In *Computing and Intelligent Systems*, S. Keerthi, Y. Narhari and N. Viswanadham, eds., Tata McGraw-Hill Publishing Co. Ltd., pp. 7–15, New Delhi, 1993.

$\mathcal{H}^\infty$ Filtering is Just Kalman Filtering in Krein Space

Thomas Kailath, Babak Hassibi, and Ali H. Sayed
Information Systems Laboratory,
Stanford University, Stanford, CA 94305

Abstract

We have recently developed a self-contained theory for linear estimation in Krein spaces. Here we show that the problem of H^∞ filtering is equivalent to Kalman filtering in Krein space. This observation further elucidates the connections between H^2 and H^∞ filtering and allows one to extend many known results on conventional Kalman filtering to the H^∞ case.

1 INTRODUCTION

Classical results in linear least-squares estimation and Kalman filtering are based on an H^2 criterion and require apriori knowledge of the statistical properties of the noise signals. In some applications, however, one is faced with model uncertainties and lack of statistical information on the exogenous signals, which led to an increasing interest in minimax estimation (see, *e.g.* [1, 2, 3] and the references therein), with the belief that the resulting so-called H^∞ algorithms will be more robust and less sensitive to parameter variations.

The H^∞ filters obtained in this fashion involve propagating a Riccati variable and bear a striking resemblance to the classical Kalman filter. In this paper we show that the H^∞ filters are indeed Kalman filters, provided we set up an appropriate estimation problem with elements not in a Hilbert space, but in an indefinite (so-called) Krein space.

We have recently developed a self-contained theory for linear estimation in Krein spaces [4]. Although Hilbert spaces and Krein spaces share many characteristics, they differ in special ways that turn out to mark the differences between the LQG or H^2 theories and the more recent H^∞ theories. The similarities, however, allow one to apply to the H^∞ setting, many of the results developed for Kalman filtering over the last three decades.

2 The H^∞ Filtering Problem

We shall begin by formulating the H^∞ filtering problem and by introducing its solution. For further details the reader may consult [1, 5] and the references therein.

Consider the time-varying state-space model

$$\begin{aligned} x_{i+1} &= F_i x_i + G_i u_i, \qquad && x_0 \\ y_i &= H_i x_i + v_i, \qquad && i \geq 0 \end{aligned} \tag{1}$$

where the $\{u_j\}$ and $\{v_j\}$ are unknown disturbances, and x_0 is the unknown initial state. Suppose we want to estimate some arbitrary linear combination of the states, *viz.*

$$z_i = L_i x_i$$

We denote the estimate and estimation error respectively as

$$\check{z}_{i|i} = \mathcal{F}(Y_i) \quad , \quad \tilde{z}_{i|i} = \check{z}_{i|i} - L_i x_i$$

where $Y_i = \begin{bmatrix} y_0^T & y_1^T & \ldots & y_i^T \end{bmatrix}^{\mathbf{T}}$. Note that since $\check{z}_{i|i}$ is a *causal* function of the $\{y_j\}$ we shall call such an estimate the *aposteriori* or *filtered* estimate of z_i.

According to Figure 2, Let T denote the transfer operator from the disturbances $\{x_0, u_i, v_i\}$ to the estimation error.

We are interested in choosing $\mathcal{F}$ so as to minimize the H^∞ norm of T where

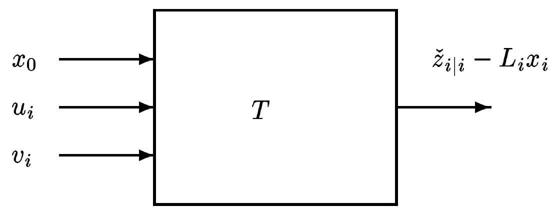

Figure 1: Transfer matrix from disturbances to estimation error

Definition 1 (Time-domain Interpretation) *The H_∞ norm of a transfer operator T is equal to*

$$\|T\|_\infty = \sup_{u \in h_2, u \neq 0} \frac{\|Tu\|_2}{\|u\|_2}$$

where $\|u\|_2$ is the h_2–norm of the causal sequence $\{u_k\}$,

$$\|u\|_2^2 = \sum_{k=0}^{\infty} u_k^* u_k$$

Note that the H^∞ norm has the interpretation of being the maximum energy gain from input to output. Our problem may now be formally stated as follows.

Problem 1 (Optimal H_∞ Problem) *Find a causal H_∞-optimal estimation strategy $\check{z}_{i|i} = \mathcal{F}(Y_i)$ that minimizes $\|T\|_\infty$, and obtain the resulting*

$$\gamma_0 = \inf_{\mathcal{F}} \quad \|T\|_\infty$$

There are very few cases where the above problem can be solved in closed form so that one normally considers the following suboptimal problem.

Problem 2 (Sub-optimal H_∞ Problem) *Given a scalar $\gamma > 0$, find an estimator $\check{z}_{i|i} = \mathcal{F}(Y_i)$ that achieves $\|T\|_\infty < \gamma$. This clearly requires checking whether $\gamma \geq \gamma_0$*

Before giving the solution to the suboptimal H_∞ problem it is illuminating to examine the structure of the problem in slightly more detail. Note that $\|T\|_\infty < \gamma$ implies that for all nonzero $\{x_0, u_i, v_i\}$

$$\frac{\sum_{k=0}^{\infty} |\check{z}_{k|k} - L_k x_k|^2}{x_0^* \Pi_0^{-1} x_0 + \sum_{k=0}^{\infty} |u_k|^2 + \sum_{k=0}^{\infty} |y_k - H_k x_k|^2} < \gamma^2$$

where Π_0 is a positive definite matrix that reflects apriori knowledge of how close x is to zero. Equivalently,

Given a scalar $\gamma > 0$, then $\|T\|_\infty < \gamma$ if, and only if, there exists $\check{z}_{k|k}$ (for all k) such that for all nonzero complex vectors x_0 and for all nonzero causal sequences $\{u_k, v_k\} \in h_2$, the scalar second-order form defined by

$$\begin{aligned} J &= x_0^* \Pi_0^{-1} x_0 + \sum_{k=0}^{\infty} u_k^* u_k \\ &\quad + \sum_{k=0}^{\infty} (y_k - H_k x_k)^* (y_k - H_k x_k) \\ &\quad - \gamma^{-2} \sum_{k=0}^{\infty} (\check{z}_{k|k} - L_k x_k)^* (\check{z}_{k|k} - L_k x_k) \end{aligned}$$

satisfies

$$J > 0.$$

In the finite horizon case we are interested in J_i, where the sums run until i. Note that the *indefinite* quadratic form J_i can be rewritten as follows:

$$\begin{aligned} J_i &= x_0^* \Pi_0^{-1} x_0 + \sum_{k=0}^{i} u_k^* u_k + \\ & \sum_{k=0}^{i} \left(\begin{bmatrix} \check{z}_{k|k} \\ y_k \end{bmatrix} - \begin{bmatrix} L_k \\ H_k \end{bmatrix} x_k \right)^* \\ & \begin{bmatrix} -\gamma^{-2} I & 0 \\ 0 & I \end{bmatrix} \left(\begin{bmatrix} \check{z}_{k|k} \\ y_k \end{bmatrix} - \begin{bmatrix} L_k \\ H_k \end{bmatrix} x_k \right) \end{aligned} \quad (2)$$

Thus the solution to the H^∞ filtering problem is intimately related to guaranteeing the positivity of an indefinite quadratic form.

Let us now quote the solution to the aposteriori H^∞ filtering problem as given in the literature (see *e.g.* [1, 2]).

Theorem 1 *For a given $\gamma > 0$, if the F_j are nonsingular, then $J_i > 0$ if and only if,*

$$P_j^{-1} + H_j^* H_j - \gamma^{-2} L_j^* L_j > 0, \qquad j = 0, \ldots, i \quad (3)$$

where $P_0 = \Pi_0$ and P_i satisfies the Riccati recursion

$$\begin{aligned} P_{i+1} &= F_i P_i F_i^* + G_i G_i^* - F_i P_i \begin{bmatrix} L_i^* & H_i^* \end{bmatrix} \times \\ & \left\{ \begin{bmatrix} -\gamma^2 I & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix} + \begin{bmatrix} L_i \\ H_i \end{bmatrix} P_i \begin{bmatrix} L_i^* & H_i^* \end{bmatrix} \right\}^{-1} \\ & \times \begin{bmatrix} L_i \\ H_i \end{bmatrix} P_i F_i^* \end{aligned}$$

If this is the case, then one possible H^∞ estimator is

$$\check{z}_{i|i} = L_i \hat{x}_{i|i}$$

where

$$\hat{x}_{i+1|i+1} = F_i\hat{x}_{i|i} + K_{f,i}(y_{i+1} - H_{i+1}F_i\hat{x}_{i|i}) \quad (4)$$

with $\hat{x}_{-1|-1} = 0$ *and*

$$K_{f,i} = P_{i+1}H^*_{i+1}(I + H_{i+1}P_{i+1}H^*_{i+1})^{-1} \quad (5)$$

Remark: This looks very much like a Kalman filter solution, except that the Riccati recursion differs from that of the Kalman filter, since

- we have indefinite *covariance* matrices, $\begin{bmatrix} -\gamma^2 I & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix}$
- the L_i (of the quantity to be estimated) enters the Riccati equation
- We have an additional condition (3), that must be satisfied for the filter to exist; in the Kalman filter problem the L_i would not appear, and the P_i would be psoitive definite, so that (3) is immediate.

We shall show that in fact the filter of Theorem 1 is a Kalman filter, but in an indefinite vector space, called a Krein space. The indefinite covariances and the appearance of L_i in the Riccati equation will be easily explained in this framework. The additional condition (3) will be seen to arise from the fact that in Krein space, unlike as in the usual Hilbert space context, there are different conditions for the existence of optimum solutions to, roughly speaking, a stochastic mean-square estimation problem and the deterministic quadratic form minimization problem. In the usual Hilbert space formulation, these two conditions are equivalent (see the discussion below Theorem 3).

3 Linear Estimation in Krein Spaces

In this section we shall review some of the basic properties of linear estimation in Krein spaces. For more details the reader may refer to [4].

3.1 General Theory

Definition 2 (Krein Spaces) *An abstract vector space $\{\mathcal{K}, <.,.>_{\mathcal{K}}\}$ that satisfies the following requirements is called a Krein Space:*

(i) $\mathcal{K}$ is a linear space over $\mathcal{C}$, the complex numbers.

(ii) There exists a bilinear form $<.,.>_{\mathcal{K}}$ such that

(a) $<\mathbf{y},\mathbf{x}>_{\mathcal{K}} = <\mathbf{x},\mathbf{y}>^*_{\mathcal{K}}$

(b) $<a\mathbf{x}+b\mathbf{y},\mathbf{z}>_{\mathcal{K}} = a<\mathbf{x},\mathbf{z}>_{\mathcal{K}} + b<\mathbf{y},\mathbf{z}>_{\mathcal{K}}$

for any $\mathbf{x},\mathbf{y},\mathbf{z} \in \mathcal{K}$, $a,b \in \mathcal{C}$: $*$ denotes complex conjugation.

(iii) The vector space $\mathcal{K}$ 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 $\mathbf{y} \in \mathcal{K}_-$.

In view of the above, depending on the sign of $<\mathbf{x},\mathbf{x}>_{\mathcal{K}}$ a vector $\mathbf{x}$ can be *positive*, *neutral* or *negative*. Correspondingly, a subspace $\mathcal{M} \subset \mathcal{K}$ can be positive, neutral or negative.

Some geometric insight into Krein spaces can be gained by considering the 3-dimensional Minkowski space of Figure 2 with (indefinite) inner product

$$<v_1,v_2>_{\mathcal{K}} = x_1.x_2 + y_1.y_2 - t_1.t_2$$

In this case the x and y axes will span $\mathcal{K}_+$ and the t axis will span $\mathcal{K}_-$. The neutral subspace is given by the cone $x^2+y^2-t^2=0$, with points inside the cone belonging to the negative subspace and points outside the cone belonging to the positive subspace.

An important notion is that of the projection.

Definition 3 (Projections) *Given* $\mathbf{z} = \begin{bmatrix} \mathbf{z}_0 \\ \vdots \\ \mathbf{z}_M \end{bmatrix}$ *where the $\{\mathbf{z}_j\}$ belong to $\mathcal{K}$, and elements $\{\mathbf{y_0},\mathbf{y_1},\ldots,\mathbf{y_N}\}$ in $\mathcal{K}$, $\hat{\mathbf{z}}$ is the projection of $\mathbf{z}$ onto the linear space spanned by $\{\mathbf{y_0},\mathbf{y_1},\ldots,\mathbf{y_N}\}$ if*

$$\mathbf{z} = \hat{\mathbf{z}} + \tilde{\mathbf{z}} \quad (6)$$

where $\hat{\mathbf{z}} \in span\{\mathbf{y}_0,...,\mathbf{y}_N\}$ and $\tilde{\mathbf{z}}$ satisfies the orthogonality condition

$$\tilde{\mathbf{z}} \perp span\{\mathbf{y}_0,...,\mathbf{y}_N\}$$

or equivalently, $<\tilde{\mathbf{z}}_j,\mathbf{y}_i>_{\mathcal{K}} = 0$ for $i=0,1,\ldots,N$ and $j=0,1,\ldots,M$.

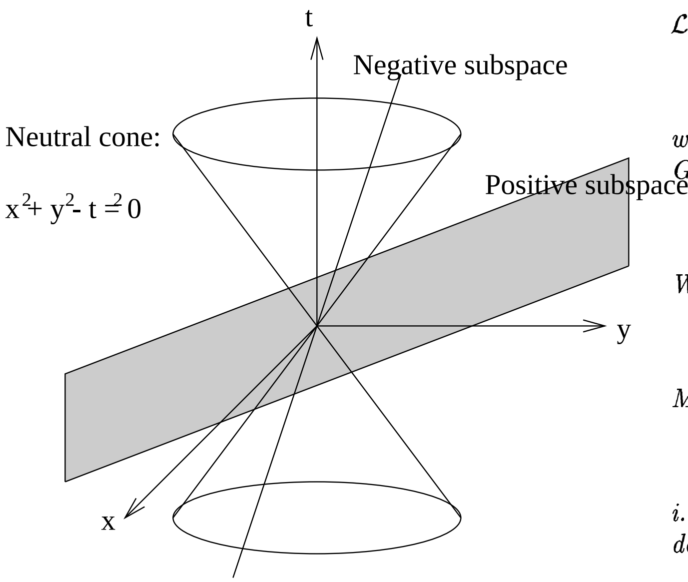

Figure 2: 3-dimensional Minkowski space

In Hilbert space projections always exist and are unique. However, in Krein space this is not always the case. Indeed we have the following result where we have defined the Gramian matrix of

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y_0} \\ \vdots \\ \mathbf{y_N} \end{bmatrix} \quad \text{as} \quad \mathbf{R_Y} = <\mathbf{Y}, \mathbf{Y}>_{\mathcal{K}}$$

with $(i,j)th$ element $<\mathbf{y_i}, \mathbf{y_j}>_{\mathcal{K}}$.

Lemma 1 (Uniqueness) *In the Hilbert space setting, projections always exist and are unique. However, in the general Krein space setting, the projection of* $\mathbf{z} \in \mathcal{K}$ *onto the linear space spanned by* $\{\mathbf{y_0}, \ldots, \mathbf{y_N}\}$ *will exist and be unique if, and only if, the Gramian matrix,* $\mathbf{R_Y} = <\mathbf{Y}, \mathbf{Y}>_{\mathcal{K}}$, *is nonsingular. In this case,* $\hat{\mathbf{z}}$ *is given by*

$$\hat{\mathbf{z}} = <\mathbf{z}, \mathbf{Y}>_{\mathcal{K}} <\mathbf{Y}, \mathbf{Y}>_{\mathcal{K}}^{-1} \ \mathbf{Y} = \mathbf{R_{zY}} \mathbf{R_Y^{-1}} \mathbf{Y} \quad (7)$$

In Hilbert space, projections always minimize the norm of a certain error Gramian. In Krein space we can only assert that projections define a stationary point. Whether the corresponding value is a minimum or not depends on further conditions being met. The precise statement follows.

Theorem 2 (Stochastic Stationary Point) *If* $\mathbf{R_Y}$ *is nonsingular, then the projection of* $\mathbf{z}$ *on* $\mathcal{L}(\mathbf{Y})$ *is* $k_o^* \mathbf{Y}$, *where*

$$\hat{\mathbf{z}} = k_o^* \mathbf{Y} \quad , \quad k_o = \mathbf{R_Y^{-1}} \mathbf{R_{Yz}}$$

where k_o *is the stationary point of the error Gramian*

$$E(k) \ \equiv <\mathbf{z} - \mathbf{k}^* \mathbf{Y}, \mathbf{z} - \mathbf{k}^* \mathbf{Y}>_{\mathcal{K}} .$$

We also have

$$E(k_o) = \mathbf{R_z} - \mathbf{R_{zY}} \mathbf{R_Y^{-1}} \mathbf{R_{Yz}}.$$

Moreover, $E(k_o)$ *is a minimum iff,*

$$\mathbf{R_Y} > 0.$$

i.e., $\mathbf{R_Y}$ *is not only nonsingular but also positive definite.*

By a stationary point we mean that for all complex vectors a

$$\left. \frac{\partial a^* E(k) a}{\partial k a} \right|_{k=k_o} = 0$$

Consider now 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 we assume the inverse exists, and where $\{z, Y\}$ are ordinary vectors of complex numbers in Euclidean space.

Theorem 3 (Deterministic Stationary Point) *If* $\mathbf{R_Y}$ *is nonsingular, then the stationary point* z_o *of* $J(z, Y)$ *is given by*

$$z_o = \mathbf{R_{zY}} \mathbf{R_Y^{-1}} Y,$$

and

$$J(z_o, Y) = Y^* \mathbf{R_Y^{-1}} Y.$$

Moreover, this stationary point is a minimum iff,

$$\mathbf{R_z} - \mathbf{R_{zY}} \mathbf{R_Y^{-1}} \mathbf{R_{Yz}} > \mathbf{0}.$$

Remark: It is worthwhile to recapitulate the key points here.

- The stationary point $z_o = \mathbf{R_{zY}} \mathbf{R_Y^{-1}} Y$ of the scalar quadratic form $J(z, Y)$ is given by a formula that is exactly the same as that of the stationary point $\hat{\mathbf{z}} = \mathbf{R_{zY}} \mathbf{R_Y^{-1}} \mathbf{Y}$ of the matrix quadratic form $E(k)$.

- Note, however, that in Theorem 3 there is no Krein space and that z and Y are just vectors in Euclidean space. What we have shown is that by appropriately defining the submatrices in $J(z,Y)$ we can set up some Krein space and use the projection to calculate the stationary point of $J(z,Y)$.

- The conditions for a minimum, i.e., $\mathbf{R_Y} > 0$ and $\mathbf{R_z} - \mathbf{R_{zY}}\mathbf{R_Y^{-1}}\mathbf{R_{Yz}} > 0$ can be different, in the sense that one may need not imply the other, unless we are in Hilbert space.

Corollary 1 (Simultaneous Minima) *For vectors* $\mathbf{z}$ *and* $\mathbf{Y}$ *of linear independent elements in a Hilbert space* $\mathcal{H}$, *the conditions* $\mathbf{R_z} - \mathbf{R_{zY}}\mathbf{R_Y^{-1}}\mathbf{R_{Yz}} > 0$ *and* $\mathbf{R_Y} > 0$ *occur simultaneously.*

Proof: In a Hilbert space setting we have,

$$< \begin{bmatrix} \mathbf{z} \\ \mathbf{Y} \end{bmatrix}, \begin{bmatrix} \mathbf{z} \\ \mathbf{Y} \end{bmatrix} >_{\mathcal{H}} = \begin{bmatrix} \mathbf{R_z} & \mathbf{R_{zY}} \\ \mathbf{R_{Yz}} & \mathbf{R_Y} \end{bmatrix} > 0 \ ,$$

which readily implies that $\mathbf{R_Y} > 0$ and $\mathbf{R_z} - \mathbf{R_{zY}}\mathbf{R_Y^{-1}}\mathbf{R_{Yz}} > 0$.

■

Remark: Corollary 1 is the reason that the usual Kalman filter can be derived directly via a deterministic maximum likelihood approach (see *e.g.* [8], chapter 12).

3.2 State-Space Structure

We now assume that the observations $\{\mathbf{y_i}\}$ arise from an underlying state-space model, viz.,

$$\begin{aligned} \mathbf{x}_{i+1} &= F_i\mathbf{x}_i + G_i\mathbf{u}_i \\ \mathbf{y}_i &= H_i\mathbf{x}_i + \mathbf{v}_i \qquad i \geq 0 \end{aligned} \tag{8}$$

where $\mathbf{x}_0$, $\{\mathbf{u}_i\}$, and $\{\mathbf{v}_i\}$ are assumed to belong to a Krein space $\mathcal{K}$ with an indefinite inner product $< .,. >_{\mathcal{K}}$, and that they satisfy the following conditions:

$$< \begin{bmatrix} \mathbf{u}_i \\ \mathbf{v}_i \\ \mathbf{x}_0 \end{bmatrix}, \begin{bmatrix} \mathbf{u}_j \\ \mathbf{v}_j \\ \mathbf{x}_0 \end{bmatrix} >_{\mathcal{K}} = \begin{bmatrix} Q_i\delta_{ij} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & R_i\delta_{ij} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \Pi_0 \end{bmatrix} \tag{9}$$

Let

$$h_{ij} = H_iF_{i-1}\dots F_{j+1}G_j$$

be the response at time i to an impulse at time $j < i$. If we define

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y_0} \\ \mathbf{y_1} \\ \mathbf{y_2} \\ \vdots \\ \mathbf{y_N} \end{bmatrix} , \quad \mathbf{U} = \begin{bmatrix} \mathbf{u_0} \\ \mathbf{u_1} \\ \mathbf{u_2} \\ \vdots \\ \mathbf{u_N} \end{bmatrix} , \quad \mathbf{V} = \begin{bmatrix} \mathbf{v_0} \\ \mathbf{v_1} \\ \mathbf{v_2} \\ \vdots \\ \mathbf{v_N} \end{bmatrix}$$

then we have

$$\mathbf{Y} = \mathcal{O}\mathbf{x}_0 + \mathbf{\Gamma}\mathbf{U} + \mathbf{V}$$

where,

$$\mathcal{O} = \mathcal{O}(0,N) = \begin{bmatrix} H_0 \\ H_1F_0 \\ H_2F_1F_0 \\ \vdots \\ H_NF_{N-1}\dots F_0 \end{bmatrix}$$

is the observablity map, and

$$\mathbf{\Gamma} = \begin{bmatrix} 0 & & & & \\ h_{10} & 0 & & & \\ h_{20} & h_{21} & 0 & & \\ h_{30} & h_{31} & h_{32} & . & \\ . & . & . & . & . \end{bmatrix}$$

is the impulse response matrix.

In the state-space context, all quantities of interest are (linear) functions of the initial state $\mathbf{x_0}$ and the vectors $\mathbf{U}$ and $\mathbf{V}$. Therefore in state-space estimation problems, one can 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 $\hat{\mathbf{x}}_{\mathbf{0|N}}$ and $\hat{\mathbf{U}}_{\mathbf{|N}}$, respectively. Moreover, we shall be interested in computing these projections in a recursive fashion.

The uniqueness condition for Krein space projection means that *recursive* estimation is unambiguous iff all the leading minors of $\mathbf{R_Y}$ are nonzero. In this case $\mathbf{R_Y}$ is said to be strongly nonsingular (or strongly regular). We shall assume this henceforth.

The standard method of recursive estimation is to introduce the *innovations*

$$\mathbf{e_i} = \mathbf{y_i} - \hat{\mathbf{y}}_\mathbf{i}$$

where $\hat{\mathbf{y}}_\mathbf{i} = \hat{\mathbf{y}}_{\mathbf{i|i-1}}$ is the projection of $\mathbf{y_i}$ onto $\{\mathbf{y_0}, \dots, \mathbf{y_{i-1}}\}$. Since these are an orthogonal basis for the linear space spanned by the $\{\mathbf{y_i}\}$, we can calculate the projection of any quantity of interest in terms of the innovations. For example,

$$\hat{\mathbf{x}}_{\mathbf{0}|\mathbf{N}} = \sum_{i=0}^{N} <\mathbf{x_0}, \mathbf{e_i}>_{\mathcal{K}} <\mathbf{e_i}, \mathbf{e_i}>_{\mathcal{K}}^{-1} \mathbf{e_i}$$

$$\hat{\mathbf{u}}_{\mathbf{j}|\mathbf{N}} = \sum_{i=0}^{N} <\mathbf{u_j}, \mathbf{e_i}>_{\mathcal{K}} <\mathbf{e_i}, \mathbf{e_i}>_{\mathcal{K}}^{-1} \mathbf{e_i}$$

where the state-space structure may be used to calculate the inner products.

In the Hilbert space setting, the innovations are computed via the Kalman filter. (For an innovations derivation of the Kalman filter see [6]). This derivation extends to the Krein space setting so that we have the following result.

Theorem 4 (Kalman Filter in Krein Space) *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*

$$\begin{aligned} \mathbf{e}_i &= \mathbf{y}_i - H_i \hat{\mathbf{x}}_i, \qquad 0 \le i \le N & (10) \\ \hat{\mathbf{x}}_{i+1} &= F_i \hat{\mathbf{x}}_i + K_{p,i}(\mathbf{y}_i - H_i \hat{\mathbf{x}}_i), \quad \hat{\mathbf{x}}_0 = 0 & (11) \\ K_{p,i} &= F_i P_i H_i^* R_{e,i}^{-1} & (12) \end{aligned}$$

where

$$R_{e,i} = <\mathbf{e}_i, \mathbf{e}_i>_{\mathcal{K}} = R_i + H_i P_i H_i^*$$

and the $\{P_i\}$ *can be recursively computed via the (Riccati) recursions*

$$P_{i+1} = F_i P_i F_i^* - K_{p,i} R_{e,i} K_{p,i}^* + G_i Q_i G_i^*, \qquad P_0 = \Pi_0 \qquad (13)$$

Note that the only difference from the conventional Kalman filter is that the matrices P_i, $R_{e,i}$, Q_i and R_i may now be indefinite.

In Kalman filter theory there are many variations of the above formulas. As a matter of fact, one was already mentioned earlier in Theorem 1, where the Kalman filter formulas are somewhat different because they are expressed in terms of the so-called *filtered* state estimates $\hat{x}_{i|i}$, rather than the predicted estimates $\hat{x}_i$.

Corollary 2 (Information Forms) *When* Π_0 *and the* $\{R_i\}$ *are all invertible, we have the formulas*

$$\begin{aligned} K_{p,i} &= F_i P_{i|i} H_i^* R_i^{-1} & (14) \\ P_{i|i}^{-1} &= P_i^{-1} + H_i^* R_i^{-1} H_i & (15) \end{aligned}$$

For numerical reasons, certain square-root algorithms are now more often used. Furthermore for constant systems, or in fact for systems where the time-variation is structured in a certain way, we may use the fast Chandrasekhar recursions [7]. Both these algorithms generalize to the Krein space setting; these will be described elsewhere.

Now that we have a method for computing the projections, we would like to identify the scalar quadratic form of Theorem 3 whose stationary point is calculated via the projection. The matrix appearing in this quadratic form is the following

$$\begin{bmatrix} \mathbf{R_{x_0}} & \mathbf{R_{x_0U}} & \mathbf{R_{x_0Y}} \\ \mathbf{R_{Ux_0}} & \mathbf{R_U} & \mathbf{R_{UY}} \\ \mathbf{R_{Yx_0}} & \mathbf{R_{YU}} & \mathbf{R_Y} \end{bmatrix} = < \begin{bmatrix} \mathbf{x_0} \\ \mathbf{U} \\ \mathbf{Y} \end{bmatrix}, \begin{bmatrix} \mathbf{x_0} \\ \mathbf{U} \\ \mathbf{Y} \end{bmatrix} >_{\mathcal{K}}$$

$$= \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ \mathcal{O} & \mathbf{\Gamma} & I \end{bmatrix} \begin{bmatrix} \Pi_0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R} \end{bmatrix} \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ \mathcal{O} & \mathbf{\Gamma} & I \end{bmatrix}^*$$

where $\mathbf{Q} = Q_0 \oplus \ldots \oplus Q_N$, $\mathbf{R} = R_0 \oplus \ldots \oplus R_N$ and, where we have made use of the state-space structure:

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

We can now identify the corresponding scalar quadratic form as follows:

$$\begin{aligned} J &= \begin{bmatrix} x_0 \\ U \\ Y \end{bmatrix}^* \begin{bmatrix} \mathbf{R_{x_0}} & \mathbf{R_{x_0U}} & \mathbf{R_{x_0Y}} \\ \mathbf{R_{Ux_0}} & \mathbf{R_U} & \mathbf{R_{UY}} \\ \mathbf{R_{Yx_0}} & \mathbf{R_{YU}} & \mathbf{R_Y} \end{bmatrix}^{-1} \begin{bmatrix} x_0 \\ U \\ Y \end{bmatrix} \\ &= \begin{bmatrix} x_0 \\ U \\ V \end{bmatrix}^* \begin{bmatrix} \Pi_0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R} \end{bmatrix}^{-1} \begin{bmatrix} x_0 \\ U \\ V \end{bmatrix} \\ &= x_0^* \Pi_0^{-1} x_0 + \sum_{j=0}^{N} (y_j - H_j x_j)^* R_j^{-1} (y_j - H_j x_j) \\ &\quad + \sum_{j=0}^{N} u_j^* Q_j^{-1} u_j \end{aligned}$$

where we have assumed that the scalar quantities satisfy the same state-space constraint, viz.,

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

Now that we have identified the scalar quadratic form we readily have the following result (compare Theorems 2 and 3).

Theorem 5 (Deterministic Interpretation) *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 $K_{p,i} = F_iP_iH_i^*R_{e,i}^{-1}$, $R_{e,i} = R_i + H_iP_iH_i^*$, $e_i = y_i - H_i\hat{x}_i$, and P_i satisfying the Riccati recursion

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

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

$$J_i = x_0^*\Pi_0^{-1}x_0 + \sum_{j=0}^{i}(y_j - H_jx_j)^*R_j^{-1}(y_j - H_jx_j) + \sum_{j=0}^{i}u_j^*Q_j^{-1}u_j$$

over $\{x_j\}$ and $\{u_j\}$, subject to the state-space constraint $x_{j+1} = F_jx_j + G_ju_j$. The value of J_i at the stationary point is equal to $\sum_{j=0}^{i}e_j^*R_{e,j}^{-1}e_j$.

However, it is important to recall from Theorem 3 that the condition for a minimum is

$$\mathbf{R_z} - \mathbf{R_{zY}}\mathbf{R_Y}^{-1}\mathbf{R_{Yz}} > 0$$

and to see how this may be simplified when there is state-space structure. Doing this gives us the following result, among several possiblities not discussed here. However, the result below allows us to link up with the results in the H^∞ literature, *e.g.* Theorem 1 quoted earlier.

Theorem 6 (Condition for Minimum) *If the F_j are nonsingular, then the stationary point of Theorem 5 will correspond to a minimum if, and only if, $Q_j > 0$, for all j, and*

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

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

4 H^∞ Filtering Revisited

Now that we have developed a background in Krein space estimation, we shall use it to study the H^∞ filtering problem and to obtain the filter of Theorem 1.

Recall from (2) that in the H^∞ estimation problem we were required to guarantee the positivity of the following quadratic form.

$$J_i = x_0^*\Pi_0^{-1}x_0 + \sum_{k=0}^{i}u_k^*u_k + \sum_{k=0}^{i}\left(\begin{bmatrix} \check{z}_{k|k} \\ y_k \end{bmatrix} - \begin{bmatrix} L_k \\ H_k \end{bmatrix}x_k\right)^* \begin{bmatrix} -\gamma^{-2}I & 0 \\ 0 & I \end{bmatrix}\left(\begin{bmatrix} \check{z}_{k|k} \\ y_k \end{bmatrix} - \begin{bmatrix} L_k \\ H_k \end{bmatrix}x_k\right) .$$

In view of Theorem 5 this is the scalar quadratic form corresponding to the Krein state-space model

$$\begin{aligned} \mathbf{x_{i+1}} &= F_i\mathbf{x_i} + G_i\mathbf{u_i} \\ \begin{bmatrix} \mathbf{\check{z}_{i|i}} \\ \mathbf{y_i} \end{bmatrix} &= \begin{bmatrix} L_i \\ H_i \end{bmatrix}\mathbf{x_i} + \mathbf{w_i} \end{aligned} \qquad (16)$$

with

$$< \begin{bmatrix} \mathbf{u_i} \\ \mathbf{w_i} \\ \mathbf{x_0} \end{bmatrix} \begin{bmatrix} \mathbf{u_j} \\ \mathbf{w_j} \\ \mathbf{x_0} \end{bmatrix} >_{\mathcal{K}} = \begin{bmatrix} I\delta_{ij} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \begin{bmatrix} -\gamma^2 I & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix}\delta_{ij} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \Pi_0 \end{bmatrix} \qquad (17)$$

Note that we need to consider a Krein space because

$$R_j = \begin{bmatrix} -\gamma^2 I & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix}$$

is indefinite.

Since the $\{y_j\}$ are fixed, and $\check{z}_{j|j}$ is a function of the $\{y_j\}$, J_i is just a function of the free variables x_0 and the $\{u_j\}$. Thus $J_i > 0$ if, and only if,

- J_i has a minimum with respect to $\{x_0, u_0, u_1, ...u_i\}$,
- The $\{\check{z}_{j|j}\}_{j=0}^{i}$ can be chosen such that the value of J_i at this minimum is positive.

We can now use our knowledge of Krein space estimation to give an outline of the proof of Theorem 1. More details are in [5].

Outline of Proof of Theorem 1:

- The H^∞ Riccati is just the Riccati recursion corresponding to Krein state-space model (16).
- Positivity of $P_{i|i}$ ensures the existence of a minimum, and yields the condition (3).
- Using $\{\check{z}_{j|j}\}_{j=0}^{i}$ as the projection ensures value at the minimum is positive.

- Calculating the projection (via the Krein space Kalman filter corresponding to (16)) yields the desired filter.

■

This example serves to illustrate the philosophy of our approach. Many problems in H^∞ and risk-sensitive control and estimation, adaptive filtering, etc, lead to indefinite quadratic forms. In each case we construct the corresponding Krein state-space model, and use the Krein space Kalman filter to study the properties of these quadratic form.

In this case this study lead to a derivation of the aposteriori H^∞ filters and the conditions for existence. One may pursue the same line of reasoning to obtain the apriori H^∞ filters, H^∞ smoothers, the full parametrization of all H^∞ estimators, and so on (see [5]).

5 Conclusion

We introduced the problem of H^∞ filtering, and showed how the concept of Krein space Kalman filtering could be used to solve this problem. This approach yields a unifying method for solving H^2 and H^∞ problems, and has enabled us to study a number of problems in H^∞ and risk-sensitive estimation and control and finite memory adaptive filtering.

The insight provided by the connection with Kalman filtering, which has a rich history of connections to other topics and also has a very extensive literature, has allowed us to develop square-root and Chandrasekhar algorithms for the H^∞ problem, find extensions of robust and adaptive filtering algorithms, get new insights into the generalized KYP (Kalman-Yacobovich-Popov) Lemma, etc. No doubt other problems may profit from the approach described here.

Acknowledgement: We would like to acknowledge the helpful influence of early discussions with Mr. Poogyeon Park, esp. on the KYP lemma.

This work was supported in part by the Air Force Office of Scientific Research, Air Force Systems Command under Contract AFOSR91-0060 and in part by a grant from Rockwell International Inc.

References

[1] U. Shaked and Y. Theodor. $H^\infty-$optimal estimation: A tutorial. In *Proc. IEEE Conference on Decision and Control*, pages 2278–2286, Tucson, AZ, Dec. 1992.

[2] P.P. Khargonekar and K. M. Nagpal. Filtering and smoothing in an $H^\infty-$ setting. *IEEE Trans. on Automatic Control*, AC-36:151–166, 1991.

[3] T. Basar. Optimum performance levels for minimax filters, predictors and smoothers. *Systems and Control Letters*, 16:309–317, 1991.

[4] B. Hassibi, A. H. Sayed, and T. Kailath. Recursive linear estimation in Krein spaces - part I: Theory. To appear in *Proc. IEEE Conference on Decision and Control*, San Antonio, TX, Dec. 1993

[5] B. Hassibi, A. H. Sayed, and T. Kailath. Recursive linear estimation in Krein spaces - Part II: Applications. To appear in *Proc. IEEE Conference on Decision and Control*, San Antonio, TX, Dec. 1993.

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

[7] 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. 19, pp. 315-323, 1974.

[8] A. E. Bryson and Y .C. Ho *Applied Optimal Control*, Blaisdell Publishing Company, 1969.