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

FUNDAMENTAL INERTIA CONDITIONS FOR THE SOLUTION OF H^{∞}–PROBLEMS

ALI H. SAYED[1], BABAK HASSIBI[2] 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 study the relation between the solutions of two minimization problems with indefinite quadratic forms. We show that a complete link between both solutions can be established by invoking a fundamental set of inertia conditions. While these inertia conditions are automatically satisfied in a standard Hilbert space setting, they nevertheless turn out to mark the differences between the two optimization problems in indefinite metric spaces. They also include, as special cases, the well-known conditions for the existence of H^{∞}-filters and controllers.

I. INTRODUCTION

Given two invertible Hermitian matrices $\{\Pi, W\}$, a column vector y, and an arbitrary matrix A of appropriate dimensions, we study the relation between the following two minimization problems:

$$\min_{z} \left[z^*\Pi^{-1}z + (y - Az)^*W^{-1}(y - Az)\right], \tag{1}$$

where z is a column vector of unknowns, and

$$\min_{K} \{\Pi - KA\Pi - \Pi A^*K^* + K[A\Pi A^* + W]K^*\}, \tag{2}$$

where K is a matrix. The symbol "$*$" stands for Hermitian conjugation. Both cost functions in (1) and (2) are quadratic in the respective independent variables z and K, and they can also be rewritten in the following revealing forms:

$$\begin{bmatrix} z^* & y^* \end{bmatrix} \begin{bmatrix} \Pi^{-1} + A^*W^{-1}A & -A^*W^{-1} \\ -W^{-1}A & W^{-1} \end{bmatrix} \begin{bmatrix} z \\ y \end{bmatrix},$$

$$\begin{bmatrix} I & -K \end{bmatrix} \begin{bmatrix} \Pi & \Pi A^* \\ A\Pi & A\Pi A^* + W \end{bmatrix} \begin{bmatrix} I \\ -K^* \end{bmatrix}, \tag{3}$$

where the central matrices are in fact the inverses of each other.

Moreover, and contrary to standard quadratic minimization problems, the weighting matrices Π and W are allowed to be indefinite. For this reason, solutions to (1) and (2) are not always guaranteed to exist. However, when they exist, we shall show that the expressions for the solutions, and the conditions for their existence, can be related via a fundamental set of inertia conditions. Here, by the inertia of an invertible Hermitian matrix X, we mean a pair of integers, denoted by $I_+(X)$ and $I_-(X)$, that are equal to the number of strictly positive and strictly negative eigenvalues of X.

The significance of the relations to be established between problems (1) and (2) is the following. It often happens in applications that one is interested in solving quadratic problems of the form (1), with indefinite weighting matrices. A particular example that has received increasing attention in the last decade is the class of H^{∞}-filtering and control problems – see, e.g., the recent book [GL95] for more details and extensive references on the topic. In this context, the Π matrix in (1) is further restricted to be positive-definite and the W matrix is indefinite but of the special diagonal form $W = \text{diag.}\{-\gamma^2 I, I\}$, for a given positive constant γ^2. Here we shall treat the general class of optimization problems suggested by (1) where both $\{\Pi, W\}$ are allowed to be arbitrary indefinite matrices.

On the other hand, problems of the form (2) are characteristic of state-space estimation formulations, where a so-called Kalman filter procedure is available as an efficient computational scheme for determining the solution in the presence of state-space structure, as pointed out in [HSK93]. By relating the solutions of (1) and (2) we shall then be able to apply Kalman-type algorithms to the solution of (1), as well as obtain a complete set of inertia conditions that will automatically test for the existence of solutions to (1), without discarding the available information from the solution of (2).

II. AN INERTIA RESULT FOR LINEAR TRANSFORMATIONS

We first establish a useful inertia result that tells us how the inertia of the matrices Π and W is affected by transformations of the form $(A\Pi A^* + W)$ and $(\Pi^{-1} + A^*W^{-1}A)$, for arbitrary matrices A of appropriate dimensions. The reason for choosing these transformations is because the positivity of these matrices will be shown later to be equivalent to necessary and sufficient conditions for the solvability of the problems (1) and (2). Hence, by studying how their inertia depends on $\{\Pi, W\}$, we shall be able to conclude how the choice of $\{\Pi, W\}$ affects the solvability of problems (1) and (2). The following three results follow by invoking the Schur decomposition of the central matrix in (3) (viz., the matrix C in (4) below), Sylvester's law of inertia [GV83], and the matrix inversion formula.

Lemma 1 *Given $\{\Pi, W\}$ Hermitian and invertible. Then, for any matrix A of appropriate dimensions, the block matrix*

$$C \triangleq \begin{bmatrix} \Pi & \Pi A^* \\ A\Pi & A\Pi A^* + W \end{bmatrix}, \tag{4}$$

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.

has the same positive and negative inertia as the block diagonal matrix $(\Pi \oplus W)$.

Lemma 2 *Given* $\{\Pi, W\}$ *Hermitian and invertible. Then, for any matrix* A *of appropriate dimensions,* $(A\Pi A^* + W)$ *is invertible if, and only if,* $(\Pi^{-1} + A^*W^{-1}A)$ *is invertible.*

Theorem 1 *Given* $\{\Pi, W\}$ *Hermitian and invertible. Then, for any matrix* A *of appropriate dimensions, the following inertia equalities hold,*

$$\begin{aligned} I_+(\Pi \oplus W) &= I_+[(\Pi^{-1} + A^*W^{-1}A) \oplus (A\Pi A^* + W)], \\ I_-(\Pi \oplus W) &= I_-[(\Pi^{-1} + A^*W^{-1}A) \oplus (A\Pi A^* + W)], \end{aligned}$$

if, and only if, $(A\Pi A^* + W)$ *is invertible.*

III. The Indefinite-Weighted Least-Squares Problem

We now focus on problem (1), which we shall refer to as the indefinite-weighted least-squares problem (IWLS, for short). The indefiniteness arises from the presence of the indefinite weighting matrices $\{\Pi, W\}$. Consequently, a bilinear form $a^*W^{-1}b$ is not guaranteed to satisfy the positivity condition $a^*W^{-1}a > 0$ for all nonzero column vectors a. We thus say that $\mathcal{C}^n$, coupled with a bilinear form $a^*W^{-1}b$ with W indefinite, is an indefinite metric space. More generally, an indefinite metric space $\{\mathcal{K}, <.,.>_{\mathcal{K}}\}$ is defined as a vector space that satisfies two simple requirements (see, e.g., [GLR83] for more details): $\mathcal{K}$ is linear over the field of complex numbers $\mathcal{C}$, and $\mathcal{K}$ possesses a bilinear form, $<.,.>_{\mathcal{K}}$, such that for any $a, b, c \in \mathcal{K}$, and for any $\alpha, \beta \in \mathcal{C}$, we have $<\alpha a + \beta b, c>_{\mathcal{K}} = \alpha <a, c>_{\mathcal{K}} + \beta <b, c>_{\mathcal{K}}$, and $<b, a>_{\mathcal{K}} = <a, b>^*_{\mathcal{K}}$.

Let $J(z)$ denote the quadratic cost function that appears in (1). Every $\hat{z}$ at which the gradient of $J(z)$ with respect to z vanishes is called a *stationary* point of $J(z)$. A stationary point $\hat{z}$ may or may not be a minimum of $J(z)$.

Theorem 2 *The stationary points* $\hat{z}$ *of* $J(z)$*, if they exist, are solutions of the linear system of equations*

$$[\Pi^{-1} + A^*W^{-1}A]\hat{z} = A^*W^{-1}y. \quad (5)$$

There exists a unique stationary point if, and only if, $[\Pi^{-1} + A^*W^{-1}A]$ *is invertible. In this case, it is given by*

$$\hat{z} = \left[\Pi^{-1} + A^*W^{-1}A\right]^{-1} A^*W^{-1}y, \quad (6)$$

and the corresponding value of the cost function is

$$J(\hat{z}) = y^*[W + A\Pi A^*]^{-1} y. \quad (7)$$

Moreover, this unique point is a minimun if, and only if, the coefficient matrix is positive-definite,

$$(\Pi^{-1} + A^*W^{-1}A) > 0. \quad (8)$$

IV. The Equivalent Estimation Problem

We next focus on problem (2), which we shall refer to as the equivalent estimation problem (or EE, for short). It arises in the following context. Consider column vectors $\{\mathbf{y}, \mathbf{v}, \mathbf{z}\}$ that are linearly related via the expression $\mathbf{y} = A\mathbf{z} + \mathbf{v}$, for some A, and where the individual entries $\{\mathbf{y}_i, \mathbf{v}_i, \mathbf{z}_i\}$ of $\{\mathbf{y}, \mathbf{v}, \mathbf{z}\}$ are all elements of an indefinite metric space, say $\mathcal{K}'$ (note that we are using boldface letters to denote the variables of the EE problem). The variables $\{\mathbf{v}, \mathbf{z}\}$ can be regarded as having Gramian matrices $\{W, \Pi\}$ and cross Gramian zero, $W = <\mathbf{v}, \mathbf{v}>_{\mathcal{K}'}$, $\Pi = <\mathbf{z}, \mathbf{z}>_{\mathcal{K}'}$, and $<\mathbf{z}, \mathbf{v}>_{\mathcal{K}'} = 0$.

Under these conditions, it follows from the linear model that the Gramian matrix of $\mathbf{y}$ is equal to $<\mathbf{y}, \mathbf{y}>_{\mathcal{K}'} = A\Pi A^* + W$. Let $J(K)$ denote the quadratic cost function that appears in (2). It is then immediate to see that $J(K)$ can be interpreted as the Gramian matrix of the vector difference $(\mathbf{z} - K\mathbf{y})$, viz., $J(K) = <\mathbf{z} - K\mathbf{y}, \mathbf{z} - K\mathbf{y}>_{\mathcal{K}'}$. Every K^o at which the gradient of $a^*J(K)a$ with respect to a^*K vanishes for all a is also called a *stationary* solution of $J(K)$. A stationary point K^o may or may not be a minimum.

Hence, solving for the stationary solutions K^o can also be interpreted as solving the problem of linearly estimating $\mathbf{z}$ from $\mathbf{y}$, denoted by $\hat{\mathbf{z}} = K^o\mathbf{y}$. This estimate is uniquely defined if K^o is unique. It is said to be the optimal linear estimate if K^o is the unique minimizing solution.

Theorem 3 *The stationary points* K^o*, if they exist, are solutions of the linear system of equations*

$$\Pi A^* = K^o[A\Pi A^* + W]. \quad (9)$$

There exists a unique stationary point K^o *if, and only if,* $(A\Pi A^* + W)$ *is invertible. In this case, it is given by*

$$K^o = \left[\Pi^{-1} + A^*W^{-1}A\right]^{-1} A^*W^{-1}, \quad (10)$$

and the corresponding value of the cost function is

$$J(K^o) = \left[\Pi^{-1} + A^*W^{-1}A\right]^{-1}. \quad (11)$$

The unique linear estimate of the corresponding $\mathbf{z}$ *is*

$$\hat{\mathbf{z}} = \left[\Pi^{-1} + A^*W^{-1}A\right]^{-1} A^*W^{-1}\mathbf{y}. \quad (12)$$

Moreover, this unique point K^o *is a minimum (and, correspondingly,* $\hat{\mathbf{z}}$ *is optimal) if, and only if, the coefficient matrix is positive-definite,* $(A\Pi A^* + W) > 0$.

V. Relations between the IWLS and EE Problems

Comparing expressions (6) and (12) we see that if we make the identifications: $\hat{\mathbf{z}} \leftrightarrow \hat{z}$ and $\mathbf{y} \leftrightarrow y$, then both expressions coincide. This means that the IWLS problem and the equivalent estimation problem have the same expressions for the stationary points, $\hat{z}$ and $\hat{\mathbf{z}}$. But while a minimum for the IWLS problem (1) exists as long as $(\Pi^{-1} + A^*W^{-1}A) > 0$, the equivalent problem (2), on the other hand, has a minimum at K^o if, and only if, $(W + A\Pi A^*) > 0$.

This indicates that both problems are not generally guaranteed to have simultaneous minima. In the special case of positive-definite matrices $\{\Pi, W\}$, both conditions

$$(\Pi^{-1} + A^*W^{-1}A) > 0 \quad \text{and} \quad (W + A\Pi A^*) > 0,$$

are simultaneously met. But this situation does not hold for general indefinite matrices Π and W. A question of interest then is the following: given that one problem has a unique stationary solution, say the EE problem, and given that this solution has been computed, is it possible to verify whether the other problem, say the IWLS problem (1), admits a minimizing solution without explicitly checking for its positivity condition $(\Pi^{-1} + A^*W^{-1}A) > 0$? The answer is positive and the next two conclusions clarify this issue.

Lemma 3 *The IWLS problem (1) has a unique stationary point* $\hat{z}$ *if, and only if, the equivalent estimation problem (2) has a unique stationary point* K^o.

Theorem 4 *Given invertible and Hermitian matrices* Π *and* W*, and an arbitrary matrix* A *of appropriate dimensions, the IWLS problem (1) has a unique minimizing solution* $\hat{z}$ *if, and only if,*

$$\begin{aligned} I_-[W + A\Pi A^*] &= I_-[\Pi \oplus W], \\ I_+[W + A\Pi A^*] &= I_+[\Pi \oplus W] - n, \end{aligned}$$

where $n \times n$ *is the size of* Π.

The importance of the above theorem is that it allows us to check whether a minimizing solution exists to the IWLS problem (1) by comparing the inertia of the Gramian matrix of the equivalent problem, viz., $(W + A\Pi A^*)$, with the inertia of $(\Pi \oplus W)$. This is relevant because, as we shall see in the next section, when state-space structure is further imposed, we can derive an efficient procedure that allows us to keep track of the inertia of $(W + A\Pi A^*)$. In particular, the procedure will produce a sequence of matrices $\{R_{e,i}\}$ such that

$$\text{Inertia}\,(W + A\Pi A^*) = \text{Inertia}\,(R_{e,0} \oplus R_{e,1} \oplus R_{e,2} \ldots).$$

The theorem then shows that "all" we need to do is compare the inertia of the given matrices Π and W with that of the matrices $\{R_{e,i}\}$ that are made available via the recursive procedure. Equally important is that this procedure will further allow us to compute the quantity $\hat{\mathbf{z}}$. But since we argued above that $\hat{\mathbf{z}}$ has the same expression as $\hat{z}$, the stationary solution of (1), then the procedure will also provide $\hat{z}$.

VI. Incorporating State-Space Structure

Now that we have established the exact relationship between the two basic optimization problems (1) and (2), we shall proceed to study an important special case of the equivalent estimation problem (2).

More specifically, we shall pose an optimization problem that will be of the same form as (2) except that the associated A matrix will have considerable structure in it. In particular, the A matrix will be block-lower triangular and its individual entries will be further parameterized in terms of matrices $\{F_i, G_i, H_i\}$ that arise from an underlying state-space assumption.

We consider vectors $\{\mathbf{y}_i, \mathbf{x}_i, \mathbf{u}_i, \mathbf{v}_i\}$, all with entries in $\mathcal{K}'$, and assume that they are related via the state-space equations

$$\begin{aligned} \mathbf{x}_{i+1} &= F_i\mathbf{x}_i + G_i\mathbf{u}_i, \\ \mathbf{y}_i &= H_i\mathbf{x}_i + \mathbf{v}_i\,, \quad i \geq 0, \end{aligned} \tag{13}$$

where F_i, H_i, and G_i are known $n \times n$, $p \times n$, and $n \times m$ matrices, respectively. It is further assumed that

$$< \begin{bmatrix} \mathbf{u}_i \\ \mathbf{v}_i \\ \mathbf{x}_0 \end{bmatrix}, \begin{bmatrix} \mathbf{u}_j \\ \mathbf{v}_j \\ \mathbf{x}_0 \end{bmatrix} >_{\mathcal{K}'} = \begin{bmatrix} Q_i\delta_{ij} & 0 & 0 \\ 0 & R_i\delta_{ij} & 0 \\ 0 & 0 & \Pi_0 \end{bmatrix},$$

where δ_{ij} is the Kronecker delta function that is equal to unity when $i = j$ and zero otherwise. The matrices $\{Q_i, R_i, \Pi_0\}$ are possibly indefinite.

The state-space structure (13) leads to a linear relation between the vectors $\{\mathbf{y}_i\}$ and the vectors $\{\mathbf{x}_0, \mathbf{u}_i\}_{i=0}^{N-1}$. Indeed, if we collect the $\{\mathbf{y}_i\}_{i=0}^{N}$ and the $\{\mathbf{v}_i\}_{i=0}^{N}$ into two column vectors, $\mathbf{y} = \text{col}\{\mathbf{y}_0, \ldots, \mathbf{y}_N\}$ and $\mathbf{v} = \text{col}\{\mathbf{v}_0, \ldots, \mathbf{v}_N\}$, and define $\mathbf{z}_N = \text{col}\{\mathbf{x}_0, \mathbf{u}_0, \ldots, \mathbf{u}_{N-1}\} = \text{col}\{\mathbf{x}_0, \mathbf{u}\}$, it then follows from the state-space equations that $\mathbf{y} = A\mathbf{z}_N + \mathbf{v}$, where A is the block-lower triangular matrix

$$A \triangleq \begin{bmatrix} H_0 & & & \\ H_1F^{[0,0]} & H_1G_0 & & \\ H_2F^{[1,0]} & H_2F^{[1,1]}G_0 & & \\ \vdots & \vdots & \ddots & \\ H_NF^{[N-1,0]} & H_NF^{[N-1,1]}G_0 & \ldots & H_NG_{N-1} \end{bmatrix}.$$

Here, the notation $F^{[i,j]}$, $i \geq j$, stands for $F_iF_{i-1}\ldots F_j$. Moreover, the Gramian matrices of the variables $\{\mathbf{z}_N, \mathbf{v}\}$ so defined are easily seen to be

$$< \mathbf{z}_N, \mathbf{z}_N >_{\mathcal{K}'} = (\Pi_0 \oplus Q_0 \ldots \oplus Q_{N-1}), \tag{14}$$

$$< \mathbf{v}, \mathbf{v} >_{\mathcal{K}'} = (R_0 \oplus R_1 \oplus \ldots \oplus R_N). \tag{15}$$

More compactly, we shall write $< \mathbf{z}_N, \mathbf{z}_N >_{\mathcal{K}'} = \Pi$ and $< \mathbf{v}, \mathbf{v} >_{\mathcal{K}'} = W$ where the $\{\Pi, W\}$ are the block diagonal matrices in (14) and (15).

We can now pose the problem of estimating $\mathbf{z}_N$ from the variables $\{\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_N\}$, as explained prior to the statement of Theorem 3. This is equivalent to a problem of the form (2). The solution is denoted by $\hat{\mathbf{z}}_{N|N}$ and may be globally expressed, in the unique case, as (*cf.* (12))

$$\hat{\mathbf{z}}_{N|N} = \left[\Pi^{-1} + A^*W^{-1}A\right]^{-1} A^*W^{-1}\mathbf{y}. \tag{16}$$

We are, however, interested in a recursive construction of the estimate $\hat{\mathbf{z}}_{N|N}$, namely one that allows us to update $\hat{\mathbf{z}}_{N|i-1}$ to $\hat{\mathbf{z}}_{N|i}$, for $i = 0, 1, \ldots, N$. Here, the notation $\hat{\mathbf{z}}_{N|i}$ denotes the linear estimate of $\mathbf{z}_N$ that is based on the data up to time i, $\{\mathbf{y}_0, \ldots, \mathbf{y}_i\}$.

Let R_y denote the Gramian matrix of the vector $\mathbf{y}$, $R_y = < \mathbf{y}, \mathbf{y} >_{\mathcal{K}'} = W + A\Pi A^*$. We have shown in [HSK93] 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 5 *Assume R_y is (block) strongly regular and start with $\hat{\mathbf{z}}_{N|-1} = 0$. Then, for $i = 0, 1, \ldots, N$,*

$$\hat{\mathbf{z}}_{N|i} = \hat{\mathbf{z}}_{N|i-1} + K_{z,i}H_i^*R_{e,i}^{-1}\mathbf{e}_i,$$

where $K_{z,0} = \begin{bmatrix} \Pi_0 \\ 0 \end{bmatrix}$ and

$$K_{z,i+1} = K_{z,i}\left[F_i - K_{p,i}H_i\right]^* + \begin{bmatrix} 0 \\ I \\ 0 \end{bmatrix} Q_iG_i^*.$$

The identity matrix in the recursion for $K_{z,i+1}$ occurs at the position that corresponds to the entry $\mathbf{u}_i$. Moreover, the $\{\mathbf{e}_i\}$ are computed via the following Kalman-type procedure: start with $\hat{\mathbf{x}}_{0|-1} = 0$, $P_0 = \Pi_0$, and repeat for $i \geq 0$:

$$\begin{aligned} \mathbf{e}_i &= \mathbf{y}_i - H_i\hat{\mathbf{x}}_{i|i-1}, \quad \hat{\mathbf{x}}_{i+1|i} = F_i\hat{\mathbf{x}}_{i|i-1} + K_{p,i}\mathbf{e}_i, \\ K_{p,i} &= F_iP_iH_i^*R_{e,i}^{-1}, \quad R_{e,i} = R_i + H_iP_iH_i^*, \\ P_{i+1} &= F_iP_iF_i^* + G_iQ_iG_i^* - K_{p,i}R_{e,i}K_{p,i}^*. \end{aligned} \tag{17}$$

Two remarks are due here. First, the successive $\hat{\mathbf{z}}_{N|i}$ have the form $\hat{\mathbf{z}}_{N|i} = \text{col}\{\hat{\mathbf{x}}_{0|i}, \hat{\mathbf{u}}_{0|i}, \ldots, \hat{\mathbf{u}}_{i-1|i}, 0, \ldots, 0\}$, where $\hat{\mathbf{x}}_{0|i}$ denotes the linear estimate of $\mathbf{x}_0$ that is based on $\{\mathbf{y}_0, \ldots, \mathbf{y}_i\}$. Likewise, $\hat{\mathbf{u}}_{j|i}$ denotes the linear estimate of $\mathbf{u}_j$ that is based on the same vectors $\{\mathbf{y}_0, \ldots, \mathbf{y}_i\}$. We shall therefore write $\hat{\mathbf{z}}_{N|i} = \text{col}\{\hat{\mathbf{z}}_{i|i}, 0, \ldots, 0\}$, where $\mathbf{z}_i = \text{col}\{\mathbf{x}_0, \mathbf{u}_0, \ldots, \mathbf{u}_{i-1}\}$.

Secondly, an important fall out of the above algorithm is that the inertia of the Gramian matrix $< \mathbf{y}, \mathbf{y} >_{\mathcal{K}'}$ is completely determined by the $\{R_{e,i}\}$:

$$\text{Inertia of } (W + A\Pi A^*) = \text{Inertia of } \left(R_{e,0} \oplus R_{e,1} \oplus \ldots \oplus R_{e,N}\right).$$

In summary, by establishing an explicit relation between both problems (1) and (2), we are capable of solving either problem via the solution of the other. In the special case of positive-definite quadratic cost functions, this point of view was exploited in [SK94, Say92] in order to establish a close link between known results in Kalman filtering theory and more recent results in adaptive filtering theory. In particular, it was shown in [SK94] 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.

The discussion in this paper, while it provides a similar connection for indefinite quadratic cost functions, it shows that a satisfactory link can be established via an additional set of inertia conditions.

VII. A Recursive IWLS Problem in State-Space Form

Now, in view of the discussion in Sec. V, the solution $\hat{\mathbf{z}}_{N|N}$ has the same expression as the solution $\hat{z}_{N|N}$ of a related minimization problem of the form (1). Indeed, it is rather immediate to write down the IWLS problem whose stationary point matches the above $\hat{\mathbf{z}}_{N|N}$:

$$\min_{z_N=\begin{bmatrix} x_0 \\ u \end{bmatrix}} \left\{ \begin{bmatrix} x_0 \\ u \end{bmatrix}^* \Pi^{-1} \begin{bmatrix} x_0 \\ u \end{bmatrix} + \left(y - A\begin{bmatrix} x_0 \\ u \end{bmatrix}\right)^* W^{-1} \left(y - A\begin{bmatrix} x_0 \\ u \end{bmatrix}\right)\right\}.$$

Equivalently, using the state-equations, this can be written as

$$\min_{\{x_0,u_0,\ldots,u_{N-1}\}} \left[x_0^*\Pi_0^{-1}x_0 + \sum_{j=0}^{N}(y_j - H_jx_j)^*R_j^{-1}(y_j - H_jx_j) + \sum_{j=0}^{N-1} u_j^*Q_j^{-1}u_j \right]$$

subject to $x_{j+1} = F_jx_j + G_ju_j$. Likewise, the IWLS problem whose stationary solution $\hat{z}_{i|i}$ matches the $\hat{\mathbf{z}}_{i|i}$ is

$$\min_{\{x_0,u_0,\ldots,u_{i-1}\}} \left[x_0^*\Pi_0^{-1}x_0 + \sum_{j=0}^{i}(y_j - H_jx_j)^*R_j^{-1}(y_j - H_jx_j) + \sum_{j=0}^{i-1} u_j^*Q_j^{-1}u_j \right]$$

subject to $x_{j+1} = F_jx_j + G_ju_j$. It is now immediate to verify that, in fact, the strong regularity assumption that we imposed earlier on the Gramian matrix R_y is not a restriction; it is a necessary requirement if we want to guarantee the existence of all the stationary solutions $\{\hat{z}_{i|i}\}$. We shall denote the above cost function that determines $\hat{z}_{i|i}$ by $J_i(x_0, u_0, \ldots, u_{i-1})$.

The following results follow as a consequence of the inertia statements of Sec. II.

Lemma 4 *Let $m \times m$ denote the size of each Q_i. Likewise, let $n \times n$ denote the size of Π_0. Define*

$$\Pi \triangleq (\Pi_0 \oplus Q_0 \ldots \oplus Q_{N-1}), \quad W \triangleq (R_0 \oplus R_1 \oplus \ldots \oplus R_N).$$

Assume $(W + A\Pi A^*)$ is (block) strongly regular (i.e., the J_i are guaranteed to have unique stationary points $\hat{z}_{i|i}$ for all $0 \le i \le N$). Then J_N has a minimum with respect to $\{x_0, u_0, \ldots, u_{N-1}\}$ (i.e., the last stationary point $\hat{z}_{N|N}$ is a minimum) if, and only if,

$$\begin{aligned} I_-[\Pi \oplus W] &= I_-\{R_{e,0} \oplus \ldots \oplus R_{e,N}\}, \\ I_+[\Pi \oplus W] &= I_+\{R_{e,0} \oplus \ldots \oplus R_{e,N}\} + n + mN, \end{aligned}$$

where the matrices $\{R_{e,i}\}$ are recursively computed as follows: $R_{e,i} = H_iP_iH_i^* + R_i$, $K_{p,i} = F_iP_iH_i^*R_{e,i}^{-1}$, and

$$P_{i+1} = F_iP_iF_i^* + G_iQ_iG_i^* - K_{p,i}R_{e,i}K_{p,i}^*, \quad P_0 = \Pi_0.$$

An immediate conclusion is the following special case where the Π matrix is itself positive-definite.

Corollary 1 *Consider the same setting of Lemma 4. Assume further that $\Pi_0 > 0$ and the $\{Q_i\}_{i=0}^{N-1}$ are positive-definite. Then J_N has a minimum with respect to z_N if, and only if,*

$$\begin{aligned} I_-\{R_0 \oplus \ldots \oplus R_N\} &= I_-\{R_{e,0} \oplus \ldots \oplus R_{e,N}\}, \\ I_+\{R_0 \oplus \ldots \oplus R_N\} &= I_+\{R_{e,0} \oplus \ldots \oplus R_{e,N}\}. \end{aligned}$$

The above results were concerned with the existence of a minimum for the last cost function J_N. More generally, we are interested in checking whether each $\hat{z}_{i|i}$ is a minimum of the corresponding J_i. This is addressed in the following statement.

Theorem 6 *Each J_i, for $0 \le i \le N$, has a minimum if, and only if,*

$$\begin{aligned} I_-[\Pi_0 \oplus R_0] &= I_-\{R_{e,0}\}, && (18) \\ I_+[\Pi_0 \oplus R_0] &= I_+\{R_{e,0}\} + n, && (19) \end{aligned}$$

and, for $i = 1, 2, \ldots, N$,

$$\begin{aligned} I_-\{Q_{i-1} \oplus R_i\} &= I_-\{R_{e,i}\}, && (20) \\ I_+\{Q_{i-1} \oplus R_i\} &= I_+\{R_{e,i}\} + m. && (21) \end{aligned}$$

Moreover, when the stationary solutions (or minima) of the J_i are uniquely defined, the value of each J_i at its unique stationary solution (or minimum) $\hat{z}_{i|i}$ is given by $J_i(\hat{z}_{i|i}) = \sum_{j=0}^{i} e_i^*R_{e,i}^{-1}e_i$, where $e_i = (y_i - H_i\hat{x}_{i|i-1})$.

It is also clear from the discussions in Sec. V. that the recursions of Theorem 5, with the proper identifications $\hat{\mathbf{z}}_{N|i} \leftrightarrow \hat{z}_{N|i}$, $\mathbf{y}_i \leftrightarrow y_i$, $\hat{\mathbf{x}}_{i|i-1} \leftrightarrow \hat{x}_{i|i-1}$, $\mathbf{u}_i \leftrightarrow u_i$, can be used to compute the stationary solutions $\{\hat{z}_{i|i}\}$ of the $\{J_i\}$. In particular, we also have that the stationary solutions $\hat{z}_{i|i}$ are related to the $\hat{z}_{N|i}$, given below in the statement of the theorem, as follows: $\hat{z}_{N|i} = \text{col}\{\hat{z}_{i|i}, 0, \ldots, 0\}$. That is, the leading entries of $\hat{z}_{N|i}$ denote the stationary solution of J_i with respect to $\{x_0, u_0, \ldots, u_{i-1}\}$.

Theorem 7 *The stationary solutions $\{\hat{z}_{i|i}\}$ of the $\{J_i\}$ can be recursively computed as follows: start with $\hat{z}_{N|-1} = 0$ and repeat for $i = 0, 1, \ldots, N$:*

$$\hat{z}_{N|i} = \hat{z}_{N|i-1} + K_{z,i}H_i^*R_{e,i}^{-1}\,(y_i - H_i\hat{x}_{i|i-1}),$$

where

$$K_{z,i+1} = K_{z,i}\left[F_i - K_iR_{e,i}^{-1}H_i\right]^* + \begin{bmatrix} 0 \\ I \\ 0 \end{bmatrix} Q_iG_i^*,$$

and $\hat{x}_{i+1|i} = F_i\hat{x}_{i|i-1} + K_{p,i}(y_i - H_i\hat{x}_{i|i-1})$, $\hat{x}_{0|-1} = 0$.

Remark. It may happen that the last term in the definition of the quadratic cost function J_i also includes the extra tem $u_i^*Q_i^{-1}u_i$, say

$$\left[x_0^*\Pi_0^{-1}x_0 + \sum_{j=0}^{i}(y_j - H_jx_j)^*R_j^{-1}(y_j - H_jx_j) + \sum_{j=0}^{i} u_j^*Q_j^{-1}u_j \right].$$

In this case, the unknown variable u_i only appears in the quadratic term $u_i^*Q_i^{-1}u_i$, and it thus follows that minimization with respect to the u_i requires the positivity of Q_i. Hence, successive minimization of the J_i would additionally require that the $\{Q_i\}$ be positive-definite, which is a special case that often arises in the context of H^∞-problems, with the additional constraint $\Pi_0 > 0$. It is thus rather immediate to handle this case. All we need to do is to simply impose a positivity condition on the $\{Q_i\}$. This motivates us to consider the following two corollaries.

Corollary 2 *Assume further that the $\{Q_i\}_{i=0}^{N-1}$ are positive-definite and that $\Pi_0 > 0$. Then each J_i has a minimum with respect to $\{x_0, u_0, \ldots, u_{i-1}\}$ if, and only if, for all i,*

$$\text{Inertia}\{R_i\} = \text{Inertia}\{R_{e,i}\}. \qquad (22)$$

In this case, it follows that $P_i \ge 0$ for $0 \le i \le N$.

The next statement further assumes that the $\{F_i\}$ are invertible.

Corollary 3 *Consider the same setting as in Corollary 2 and assume further that the $\{F_i\}$ are invertible. Then the following two statements provide equivalent necessary and sufficient conditions for each J_i to have a minimum with respect to $\{x_0, u_0, \ldots, u_{i-1}\}$.*

(i) All $\{J_i\}$ have minima iff, for $0 \leq i \leq N$,

$$P_i^{-1} + H_i^* R_i^{-1} H_i > 0. \tag{23}$$

(ii) All $\{J_i\}$ have minima iff, for $0 \leq i \leq N$,

$$P_{i+1} - G_i Q_i G_i^* > 0\,. \tag{24}$$

It follows in the minimum case that, for all i, $P_{i+1} > 0$.

Conditions of the form (23) are the ones most cited in H^∞–applications (see, e.g., [YS91] and the next section). Here we see that they are related to the inertia conditions (22) and, more generally, to the conditions of Theorem 6. The inertia conditions (22) also arise in the H^∞–context (see, e.g., [GL95][p. 495] and the next section), where R_i has the additional structure $R_i = (-\gamma^2 I \oplus I)$. Here, we have derived these conditions as special cases of the general statement of Theorem 6, which holds for arbitrary indefinite matrices $\{\Pi_0, Q_i, R_i\}$, while the H^∞–results hold only for positive-definite matrices $\{\Pi_0, Q_i\}$ and for matrices R_i of the above form. Note also that testing for (23) not only requires that we compute the P_i (via a Riccati recursion (17)), but also that we invert P_i and R_i at each step and then check for the positivity of $P_i^{-1} + H_i^* R_i^{-1} H_i$. The inertia tests given by (22), on the other hand, employ the quantities $R_{e,i}$ and R_i, which are $p \times p$ matrices (as opposed to P_i which is $n \times n$). These tests can be used as the basis for alternative computational variants that are based on square-root ideas, as pursued in [HSK94] for the case of H^∞–filters.

VIII. An Application to H^∞-Filtering

We now illustrate the applicability of the earlier results to a particular problem in H^∞-filtering. For this purpose, we consider a state-space model of the form

$$x_{i+1} = F_i x_i + G_i u_i\,, \quad y_i = H_i x_i + v_i\,, \tag{25}$$

where $\{x_0, u_i, v_i\}$ are unknown deterministic signals and $\{y_i\}_{i=0}^N$ are known (or measured) signals. Let $s_j = L_j x_j$ be a linear transformation of the state-vector x_j, where L_j is a known matrix.

Let $\hat{s}_{j|j}$ denote a function of the $\{y_k\}$ up to and including time j. For every time instant i, we define the quadratic cost function $J_i(x_0, u_0, \ldots, u_i) = x_0^* \Pi_0^{-1} x_0 +$

$$\sum_{j=0}^{i} u_j^* Q_j^{-1} u_j + \sum_{j=0}^{i} v_j^* v_j - \gamma^{-2} \sum_{j=0}^{i} (\hat{s}_{j|j} - L_j x_j)^* (\hat{s}_{j|j} - L_j x_j),$$

where $\{\Pi_0, Q_j\}$ are given positive-definite matrices, and γ is a given positive real number. We would like to determine the existence or not of functions $\{\hat{s}_{0|0}, \hat{s}_{1|1}, \ldots, \hat{s}_{N|N}\}$ that would guarantee $J_i > 0$ for $0 \leq i \leq N$.

The expression for J_i can be rewritten in the equivalent form

$$J_i = x_0^* \Pi_0^{-1} x_0 + \sum_{j=0}^{i} u_j^* Q_j^{-1} u_j +$$

$$\sum_{j=0}^{i} \left(\begin{bmatrix} \hat{s}_{j|j} \\ y_j \end{bmatrix} - \begin{bmatrix} L_j \\ H_j \end{bmatrix} x_j \right)^* R_j^{-1} \left(\begin{bmatrix} \hat{s}_{j|j} \\ y_j \end{bmatrix} - \begin{bmatrix} L_j \\ H_j \end{bmatrix} x_j \right),$$

where $R_j = (-\gamma^2 I \oplus I)$. This is a quadratic cost function in the unknowns $\{x_0, u_0, \ldots, u_i\}$. Therefore, each J_i will be positive if, and only if, it has a minimum with respect to $\{x_0, u_0, \ldots, u_i\}$ and, moreover, the value of J_i at its minimum is positive.

We then conclude from Corollary 2, and according to the remark after Theorem 7, that each J_i will admit a minimizing solution if, and only if, the corresponding $R_{e,i}$ and R_i have the same inertia. In the present context, we have $R_i = (-\gamma^2 I \oplus I)$ and

$$R_{e,i} = \begin{bmatrix} -\gamma^2 I & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} L_i \\ H_i \end{bmatrix} P_i \begin{bmatrix} L_i \\ H_i \end{bmatrix}^*,$$

where P_i satisfies the Riccati difference equation

$$P_{i+1} \;=\; F_i \left[P_i^{-1} + H_i^* H_i - \gamma^{-2} L_i^* L_i \right]^{-1} F_i^* + G_i Q_i G_i^*.$$

Equivalently, we require

$$\begin{aligned} I + H_i P_i H_i^* &> 0, \\ (-\gamma^2 I + L_i P_i L_i^*) - L_i P_i H_i^* (I + H_i P_i H_i^*)^{-1} H_i P_i L_i^* &< 0. \end{aligned}$$

If the F_i are further assumed invertible, then we also conclude from Corollary 3 that the following alternative conditions can be used to guarantee the existence of minima for the J_i:

$$P_i^{-1} + H_i^* H_i - \gamma^{-2} L_i^* L_i > 0, \ \text{ for } \ 0 \leq i \leq N. \tag{26}$$

We may proceed and show how to determine estimates $\hat{s}_{j|j}$ once the existence of minima for the J_i are guaranteed. These estimates have to be chosen so as to guarantee that the values of the successive J_i at their minima are positive. We shall omit the details here and only state the final recursions:

$$\hat{s}_{i|i} = L_i \left[\hat{x}_{i|i-1} + P_i H_i^* (I + H_i P_i H_i^*)^{-1} (y_i - H_i \hat{x}_{i|i-1}) \right],$$

where $\hat{x}_{i|i-1}$ is constructed recursively via $\hat{x}_{0|-1} = 0$,

$$\hat{x}_{i+1|i} = F_i \left[\hat{x}_{i|i-1} + P_i H_i^* (I + H_i P_i H_i^*)^{-1} (y_i - H_i \hat{x}_{i|i-1}) \right].$$

References

[GL95] M. Green and D. J. N. Limebeer. *Linear Robust Control.* Prentice Hall, NJ, 1995.

[GLR83] I. Gohberg, P. Lancaster, and L. Rodman. *Matrices and Indefinite Scalar Products.* Birkhäuser Verlag, Basel, 1983.

[GV83] G. Golub and C. F. Van Loan. *Matrix Computations.* The Johns Hopkins University Press, second edition, 1989.

[HSK93] B. Hassibi, A. H. Sayed, and T. Kailath. Recursive linear estimation in Krein spaces - Part I: Theory. In *Proc. Conference on Decision and Control*, vol. 4, pages 3489–3494, San Antonio, Texas, December 1993.

[HSK94] B. Hassibi, A. H. Sayed, and T. Kailath. Square-root arrays and Chandrasekhar recursions for H^∞ problems. In *Proc. Conference on Decision and Control*, Orlando, FL, December 1994.

[SK94] A. H. Sayed and T. Kailath. A state-space approach to adaptive RLS filtering. *IEEE Signal Processing Magazine*, 11(3):18–60, July 1994.

[Say92] A. H. Sayed. *Displacement Structure in Signal Processing and Mathematics.* PhD thesis, Stanford University, August 1992.

[YS91] I. Yaesh and U. Shaked. H^∞–optimal estimation: The discrete time case. In *Proc. Inter. Symp. on MTNS*, pages 261–267, Kobe, Japan, June 1991.