Proceedings of the 32nd Conference
on Decision and Control
San Antonio, Texas • December 1993

RECURSIVE LINEAR ESTIMATION IN KREIN SPACES - PART I: THEORY *

Babak Hassibi, Ali H. Sayed† and Thomas Kailath

Information Systems Laboratory, Stanford University, CA 94305, USA.

ABSTRACT

We develop a self-contained theory for linear estimation in Krein spaces. The theory is based on simple concepts such as projections and matrix factorizations, and leads to an interesting connection between Krein space projection and the computation of the stationary points of certain second order (or quadratic) forms. We use the innovations process to obtain a rather general recursive linear estimation algorithm, which when specialized to a state space model yields a Krein space generalization of the celebrated Kalman filter with applications in several areas such as H^{∞}-filtering and control, game problems, risk sensitive control, and adaptive filtering.

I. INTRODUCTION

We have recently shown that H^{∞}−estimation and control, and several related problems such as risk-sensitive estimation and control, finite memory adaptive filtering, and others, can be studied in a simple and unified way by relating them to Kalman filtering problems not in the usual (stochastic) Hilbert space but in a special kind of indefinite metric space known as a Krein space. Though the two types of spaces share many characteristics, they differ in special ways that turn out to mark the difference between the LQG or H^2 theories and the more recent H^{∞} theories. In this paper we develop a self-contained theory for linear estimation in Krein spaces. The ensuing theory is much more rich than the conventional Hilbert space case, which is why it yields a unified approach to the above mentioned problems. We shall, for brevity, omit several details and most of the proofs, which will be published elsewhere. For alternative points of view and for results on H^{∞}−filtering and related questions, the reader may consult [1]-[6], and the references therein.

The remainder of the paper is organized as follows. We introduce Krein spaces and define the corresponding notion of projection in Section I.1. Contrary to the Hilbert space case where projections are always unique, the Krein space projection is unique (unambiguous) iff a certain Gramian matrix is nonsingular. In Section II, we first remark that while quadratic forms in Hilbert space always have minima (or maxima), in Krein spaces one can only assert that they will always have stationary points. Further conditions will have to be met for these to be minima or maxima. We explore this by first considering the problem of finding a vector k to *stationarize* a quadratic form $< \mathbf{z}-k^*\mathbf{Y}, \mathbf{z}-k^*\mathbf{Y} >_{\mathcal{K}}$ where $< .,. >_{\mathcal{K}}$ is an indefinite inner product and $*$ denotes conjugate transpose. $\mathbf{Y}$ is a collection of vectors in the Krein space and $\mathbf{z}$ is a vector outside the linear space spanned by the $\mathbf{Y}$. If the Gramian matrix $\mathbf{R_Y} =< \mathbf{Y}, \mathbf{Y} >_{\mathcal{K}}$ is nonsingular, then there is a unique stationary $k^*\mathbf{Y}$, which is given by the projection of $\mathbf{z}$ onto the linear space spanned by the $\mathbf{Y}$; the stationary point will be a minimum iff $\mathbf{R_Y}$ is strictly positive definite as well. In a Hilbert space of course, the nonsingularity of $\mathbf{R_Y}$ and its strict positive definiteness are equivalent properties, but this is not true with $\mathbf{Y}$ in a Krein space.

Now in the Hilbert space theory it is well known that a certain deterministic quadratic form (arising from a Bayesian approach to the problem) is also minimized by the same element k^*Y. In the Krein space case, k^*Y also yields a stationary point of the corresponding deterministic quadratic form, but now this point will be a minimum iff $\mathbf{R_z} - \mathbf{R_{zY}}\mathbf{R_Y}^{-1}\mathbf{R_{Yz}}$ is positive definite, where the $\mathbf{R_z}$,$\mathbf{R_{zY}}$, and $\mathbf{R_{Yz}}$ are the usual Gramians and cross Gramians of $\mathbf{z}$ and $\mathbf{Y}$. In Krein space, unlike Hilbert space, the conditions $\mathbf{R_Y} > 0$ and $\mathbf{R_z} - \mathbf{R_{zY}}\mathbf{R_Y}^{-1}\mathbf{R_{Yz}} > 0$ need not hold simulta-

†Dr. Sayed is currently with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106, USA. This work was performed while he was a Research Associate in the Information Systems Laboratory, Stanford University, and on Leave from Escola Politécnica da Universidade de São Paulo, Brazil.

* 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.

neously.

This simple distinction turns out to be crucial in understanding the difference between H^2 and H^∞ estimation as we show in the companion papers [7, 8]. In the present paper, we continue with the general theory, by exploring the consequences of assuming that $\{\mathbf{z}, \mathbf{Y}\}$ are based on some underlying state-space model. This will then lead us to recursions of the same form as those of the celebrated Kalman filter, except that in Krein space the *Riccati* variable P_i and the innovations Gramian $R_{e,i}$ are not necessarily positive (semi)definite. Moreover, the distinctions between the results for the stochastic and deterministic quadratic forms stand out quite clearly. These results are developed in Sections II and III.

I.1 Krein Spaces and Projections

A Krein space $\{\mathcal{K}, <.,.>_{\mathcal{K}}\}$ is a vector space that satisfies the following: (i) $\mathcal{K}$ is linear over $\mathcal{C}$, the complex numbers, (ii) $\mathcal{K}$ possesses a bilinear form $<.,.>_{\mathcal{K}}$ such that for any $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{K}$, and for any $a, b \in \mathcal{C}$, we have $<\mathbf{y}, \mathbf{x}>_{\mathcal{K}} = <\mathbf{x}, \mathbf{y}>_{\mathcal{K}}^*$ and $<a\mathbf{x} + b\mathbf{y}, \mathbf{z}>_{\mathcal{K}} = a <\mathbf{x}, \mathbf{z}>_{\mathcal{K}} + b <\mathbf{y}, \mathbf{z}>_{\mathcal{K}}$, and (iii) $\mathcal{K}$ admits a direct orthogonal sum decomposition $\mathcal{K} = \mathcal{K}_+ \oplus \mathcal{K}_-$ such that $\{\mathcal{K}_+, <.,.>_{\mathcal{K}}\}$ and $\{\mathcal{K}_-, - <.,.>_{\mathcal{K}}\}$ are Hilbert spaces, and $<\mathbf{x}, \mathbf{y}>_{\mathcal{K}} = 0$ for any $\mathbf{x} \in \mathcal{K}_+$ and $\mathbf{y} \in \mathcal{K}_-$. The symbol $*$ denotes complex conjugation. Also, elements in a Krein space will be denoted by bold face letters.

Given $\mathbf{z} \in \mathcal{K}$ and the elements $\{\mathbf{y_0}, \mathbf{y_1}, \ldots, \mathbf{y_N}\}$ also in $\mathcal{K}$, we define $\hat{\mathbf{z}}$ to be the projection of $\mathbf{z}$ onto the linear space spanned by $\{\mathbf{y_0}, \mathbf{y_1}, \ldots, \mathbf{y_N}\}$ iff $\mathbf{z} = \hat{\mathbf{z}} + \tilde{\mathbf{z}}$ where $\hat{\mathbf{z}} \in span\{\mathbf{y}_0, \ldots, \mathbf{y}_N\}$ and $\tilde{\mathbf{z}}$ satisfies the orthogonality condition $<\tilde{\mathbf{z}}, \mathbf{y}_i>_{\mathcal{K}} = 0$ for $i = 0, 1, \ldots, N$. It is straightforward to verify that in a Hilbert space setting, projections always exist and are unique, whereas in the general Krein space setting, $\hat{\mathbf{z}}$ will exist and be unique iff the Gramian matrix, $\mathbf{R_Y} = <\mathbf{Y}, \mathbf{Y}>_{\mathcal{K}}$, is nonsingular. In this case, $\hat{\mathbf{z}}$ is given by $\hat{\mathbf{z}} = <\mathbf{z}, \mathbf{Y}>_{\mathcal{K}} <\mathbf{Y}, \mathbf{Y}>_{\mathcal{K}}^{-1} \mathbf{Y} = \mathbf{R_{zY}} \mathbf{R_Y^{-1}} \mathbf{Y}$.

II. QUADRATIC FORMS

It turns out that projections determine the stationary point of certain second-order forms. To clarify this point, consider an arbitrary linear combination $\mathbf{z_e}(k) = k^* \mathbf{Y}$ of the observation vectors $\{\mathbf{y_i}\}$. A natural object to study is the error Gramian $E(k) = <\mathbf{z} - \mathbf{z_e}(k), \mathbf{z} - \mathbf{z_e}(k)>_{\mathcal{K}} = \mathbf{R_z} - \mathbf{R_{zY}} k - k^* \mathbf{R_{Yz}} + k^* \mathbf{R_Y} k$. In a Hilbert space, where the inner product is positive definite, $E(k)$ is a norm and is a measure of the *closeness* of $\mathbf{z}$ and $\mathbf{z_e}(k)$. That interpretation does not extend to Krein spaces, where a minimum of $E(k)$ does not necessarily exist. However, a stationary point is always guaranteed to exist. The matrix k_o is said to be the stationary point of $E(k)$ iff $k_o a$ is a stationary point of the scalar second-order form $a^* E(k) a$ for all complex column vectors a, viz., $\partial(a^* E(k) a)/\partial(ka) = 0$ at $k = k_o$. Moreover, the stationary point of $E(k)$ is a minimum iff for all k we have $\partial^2(a^* E(k) a)/\partial(ka)^2 > 0$.

Calculating the stationary point of $E(k)$, and the corresponding condition for a minimum, is now straightforward and leads to our first result on the significance of the Krein space projection of $\mathbf{z}$ on the span of $\mathbf{Y}$. If $\mathbf{R_Y}$ is nonsingular, then the projection $\hat{\mathbf{z}}$ defined in the previous section is equal to $k_o^* \mathbf{Y}$, where $k_o = \mathbf{R_Y^{-1}} \mathbf{R_{Yz}}$ is the stationary point of $E(k)$. Moreover, $E(k_o) = \mathbf{R_z} - \mathbf{R_{zY}} \mathbf{R_Y^{-1}} \mathbf{R_{Yz}}$ is a minimum iff $\mathbf{R_Y} > 0$.

We refer to this as a vector, or stochastic, interpretation of the notion of projection since $\{\mathbf{y_0}, \ldots, \mathbf{y_N}\}$ and $\mathbf{z}$ can be viewed as random quantities in the special case of a Hilbert space of stochastic variables. We can however consider a second interpretation for the projection, which we shall refer to as deterministic, because it involves computing the stationary point of a certain *scalar* second order form. To this end, consider the following scalar second order form

$$J(z, Y) \equiv \begin{bmatrix} z^* & Y^* \end{bmatrix} \begin{bmatrix} \mathbf{R_z} & \mathbf{R_{zY}} \\ \mathbf{R_{Yz}} & \mathbf{R_Y} \end{bmatrix}^{-1} \begin{bmatrix} z \\ Y \end{bmatrix},$$

where z and Y are no longer bold face, meaning that they are to be regarded as (ordinary) vectors of complex numbers, and the quantities $\mathbf{R_z}, \mathbf{R_{zY}}, \mathbf{R_{Yz}}, \mathbf{R_Y}$ are as before. We can also verify that if $\mathbf{R_Y}$ is nonsingular, then the stationary point z_o of $J(z, Y)$ is given by $z_o = \mathbf{R_{zY}} \mathbf{R_Y^{-1}} Y$, with $J(z_o, Y) = Y^* \mathbf{R_Y^{-1}} Y$, and this point is a minimum iff $\mathbf{R_z} - \mathbf{R_{zY}} \mathbf{R_Y^{-1}} \mathbf{R_{Yz}} > \mathbf{0}$.

This last result implies that the stationary point z_o is given by the exact same expression as the projection $\hat{\mathbf{z}}$ defined above. It is important to clearly understand what is shown here. The stationary point z_0 of the scalar quadratic form $J(z, Y)$ is given by a formula similar to that for the projection of a Krein space vector $\mathbf{z}$ on the linear span of Krein space vectors $\mathbf{Y}$. However, z and Y are just vectors in Euclidean space, with no Krein space involved, and z_0 is not the "projection" of z on the vector Y. What the above result shows is that by properly defining the scalar quadratic form $J(z, Y)$ using coefficient matrices $\mathbf{R_z}, \mathbf{R_Y}, \mathbf{R_{zY}}, \mathbf{R_{Yz}}$ that are arbitrary, but that can be regarded as being obtained from Gramians and cross-Gramians of some vectors $\mathbf{z}, \mathbf{Y}$ in Krein space, we can get a result similar to that of minimizing a matrix quadratic forms via projections in Krein spaces. This distinction does not happen in the Hilbert space setting since the conditions $\mathbf{R_z} - \mathbf{R_{zY}} \mathbf{R_Y^{-1}} \mathbf{R_{Yz}} > 0$ and $\mathbf{R_Y} > 0$ occur simultaneously.

III. STATE-SPACE STRUCTURE

We now assume that the components $\{\mathbf{y}_j\}$ of $\mathbf{Y}$ arise from an underlying state-space model: $\mathbf{x}_{j+1} = F_j \mathbf{x}_j + G_j \mathbf{u}_j$, $\mathbf{y}_j = H_j \mathbf{x}_j + \mathbf{v}_j$, where $\mathbf{x}_0$, $\{\mathbf{u_i}\}$, and $\{\mathbf{v_i}\}$ are assumed to be uncorrelated elements in a Krein space $\mathcal{K}$ and such that $< \mathbf{u_i}, \mathbf{u_j} >_{\mathcal{K}} = Q_i \delta_{ij}$, $< \mathbf{v_i}, \mathbf{v_j} >_{\mathcal{K}} = R_i \delta_{ij}$, $< \mathbf{x_0}, \mathbf{x_0} >_{\mathcal{K}} = \Pi_0$, where δ_{ij} denotes the Kronecker delta function. Let $h_{jk} = H_j F_{j-1} \ldots F_{k+1} G_k$ be the response at time j to an impulse at time $k < j$ (assuming $\mathbf{x}_0 = 0$ and $\mathbf{v}_k = 0$ for all k), and define

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y_0} \\ \vdots \\ \mathbf{y_N} \end{bmatrix}, \mathbf{U} = \begin{bmatrix} \mathbf{u_0} \\ \vdots \\ \mathbf{u_N} \end{bmatrix}, \mathbf{V} = \begin{bmatrix} \mathbf{v_0} \\ \vdots \\ \mathbf{v_N} \end{bmatrix}.$$

Then $\mathbf{Y} = \mathcal{O}\mathbf{x_0} + \mathbf{\Gamma U} + \mathbf{V}$, where

$$\mathcal{O} = \begin{bmatrix} H_0 \\ H_1 F_0 \\ \vdots \\ H_N F_{N-1} \ldots F_0 \end{bmatrix}, \quad \mathbf{\Gamma} = \begin{bmatrix} 0 & & & \\ h_{10} & 0 & & \\ h_{20} & h_{21} & 0 & \\ & & & \ddots \end{bmatrix}.$$

In state-space estimation problems, one may regard as fundamental quantities the projections of $\mathbf{x_0}$ and $\mathbf{U}$ onto the space spanned by $\mathbf{Y}$, which we shall denote by $\hat{\mathbf{x}}_{\mathbf{0}|\mathbf{N}}$ and $\hat{\mathbf{U}}_{|\mathbf{N}}$, respectively. Note that the span of $\{\mathbf{x_0}, \mathbf{U}, \mathbf{V}\}$ is equivalent to the span of $\{\mathbf{x_0}, \mathbf{U}, \mathbf{Y}\}$ and hence, due to linearity, the projection of any other quantity of interest on $\mathbf{Y}$ can be obtained as a linear combination of the $\hat{\mathbf{x}}_{\mathbf{0}|\mathbf{N}}$, $\hat{\mathbf{U}}_{|\mathbf{N}}$ and $\mathbf{Y}$. Also, the Gramian of $\mathbf{Y}$ has the form $\mathbf{R_Y} = \mathcal{O}\Pi_0\mathcal{O}^* + \mathbf{\Gamma Q \Gamma^*} + \mathbf{R}$, where we have defined $\mathbf{Q} = Q_0 \oplus Q_1 \oplus \ldots \oplus Q_N$ and $\mathbf{R} = R_0 \oplus R_1 \oplus \ldots \oplus R_N$. We are then led to the following stochastic interpretation.

Lemma 1 *Suppose $\mathbf{x_0}$, $\mathbf{U}$, and $\mathbf{Y}$ are related through a state space model as above, and that $\mathbf{R_Y}$ is nonsingular. Then the stationary point of the error Gramian*

$$< \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{U} \end{bmatrix} - \begin{bmatrix} \mathbf{x_0^e} \\ \mathbf{U^e} \end{bmatrix}, \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{U} \end{bmatrix} - \begin{bmatrix} \mathbf{x_0^e} \\ \mathbf{U^e} \end{bmatrix} >_{\mathcal{K}}$$

over all $\begin{bmatrix} \mathbf{x_0^e} \\ \mathbf{U^e} \end{bmatrix}$ in the span of $\mathbf{Y}$ is given by

$$\begin{bmatrix} \hat{\mathbf{x}}_{\mathbf{0}|\mathbf{N}} \\ \hat{\mathbf{U}}_{|\mathbf{N}} \end{bmatrix} = \begin{bmatrix} \Pi_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{bmatrix} \begin{bmatrix} \mathcal{O}^* \\ \mathbf{\Gamma}^* \end{bmatrix} \mathbf{R_Y}^{-1}\mathbf{Y},$$

which is the projection of $\begin{bmatrix} \mathbf{x_0} \\ \mathbf{U} \end{bmatrix}$ on the span of $\mathbf{Y}$. Moreover, this stationary point is a minimum if, and only if, $\mathbf{R_Y} > 0$.

Now let us consider the (deterministic) scalar form analog of Lemma 1. We shall first identify the corresponding scalar quadratic form. For this, we recall that the projection of $\mathbf{x_0}$ and $\mathbf{U}$ onto $\mathbf{Y}$ also yields the stationary point (over x_0 and U) of the following second order form,

$$\begin{bmatrix} x_0^* & U^* & Y^* \end{bmatrix} \begin{bmatrix} \mathbf{R_{x_0}} & \mathbf{R_{x_0 U}} & \mathbf{R_{x_0 Y}} \\ \mathbf{R_{U x_0}} & \mathbf{R_U} & \mathbf{R_{UY}} \\ \mathbf{R_{Y x_0}} & \mathbf{R_{YU}} & \mathbf{R_Y} \end{bmatrix}^{-1} \begin{bmatrix} x_0 \\ U \\ Y \end{bmatrix},$$

where we assume that the (ordinary) vectors x_0, U, and Y obey the (state-space) constraints: $x_{j+1} = F_j x_j + G_j u_j$ and $y_j = H_j x_j + v_j$, and where U and Y are defined accordingly. To compute the actual coefficient matrix in the above expression, it will be convenient to make a change of variables through the following easily verifiable relation:

$$\begin{bmatrix} \mathbf{x_0} \\ \mathbf{U} \\ \mathbf{Y} \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ \mathcal{O} & \mathbf{\Gamma} & I \end{bmatrix} \begin{bmatrix} \mathbf{x_0} \\ \mathbf{U} \\ \mathbf{V} \end{bmatrix},$$

which implies that the above quadratic form is equal to

$$\begin{bmatrix} x_0^* & U^* & V^* \end{bmatrix} \begin{bmatrix} \Pi_0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R} \end{bmatrix}^{-1} \begin{bmatrix} x_0 \\ U \\ V \end{bmatrix}.$$

We are then led to the following deterministic interpretation.

Lemma 2 *The element*

$$\begin{bmatrix} \hat{x}_{0|N} \\ \hat{U}_{|N} \end{bmatrix} = \begin{bmatrix} \Pi_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{bmatrix} \begin{bmatrix} \mathcal{O}^* \\ \mathbf{\Gamma}^* \end{bmatrix} \mathbf{R_Y}^{-1} Y$$

is the stationary point of the following second order form:

$$\begin{aligned} J_N &= x_0^* \Pi_0^{-1} x_0 + \sum_{j=0}^{N} u_j^* Q_j^{-1} u_j \\ &\quad + \sum_{j=0}^{N} (y_j - H_j x_j)^* R_j^{-1} (y_j - H_j x_j). \end{aligned}$$

The value of J_N at the stationary point is $Y^* \mathbf{R_Y^{-1}} Y$.

III.1 Conditions for a Minimum

It remains to further explore the effect of the state-space assumption on the condition that the above stationary point is actually a minimum. We recall that the condition is that the Gramian matrix

$$\mathbf{M} = \mathbf{R_z} - \mathbf{R_{zY}}\mathbf{R_Y}^{-1}\mathbf{R_{Yz}} \text{ be positive,}$$

where $\mathbf{z} = \begin{bmatrix} \mathbf{x_0^T} & \mathbf{U^T} \end{bmatrix}^{\mathbf{T}}$. The assumption of state-space structure gives

$$\mathbf{M} = \begin{bmatrix} \Pi_0 - \Pi_0\mathcal{O}^*\mathbf{R_Y}^{-1}\mathcal{O}\Pi_0 & -\Pi_0\mathcal{O}^*\mathbf{R_Y}^{-1}\mathbf{\Gamma Q} \\ -\mathbf{Q\Gamma}^*\mathbf{R_Y}^{-1}\mathcal{O}\Pi_0 & \mathbf{Q} - \mathbf{Q\Gamma}^*\mathbf{R_Y}^{-1}\mathbf{\Gamma} \end{bmatrix}.$$

One characterization of the fact that $\mathbf{M} > 0$ is that the $(1,1)$ block entry and its Schur complement must be positive definite. The $(1,1)$ is by definition the norm of the error $\mathbf{x_0} - \hat{\mathbf{x}}_{\mathbf{0|N}}$, *i.e.*,

$$\Pi_0 - \Pi_0\mathcal{O}^*\mathbf{R_Y}^{-1}\mathcal{O}\Pi_0 = < \mathbf{x_0} - \hat{\mathbf{x}}_{\mathbf{0|N}}, \mathbf{x_0} - \hat{\mathbf{x}}_{\mathbf{0|N}} >_{\mathcal{K}} = P_{0|N}$$

So one conclusion is $P_{0|N} > 0$. As for the Schur complement of the $(1,1)$ block entry of $\mathbf{M}$, say $\mathbf{\Delta}$, it can be shown that $\mathbf{\Delta^{-1}} = \mathbf{Q}^{-1} + \mathbf{\Gamma}^*\mathbf{R}^{-1}\mathbf{\Gamma}$, so that we have the following result.

Lemma 3 *A necessary and sufficient condition for the stationary point of Lemma 2 to be a minimum is that* $P_{0|N} > 0$ *and* $\mathbf{Q}^{-1} + \mathbf{\Gamma}^*\mathbf{R^{-1}}\mathbf{\Gamma} > 0$.

Now Lemma 3 gives just one set of conditions for testing the positivity of $\mathbf{M}$. Many others can be obtained by various congruence transformations on $\mathbf{M}$. We see that Lemma 3 involves the *smoothing error* $P_{0|N}$. In ordinary least-squares theory, more often computed is the one-step prediction error (the innovations variance),

$$P_{N+1} = < \mathbf{x}_{N+1} - \hat{\mathbf{x}}_{N+1}, \mathbf{x}_{N+1} - \hat{\mathbf{x}}_{N+1} >_{\mathcal{K}},$$

where $\hat{\mathbf{x}}_{N+1}$ denotes the estimate of $\mathbf{x}_{N+1}$ based on $\{\mathbf{y}_0, \ldots, \mathbf{y}_N\}$. Exploration of such facts leads us to the result of Lemma 4, which will turn out to be more useful than Lemma 3, for reasons that will appear later. This does not mean, however, that other sets of conditions may not be as or even more useful in certain circumstances.

Lemma 4 *Assuming that the* $\{F_j\}$ *are invertible, then a necessary and sufficient condition for the stationary point of Lemma 2 to be a minimum is that* $P_{N+1} > 0$ *and* $\mathbf{Q}^{-1} + \mathbf{\Gamma}^*\mathbf{R^{-1}}\mathbf{\Gamma} - \mathcal{C}^*P_{N+1}^{-1}\mathcal{C} > 0$, *where*

$$\mathcal{C} = \begin{bmatrix} F_NF_{N-1}\ldots F_1G_0 & F_NF_{N-1}\ldots F_2G_1 & \ldots & G_N \end{bmatrix}.$$

IV. RECURSIVE FORMULAS

The key consequence of state-space structure in Hilbert space is that the computational burden of finding the estimates and the error variances can be significantly reduced by using the celebrated Kalman filter recursions for the quantities $\{\hat{\mathbf{x}}_{i+1}, P_i\}$. We shall see that similar recursions hold in Krein space as well, provided we have the additional assumption that $\mathbf{R_Y}$ is strongly nonsingular (or strongly regular), in the sense that all its leading minors are nonzero. From the properties of the Krein space projection described earlier, recursive projection onto the $\{\mathbf{y}_j\}$ is possible only if all leading submatrices of $\mathbf{R_Y}$ are nonsingular. The condition of strong regularity is well known to imply that $\mathbf{R_Y}$ has a unique triangular decomposition $\mathbf{R_Y} = \mathbf{LDL}^*$, where $\mathbf{D}$ is diagonal and $\mathbf{L}$ is lower triangular with unit diagonal. A very useful geometric insight into this factorization is that it is an easy consequence of a Gram-Schmidt orthogonalization on the vectors $\{\mathbf{y}_j\}$. Namely, if we define the innovations $\mathbf{e}_j = \mathbf{y}_j - \hat{\mathbf{y}}_j$, where $\hat{\mathbf{y}}_j \equiv \hat{\mathbf{y}}_{j|j-1}$ = the projection of $\mathbf{y}_j$ onto the linear space spanned by $\{\mathbf{y}_0, \ldots, \mathbf{y}_{j-1}\}$, then the uniqueness of the triangular decomposition implies that

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y_0} \\ \vdots \\ \mathbf{y_N} \end{bmatrix} = \mathbf{L} \begin{bmatrix} \mathbf{e_0} \\ \vdots \\ \mathbf{e_N} \end{bmatrix} = \mathbf{LE}$$

so that $\mathbf{D} = < \mathbf{E}, \mathbf{E} >_{\mathcal{K}} = \mathbf{R_e}$.

We should also point out that the value at the stationary point of J_N in terms of the innovations is given by $Y^*\mathbf{R_Y}^{-1}Y = \sum_{j=0}^{N} e_j^* R_{e,j}^{-1} e_j$. Now the state-space structure allows us to compute the innovations recursively and efficiently. Moreover, once the innovations have been found, many other estimates such as $\hat{\mathbf{x}}_{\mathbf{0|N}}$ and $\hat{\mathbf{U}}_{\mathbf{|N}}$ can also be readily computed, as described at the end of this section. The derivation leads to an extension of the Kalman filter algorithm to Krein spaces.

Theorem 1 *Consider the Krein-space state equations*

$$\begin{aligned} \mathbf{x}_{i+1} &= F_i\mathbf{x}_i + G_i\mathbf{u}_i, \qquad 0 \le i \le N \\ \mathbf{y}_i &= H_i\mathbf{x}_i + \mathbf{v}_i \end{aligned}$$

with

$$< \begin{bmatrix} \mathbf{u}_j \\ \mathbf{v}_j \\ \mathbf{x}_0 \end{bmatrix}, \begin{bmatrix} \mathbf{u}_k \\ \mathbf{v}_k \\ \mathbf{x}_0 \end{bmatrix} >_{\mathcal{K}} = \begin{bmatrix} Q_j\delta_{jk} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & R_j\delta_{jk} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \Pi_0 \end{bmatrix}.$$

Assume that $\mathbf{R_Y} = [< \mathbf{y}_i, \mathbf{y}_j >_{\mathcal{K}}]$ *is strongly regular. Then the innovations can be computed via the* *formulas*

$$\begin{aligned}
\mathbf{e}_i &= \mathbf{y}_i - H_i\hat{\mathbf{x}}_i, \qquad 0 \le i \le N \\
\hat{\mathbf{x}}_{i+1} &= F_i\hat{\mathbf{x}}_i + K_{p,i}(\mathbf{y}_i - H_i\hat{\mathbf{x}}_i), \qquad \hat{\mathbf{x}}_0 = 0 \\
K_{p,i} &= F_iP_iH_i^*R_{e,i}^{-1}
\end{aligned}$$

where $R_{e,i} =< \mathbf{e}_i, \mathbf{e}_i >_{\mathcal{K}} = R_i + H_iP_iH_i^*$, *and the* $\{P_i\}$ *can be recursively computed via the (Riccati) recursions:* $P_0 = \Pi_0$,

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

We note that $\hat{\mathbf{x}}_{i+1}$ is the projection of $\mathbf{x}_{i+1}$ on the linear span of $\{\mathbf{y}_0, \ldots, \mathbf{y}_i\}$, and that $P_i =< \mathbf{x}_i - \hat{\mathbf{x}}_i, \mathbf{x}_i - \hat{\mathbf{x}}_i >_{\mathcal{K}}$. The derivation of the above algorithm follows the usual pattern as in the Kalman filter theory (see, *e.g.*, [9]) and is therefore not repeated here. The only remark to be made is that unlike the usual theory, none of the matrices $\{Q_i, R_i, R_{e,i}, P_i\}$ need be positive, and that the invertibility of $R_{e,i}$ follows form the assumption that $\mathbf{R_Y}$ is strongly regular. Moreover, in Kalman filter theory there are many variations of the above formulas such as the measurement update formulas, the time update formulas, the information filter form, and the square-root forms. These will be discussed elsewhere. Furthermore, for constant systems, or in fact for systems where the time-variation is structured in a certain way, the Riccati recursions and the square-root recursions, both of which take $O(n^3)$ elementary computations (flops) per iteration (where n is the dimension of the state-space), can be replaced by the more efficient Chandrasekhar recursions, which require only $O(n^2)$ flops per iteration [10, 11]. We shall not present these square-root and Chandrasekhar equations here, because, despite their formal similarity with the Hilbert space case, these recursions will only exist if the projection performs a certain minimization. In this case even though the $R_{e,i}$ are not necessarily non-negative definite, it turns out that their inertia has a certain property that allows generalized square-roots to be defined, and appropriate square-root and Chandrasekhar recursions to be devised. This will be done elsewhere.

Before closing this section we note that it is not difficult to verify that the innovations can also be used to recursively compute the projections $\hat{\mathbf{x}}_{\mathbf{0}|\mathbf{N}}$ and $\hat{\mathbf{U}}_{|\mathbf{N}}$ as follows:

$$\begin{aligned}
\hat{\mathbf{x}}_{\mathbf{0}|\mathbf{i}} &= \sum_{k=0}^{i} < \mathbf{x_0}, \tilde{\mathbf{x}}_\mathbf{k} >_{\mathcal{K}} H_k^*R_{e,k}^{-1}\mathbf{e_k} \\
\hat{\mathbf{u}}_{\mathbf{j}|\mathbf{i}} &= \sum_{k=0}^{i} < \mathbf{u_j}, \tilde{\mathbf{x}}_\mathbf{k} >_{\mathcal{K}} H_k^*R_{e,k}^{-1}\mathbf{e_k}
\end{aligned}$$

where the inner products satisfy the recursive formulas

$$< \mathbf{x_0}, \tilde{\mathbf{x}}_{\mathbf{i+1}} >_{\mathcal{K}} = (F_i - K_{p,i}H_i) < \mathbf{x_0}, \tilde{\mathbf{x}}_\mathbf{i} >_{\mathcal{K}} ,$$

$$< \mathbf{u_j}, \tilde{\mathbf{x}}_{\mathbf{i+1}} >_{\mathcal{K}} = (F_i - K_{p,i}H_i) < \mathbf{u_j}, \tilde{\mathbf{x}}_\mathbf{i} >_{\mathcal{K}} ,$$

with $< \mathbf{x_0}, \tilde{\mathbf{x}}_\mathbf{0} >_{\mathcal{K}} = \Pi_0$, $< \mathbf{u_j}, \tilde{\mathbf{x}}_{\mathbf{j+1}} >_{\mathcal{K}} = Q_jG_j^*$, $K_{p,i} = F_iP_{i|i}H_i^*R_i^{-1}$, and $P_{i|i}^{-1} = P_i^{-1} + H_i^*R_i^{-1}H_i$.

IV.1 Recursive Estimation and Second-Order Forms

The scalar quadratic form associated with projecting $\mathbf{x_0}$ and $\mathbf{U}$ onto $\mathbf{Y}$ has already been identified in Lemma 2. We shall presently study the condition for the existence of a minimum in more detail, and present a recursive procedure for testing this condition.

Recall from Lemma 4 that if the $\{F_j\}$ are nonsingular then the condition for a minimum for the scalar quadratic form can be expressed as $P_{N+1} > 0$ and

$$\mathbf{Q}^{-1} + \mathbf{\Gamma}^*\mathbf{R}^{-1}\mathbf{\Gamma} - \mathcal{C}^*P_{N+1}^{-1}\mathcal{C} > 0. \qquad (1)$$

In order to identify the above condition, the following result will be useful.

Lemma 5 (Complementary Model) *Consider the backwards state-space model*

$$\begin{aligned}
\mathbf{x_j^c} &= F_j^*\mathbf{x_{j+1}^c} + H_j^*\mathbf{u_j^c}, \qquad N \ge j \ge 0 \\
\mathbf{y_j^c} &= G_j^*\mathbf{x_{j+1}^c} + \mathbf{v_j^c}
\end{aligned} \qquad (2)$$

with $< \mathbf{u}_j^c, \mathbf{u}_k^c >_{\mathcal{K}} = R_j^{-1}\delta_{jk}$, $< \mathbf{v}_j^c, \mathbf{v}_k^c >_{\mathcal{K}} = Q_j^{-1}\delta_{jk}$, *and* $< \mathbf{x}_{N+1}^c, \mathbf{x}_{N+1}^c >_{\mathcal{K}} = -P_{N+1}^{-1}$, *and define the corresponding column vector* $\mathbf{Y}^c = \begin{bmatrix} \mathbf{y}_0^{c\mathbf{T}} & \ldots & \mathbf{y}_N^{c\mathbf{T}} \end{bmatrix}^\mathbf{T}$. *Then* $\mathbf{R_{Yc}} = \mathbf{Q}^{-1} + \mathbf{\Gamma}^*\mathbf{R}^{-1}\mathbf{\Gamma} - \mathcal{C}^*P_{N+1}^{-1}\mathcal{C}$.

There are a variety of ways to motivate this state-space model (for example via duality), however, we shall pursue this line of thought elsewhere. Now that we have identified the matrix in (1) as the Gramian of a state-space model, we can recursively test condition (1) by checking the positivity of the innovations of this model. This is obtained by developing the corresponding Krein Kalman filter for the backwards model.

Lemma 6 *If the* $\{F_j\}$ *are nonsingular, then the Gramian of the innovations of the complementary state-space model is given by:*

$$Q_j^{-1} - G_j^*P_{j+1}^{-1}G_j, \qquad N \ge j \ge 0 ,$$

where P_j satisfies the usual Riccati recursion:

$$P_{j+1} = F_j P_j F_j^* - K_{p,j} R_{e,j} K_{p,j}^* + G_j Q_j G_j^*$$

with $K_{p,j} = F_j P_j R_{e,j}^{-1}$, $R_{e,j} = R_j + H_j P_j H_j^*$ *and* $P_0 = \Pi_0$.

The condition for a minimum can now be written as $P_{N+1} > 0$ and $Q_j^{-1} - G_j^* P_{j+1}^{-1} G_j > 0$ for $j = 0, \ldots, N$. But it can be verified that the matrices $(P_{i+1} \oplus Q_i^{-1} - G_i^* P_{i+1}^{-1} G_i)$ and $((P_i^{-1} + H_i^* R_i^{-1} H_i)^{-1} \oplus Q_i^{-1})$ are congruent. Hence, we obtain the following result, which is one of the most important in this presentation, and will be of use in a variety of applications.

Theorem 2 *If $\mathbf{R_Y}$ is strongly regular, then the Krein space Kalman filter*

$$\hat{x}_{i+1} = F_i \hat{x}_i + K_{p,i} e_i, \qquad \hat{x}_0 = 0,$$

with $K_{p,i} = F_i P_i H_i^* R_{e,i}^{-1}$, $R_{e,i} = R_i + H_i P_i H_i^*$, $e_i = y_i - \hat{y}_i = y_i - H_i \hat{x}_i$, *and P_i satisfying the Riccati recursion*

$$P_{i+1} = F_i P_i F_i^* + G_i Q_i G_i^* - K_{p,i} R_{e,i} K_{p,i}^* \ , \ \ P_0 = \Pi_0,$$

recursively computes the stationary point of the following second-order form

$$\begin{aligned} J_i &= x_0^* \Pi_0^{-1} x_0 + \sum_{j=0}^{i} u_j^* Q_j^{-1} u_j \\ &\quad + \sum_{j=0}^{i} (y_j - H_j x_j)^* R_j^{-1} (y_j - H_j x_j) \end{aligned}$$

over $\{x_j\}$ and $\{u_j\}$, subject to the state-space constraint $x_{j+1} = F_j x_j + G_j u_j$. The value of J_i at the stationary point is equal to $\sum_{j=0}^{i} e_j^* R_{e,j}^{-1} e_j$. If the $\{F_j\}_{j=0}^{i}$ are nonsingular, then the stationary point will correspond to a minimum iff, $Q_j > 0$ and

$$P_{j|j}^{-1} = P_j^{-1} + H_j^* R_j^{-1} H_j > 0 \; for \; j = 0, 1, \ldots, i$$

It also follows in the minimum case that $P_{j+1} > 0$ for $j = 0, 1, \ldots, i$.

IV. CONCLUDING REMARKS

Briefly, the major conclusion is that given a deterministic quadratic form in Krein space, (to which H^∞ problems lead almost by inspection), one can relate them to a corresponding stochastic problem for which the Kalman filter solution can be written down immediately; moreover, the condition for a minimum can also be expressed in terms of quantities easily related to the basic Riccati equations of the Kalman filter. Several applications are described in the companion papers [7, 8].

REFERENCES

[1] G. Tadmor, "H^∞ in the time domain: The standard problem," In *American Control Conference*, pp. 772-773, 1989.

[2] P.P. Khargonekar and K. M. Nagpal, "Filtering and smoothing in an H^∞– setting," *IEEE Trans. on Automatic Control*, vol. AC-36, pp. 151–166, 1991

[3] T. Basar and P. Bernhard, *H^∞-Optimal Control and Related Minimax Design Problems - A Dynamic Game Approach*, Birkhauser, Boston, Berlin, Basel, 1991.

[4] D. Limebeer, B.D.O. Anderson, P.P. Khargonekar, and M. Green, "A game theoretic approach to H^∞ control for time varying systems," *SIAM Journal on Control and Optimization*, vol. 30, pp. 262-283, 1992.

[5] U. Shaked and Y. Theodor, "H^∞–optimal estimation: A tutorial," In *Proc. IEEE Conference on Decision and Control*, pp. 2278–2286, Tucson, AZ, Dec. 1992.

[6] M. J. Grimble, "Polynomial matrix solution of the H^∞ filtering problem and the relationship to Riccati equation state-space results," *IEEE Trans. on Signal Processing*, vol. SP-41, no. 1, pp. 67–81, January 1993.

[7] B. Hassibi, A. H. Sayed, and T. Kailath, "Recursive linear estimation in Krein spaces - Part II: Applications," *in Proc. of this Conference.*

[8] B. Hassibi, A. H. Sayed, and T. Kailath, "LMS is H^∞–optimal," *in Proc. of this Conference.*

[9] T. Kailath, *Lecture Notes in Wiener and Kalman Filtering.* Springer-Verlag, Berlin, 1980.

[10] M. Morf, G. S. Sidhu, and T. Kailath, "Some new algorithms for recursive estimation in constant, linear, discrete-time systems," *IEEE Transactions on Automatic Control*, vol. AC-19, pp. 315-323, 1974.

[11] A. H. Sayed and T. Kailath, "Extended Chandrasekhar recursions," To appear in *IEEE Transactions on Automatic Control*, 1994.