b} >^*_{\mathcal{M}}$. The space of random variables with the bilinear form $< \mathbf{p}, \mathbf{q} >_{\mathcal{M}} = E\mathbf{p}\mathbf{q}^*$ is a prominent example. [Note that we are using boldface letters to denote the variables of an EE problem].

The EE problem can be formulated as follows. Consider column vectors $\{\mathbf{y}_1, \mathbf{y}_2, \mathbf{v}_1, \mathbf{v}_2, \mathbf{z}_1, \mathbf{z}_2\}$ that are linearly related via the expressions $\mathbf{y}_1 = A_1\mathbf{z}_1 + \mathbf{v}_1$ and $\mathbf{y}_2 = A_2\mathbf{z}_2 + \mathbf{v}_2$, for some matrices (A_1, A_2), and where the individual entries of the column vectors $\{\mathbf{y}_1, \mathbf{y}_2, \mathbf{v}_1, \mathbf{v}_2, \mathbf{z}_1, \mathbf{z}_2\}$ are all elements of $\mathcal{M}$.

The variables $\{\mathbf{v}_1, \mathbf{v}_2\}$ and $\{\mathbf{z}_1, \mathbf{z}_2\}$ can be regarded as having cross-Gramian matrices $\{W, \Pi\}$, $W = < \mathbf{v}_2, \mathbf{v}_1 >_{\mathcal{M}}$, $\Pi = < \mathbf{z}_2, \mathbf{z}_1 >_{\mathcal{M}}$, along with the conditions $< \mathbf{z}_1, \mathbf{v}_2 >_{\mathcal{M}} = 0$ and $< \mathbf{z}_2, \mathbf{v}_1 >_{\mathcal{M}} = 0$. In a stochastic setting, this means that the noise variables of one experiment are uncorrelated with the parameters of the other experiment. Under these conditions, it follows from the linear models that the cross-Gramian matrix of $(\mathbf{y}_2, \mathbf{y}_1)$ is $< \mathbf{y}_2, \mathbf{y}_1 >_{\mathcal{M}} = A_2 \Pi A_1^* + W$.

Let $J(K)$ denote the two-variable cost function $J[K_1, K_2] = < \mathbf{z}_1 - K_1\mathbf{y}_1, \mathbf{z}_2 - K_2\mathbf{y}_2 >_{\mathcal{M}}$. The equivalent problem is one that estimates $\mathbf{z}_1$ from $\{\mathbf{y}_1, A_1\}$ and $\mathbf{z}_2$ from $\{\mathbf{y}_2, A_2\}$ via linear relations of the form

$$\hat{\mathbf{z}}_1 = K_1^o \mathbf{y}_1, \;\; \hat{\mathbf{z}}_2 = K_2^o \mathbf{y}_2,$$

where (K_1^o, K_2^o) are chosen as a stationary solution of $J[K_1, K_2]$.

It is immediate to verify, under the assumption of an invertible cross-Gramian matrix $(A_2 \Pi A_1^* + W)$, that the unique estimates $\hat{\mathbf{z}}_1$ and $\hat{\mathbf{z}}_2$ are given by the expressions

$$\hat{\mathbf{z}}_1 = \left[\Pi^{-*} + A_2^* W^{-*} A_1\right]^{-1} A_2^* W^{-*} \mathbf{y}_1, \qquad (7)$$

$$\hat{\mathbf{z}}_2 = \left[\Pi^{-1} + A_1^* W^{-1} A_2\right]^{-1} A_1^* W^{-1} \mathbf{y}_2. \qquad (8)$$

IV. Relations between the WOP and the EE Problems

Comparing expressions (7) and (8) with (5) and (6), we see that if we make the identifications: $\hat{\mathbf{z}}_1 \leftrightarrow \hat{z}_1$, $\hat{\mathbf{z}}_2 \leftrightarrow \hat{z}_2$, $\mathbf{y}_1 \leftrightarrow y_1$, $\mathbf{y}_2 \leftrightarrow y_2$, then the expressions coincide. This means that the WOP problem and the equivalent problem have the same expressions for the solutions, $(\hat{z}_1, \hat{z}_2)$ and $(\hat{\mathbf{z}}_1, \hat{\mathbf{z}}_2)$.

The significance of this fact is the following. It often happens in applications that one is interested in solving estimation problems of either forms (3) or (4). Particular examples arise in instrumental-variable methods and in higher-order spectra analysis [2, 6] (though not in such an explicit form – see, e.g., Sec. VIII. ahead). On the other hand, problems of the EE type will be shown here to lead, in the presence of state-space structure, to what we shall call *an oblique Kalman filter*. By relating the solutions of the WOP and EE problems, we shall then be able to apply Kalman-type algorithms to the solution of the WOP problem. This is particularly useful since it not only allows us to interpret results in other areas as special instances of the oblique Kalman filter, but it also allows us to exploit the wealth of results that are available in the standard Kalman filtering literature. An example to this effect was given in [8] for the special case of positive-definite quadratic cost functions, where a close link was established between known results in Kalman filtering theory and more recent results in adaptive filtering theory. In particular, it was shown in [8] that once such an equivalence relation is established, the varied forms of adaptive filtering algorithms can be obtained by writing down different variants of the Kalman-filter. One may therefore expect a similar situation in the oblique case in view of the relation between the WOP and EE problems.

V. Incorporating State-Space Structure

Now that we have established a relation betwee the two basic estimation problems, WOP and EE, we shall proceed to study an important special case of the equivalent problem.

More specifically, we shall pose an equivalent estimation problem that will be of the same form as the one studied in Sec. III., except that the associated matrices A_1 and A_2 will have considerable structure in them. In particular, they will be block-lower triangular and their individual entries will be further parameterized in terms of matrices

$$\{F_{i,1}, G_{i,1}, H_{i,1}, F_{i,2}, G_{i,2}, H_{i,2}\}$$

that arise from an underlying state-space assumption.

Consider vectors $\{\mathbf{y}_{i,1}, \mathbf{x}_{i,1}, \mathbf{u}_{i,1}, \mathbf{v}_{i,1}, \mathbf{y}_{i,2}, \mathbf{x}_{i,2}, \mathbf{u}_{i,2}, \mathbf{v}_{i,2}\}$, all with entries in $\mathcal{M}$, and assume that they satisfy the state-space equations

$$\begin{aligned} \mathbf{x}_{i+1,1} &= F_{i,1}\mathbf{x}_{i,1} + G_{i,1}\mathbf{u}_{i,1}, \\ \mathbf{y}_{i,1} &= H_{i,1}\mathbf{x}_{i,1} + \mathbf{v}_{i,1}, \;\; i \geq 0, \end{aligned} \qquad (9)$$

$$\begin{aligned} \mathbf{x}_{i+1,2} &= F_{i,2}\mathbf{x}_{i,2} + G_{i,2}\mathbf{u}_{i,2}, \\ \mathbf{y}_{i,2} &= H_{i,2}\mathbf{x}_{i,2} + \mathbf{v}_{i,2}, \;\; i \geq 0, \end{aligned} \qquad (10)$$

where $\{F_{i,1}, H_{i,1}, G_{i,1}\}$ are known $n \times n$, $p \times n$, and $n \times m$ matrices, respectively. Likewise for $\{F_{i,2}, H_{i,2}, G_{i,2}\}$. It is further assumed that

$$< \begin{bmatrix} \mathbf{u}_{i,2} \\ \mathbf{v}_{i,2} \\ \mathbf{x}_{0,2} \end{bmatrix}, \begin{bmatrix} \mathbf{u}_{j,1} \\ \mathbf{v}_{j,1} \\ \mathbf{x}_{0,1} \end{bmatrix} >_{\mathcal{M}} = \begin{bmatrix} Q_i\delta_{ij} & 0 & 0 \\ 0 & R_i\delta_{ij} & 0 \\ 0 & 0 & \Pi_0 \end{bmatrix},$$

where δ_{ij} is the Kronecker delta function that is equal to unity when $i = j$ and zero otherwise.

The state-space structure (9) leads to a linear relation between the vectors $\{\mathbf{y}_{i,1}\}$ and the vectors $\{\mathbf{x}_{0,1}, \mathbf{u}_{i,1}\}_{i=0}^{N-1}$. Indeed, if we collect the $\{\mathbf{y}_{i,1}\}_{i=0}^{N}$ and the $\{\mathbf{v}_{i,1}\}_{i=0}^{N}$ into two column vectors, $\mathbf{y}_1 = \text{col}\{\mathbf{y}_{0,1}, \ldots, \mathbf{y}_{N,1}\}$ and $\mathbf{v}_1 = \text{col}\{\mathbf{v}_{0,1}, \ldots, \mathbf{v}_{N,1}\}$, and define

$$\mathbf{z}_{N,1} = \text{col}\{\mathbf{x}_{0,1}, \mathbf{u}_{0,1}, \ldots, \mathbf{u}_{N-1,1}\} = \text{col}\{\mathbf{x}_{0,1}, \mathbf{u}_1\},$$

it then follows from the state-space equations (9) that $\mathbf{y}_1 = A_1\mathbf{z}_{N,1} + \mathbf{v}_1$, where A_1 is the block-lower triangular matrix

$$\begin{bmatrix} H_{0,1} & & & \\ H_{1,1}F_1^{[0,0]} & H_{1,1}G_{0,1} & & \\ H_{2,1}F_1^{[1,0]} & H_{2,1}F_1^{[1,1]}G_{0,1} & & \\ \vdots & \vdots & \ddots & \\ H_{N,1}F_1^{[N-1,0]} & H_{N,1}F_1^{[N-1,1]}G_{0,1} & \ldots & H_{N,1}G_{N-1,1} \end{bmatrix}.$$

Here, the notation $F_1^{[i,j]}$, $i \geq j$, stands for $F_{i,1}F_{i-1,1}\ldots F_{j,1}$. A similar remark holds for the second state-space equations (10), thus leading to $\mathbf{y}_2 = A_2\mathbf{z}_{N,2} + \mathbf{v}_2$.

Moreover, the cross-Gramian matrices of the variables $\{\mathbf{z}_{N,1}, \mathbf{v}_1, \mathbf{y}_1\}$ and $\{\mathbf{z}_{N,2}, \mathbf{v}_2, \mathbf{y}_2\}$, so defined, are easily seen to be $< \mathbf{z}_{N,2}, \mathbf{z}_{N,1} >_{\mathcal{M}} = (\Pi_0 \oplus Q_0 \ldots \oplus Q_{N-1})$ and $< \mathbf{v}_2, \mathbf{v}_1 >_{\mathcal{M}} = (R_0 \oplus R_1 \oplus \ldots \oplus R_N)$. More compactly, we shall write $< \mathbf{z}_{N,2}, \mathbf{z}_{N,1} >_{\mathcal{M}} = \Pi$ and $< \mathbf{v}_2, \mathbf{v}_1 >_{\mathcal{M}} = W$, where the $\{\Pi, W\}$ are the above block diagonal matrices.

We can now pose the problem of estimating $\mathbf{z}_{N,1}$ from the variables $\{\mathbf{y}_{0,1}, \mathbf{y}_{1,1}, \ldots, \mathbf{y}_{N,1}\}$, as explained at the end of Sec. III. The solution is denoted by $\hat{\mathbf{z}}_{N,1|N}$ and may be globally expressed, in the unique case, as (*cf.* (5))

$$\hat{\mathbf{z}}_{N,1|N} = \left[\Pi^{-*} + A_2^* W^{-*} A_1\right]^{-1} A_2^* W^{-*} \mathbf{y}_1. \tag{11}$$

Likewise,

$$\hat{\mathbf{z}}_{N,2|N} = \left[\Pi^{-1} + A_1^* W^{-1} A_2\right]^{-1} A_1^* W^{-1} \mathbf{y}_2. \tag{12}$$

We are, however, interested in a recursive construction of the estimates $\hat{\mathbf{z}}_{N,1|N}$ and $\hat{\mathbf{z}}_{N,2|N}$, namely one that allows us to update $\hat{\mathbf{z}}_{N,1|i-1}$ to $\hat{\mathbf{z}}_{N,1|i}$, and $\hat{\mathbf{z}}_{N,2|i-1}$ to $\hat{\mathbf{z}}_{N,2|i}$, for $i = 0, 1, \ldots, N$. Here, the notation $\hat{\mathbf{z}}_{N,1|i}$ denotes the linear estimate of $\mathbf{z}_{N,1}$ that is based on the data up to time i, $\{\mathbf{y}_{0,1}, \ldots, \mathbf{y}_{i,1}\}$.

Let R_y denote the cross-Gramian matrix $< \mathbf{y}_2, \mathbf{y}_1 >_{\mathcal{M}} = W + A_2\Pi A_1^*$. It can be shown that a recursive algorithm is possible as long as R_y is a (block) strongly regular matrix. In this case, the resulting (smoothing) algorithm is the following.

Theorem 1 *Assume R_y is (block) strongly regular and start with $\hat{\mathbf{z}}_{N,1|-1} = \hat{\mathbf{z}}_{N,2|-1} = 0$. Then, for $i = 0, \ldots, N$,*

$$\begin{aligned}
\hat{\mathbf{z}}_{N,1|i} &= \hat{\mathbf{z}}_{N,1|i-1} + K_{z,i,1} H_{i,2}^* R_{e,i}^{-*}\, \mathbf{e}_{i,1}, \\
\hat{\mathbf{z}}_{N,2|i} &= \hat{\mathbf{z}}_{N,2|i-1} + K_{z,i,2} H_{i,1}^* R_{e,i}^{-1}\, \mathbf{e}_{i,2},
\end{aligned}$$

where $K_{z,0,1} = \begin{bmatrix} \Pi_0^* \\ 0 \end{bmatrix} = K_{z,0,2}$ *and*

$$K_{z,i+1,1} = K_{z,i,1}\left[F_{i,2} - K_{i,2} R_{e,i}^{-1} H_{i,2}\right]^* + \begin{bmatrix} 0 \\ I \\ 0 \end{bmatrix} Q_i^* G_{i,2}^*,$$

$$K_{z,i+1,2} = K_{z,i,2}\left[F_{i,1} - K_{i,1} R_{e,i}^{-*} H_{i,1}\right]^* + \begin{bmatrix} 0 \\ I \\ 0 \end{bmatrix} Q_i G_{i,1}^*.$$

The identity matrices in the recursion for either $K_{z,i+1,1}$ or $K_{z,i+1,2}$ occur at the positions that correspond to the entries $\mathbf{u}_{i,1}$ and $\mathbf{u}_{i,2}$. Moreover, the $\{\mathbf{e}_{i,1}, \mathbf{e}_{i,2}\}$ are computed via the following so-called oblique Kalman filter: start with $\hat{\mathbf{x}}_{0,1|-1} = \hat{\mathbf{x}}_{0,2|-1} = 0$, $P_0 = \Pi_0$, and repeat for $i \geq 0$:

$$\begin{aligned}
\mathbf{e}_{i,1} &= \mathbf{y}_{i,1} - H_{i,1}\hat{\mathbf{x}}_{i,1}, \quad \mathbf{e}_{i,2} = \mathbf{y}_{i,2} - H_{i,2}\hat{\mathbf{x}}_{i,2}, \\
\hat{\mathbf{x}}_{i+1,1} &= F_{i,1}\hat{\mathbf{x}}_{i,1} + K_{i,1} R_{e,i}^{-*} \mathbf{e}_{i,1}, \\
\hat{\mathbf{x}}_{i+1,2} &= F_{i,2}\hat{\mathbf{x}}_{i,2} + K_{i,2} R_{e,i}^{-1} \mathbf{e}_{i,2}, \\
R_{e,i} &= R_i + H_{i,2} P_i H_{i,1}^*, \\
K_{i,1} &= F_{i,1} P_i^* H_{i,2}^*, \quad K_{i,2} = F_{i,2} P_i H_{i,1}^*, \\
P_{i+1} &= F_{i,2} P_i F_{i,1}^* + G_{i,2} Q_i G_{i,1}^* - K_{i,2} R_{e,i}^{-1} K_{i,1}^*.
\end{aligned}$$

Two remarks are due here. First, the successive $\hat{\mathbf{z}}_{N,1|i}$ have the form $\hat{\mathbf{z}}_{N,1|i} = \mathrm{col}\{\hat{\mathbf{x}}_{0,1|i}, \hat{\mathbf{u}}_{0,1|i}, \ldots, \hat{\mathbf{u}}_{i-1,1|i}, 0, \ldots, 0\}$, where $\hat{\mathbf{x}}_{0,1|i}$ denotes the linear estimate of $\mathbf{x}_{0,1}$ that is based on $\{\mathbf{y}_{0,1}, \ldots, \mathbf{y}_{i,1}\}$. Likewise, $\hat{\mathbf{u}}_{j,1|i}$ denotes the linear estimate of $\mathbf{u}_{j,1}$ that is based on the same vectors $\{\mathbf{y}_{0,1}, \ldots, \mathbf{y}_{i,1}\}$. We shall therefore write $\hat{\mathbf{z}}_{N,1|i} = \mathrm{col}\{\hat{\mathbf{z}}_{i,1|i}, 0, \ldots, 0\}$, where $\mathbf{z}_{i,1} = \mathrm{col}\{\mathbf{x}_{0,1}, \mathbf{u}_{0,1}, \ldots, \mathbf{u}_{i-1,1}\}$. A similar remark holds for $\hat{\mathbf{z}}_{N,2|i}$ and we write $\hat{\mathbf{z}}_{N,2|i} = \mathrm{col}\{\hat{\mathbf{z}}_{i,2|i}, 0, \ldots, 0\}$.

Secondly, the oblique Kalman filter can be regarded as an instrument that "uncorrelates" the input data: it transforms the given "correlated" data $\{\mathbf{y}_{i,1}, \mathbf{y}_{i,2}\}$ into two sets of "*uncorrelated*" data $\{\mathbf{e}_{i,1}, \mathbf{e}_{i,2}\}$. Moreover, it consists of two blocks: one block processes $\{\mathbf{y}_{i,1}\}$ and provides $\{\mathbf{e}_{i,1}\}$ and the other processes $\{\mathbf{y}_{i,2}\}$ and provides $\{\mathbf{e}_{i,2}\}$. The two blocks do not share the observations vectors $\{\mathbf{y}_{i,1}, \mathbf{y}_{i,2}\}$ but rather the underlying state-space structures.

VI. A Recursive WOP Problem in State-Space Form

Now, in view of the discussion in Sec. IV., the solutions $\hat{\mathbf{z}}_{N,1|N}$ and $\hat{\mathbf{z}}_{N,2|N}$ have the same expression as the solutions $\hat{z}_{N,1|N}$ and $\hat{z}_{N,2|N}$ of a related WOP problem. Indeed, it is rather immediate to write down the WOP problem whose stationary solution $(\hat{z}_{N,1|N}, \hat{z}_{N,2|N})$ matches the above $(\hat{\mathbf{z}}_{N,1|N}, \hat{\mathbf{z}}_{N,2|N})$; its cost function takes the form

$$J_N \equiv J\left[z_{N,1}, z_{N,2}\right] = \begin{bmatrix} x_{0,1} \\ u_1 \end{bmatrix}^* \Pi^{-1} \begin{bmatrix} x_{0,2} \\ u_2 \end{bmatrix} +$$

$$\left(y_1 - A_1 \begin{bmatrix} x_{0,1} \\ u_1 \end{bmatrix}\right)^* W^{-1} \left(y_2 - A_2 \begin{bmatrix} x_{0,2} \\ u_2 \end{bmatrix}\right),$$

subject to the state-space constrains $x_{i+1,1} = F_{i,1}x_{i,1} + G_{i,1}u_{i,1}$ and $x_{i+1,2} = F_{i,2}x_{i,2} + G_{i,2}u_{i,2}$. Here

$$z_{N,1} = \begin{bmatrix} x_{0,1} \\ u_{0,1} \\ \vdots \\ u_{N-1,1} \end{bmatrix}, \quad z_{N,2} = \begin{bmatrix} x_{0,2} \\ u_{0,2} \\ \vdots \\ u_{N-1,2} \end{bmatrix},$$

and $\Pi = (\Pi_0 \oplus Q_0 \ldots \oplus Q_{N-1})$, $W = (R_0 \oplus \ldots \oplus R_N)$. Equivalently, using the state-equations, we can rewrite J_N as

$$J_N = x_{0,1}^* \Pi_0^{-1} x_{0,2} + \sum_{j=0}^{N-1} u_{j,1}^* Q_j^{-1} u_{j,2} +$$

$$\sum_{j=0}^{N} (y_{j,1} - H_{j,1}x_{j,1})^* R_j^{-1} (y_{j,2} - H_{j,2}x_{j,2}),$$

subject to the same state-space constraints. Likewise, the cost function of the WOP problem whose stationary solution $\{\hat{z}_{i,1|i}, \hat{z}_{i,2|i}\}$ matches the $\{\hat{\mathbf{z}}_{i,1|i}, \hat{\mathbf{z}}_{i,2|i}\}$ is given by

$$J_i = x_{0,1}^* \Pi_0^{-1} x_{0,2} + \sum_{j=0}^{i-1} u_{j,1}^* Q_j^{-1} u_{j,2} +$$

$$\sum_{j=0}^{i} (y_{j,1} - H_{j,1}x_{j,1})^* R_j^{-1} (y_{j,2} - H_{j,2}x_{j,2}),$$

subject to the same state-space constraints.

It is also clear from Sec. IV. that the recursions of Theorem 1, with the proper identifications $\hat{\mathbf{z}}_{N,1|i} \leftrightarrow \hat{z}_{N,1|i}$, $\hat{\mathbf{z}}_{N,2|i} \leftrightarrow \hat{z}_{N,2|i}$, $\mathbf{y}_{i,1} \leftrightarrow y_{i,1}$, $\mathbf{y}_{i,2} \leftrightarrow y_{i,2}$, $\mathbf{u}_{i,1} \leftrightarrow u_{i,1}$, $\mathbf{u}_{i,2} \leftrightarrow u_{i,2}$, can be used to compute the stationary solutions $\{\hat{z}_{i,1|i}, \hat{z}_{i,2|i}\}$ of J_i. In particular, we also have that the stationary solutions $\{\hat{z}_{i,1|i}, \hat{z}_{i,2|i}\}$ are related to the $\{\hat{z}_{N,1|i}, \hat{z}_{N,2|i}\}$ as follows: $\hat{z}_{N,1|i} = \mathrm{col}\{\hat{z}_{i,1|i}, 0\}$ and $\hat{z}_{N,2|i} = \mathrm{col}\{\hat{z}_{i,2|i}, 0\}$.

VII. Array Algorithms for Oblique Estimation

Many computational variants to the oblique Kalman filter of Theorem 1 can be developed, by borrowing on the wealth of material that is available in the literature on the classical Kalman filtering problem. In particular, an attractive variant is the class of array algorithms. In this case, one forms a prearray of numbers and proceeds to triangularize it using elementary rotations; the rotations may be unitary, hyperbolic, or coupled. The relevant details will be pursued elsewhere. Here, we shall only provide an illustrative example that assumes $G_{i,1} = G_{i,2} = 0$ in the state-space models (9) and (10).

To begin with, we first note that the oblique Kalman recursions of Theorem 1 require that a Riccati variable P_i be explicitly propagated via a Riccati recursion. An alternative algorithm that avoids this step, and which is more amenable to parallelizable implementations, is to develop an array variant. One such variant is what we shall call the *extended oblique information filter*. It can be derived as follows: introduce lower-upper triangular factorizations of the matrices P_i, Q_i, R_i, and $R_{e,i}$, viz.,

$$P_i = L_{p,i}U_{p,i}\ ,\ \ Q_i = L_{q,i}U_{q,i},$$
$$R_i = L_{r,i}U_{r,i}\ ,\ \ R_{e,i} = L_{e,i}U_{e,i},$$

and form the two prearrays that appear on the left-hand side of the equations shown below. Then choose rotation matrices Θ_i and Γ_i that satisfy two constraints: (i) they are coupled, i.e., they satisfy $\Theta_i\Gamma_i^* = I$ and (ii) they annihilate the $(1,2)$ block entries of the prearrays. Examples to this effect can be found in [9].

The claim is that once the prearray entries are processed by these rotation matrices, the quantities that appear in the postarrays on the right-hand side are the ones indicated below:

$$\begin{bmatrix} F_{i,2}^{-*}L_{p,i}^{-*} & F_{i,2}^{-*}H_{i,2}^{*}L_{r,i}^{-*} \\ \hat{\mathbf{x}}_{i,2}^{*}L_{p,i}^{-*} & \mathbf{y}_{i,2}^{*}U_{r,i}^{-1} \\ 0 & L_{r,i}^{-*} \\ F_{i,1}U_{p,i}^{*} & 0 \end{bmatrix}\Gamma_i =$$
$$\begin{bmatrix} L_{p,i+1}^{-*} & 0 \\ \hat{\mathbf{x}}_{i+1,2}^{*}L_{p,i+1}^{-*} & \mathbf{e}_{i,2}^{*}U_{e,i}^{-1} \\ R_i^{-1}H_{i,1}F_{i,1}^{-1}U_{p,i+1}^{*} & U_{e,i}^{-1} \\ U_{p,i+1}^{*} & -\bar{K}_{p,i,1} \end{bmatrix},$$

and

$$\begin{bmatrix} F_{i,1}^{-*}U_{p,i}^{-1} & F_{i,1}^{-*}H_{i,1}^{*}U_{r,i}^{-1} \\ \hat{\mathbf{x}}_{i,1}^{*}U_{p,i}^{-1} & \mathbf{y}_{i,1}^{*}L_{r,i}^{-*} \\ 0 & U_{r,i}^{-1} \\ F_{i,2}L_{p,i} & 0 \end{bmatrix}\Theta_i =$$
$$\begin{bmatrix} U_{p,i+1}^{-1} & 0 \\ \hat{\mathbf{x}}_{i+1,1}^{*}U_{p,i+1}^{-1} & \mathbf{e}_{i,1}^{*}L_{e,i}^{-*} \\ R_i^{-*}H_{i,2}F_{i,2}^{-1}L_{p,i+1} & L_{e,i}^{-*} \\ L_{p,i+1} & -\bar{K}_{p,i,2} \end{bmatrix}.$$

Moreover, the matrices $(\bar{K}_{p,i,1}, \bar{K}_{p,i,2})$ are normalized gain matrices that allow us to update the state-estimates as follows

$$\hat{\mathbf{x}}_{i+1,1} = F_{i,1}\hat{\mathbf{x}}_{i,1} + \bar{K}_{p,i,1}L_{e,i}^{-1}\mathbf{e}_{i,1},$$
$$\hat{\mathbf{x}}_{i+1,2} = F_{i,2}\hat{\mathbf{x}}_{i,2} + \bar{K}_{p,i,2}U_{e,i}^{-*}\mathbf{e}_{i,2}.$$

In other words, the array equations allow us to update the gain matrices without explicitly requiring the Riccati variable P_i. Alternatively, the state-estimates can also be evaluated from other entries in the postarrays such as

$$\hat{\mathbf{x}}_{i+1,1} = U_{p,i+1}^{*}\left[\hat{\mathbf{x}}_{i+1,1}^{*}U_{p,i+1}^{-1}\right]^{*},$$
$$\hat{\mathbf{x}}_{i+1,2} = L_{p,i+1}\left[\hat{\mathbf{x}}_{i+1,2}^{*}L_{p,i+1}^{-*}\right]^{*}$$

VIII. An Application to Instrumental Variable Methods

Consider a collection of observations that are linearly related to an unknown vector z_1, say $y_1 = A_1z_1 + v_1$. The least-squares estimate of z_1 is known to be (assuming A_1 is full rank)

$$\begin{aligned} \hat{z}_1 &= [A_1^*A_1]^{-1}A_1^*y_1\ , \\ &= z_1 + [A_1^*A_1]^{-1}A_1^*v_1 . \end{aligned} \tag{13}$$

The above solution does not provide a consistent estimate since, in general, i.e., $\lim_{N\to\infty}\hat{z}_1 \neq z_1$. Consistency would often require a condition of the form (e.g., [1, pp.23–24] and [3, Ch. 7]) $\lim_{N\to\infty}\frac{1}{N}A_1^*v_1 = 0$, which can also mean that the noise sequence must be white.

A classical procedure that provides consistent estimators is to employ instrumental variable methods. In this context, the estimate for z_1 is computed by an expression of the form (compare with (13))

$$\begin{aligned} \hat{z}_1 &= [A_2^*A_1]^{-1}A_2^*y_1, \\ &= z_1 + [A_2^*A_1]^{-1}A_2^*v_1\ , \end{aligned} \tag{14}$$

where the new matrix A_2 is chosen so as to result in an invertible matrix $A_2^*A_1$ and such that $\lim_{N\to\infty}\frac{1}{N}A_2^*v_1 = 0$. Expressions similar to (14) also arise in HOS analysis, where the choice of A_2 is suggested by the problem formulation [6, Chs. 7, 9].

Comparing (14) with our earlier expression (5) we see that the instrumental variable estimate can be interpreted as an oblique estimate with a weighting matrix $W = I$ and $\Pi \to \infty I$. This interpretation is not new. It has been noted earlier in the literature, though from a very different point of view (e.g., [4]). Here we shall pursue this connection from a state-space perspective (along the same lines of [8]). In particular, we shall derive a new coupled array implementation for the IV method that avoids the need for backsubstitution steps.

We can also incorporate apriori information, as well as weighting, into the statement of the problem in much the same as we did earlier in Sec. II. We would then require the invertibility of a matrix of the form $\left[\Pi^{-*} + A_2^*W^{-*}A_1\right]$ instead of $A_2^*A_1$.

The IV (and the related HOS) problem is therefore related to the computation of one of the entries of the stationary solution of a two-variable cost function. It can also be stated, with general apriori information, in state-space form as follows: given the state-space constraints

$$\begin{aligned} x_{i+1,1} &= x_{i,1},\ \ y_{i,1} = A_{i,1}x_{i,1} + v_{i,1},\ x_{0,1} = z_1, \\ x_{i+1,2} &= x_{i,2},\ \ y_{i,2} = A_{i,2}x_{i,2} + v_{i,2},\ x_{0,2} = z_2, \end{aligned}$$

with known $\{A_{i,1}, A_{i,2}, y_{i,1}, \Pi_0\}$, determine the estimate of z_1 that corresponds to an entry of the saddle solution of the cost function $J[z_1, z_2] =$

$$x_{0,1}^*\Pi_0^{-1}x_{0,2} + \sum_{j=0}^{N}[y_{j,1} - A_{j,1}x_{j,1}]^*[y_{j,2} - A_{j,2}x_{j,2}]. \tag{15}$$

Here, the $\{y_{i,2}, x_{i,2}\}$ are simply auxiliary quantities. But in other contexts, such as in higher-order spectral analysis methods, the instruments are often known.[2]

Specializing Theorem 1 to the above state-space constraints leads to the standard recursive IV procedure for estimating

[2] More general cost functions, say with exponential forgetting factors, can also be handled as in [8].

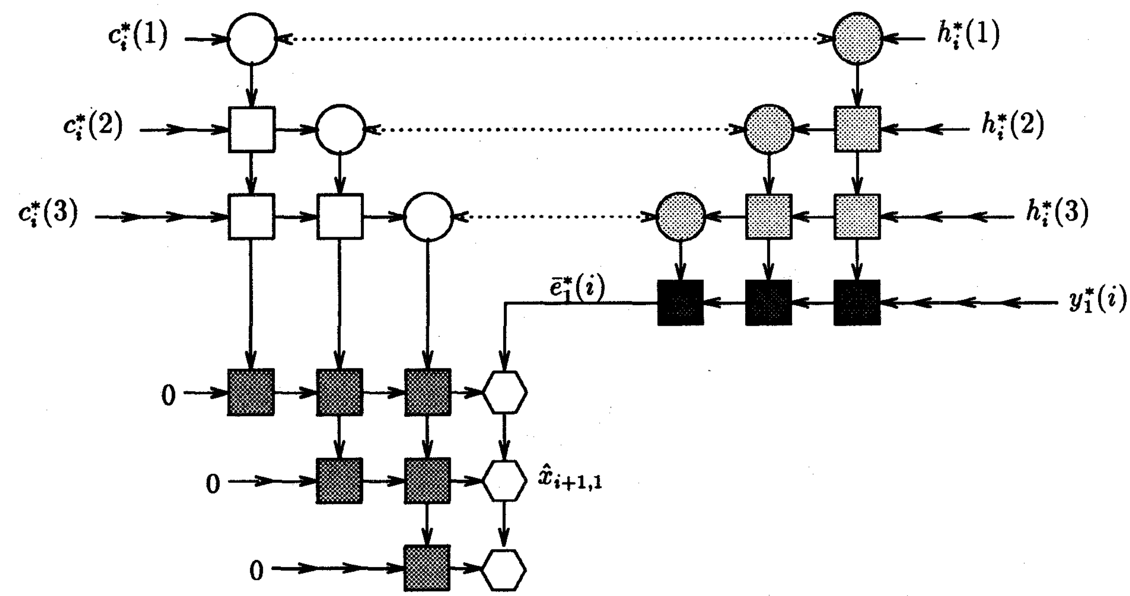

Figure 1: *A coupled array for instrumental-variable and HOS parameter estimation.*

z_1, where we now have $\hat{z}_1 = \hat{x}_{N+1,1}$: start with $\hat{x}_{0,1} = 0$, $P_0 = \Pi_0$, and repeat: $e_{i,1} = y_{i,1} - A_{i,1}\hat{x}_{i,1}$, $\hat{x}_{i+1,1} = \hat{x}_{i,1} + P_i^* A_{i,2}^* R_{e,i}^{-*} e_{i,1}$, $P_{i+1} = P_i - P_i A_{i,1}^* R_{e,i}^{-1} A_{i,2} P_i$, and $R_{e,i} = I + A_{i,2} P_i A_{i,1}^*$.

Once a connection with the state-space framework is made explicit, many algorithmic variants can now be applied to the IV and HOS contexts, in much the same way as was developed in the standard RLS problem in [8].

Indeed, translating the (information) array equations of Sec. VII. to the IV problem leads to the following alternative scheme: form the prearrays shown below and then triangularize them by choosing coupled rotations Θ_i and Γ_i (i.e., $\Theta_i \Gamma_i^* = I$):

$$\begin{bmatrix} L_{p,i}^{-*} & A_{i,2}^* \\ U_{p,i}^* & 0 \end{bmatrix} \Gamma_i = \begin{bmatrix} L_{p,i+1}^{-*} & 0 \\ U_{p,i+1}^* & -\bar{K}_{p,i,1} \end{bmatrix},$$

$$\begin{bmatrix} U_{p,i}^{-1} & A_{i,1}^* \\ \hat{x}_{i,1}^* U_{p,i}^{-1} & y_{i,1}^* \end{bmatrix} \Theta_i = \begin{bmatrix} U_{p,i+1}^{-1} & 0 \\ \hat{x}_{i+1,1}^* U_{p,i+1}^{-1} & e_{i,1}^* L_{e,i}^{-*} \end{bmatrix}.$$

The quantities obtained in the postarray can be used to form the prearrays for the next time instant, as well as for the update of the IV estimate,

$$\hat{x}_{i+1,1} \quad = \quad \hat{x}_{i,1} + \bar{K}_{p,i,1} L_{e,i}^{-1} e_{i,1}. \tag{16}$$

This solution avoids the need for a backsubstitution step, which is usually needed to solve a linear system of equations of the form $(A_2^* A_1)\hat{z}_1 = A_2^* y_1$, as in (14). An alternative presentation of this result in the context of (structured) matrix factorization can be found in [9].

For the sake of illustration, assume the $A_{i,1}$ and $A_{i,2}$ are row vectors, say 1×3, $A_{i,1} = \begin{bmatrix} h_i(1) & h_i(2) & h_i(3) \end{bmatrix}$ and $A_{i,2} = \begin{bmatrix} c_i(1) & c_i(2) & c_i(3) \end{bmatrix}$. Assume, accordingly, that $y_{i,1}$ and $L_{e,i}^{-1} e_{i,1}$ are scalar quantities, denoted by $y_1(i)$ and $\bar{e}_1(i)$, respectively. Then the above array equations admit a coupled implementation as depicted in Figure 1 for the special case $n = 3$. The figure consists of three triangular arrays and one linear array (see also [9]).

REFERENCES

[1] T. Söderström and P. G. Stoica. *Instrumental Variable Methods for System Identification.* Springer Verlag, Berlin Heidelberg, 1983.

[2] P. Young, *Recursive Estimation and Time-Series Analysis*, Springer-Verlag, 1984.

[3] L. Ljung. *System Identification: Theory for the User.* Prentice Hall, NJ, 1987.

[4] C. Samson. A unified treatment of fast algorithms for identification. *Int. J. Control*, 35(5):909–934, 1982.

[5] S. Kayalar and H. L. Weinert. Oblique projections: Formulas, algorithms and error bounds. *Math. Contr. Signals Syst.*, 2(1):33–45, 1989.

[6] C. L. Nikias and A. P. Petropulu. *Higher-Order Spectra Analysis: A Nonlinear Signal Processing Framework*, Prentice Hall, NJ, 1993.

[7] R. T. Behrens and L. L. Scharf. Signal processing applications of oblique projection operators. *IEEE Trans. on Signal Processing*, 42(6):1413–1424, June 1994.

[8] A. H. Sayed and T. Kailath. A state-space approach to adaptive RLS filtering. *IEEE Signal Processing Magazine*, 11(3):18–60, July 1994.

[9] A. H. Sayed. Time-variant structured matrices: An application to instrumental variable methods. *Proc. SPIE Conference on Advanced Signal Processing: Algorithms, Architectures, and Implementations*, vol. 2296, pp. 516–527, San Diego, CA, July 1994.