Chapter 3

Finite Horizon H^∞ Filtering

In this chapter we study the problem of finite horizon H^∞ filtering. We consider the two cases of a posteriori a priori filtering, as well as smoothing and l-step prediction. We essentially show that all these problems can be cast into the problem of calculating the stationary point of certain indefinite quadratic forms. This allows us to exploit (to great effect) the machinery developed in Chapter 2. In particular, we show that by considering the appropriate state space models and error Gramians, we can use the Krein space estimation theory to calculate the stationary points and study their properties and thereby solve the various H^∞ filtering problems. The approach discussed here allows for interesting generalizations and extensions, which we will discuss in Chapters 4 and 5.

3.1 Introduction

Classical results in linear least-squares estimation and Kalman filtering are based on an L^2 or H^2 criterion and require a-priori 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 has led to an increasing interest in minimax estimation (see, *e.g.,* [Kwa86, DGKF89, KN91, Bas91, LS91, ST92, YS91, Gri93] and the references therein), with the belief that the resulting so-called H^∞ algorithms will be more robust and less sensitive to parameter variations.

H^∞ filtering problems (as well as many other problems in risk-sensitive estimation and control problems, quadratic games and finite memory adaptive filtering) lead almost by inspection to indefinite deterministic quadratic forms. Following Chapter 2, we solve these problems by constructing the corresponding Krein space "stochastic" problems for which the Kalman filter solutions can be written down immediately; moreover, the conditions for a minimum (related to the existence of solutions to the H^∞ problem) can also be expressed in terms of quantities easily related to the basic Riccati equations of the Kalman filter. This approach also explains the many similarities between say, the H^∞ solutions and the classical LQ solutions, and in addition marks out their key differences.

The chapter is organized as follows. In Section 3.2 we introduce the H^∞ estimation problem, state the conventional solution and discuss its similarities with and differences from the conventional Kalman filter. In Section 3.3 we reduce the H^∞ estimation problem to guaranteeing the positivity of a certain indefinite quadratic form. We then relate this quadratic form to a certain Krein state-space model, which allows us to use the results of Chapter 2 to derive conditions for its positivity, and to show that projection in the Krein space allows us to solve the H^∞ estimation problem. In this context we derive the H^∞ a posteriori, a priori, and smoothing filters, as well as l-step predictors,[1] and show that H^∞ estimation is essentially Kalman filtering in Krein space; we also obtain a natural parametrization of all H^∞ estimators. One advantage of our approach is that it suggests how well known conventional Kalman filtering algorithms, such as square root arrays and Chandrasekhar recursions, can be extended to the H^∞ setting. Much more will be said about this in Chapter 5.

As was done in Chapter 2, we shall use bold letters for elements in a Krein space, and normal letters for corresponding complex numbers. Also, we shall use $\check{s}$ to denote the estimate of s (according to some criterion), and $\hat{s}$ to denote the Krein space projection, thereby stressing the fact that they need not coincide. Many of the results discussed here were obtained earlier by several other authors, using different methods and arguments. Our approach, we believe, provides a powerful unification,

[1]The solution to the l-step prediction problem presented here is, to the best of our knowledge, new. We have not encountered general l-step H^∞ filters in the literature.

with immediate insights to various extensions.

3.2 H^∞ Estimation

Several H^∞-filtering algorithms have been recently derived by a variety of methods in both the continuous and discrete-time cases (see, *e.g.,* [Kwa86, DGKF89, KN91, Bas91, LS91, ST92, YS91, Gri93] and the references therein).

Many authors have noticed some formal similarities between the H^∞ filters and the conventional Kalman filter, however we shall further clarify this connection by showing that H^∞ filters are nothing more than certain Krein space Kalman filters. In other words, the H^∞ filters can be viewed as recursively performing a (Gram-Schmidt) orthogonalization (or projection) procedure on a convenient set of observation data that obey a state-space model whose state evolves in an indefinite metric space. This is of significance since it yields a geometric derivation of the H^∞ filters, and because it unifies H^2 and H^∞ estimation in a simple framework. Moreover, once this connection has been made explicit, many known alternative and more efficient algorithms, such as square-root arrays and Chandrasekhar equations[HSK94c], can be applied to the $H^\infty-$setting as well. Also our results deal directly with the time-varying scenario. Finally, we note that although we restrict ourselves here, to the discrete-time case, however, the continuous time analogs follow the same principles.

3.2.1 Formulation of the H^∞ Filtering Problem

Consider a time-variant state-space model of the form

$$\left\{ \begin{array}{rcll} x_{i+1} & = & F_i x_i + G_i u_i, & x_0 \\ y_i & = & H_i x_i + v_i, & i \geq 0 \end{array} \right. \tag{3.2.1}$$

where $F_i \in \mathcal{C}^{n\times n}$, $G_i \in \mathcal{C}^{n\times m}$ and $H_i \in \mathcal{C}^{p\times n}$ are known matrices, x_0, $\{u_i\}$, and $\{v_i\}$ are *unknown* quantities and y_i is the measured output. We can regard v_i as a measurement noise and u_i as a process noise or driving disturbance. We make no assumption on the nature of the disturbances (*e.g.* , normally distributed, uncorrelated, etc). In general, we would like to estimate some arbitrary linear combination of the states, say

$$s_i = L_i x_i,$$

where $L_i \in \mathcal{C}^{q\times n}$ is given, using the observations $\{y_j\}$. Let $\check{s}_{i|i} = \mathcal{F}_f(y_0, y_1, \ldots, y_i)$ denote the estimate of s_i given observations $\{y_j\}$ from time 0 up to and including time i, and $\check{s}_i = \mathcal{F}_p(y_0, y_1, \ldots, y_{i-1})$ denote the estimate of s_i given observations $\{y_j\}$ from time 0 to time $i-1$. We then have the following two estimation errors: the *filtered* error

$$e_{f,i} = \check{s}_{i|i} - L_i x_i, \tag{3.2.2}$$

and the *predicted* error

$$e_{p,i} = \check{s}_i - L_i x_i. \tag{3.2.3}$$

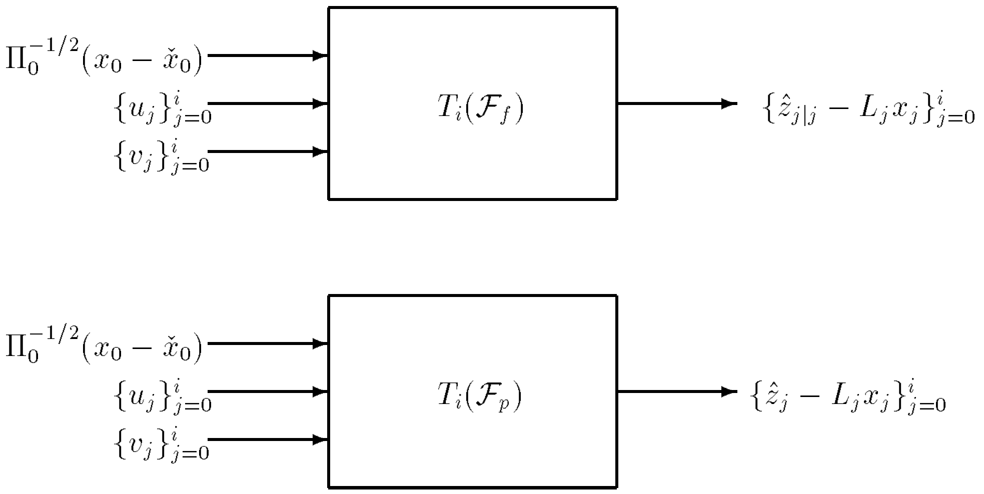

Figure 3.1: Transfer matrix from disturbances to filtered and predicted estimation errors.

As depicted in Fig. 3.1, let $T_i(\mathcal{F}_f)$ and $T_i(\mathcal{F}_p)$ denote the transfer operators that map the unknown disturbances $\{\Pi_0^{-1/2}(x_0 - \check{x}_0), \{u_j\}_{j=0}^i, \{v_j\}_{j=0}^i\}$ (where $\check{x}_0$ denotes an initial guess for x_0, and Π_0 is a given positive definite matrix) to the filtered and predicted errors $\{e_{f,j}\}_{j=0}^{i}$ and $\{e_{p,j}\}_{j=0}^{i}$, respectively. The problem is to choose the functionals $\mathcal{F}_f(\cdot)$ and $\mathcal{F}_p(\cdot)$ so as to respectively minimize the H^∞ norm of the transfer operators $T_i(\mathcal{F}_f)$ and $T_i(\mathcal{F}_p)$.

Definition 3.2.1 *The H^∞ norm of a transfer operator T is defined as*

$$\|T\|_\infty = \sup_{u\in h^2, u\neq 0} \frac{\|Tu\|_2}{\|u\|_2}$$

where $\|u\|_2$ is the h^2-norm of the causal sequence $\{u_k\}$, i.e. $\|u\|_2^2 = \sum_{k=0}^{\infty} u_k^* u_k$.

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

Problem 3.2.1 (Optimal H^∞ Problem) *Find filtered and predicted H^∞-optimal estimation strategies $\check{s}_{i|i} = \mathcal{F}_f(y_0, y_1, \ldots, y_i)$ and $\check{s}_i = \mathcal{F}_p(y_0, y_1, \ldots, y_{i-1})$ that respectively minimize $\| T_i(\mathcal{F}_f) \|_\infty$ and $\| T_i(\mathcal{F}_p) \|_\infty$, and obtain the resulting*

$$\begin{aligned} \gamma_{f,o}^2 &= \inf_{\mathcal{F}_f} \| T_i(\mathcal{F}_f) \|_\infty^2 \\ &= \inf_{\mathcal{F}_f} \sup_{x_0, u\in h^2, v\in h^2} \frac{\sum_{j=0}^{i} e_{f,j}^* e_{f,j}}{(x_0-\check{x}_0)^*\Pi_0^{-1}(x_0-\check{x}_0) + \sum_{j=0}^{i} u_j^* u_j + \sum_{j=0}^{i} v_j^* v_j} \end{aligned} \tag{3.2.4}$$

and

$$\begin{aligned} \gamma_{p,o}^2 &= \inf_{\mathcal{F}_p} \| T_i(\mathcal{F}_p) \|_\infty^2 \\ &= \inf_{\mathcal{F}_p} \sup_{x_0, u\in h^2, v\in h^2} \frac{\sum_{j=0}^{i} e_{p,j}^* e_{p,j}}{(x_0-\check{x}_0)^*\Pi_0^{-1}(x_0-\check{x}_0) + \sum_{j=0}^{i-1} u_j^* u_j + \sum_{j=0}^{i-1} v_j^* v_j} \end{aligned} \tag{3.2.5}$$

where Π_0 is a positive definite matrix that reflects a priori knowledge as to how close x_0 is to the initial guess $\check{x}_0$.

Note that the infimum in (3.2.5) is taken over all *strictly* causal estimators $\mathcal{F}_p$, whereas in (3.2.4) the estimators $\mathcal{F}_f$ are only causal, since they have additional access to y_i. This is relevant since the solution to the H^∞ problem, as we shall see, depends on the structure of the information available to the estimator.

The above problem formulation shows that H^∞ optimal estimators guarantee the smallest estimation error energy over all possible disturbances of fixed energy. They are therefore conservative (and at times over-conservative), which results in a better robust behaviour to disturbance variation.

A closed form solution to the optimal H^∞ estimation problem is available only in some special cases (see *e.g.* , [HSK96a] and Chapter 9), and so it is common in the literature to settle for a suboptimal solution.

Problem 3.2.2 (Sub-optimal H^∞ Problem) *Given scalars $\gamma_f > 0$ and $\gamma_p > 0$, H^∞ suboptimal estimation strategies $\check{s}_{i|i} = \mathcal{F}_f(y_0, y_1, \ldots, y_i)$ (known as an a posteriori filter) and $\check{s}_i = \mathcal{F}_p(y_0, y_1, \ldots, y_{i-1})$ (known as an a priori filter) that respectively achieve $\| T_i(\mathcal{F}_f) \|_\infty < \gamma_f$ and $\| T_i(\mathcal{F}_p) \|_\infty < \gamma_p$. In other words, find strategies that respectively achieve*

$$\sup_{x_0, u \in h^2, v \in h^2} \frac{\sum_{j=0}^{i} e_{f,j}^* e_{f,j}}{(x_0 - \check{x}_0)^* {\Pi_0}^{-1} (x_0 - \check{x}_0) + \sum_{j=0}^{i} u_j^* u_j + \sum_{j=0}^{i} v_j^* v_j} < \gamma_f^2 \tag{3.2.6}$$

and

$$\sup_{x_0, u \in h^2, v \in h^2} \frac{\sum_{j=0}^{i} e_{p,j}^* e_{p,j}}{(x_0 - \check{x}_0)^* {\Pi_0}^{-1} (x_0 - \check{x}_0) + \sum_{j=0}^{i-1} u_j^* u_j + \sum_{j=0}^{i-1} v_j^* v_j} < \gamma_p^2. \tag{3.2.7}$$

This clearly requires checking whether $\gamma_f \geq \gamma_{f,o}$ and $\gamma_p \geq \gamma_{p,o}$.

Note that the solutions to Problem 3.2.1 can be obtained to desired accuracy by iterating on the γ_f and γ_p of Problem 3.2.2. From here on we shall be only dealing with Problem 3.2.2.

Note that since the upper limits in the aforementioned sums are the finite number i, the problems defined above are *finite-horizon* problems. So-called *infinite-horizon* problems can be considered if we define $T(\mathcal{F}_f)$ and $T(\mathcal{F}_p)$ as the transfer operators that map $\{x_0 - \check{x}_0, \{u_j\}_{j=0}^\infty, \{v_j\}_{j=0}^\infty\}$ to $\{e_{f,j}\}_{j=0}^\infty$ and $\{e_{p,j}\}_{j=0}^\infty$, respectively. Then by guaranteeing $\| T_i(\mathcal{F}_f) \|_\infty < \gamma_f$ and $\| T_i(\mathcal{F}_p) \|_\infty < \gamma_p$ for all i, we can solve the infinite-horizon problems $\| T(\mathcal{F}_f) \|_\infty \leq \gamma_f$ and $\| T(\mathcal{F}_p) \|_\infty \leq \gamma_p$, respectively. However, as we have seen in Sec. 1.4.1, direct solutions are also possible. We shall study the infinite horizon problem in more detail in Chapter 7.

3.2.2 Solution of the Suboptimal H^∞ Filtering Problem

We now present the existing solutions (see *e.g.*, [KN91, ST92]) to the suboptimal H^∞ filtering problem and note that they are intriguingly similar, in several ways, to the conventional Kalman filter. It was this similarity in structure that led us to extend Kalman filtering to Krein spaces (see Chapter 2): in effect, H^∞ filters are just Kalman filters in Krein space.

Theorem 3.2.1 (A Finite Horizon H^∞ A Posteriori Filter) *[ST92] For a given $\gamma > 0$, if the $\begin{bmatrix} F_j & G_j \end{bmatrix}$ have full rank, then an estimator that achieves $\|T_i(\mathcal{F}_f)\|_\infty < \gamma$ exists if, and only if,*

$$P_j^{-1} + H_j^* H_j - \gamma^{-2} L_j^* L_j > 0, \qquad j = 0, \ldots, i \tag{3.2.8}$$

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

$$P_{j+1} = F_j P_j F_j^* + G_j G_j^* - F_j P_j \begin{bmatrix} H_j^* & L_j^* \end{bmatrix} R_{e,j}^{-1} \begin{bmatrix} H_j \\ L_j \end{bmatrix} P_j F_j^*, \tag{3.2.9}$$

with

$$R_{e,j} = \begin{bmatrix} I & 0 \\ 0 & -\gamma^2 I \end{bmatrix} + \begin{bmatrix} H_j \\ L_j \end{bmatrix} P_j \begin{bmatrix} H_j^* & L_j^* \end{bmatrix}. \tag{3.2.10}$$

If this is the case, then one possible level-γ H^∞ filter is given by

$$\check{s}_{j|j} = L_j \hat{x}_{j|j},$$

where $\hat{x}_{j|j}$ is recursively computed as

$$\hat{x}_{j+1|j+1} = F_j \hat{x}_{j|j} + K_{s,j+1}(y_{j+1} - H_{j+1} F_j \hat{x}_{j|j}) \quad , \quad \hat{x}_{-1|-1} = \textit{initial guess} \tag{3.2.11}$$

and

$$K_{s,j+1} = P_{j+1} H_{j+1}^* (I + H_{j+1} P_{j+1} H_{j+1}^*)^{-1}. \tag{3.2.12}$$

Theorem 3.2.2 (A Finite Horizon H^∞ A Priori Filter) *[ST92] Given $\gamma > 0$, if the $\begin{bmatrix} F_j & G_j \end{bmatrix}$ have full rank, then an estimator that achieves $\|T_i(\mathcal{F}_p)\|_\infty < \gamma$ exists if, and only if,*

$$\tilde{P}_j^{-1} = P_j^{-1} - \gamma^{-2} L_j^* L_j > 0, \qquad j = 0, \ldots, i \tag{3.2.13}$$

where P_j is the same as in Theorem 3.2.1. If this is the case, then one possible level-γ H^∞ filter is given by

$$\check{s}_j = L_j \hat{x}_j, \tag{3.2.14}$$

$$\hat{x}_{j+1} = F_j \hat{x}_j + K_{a,j}(y_j - H_j \hat{x}_j) \quad , \quad \hat{x}_0 = \textit{initial guess} \tag{3.2.15}$$

where

$$K_{a,j} = F_j \tilde{P}_j H_j^* (I + H_j \tilde{P}_j H_j^*)^{-1}. \tag{3.2.16}$$

Comparisons with the Kalman Filter

The Kalman filter algorithm for estimating the states in (3.2.1), assuming that the $\{u_i\}$ and $\{v_i\}$ are uncorrelated unit variance white noise processes is, in the a posteriori case,

$$\hat{x}_{j+1|j+1} = F_j \hat{x}_{j|j} + P_{j+1} H_{j+1}^* (I + H_{j+1} P_{j+1} H_{j+1}^*)^{-1} (y_{j+1} - H_{j+1} \hat{x}_{j+1}), \tag{3.2.17}$$

and, in the a priori case,

$$\hat{x}_{j+1|j+1} = F_j \hat{x}_{j|j} + P_{j+1} H_{j+1}^* (I + H_{j+1} P_{j+1} H_{j+1}^*)^{-1} (y_{j+1} - H_{j+1} \hat{x}_{j+1}), \tag{3.2.18}$$

where

$$P_{j+1} = F_j P_j F_j^* + G_j G_j^* - F_j P_j (I + H_j P_j H_j^*)^{-1} P_j F_j^*, \quad P_0 = \Pi_0. \tag{3.2.19}$$

As several authors have noted, the H^∞ solutions are very similar to the conventional Kalman filter. The major differences are the following:

- The structure of the H^∞ estimators depends, via the Riccati recursion (3.2.9), on the linear combination of the states that we intend to estimate (*i.e.* the L_i). This is as opposed to the Kalman filter, where the estimate of any linear combination of the state is given by that linear combination of the state estimate. Intuitively, this means that the H^∞ filters are specifically tuned towards the linear combination $L_i x_i$.

- We have additional conditions, (3.2.8) or (3.2.13), 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 positive definite, so that (3.2.8) and (3.2.13) would be immediate.

- We have indefinite (*covariance*) matrices, *e.g.* , $\begin{bmatrix} I & \mathbf{0} \\ \mathbf{0} & -\gamma^2 I \end{bmatrix}$ vs. just I in the Kalman filter.

- As $\gamma \to \infty$, the Riccati recursion (3.2.9) reduces to the Kalman filter recursion (3.2.19). [This suggests that the H^∞ norm of the conventional Kalman filter may be quite large, and that it may have poor robustness properties.[2] Note also that condition (3.2.13) is more stringent than condition (3.2.8), indicating that the existence of an a priori filter of level γ implies the existence of an a posteriori filter of the same level, but not necessarily vice versa.]

Despite these differences, we shall show by applying the results of Chapter 2, that the filters of Theorems 3.2.1 and 3.2.2 can in fact be obtained as certain Kalman filters, not in an H^2 (Hilbert) space, but in a certain 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 (3.2.8) will be seen to arise from the fact that in Krein space, unlike as in the usual Hilbert space context, quadratic forms need not always have minima or maxima, unless certain additional conditions are met. Moreover, our approach will provide a simpler and more general alternative to the tests (3.2.8) and (3.2.13).

[2]We shall have more to say about this in Chapter 10.

3.3 Derivation of the H^∞ Filters

In Sec. 1.4.1 we derived the H^∞ suboptimal filters using the canonical factorization of a certain indefinite transfer operator. Here we can take the same approach: the indefinite transfer operator will now be a finite indefinite matrix (or Gramian) and the canonical factorization will correspond to a (block) LDU (lower-diagonal-factorization). As we have seen in Chapter 2, when the Gramian has state-space structure (as it does in the H^∞ state-space estimation problem under consideration) the triangular factorization can be obtained from the Krein space Kalman filter. This route will then lead us both to the solution of the H^∞ filtering problems (via the Krein space Kalman filter, which incidentally also explains the connections with the conventional Kalman filter) and to the existence conditions for such filters (since an LDU factorization with certain inertia properties requires additional conditions). However, we shall here take a slightly different (and closer to the general spirit of Chapter 2) route to the solution, *i.e.*, one that is based upon associating an indefinite quadratic form (instead of an indefinite transfer operator) with the H^∞ filtering problem. There are, of course, close connections between these two approaches and we shall comment on them wherever deemed appropriate.

As shown Chapter 2, the first step is to associate an indefinite quadratic form with each of the (level γ) a posteriori and a priori filtering problems. This will lead us to construct an appropriate (so-called partially equivalent) Krein space state-space model, the Kalman filter for which will allow us to compute the stationary points for the H^∞ quadratic forms; conditions that these are actually minima will be deduced from the general results of Chapter 2, and shown to be just (3.2.8) and (3.2.13); simpler equivalent conditions will also be noted.

Therefore we begin by examining the structure of the H^∞ problem in more detail. The goal will be to relate the problem to an indefinite quadratic form. We shall first consider the a posteriori filtering problem.

3.3.1 The Suboptimal H^∞ Problem and Quadratic Forms

Referring to Problem 3.2.2, we first note that $\|T_i(\mathcal{F}_f)\|_\infty < \gamma_f$, implies that for all nonzero $\{x_0, \{u_j\}_{j=0}^i, \{v_j\}_{j=0}^i\}$

$$\frac{\sum_{j=0}^i |\check{s}_{j|j} - L_j x_j|^2}{(x_0 - \check{x}_0)^* \Pi_0^{-1}(x_0 - \check{x}_0) + \sum_{j=0}^i |u_j|^2 + \sum_{j=0}^i |y_j - H_j x_j|^2} < \gamma_f^2. \tag{3.3.1}$$

Moreover, (3.3.1) implies that for *all* $k \leq i$, we must have

$$\frac{\sum_{j=0}^k |\check{s}_{j|j} - L_j x_j|^2}{(x_0 - \check{x}_0)^* \Pi_0^{-1}(x_0 - \check{x}_0) + \sum_{j=0}^k |u_j|^2 + \sum_{j=0}^k |y_k - H_i x_i|^2} < \gamma_f^2. \tag{3.3.2}$$

We remark that if the $\{y_j\}_{j=0}^i$ are all zero then it is easy to see that the $\{\check{s}_{j|j}\}$ must all be zero as well. Therefore we need only consider the case where $\{y_j\}_{j=0}^i$ is a nonzero sequence. We shall then prove the following result, relating the condition $\|T_i(\mathcal{F}_f)\|_\infty < \gamma_f$ to the positivity of a certain indefinite quadratic form. Incidentally from now on, without loss of generality, we assume $\check{x}_0 = 0$; a nonzero $\check{x}_0 = 0$ will only change the initial condition of the filter.

Lemma 3.3.1 (Indefinite Quadratic Form) *Given $\gamma_f > 0$, then $\|T_i(\mathcal{F}_f)\|_\infty < \gamma_f$ if, and only if, there exists $\check{s}_{k|k} = \mathcal{F}_f(y_0, \ldots, y_k)$ (for all $0 \leq k \leq i$) such that for all complex vectors x_0, for all causal sequences $\{u_j\}_{j=0}^i$ and for all nonzero causal sequences $\{y_j\}_{j=0}^i$, the scalar quadratic form*

$$\begin{aligned} J_{f,k}(x_0, u_0, \ldots, u_k, y_0, \ldots, y_k) = x_0^* \Pi_0^{-1} x_0 + \textstyle\sum_{j=0}^k u_j^* u_j + \\ \textstyle\sum_{j=0}^k (y_j - H_j x_j)^*(y_j - H_j x_j) - \gamma_f^{-2} \sum_{j=0}^k (\check{s}_{j|j} - L_j x_j)^*(\check{s}_{j|j} - L_j x_j) \end{aligned} \tag{3.3.3}$$

satisfies

$$J_{f,k}(x_0, u_0, \ldots, u_k, y_0, \ldots, y_k) > 0 \quad \textit{for all} \quad 0 \leq k \leq i. \tag{3.3.4}$$

Proof: For one direction, assume there exists a solution $\check{s}_{k|k}$ (for all $k \leq i$) that achieves $\|T_i(\mathcal{F}_f)\|_\infty < \gamma_f$. Then if we multiply both sides of (3.3.2) by the positive denominator on the LHS, we obtain (3.3.4).

Conversely, if there exists a solution $\check{s}_{k|k}$ (for all $k \leq i$) that achieves (3.3.4), we can divide both sides of (3.3.4) by the positive quantity

$$x_0^*\Pi_0^{-1}x_0 + \sum_{j=0}^{k} u_j^* u_j + \sum_{j=0}^{k}(y_j - H_j x_j)^*(y_j - H_j x_j)$$

to obtain (3.3.2), and thereby $\|T_i(\mathcal{F}_f)\|_\infty < \gamma_f$.

■

Remark: Lemma 3.3.1 is a straightforward restatement of the inequality (3.3.2) which is required of all suboptimal H^∞ a posteriori filters with level γ_f. However, the statement of Lemma 3.3.2, given below, is a key result, since it shows how to check the conditions of Lemma 3.3.1 by computing the stationary point of the indefinite quadratic form $J_{f,k}(x_0, u_0, \ldots, u_k, y_0, \ldots, y_k)$ and checking its condition for a minimum. This is in the spirit of the approach taken in Chapter 2.

Note that since the $\check{s}_{k|k}$ are functions of the $\{y_j\}_{j=0}^{k}$, $J_{f,k}(x_0, u_0, \ldots, u_k, y_0, \ldots, y_k)$ is really a function of *only* $\{x_0, u_0, \ldots, u_k, y_0, \ldots, y_k\}$. Moreover, since the $\{y_j\}$ are fixed observations, the only *free variables* in $J_{f,k}(x_0, u_0, \ldots, u_k, y_0, \ldots, y_k)$ are the disturbances $\{x_0, u_0, \ldots, u_k\}$. We then have the following result.

Lemma 3.3.2 (Positivity Condition) *The scalar quadratic forms*

$$J_{f,k}(x_0, u_0, \ldots, u_k, y_0, \ldots, y_k),$$

satisfy the conditions (3.3.4), if, and only if, for all $0 \leq k \leq i$,

(i) $J_{f,k}(x_0, u_0, \ldots, u_k, y_0, \ldots, y_k)$ *has a minimum with respect to* $\{x_0, u_0, u_1, ...u_k\}$.

(ii) The $\{\check{s}_{k|k}\}_{k=0}^{i}$ *can be chosen such that the value of* $J_{f,k}(x_0, u_0, \ldots, u_k, y_0, \ldots, y_k)$ *at this minimum is positive, viz.,*

$$\min_{\{x_0, u_0, \ldots, u_k\}} J_{f,k}(x_0, u_0, \ldots, u_k, y_0, \ldots, y_k, \check{s}_{0|0}, \ldots, \check{s}_{k|k}) > 0.$$

Proof: Assume $J_{f,k}(x_0, \{u_j\}_{j=0}^{k}, \{y_j\}_{j=0}^{k}) > 0$, then condition (i) is clearly satisfied because if $J_{f,k}(x_0, \{u_j\}_{j=0}^{k}, \{y_j\}_{j=0}^{k})$ does not have a minimum over $\{x_0, u_0, \ldots, u_k\}$, then it is always possible to choose $\{x_0, \{u_j\}_{j=0}^k\}$ so as to make $J_{f,k}(x_0, \{u_j\}_{j=0}^k, \{y_j\}_{j=0}^k)$ arbitrarily small and negative. Moreover, the existence of a minimum, along with $J_{f,k}(x_0, \{u_j\}_{j=0}^k, \{y_j\}_{j=0}^k) > 0$, guarantees condition (ii) since the value at the minimum must be positive.

Conversely, if (i) and (ii) hold, then it follows that $J_{f,k}(x_0, \{u_j\}_{j=0}^k, \{y_j\}_{j=0}^k) > 0$. ∎

3.3.2 A Krein Space State-Space Model

To apply the methodology of Chapter 2, we first identify the indefinite quadratic form $J_{f,k}$ as a special case of the general form studied in Theorem 2.7.4 by rewriting it as

$$J_{f,k}(x_0, u_0, \ldots, u_k, y_0, \ldots, y_k) =$$
$$x_0^* \Pi_0^{-1} x_0 + \sum_{j=0}^k u_j^* u_j + \sum_{j=0}^k \left(\begin{bmatrix} y_j \\ \check{s}_{j|j} \end{bmatrix} - \begin{bmatrix} H_j \\ L_j \end{bmatrix} x_j \right)^* \begin{bmatrix} I & 0 \\ 0 & -\gamma_f^{-2} I \end{bmatrix} \left(\begin{bmatrix} y_j \\ \check{s}_{j|j} \end{bmatrix} - \begin{bmatrix} H_j \\ L_j \end{bmatrix} x_j \right). \tag{3.3.5}$$

Then by Lemmas 2.6.1 and 2.6.2, we can introduce the following Krein space state-space model,

$$\left\{ \begin{array}{rcl} \mathbf{x}_{j+1} & = & F_j \mathbf{x}_j + G_j \mathbf{u}_j \\ \begin{bmatrix} \mathbf{y}_j \\ \check{\mathbf{s}}_{j|j} \end{bmatrix} & = & \begin{bmatrix} H_j \\ L_j \end{bmatrix} \mathbf{x}_j + \mathbf{v}_j \end{array} \right. \tag{3.3.6}$$

with

$$\left\langle \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{u}_j \\ \mathbf{v}_j \end{bmatrix}, \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{u}_k \\ \mathbf{v}_k \end{bmatrix} \right\rangle = \begin{bmatrix} \Pi_0 & 0 & 0 \\ 0 & I\delta_{jk} & 0 \\ 0 & 0 & \begin{bmatrix} I & 0 \\ 0 & -\gamma_f^2 I \end{bmatrix} \delta_{jk} \end{bmatrix} \tag{3.3.7}$$

Note that $Q_j = I$, $S_j = 0$, $\Pi_0 > 0$, and that we must consider a Krein space since

$$R_j = \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0} & -\gamma_f^2 I \end{bmatrix} \tag{3.3.8}$$

is indefinite.[3]

---

[3]We should also mention that the output Gramian of $\begin{bmatrix} \mathbf{y}_j \\ \check{\mathbf{s}}_{j|j} \end{bmatrix}$ (in the state-space model (3.3.5-3.3.7) is simply the indefinite Gramian of Theorem 1.4.1 whose canonical factorization is needed in

3.3.3 Proof of Theorem 3.2.1

In order to focus the discussion we briefly review the procedure of the proof.

- Referring to Lemma 3.3.2, we first need to check the whether $J_{f,k}(x_0, \{u_j, y_j\}_{j=0}^k)$ has a minimum with respect to $\{x_0, \{u_j\}_{j=0}^k\}$. This is done via the Krein space Kalman filter corresponding to (3.3.6-3.3.7), and yields the condition (3.2.8) along with several equivalent conditions.

- Next we need to choose the $\{\check{s}_{k|k}\}_{k=0}^i$ such that the value of $J_{f,k}(x_0, \{u_j, y_j\}_{j=0}^k)$ is positive at its minimum. Now according to Theorem 2.7.4, the value at the minimum is $J_{f,k}(\min) = \sum_{j=0}^k e_j R_{e,j}^{-1} e_j$, where e_j is the innovations corresponding to (3.3.6-3.3.7). We can then compute the $\{e_j\}$ using the Krein space Kalman filter, and thereby choose the appropriate $\{\check{s}_{k|k}\}_{k=0}^i$ which yields the desired a posteriori filter.

A remark on the strong regularity of the model (3.3.6-3.3.7): In what follows we would like to use the Krein space Kalman filter corresponding to the state-space model (3.3.6-3.3.7). This of course requires the strong regularity of its output Gramian matrix, which we denote by $R_{\substack{y\\s}}$ (since the output of (3.3.6) consists of both a $\mathbf{y}$ and a $\mathbf{s}$ component).

If $R_{\substack{y\\s}}$ is strongly regular, then the Krein space Kalman filter may be applied to check for the positivity of $J_{f,k}$ for each $0 \leq k \leq i$. But what if $R_{\substack{y\\s}}$ is not strongly regular? Then it turns out that $J_{f,k}$ cannot be positive for all $0 \leq k \leq i$. To see why suppose that $J_{f,k} > 0$ for some arbitrary k. Then $J_{f,k}$ must have a minimum, and according to Lemma 2.6.4, the leading $k \times k$ block submatrices of $R_{\substack{y\\s}}$ and $R - S^*QS = R$ must have the same inertia. Now due to (3.3.8), all leading submatrices of R are nonsingular, and since k was arbitrary, the same will be true of $R_{\substack{y\\s}}$. Therefore $R_{\substack{y\\s}}$ will be strongly regular.

To summarize, we may use the Krein space Kalman filter to check the positivity of $J_{f,k}$. If one of the $R_{e,k}$ becomes singular (so that $R_{\substack{y\\s}}$ is no longer strongly regular),

the solution of the H^∞ estimation problem. Since the Krein space Kalman filter corresponding to (3.3.5) can be used to perform this canonical factorization, we will shortly (albeit via different route) see that it plays a crucial role in the solution.

$J_{f,k}$ will lose its positivity by default.[4]

Proof of existence condition (3.2.8):

The Riccati recursion corresponding to (3.3.6) is the exact same Riccati recursion that was given by (3.2.9) in Theorem 3.2.1. We can now apply *any* of the conditions for a minimum developed in Chapter 2, to check whether a minimum exists for $J_{f,k}(x_0, u_0, \ldots, u_k, y_0 \ldots, y_k)$ for all $0 \le k \le i$. If we assume that the $\begin{bmatrix} F_k & G_k \end{bmatrix}$ have full rank, then according to Lemma 2.7.4, $J_{f,k}(x_0, u_0, \ldots, u_k, y_0 \ldots, y_k)$ will have a minimum for all $0 \le k \le i$, if, and only if,

$$P_{j|j}^{-1} = P_j^{-1} + \begin{bmatrix} H_j^* & L_j^* \end{bmatrix} \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0} & -\gamma_f^2 I \end{bmatrix}^{-1} \begin{bmatrix} H_j \\ L_j \end{bmatrix} > 0$$

which yields the condition (3.2.8).

Since we still need to satisfy the second condition of Lemma 3.3.2, this, of course, only shows that (3.2.8) is a necessary condition for the existence of an H^∞ a posteriori filter of level γ_f. However, we shall later show that if the minimum condition is satisfied then the second condition of Lemma 3.3.2 can also be satisfied. Therefore (3.2.8) is indeed necessary and sufficient for the existence of the filter.

∎

Other Existence Conditions

Using the results Chapter 2, we can obtain alternative conditions for the existence of H^∞ a posteriori filters of level γ_f. If we use Lemma 2.7.3, we have the following condition.

[4]Another way to show the strong regularity of $R_{\underset{s}{y}}$ is to appeal to Theorem 1.4.1. There it was shown (in a more general setting) that an H^∞ filter exists if, and only if, $R_{\underset{s}{y}}$ admits a canonical factorization with a certain prescribed inertia. It is easy to check that the existence of this canonical factorization requires, at the very least, the strong regularity of $R_{\underset{s}{y}}$.

Lemma 3.3.3 (Alternative Test for Existence) *The condition (3.2.8) can be replaced by the condition that*

$$R_j = \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0} & -\gamma_f^2 I \end{bmatrix} \quad \text{and} \quad R_{e,j} = \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0} & -\gamma_f^2 I \end{bmatrix} + \begin{bmatrix} H_j \\ L_j \end{bmatrix} P_j \begin{bmatrix} H_j^* & L_j^* \end{bmatrix}$$

have the same inertia for all $0 \leq j \leq i$. We no longer require that $\begin{bmatrix} F_j & G_j \end{bmatrix}$ have full rank, and the size of the matrices involved is generally smaller than in (3.2.8).

Using a block triangular factorization of $R_{e,j}$, and the fact that when we have a minimum, P_j is positive definite, we can show the following result.

Corollary 3.3.1 (Alternative Test for Existence) *The condition of Lemma 3.3.3 is equivalent to*

$$I + H_j P_j H_j^* > 0 \quad \text{and} \quad -\gamma_f^2 I + L_j (P_j^{-1} + H_j^* H_j)^{-1} L_j^* < 0, \tag{3.3.9}$$

for all $0 \leq j \leq i$.

The test of Lemma 3.3.3 has various advantages over (3.2.8) that are mentioned in the discussions following Lemma 2.7.4. In particular, Lemma 3.3.3 allows us to go to a square-root form of the H^∞ filtering algorithm, where there is no need to explicitly check for the existence condition - these conditions are built into the the square-root recursions themselves, so that a solution exists if, and only if, the algorithm can be performed (see [HSK94c] and Chapter 5).

Many alternative existence conditions can also be obtained. Here is one that follows Lemma 2.7.5.

Lemma 3.3.4 (Alternative Test for Existence) *If the $\{F_j\}$ are nonsingular, an H^∞ a posteriori filter of level γ_f exists, if, and only if,*

$$P_{i+1} > 0,$$

and

$$I - G_j^* P_{j+1}^{-1} G_j > 0 \quad j = 0, 1, \ldots, i.$$

Construction of the H^∞ a posteriori filters

To complete the proof of Theorem 3.2.1 we still need to show that if a minimum over $\{x_0, u_0, \ldots, u_k\}$ exists for all $0 \le k \le i$, then we can find the estimates $\{\check{s}_{k|k}\}_{k=0}^{i}$ such that the value of $J_{f,k}(x_0, u_0, \ldots, u_k, y_0, \ldots, y_k)$ at its minimum is positive.

According to Theorem 2.7.4, the minimum value of $J_{f,k}(x_0, u_0, \ldots, u_k, y_0, \ldots, y_k)$ is

$$\sum_{j=0}^{k} \begin{bmatrix} e_{y,j}^* & e_{s,j}^* \end{bmatrix} R_{e,j}^{-1} \begin{bmatrix} e_{y,j} \\ e_{s,j} \end{bmatrix} =$$

$$\sum_{j=0}^{k} \begin{bmatrix} y_j - \hat{y}_{j|j-1} \\ \check{s}_{j|j} - \hat{s}_{j|j-1} \end{bmatrix}^* \begin{bmatrix} I + H_j P_j H_j^* & H_j P_j L_j^* \\ L_j P_j H_j^* & -\gamma_f^2 I + L_j P_j L_j^* \end{bmatrix}^{-1} \begin{bmatrix} y_j - \hat{y}_{j|j-1} \\ \check{s}_{j|j} - \hat{s}_{j|j-1} \end{bmatrix}$$

where $\hat{y}_{j|j-1}$ and $\hat{s}_{j|j-1}$ are obtained from the Krein space projections of $\mathbf{y}_j$ and $\check{\mathbf{s}}_{j|j}$ onto $\mathcal{L}\left\{\{\mathbf{y}_l\}_{l=0}^{j-1}, \{\check{\mathbf{s}}_{l|l}\}_{l=0}^{j-1}\right\}$, respectively. Thus $\hat{s}_{j|j-1}$ is a linear function of $\{y_l\}_{l=0}^{j-1}$.

Using the block triangular factorization of the $R_{e,j}$ we may rewrite the above as

$$\sum_{j=0}^{k} \begin{bmatrix} y_j - \hat{y}_{j|j-1} \\ \check{s}_{j|j} - \hat{s}_{j|j} \end{bmatrix}^* \begin{bmatrix} I + H_j P_j H_j^* & 0 \\ 0 & -\gamma_f^2 I + L_j (P_j^{-1} + H_j^* H_j)^{-1} L_j^* \end{bmatrix}^{-1} \begin{bmatrix} y_j - \hat{y}_{j|j-1} \\ \check{s}_{j|j} - \hat{s}_{j|j} \end{bmatrix}$$

where

$$\begin{bmatrix} y_j - \hat{y}_{j|j-1} \\ \check{s}_{j|j} - \hat{s}_{j|j} \end{bmatrix} \triangleq \begin{bmatrix} I & 0 \\ -L_j P_j H_j^* (I + H_j P_j H_j^*)^{-1} & I \end{bmatrix} \begin{bmatrix} y_j - \hat{y}_{j|j-1} \\ \check{s}_{j|j} - \hat{s}_{j|j-1} \end{bmatrix}. \tag{3.3.10}$$

Or in other words,

$$\hat{s}_{j|j} = \hat{s}_{j|j-1} + L_j P_j H_j^* (I + H_j P_j H_j^*)^{-1} (y_j - \hat{y}_{j|j-1}). \tag{3.3.11}$$

Note that $\hat{s}_{j|j}$ is obtained from the Krein space projection of $\check{\mathbf{s}}_{j|j}$ onto $\mathcal{L}\left\{\{\mathbf{y}_l\}_{l=0}^{j}, \{\check{\mathbf{s}}_{l|l}\}_{l=0}^{j-1}\right\}$, and is therefore a linear function of $\{y_l\}_{l=0}^{j}$. Recall from Corollary 3.3.1 that

$$I + H_j P_j H_j^* > 0 \quad \text{and} \qquad -\gamma_f^2 I + L_j (P_j^{-1} + H_j^* H_j)^{-1} L_j^* < 0. \tag{3.3.12}$$

Therefore all we must do is choose some $\check{s}_{j|j}$ such that

$$\begin{aligned} &\textstyle\sum_{j=0}^{k} (y_j - \hat{y}_{j|j-1})^* (I + H_j P_j H_j^*)^{-1} (y_j - \hat{y}_{j|j-1}) \\ &+ \textstyle\sum_{j=0}^{k} (\check{s}_{j|j} - \hat{s}_{j|j})^* \left(-\gamma_f^2 I + L_j (P_j^{-1} + H_j H_j^*)^{-1} L_j^*\right)^{-1} (\check{s}_{j|j} - \hat{s}_{j|j}) > 0. \end{aligned} \tag{3.3.13}$$

There are many such choices, but in view of (3.3.12), the simplest is

$$\begin{aligned}\check{s}_{j|j} = \hat{s}_{j|j} &= L_j\hat{x}_{j|j} \\ &= L_j\hat{x}_j + L_jP_jH_j^*(I+H_jP_jH_j^*)^{-1}(y_j - H_j\hat{x}_j)\end{aligned}$$

where $\hat{x}_j$ is the usual predicted state-variable of the Krein space Kalman filter, and where $\hat{x}_{j|j}$ is given by the Krein space projection of the state $\mathbf{x_j}$ onto $\left\{\{\mathbf{y}_l\}_{l=0}^{j}, \{\check{\mathbf{s}}_{l|l}\}_{l=0}^{j-1}\right\}$. Using the (predicted form) of the Krein space Kalman filter allows us to write

$$\left\{\begin{array}{rcl}\hat{x}_{j+1} &=& F_j\hat{x}_j + K_{p,j}\begin{bmatrix} y_j - H_j\hat{x}_j \\ \check{s}_{j|j} - L_j\hat{x}_j\end{bmatrix} \\ \check{s}_{j|j} &=& L_j\hat{x}_j + L_jP_jH_j^*(I+H_jP_jH_j^*)^{-1}(y_j - H_j\hat{x}_j)\end{array}\right. . \tag{3.3.14}$$

Replacing the second of the above equations into the first will allow us to obtain a more explicit recursion for $\hat{x}_j$, and hence a more explicit formula for $\check{s}_{j|j}$. However, an easier route is to utilize the filtered form of the Krein space Kalman filter corresponding to the state-space model (3.3.6) to recursively compute $\hat{x}_{j|j}$ (see Corollary 2.7.1),

$$\hat{x}_{j+1|j+1} = F_j\hat{x}_{j|j} + P_{j+1}\begin{bmatrix} H_{j+1}^* & L_{j+1}^*\end{bmatrix} R_{e,j+1}^{-1}\begin{bmatrix} y_{j+1} - \hat{y}_{j+1|j} \\ \check{s}_{j+1|j+1} - \hat{s}_{j+1|j}\end{bmatrix} \tag{3.3.15}$$

Using $\hat{y}_{j+1|j} = H_{j+1}F_j\hat{x}_{j|j}$ and the above mentioned triangular factorization of $R_{e,j+1}$ we have

$$\begin{aligned}\hat{x}_{j+1|j+1} &= F_j\hat{x}_{j|j} + P_{j+1}\begin{bmatrix} H_{j+1}^* & L_{j+1}^*\end{bmatrix}\begin{bmatrix} I & -(I+H_{j+1}P_{j+1}H_{j+1}^*)^{-1}H_{j+1}P_{j+1}L_{j+1}^* \\ 0 & I\end{bmatrix} \\ &\quad \begin{bmatrix} I + H_{j+1}P_{j+1}H_{j+1}^* & 0 \\ 0 & -\gamma_f^2 I + L_{j+1}(P_{j+1}^{-1} + H_{j+1}^*H_{j+1})^{-1}L_{j+1}^*\end{bmatrix}^{-1}\begin{bmatrix} y_{j+1} - H_{j+1}F_j\hat{x}_{j|j} \\ \check{s}_{j+1|j+1} - \hat{s}_{j+1|j+1}\end{bmatrix}\end{aligned}$$

Choosing $\check{s}_{j+1|j+1} = \hat{s}_{j+1|j+1}$ yields the desired recursion of Theorem 3.2.1

$$\hat{x}_{j+1|j+1} = F_j\hat{x}_{j|j} + P_{j+1}H_{j+1}^*(I + H_{j+1}P_{j+1}H_{j+1}^*)^{-1}(y_{j+1} - H_{j+1}F_j\hat{x}_{j|j}).$$

■

We mention in passing that using the canonical factorization of $R_{\underset{s}{y}}$ (via the Krein space Kalman filter), it is straightforward to show that the transfer operators $\mathcal{L}_{ij}$ of Theorem 1.4.1 have the following state-space representations

$$\begin{bmatrix}\mathcal{L}_{11} & \mathcal{L}_{12} \\ \mathcal{L}_{21} & \mathcal{L}_{22}\end{bmatrix} : \left\{\begin{array}{rcl} X_{i+1} &=& A_iX_i + B_iU_i \\ Y_i &=& C_iX_i + D_iU_i\end{array}\right. \tag{3.3.16}$$

where

$$\left\{\begin{array}{lcl} A_i &=& F_i \\ G_i &=& K_{p,i}\left[\begin{array}{cc} (I_p+H_iP_iH_i^*)^{1/2} & 0 \\ -L_iP_iH_i^*(I_p+H_iP_iH_i^*)^{-*/2} & (\gamma^2 I_q - L_iP_i(I+H_i^*H_iP_i)^{-1}L^*)^{1/2} \end{array}\right] \\ C_i &=& \left[\begin{array}{c} H_i \\ -L_i \end{array}\right] \\ D_i &=& \left[\begin{array}{cc} I_p & 0 \\ 0 & I_q \end{array}\right] \end{array}\right. . \tag{3.3.17}$$

Using the above state-space representation, it is straightforward to show (though we shall not provide the algebraic manipulations here) that the central filter, $\mathcal{K}_{cen} = -\mathcal{L}_{21}\mathcal{L}_{11}^{-1}$, has state-space model given by the recursions of Theorem 3.2.1. It is also possible to use the $\mathcal{L}_{ij}$ (obtained above) to parametrize all possible H^∞ filters of level γ along the lines of Theorem 1.4.1. Here, however, we shall take a different route to this parametrization which has the benefit of allowing for nonlinear estimators.

3.3.4 Parametrization of All H^∞ A Posteriori Filters

The filter of Theorem 3.2.1 is one among *many possible* filters with level γ. All filters that guarantee $J_{f,k}(x_0, u_0, \ldots, u_k, y_0, \ldots, y_k) > 0$ are represented by (3.3.12) and (3.3.13). It is useful to formalize this representation in the following result.

Lemma 3.3.5 (All H^∞ A Posteriori Estimators) *All H^∞ a posteriori estimators that achieve a level γ_f (assuming they exist) are given by any $\check{s}_{j|j} = \mathcal{F}_{f,j}(y_0, \ldots y_j)$ that satisfy*

$$\sum_{j=0}^{k}\left[\begin{array}{c} y_j - H_j\hat{x}_j \\ \check{s}_{j|j} - L_j\hat{x}_j \end{array}\right]^* R_{e,j}^{-1}\left[\begin{array}{c} y_j - H_j\hat{x}_j \\ \check{s}_{j|j} - L_j\hat{x}_j \end{array}\right] \geq 0, \quad 0 \leq k \leq i \tag{3.3.18}$$

where $\hat{x}_j$ satisfies the recursion,

$$\hat{x}_{j+1} = F_j\hat{x}_j + K_{p,j}\left[\begin{array}{c} y_j - H_j\hat{x}_j \\ \check{s}_{j|j} - L_j\hat{x}_j \end{array}\right], \qquad \hat{x}_0 \tag{3.3.19}$$

with

$$K_{p,j} = F_j P_j \begin{bmatrix} H_j^* & L_j^* \end{bmatrix} R_{e,j}^{-1} \quad \text{and} \quad R_{e,j} = \begin{bmatrix} I_p & 0 \\ 0 & -\gamma_f^2 I_q \end{bmatrix} + \begin{bmatrix} H_j \\ L_j \end{bmatrix} P_j \begin{bmatrix} H_j^* & L_j^* \end{bmatrix} \tag{3.3.20}$$

and

$$P_{j+1} = F_j P_j F_j^* + G_j Q_j G_j^* - K_{p,j} R_{e,j} K_{p,j}^*, \quad P_0 = \Pi_0. \tag{3.3.21}$$

We may also use these expressions to obtain a more explicit characterization of all possible estimators. Similar results appear in [KN91, ST92, GL95].

Theorem 3.3.1 (All H^∞ A Posteriori Estimators) *All H^∞ a posteriori estimators that achieve a level γ_f (assuming they exist) are given by*

$$\begin{aligned} \check{s}_{j|j} &= L_j \hat{x}_{j|j} + [\gamma_f^2 I - L_j (P_j^{-1} + H_j^* H_j)^{-1} L_j^*]^{\frac{1}{2}} \\ &\quad \mathcal{S}_j \left((I + H_j P_j H_j^*)^{\frac{1}{2}} (y_j - H_j \hat{x}_{j|j}), \ldots, (I + H_0 P_0 H_0^*)^{\frac{1}{2}} (y_0 - H_0 \hat{x}_{0|0}) \right) \end{aligned} \tag{3.3.22}$$

where $\hat{x}_{j|j}$ satisfies the recursion

$$\hat{x}_{j+1|j+1} = F_j \hat{x}_{j|j} + K_{s,j+1}(y_{j+1} - H_{j+1} F_j \hat{x}_{j|j}) - K_{c,j}(\check{s}_{j|j} - L_j \hat{x}_{j|j}) \tag{3.3.23}$$

with $K_{s,j+1}$ the same as in Theorem 3.2.1,

$$K_{c,j} = (I + P_{j+1} H_{j+1} H_{j+1}^*)^{-1} F_j (P_j^{-1} + H_j H_j^* - \gamma_f^{-2} L_j L_j^*)^{-1} L_j^*, \tag{3.3.24}$$

and

$$\mathcal{S}(a_j, \ldots, a_0) = \begin{bmatrix} \mathcal{S}_0(a_0) \\ \mathcal{S}_1(a_1, a_0) \\ \vdots \\ \mathcal{S}_j(a_j, \ldots, a_0) \end{bmatrix}$$

is any (possibly nonlinear) contractive causal mapping, i.e.,

$$\sum_{j=0}^{k} |\mathcal{S}_j(a_j, \ldots, a_0)|^2 < \sum_{j=0}^{k} |a_j|^2 \quad \text{for all} \quad k = 0, 1, \ldots, i.$$

Remark: Note that when the contraction of Theorem 3.3.1 is chosen as $\mathcal{S}=0$, then we have $\check{s}_{j|j}=L_j\hat{x}_{j|j}$, and (3.3.23) reduces to the recursion of Theorem 3.2.1.

Proof of Theorem 3.3.1: Expression (3.3.13) may be rewritten as:

$$\begin{array}{c}\sum_{j=0}^{k}(y_j-H_j\hat{x}_j)^*(I+H_jP_jH_j^*)^{-1}(y_j-H_j\hat{x}_j)\\ +\sum_{j=0}^{k}(\check{s}_{j|j}-L\hat{x}_{j|j})^*\left(-\gamma_f^2I+L_j(P_j^{-1}+H_jH_j^*)^{-1}L_j^*\right)^{-1}(\check{s}_{j|j}-L_j\hat{x}_{j|j})>0\end{array} \tag{3.3.25}$$

where $\hat{x}_j$ and $\hat{x}_{j|j}$ denote the Krein space projections of $\mathbf{x_j}$ onto $\left\{\{\mathbf{y_l}\}_{l=0}^{j-1},\{\mathbf{\check{s}_{l|l}}\}_{l=0}^{j-1}\right\}$ and $\left\{\{\mathbf{y_l}\}_{l=0}^{j},\{\mathbf{\check{s}_{l|l}}\}_{l=0}^{j-1}\right\}$, respectively. Therefore $\hat{x}_j$ and $\hat{x}_{j|j}$ are related through one additional projection onto y_j, and we may write

$$\hat{x}_{j|j}=\hat{x}_j+P_jH_j^*(I+H_jP_jH_j^*)^{-1}(y_j-H_j\hat{x}_j). \tag{3.3.26}$$

Therefore

$$\begin{aligned} y_j-H_j\hat{x}_{j|j} &= \left(I-H_jP_jH_j^*(I+H_jP_jH_j^*)^{-1}\right)(y_j-H_j\hat{x}_j)\\ &= (I+H_jP_jH_j^*)^{-1}(y_j-H_j\hat{x}_j)\end{aligned}$$

so that

$$(y_j-H_j\hat{x}_j)=(I+H_jP_jH_j^*)(y_j-H_j\hat{x}_{j|j}).$$

Now (3.3.25) can be written as:

$$\begin{array}{c}\sum_{j=0}^{k}(y_j-H_j\hat{x}_{j|j})^*(I+H_jP_jH_j^*)(y_j-H_j\hat{x}_{j|j})\\ +\sum_{j=0}^{k}(\check{s}_{j|j}-L_j\hat{x}_{j|j})^*\left(-\gamma_f^2I+L_j(P_j^{-1}+H_jH_j^*)^{-1}L_j^*\right)^{-1}(\check{s}_{j|j}-L_j\hat{x}_{j|j})>0\end{array}$$

or equivalently,

$$\|\,(\gamma_f^2I-L_j(P_j^{-1}+H_jH_j^*)^{-1}L_j^*)^{-\frac{1}{2}}(\check{s}_{j|j}-\hat{s}_{j|j})\,\|_2^2\ <\ \|\,(I+H_jP_jH_j^*)^{\frac{1}{2}}(y_j-H_j\hat{x}_{j|j})\,\|_2^2\,.$$

Since $\check{s}_{j|j}$ is a causal function of the observations y_j, then $(\check{s}_{j|j}-\hat{s}_{j|j})$ will also be a causal function of $(y_j-H_j\hat{x}_{j|j})$. Therefore using the above expression, we can write

$$\begin{bmatrix}(\gamma_f^2I-L_0(P_0^{-1}+H_0H_0^*)^{-1}L_0^*)^{-\frac{1}{2}}(\check{s}_{0|0}-\hat{s}_{0|0})\\ \vdots\\ (\gamma_f^2I-L_i(P_i^{-1}+H_iH_i^*)^{-1}L_i^*)^{-\frac{1}{2}}(\check{s}_{i|i}-\hat{s}_{i|i})\end{bmatrix}=\mathcal{S}\begin{bmatrix}(I+H_0P_0H_0^*)^{\frac{1}{2}}(y_0-H_0\hat{x}_{0|0})\\ \vdots\\ (I+H_iP_iH_i^*)^{\frac{1}{2}}(y_i-H_i\hat{x}_{i|i})\end{bmatrix}$$

for some causal contractive mapping $\mathcal{S}$. Eq. (3.3.22) now readily follows.

Finally, we must show (3.3.23). To this end, recall from the proof of Theorem 3.2.1 (see (3.3.15)) that the recursion for $\hat{x}_j$ is given by

$$\hat{x}_{j+1|j+1} = F_j \hat{x}_{j|j} + P_{j+1} \begin{bmatrix} H_{j+1}^* & L_{j+1}^* \end{bmatrix} R_{e,j+1}^{-1} \begin{bmatrix} y_{j+1} - \hat{y}_{j+1|j} \\ \check{s}_{j+1|j+1} - \hat{s}_{j+1|j} \end{bmatrix}.$$

Using (3.3.26) to replace $\hat{x}_j$ by $\hat{x}_{j|j}$ yields, after some algebra, the desired recursion (3.3.23).

∎

Note that although the filter obtained in Theorem 3.2.1 is linear, the full parametrization of all H^∞ filters with level γ_f is given by a *nonlinear* causal contractive mapping $\mathcal{S}$.[5] The filter of Theorem 3.2.1 is known as the *central* filter and as we have seen, corresponds to $\mathcal{S} = 0$. This central filter has a number of other interesting properties. It corresponds, as we shall see in the next chapter, to the risk-sensitive optimal filter [Whi90], and, as we have seen in Chapter 1, to the *maximum entropy* filter [GM89]. Moreover, in the game theoretic formulation of the H^∞ problem, the central filter corresponds to the solution of the game [BB95]. In our context, the central filter is recognized as the Krein space Kalman filter corresponding to the state-space model (3.3.6).

3.3.5 Derivation of the A Priori H^∞ Filter

We shall now turn to the H^∞ a priori filter of Problem 3.2.2 and our main goal will be to prove the results of Theorem 3.2.2. Our approach will follow the one used for the a posteriori case, namely we will relate an indefinite quadratic form to the a priori problem, construct its corresponding Krein space state-space model, and use

[5]We should remark that, from the outset, we did not restrict ourselves to linear estimators. Therefore Theorem 3.2.1 states that whenever the H^∞ filtering problem of level γ is feasible there exists a linear filter that achieves it. This has led various authors to claim that nonlinear estimators and controllers have no advantage over linear ones when estimating or controlling linear plants (see *e.g.*, [GL95] page 216). [Incidentally, LQG-optimal controllers are also linear.] However, including nonlinear estimators in the above parametrization may prove to be of use when attempting to optimize some additional cost criterion over the set of all possible H^∞ filters achieving a certain level γ.

the Krein space Kalman filter to obtain the solution. Since our derivations parallel the ones given earlier, we shall omit several details.[6]

The Suboptimal H^∞ A Priori Problem and Quadratic Forms

Referring to Problem 3.2.2 we first note that $\|T_i(\mathcal{F}_p)\|_\infty < \gamma_p$, implies that for all nonzero $\{x_0, \{u_j\}_{j=0}^{i-1}, \{v_j\}_{j=0}^{i-1}\}$

$$\frac{\sum_{j=0}^{i} |\check{s}_j - L_j x_j|^2}{x_0^* \Pi_0^{-1} x_0 + \sum_{j=0}^{i-1} |u_j|^2 + \sum_{j=0}^{i-1} |y_j - H_j x_j|^2} < \gamma_p^2, \tag{3.3.28}$$

where, without loss of generality, we have assumed $\check{x}_0 = 0$. Moreover, (3.3.28) implies that for *all* $k \leq i$, we must have

$$\frac{\sum_{j=0}^{k} |\check{s}_j - L_j x_j|^2}{x_0^* \Pi_0^{-1} x_0 + \sum_{j=0}^{k-1} |u_j|^2 + \sum_{j=0}^{k-1} |y_j - H_j x_j|^2} < \gamma_p^2. \tag{3.3.29}$$

As before, we may easily show the following result.

Lemma 3.3.6 (Indefinite Quadratic Form) *Given $\gamma_p > 0$, then $\|T_i(\mathcal{F}_p)\|_\infty < \gamma_p$ if, and only if, there exists $\check{s}_k = \mathcal{F}_p(y_0, \ldots, y_{k-1})$ (for all $0 \leq k \leq i$) such that for*

[6] Returning to the general estimation problem of Sec. 1.2, and using the canonical factorization approach of Sec. 1.4, it is possible to show that a *strictly* causal estimator of level γ exists if, and only if, there exists a canonical factorization of the form

$$\begin{bmatrix} I + \mathcal{H}\mathcal{H}^* & -\mathcal{H}\mathcal{L}^* \\ -\mathcal{L}\mathcal{H}^* & -\gamma^2 I + \mathcal{L}\mathcal{L}^* \end{bmatrix} = \begin{bmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} \\ \mathcal{L}_{21} & \mathcal{L}_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} \mathcal{L}_{11}^* & \mathcal{L}_{21}^* \\ \mathcal{L}_{12}^* & \mathcal{L}_{22}^* \end{bmatrix},$$

with $\begin{bmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} \\ \mathcal{L}_{21} & \mathcal{L}_{22} \end{bmatrix}$ and $\mathcal{L}_{11}$ causal and causally invertible, and $\mathcal{L}_{21}$ strictly causal. [Recall that in Theorem 1.4.1 we required $\mathcal{L}_{21}$ to be strictly causal.] If this is the case, then all possible strictly causal H^∞ estimators of level γ are given by

$$\mathcal{K} = (\mathcal{L}_{22}\mathcal{Q} - \mathcal{L}_{21})(\mathcal{L}_{11} - \mathcal{L}_{12}\mathcal{Q})^{-1}, \tag{3.3.27}$$

where $\mathcal{Q}$ is any strictly causal and strictly contractive operator. The central solution, $\mathcal{K}_{cen} = -\mathcal{L}_{21}\mathcal{L}_{11}^{-1}$, results from the choice, $\mathcal{Q} = 0$.

As before, when we have state-space structure, the canonical factorization can be performed via an appropriate Krein space Kalman filter. We shall not present this approach to the H^∞ a priori filtering problem here, but shall instead present the (closely related) approach that relies on indefinite quadratic forms.

all complex vectors x_0, for all causal sequences $\{u_j\}_{j=0}^{i-1}$ and for all nonzero causal sequences $\{y_j\}_{j=0}^{i-1}$ the scalar quadratic form

$$\begin{array}{c} J_{p,k}(x_0, u_0, \ldots, u_{k-1}, y_0, \ldots, y_{k-1}) = x_0^*\Pi_0^{-1}x_0 + \sum_{j=0}^{k-1} u_j^* u_j + \\ \sum_{j=0}^{k-1}(y_j - H_j x_j)^*(y_j - H_j x_j) - \gamma_p^{-2} \sum_{j=0}^{k}(\check{s}_j - L_j x_j)^*(\check{s}_j - L_j x_j) \end{array} \tag{3.3.30}$$

satisfies

$$J_{p,k}(x_0, u_0, \ldots, u_{k-1}, y_0, \ldots, y_{k-1}) > 0 \quad \textit{for all} \quad 0 \le k \le i. \tag{3.3.31}$$

We can also readily obtain the analog of Lemma 3.3.2.

Lemma 3.3.7 (Positivity Condition) *The scalar quadratic forms*

$$J_{p,k}(x_0, u_0, \ldots, u_{k-1}, y_0, \ldots, y_{k-1}),$$

satisfy the conditions (3.3.31), if, and only if, for all $0 \le k \le i$,

(i) $J_{p,k}(x_0, u_0, \ldots, u_{k-1}, y_0, \ldots, y_{k-1})$ has a minimum with respect to $\{x_0, u_0, u_1, \ldots u_{k-1}\}$.

(ii) The $\{\check{s}_k\}_{k=0}^{i}$ can be chosen such that the value of $J_{p,k}(x_0, u_0, \ldots, u_{k-1}, y_0, \ldots, y_{k-1})$ at this minimum is positive, viz.,

$$\min_{\{x_0, u_0, \ldots, u_{k-1}\}} J_{p,k}(x_0, u_0, \ldots, u_{k-1}, y_0, \ldots, y_{k-1}, \check{s}_0, \ldots, \check{s}_k) > 0.$$

A Krein Space State-Space Model

Due to the fact that the summations in $J_{p,k}$ go up to both k and $k-1$ (see (3.3.30)), it is slightly more difficult to come up with a Krein state-space model whose corresponding quadratic form is $J_{p,k}$. However, with some effort, we see that the appropriate Krein state-space model is

$$\begin{cases} \xi_{2j+1} &= \xi_{2j}, \quad \xi_0 = \mathbf{x}_0 \\ \hat{\mathbf{s}}_j &= L_j \xi_{2j} + \mathbf{v}_{2j} \end{cases} \qquad j \le 0 \tag{3.3.32}$$
$$\begin{cases} \xi_{2j+2} &= F_j \xi_{2j+1} + G_j \bar{\mathbf{u}}_{2j+1} = \mathbf{x}_{2i+2} \\ \mathbf{y}_j &= H_j \xi_{2j+1} + \mathbf{v}_{2j+1} \end{cases}$$

where $\Pi_0 > 0$, $Q_{2j} = 0$, $Q_{2j+1} = I$, $R_{2j} = -\gamma_p^2 I$, $R_{2j+1} = I$, and $S_j = 0$. To see why, let us construct the deterministic quadratic form corresponding to (3.3.32). Thus:

$$\begin{aligned} J_{\xi,2k} &= \xi_0^* \Pi_0^{-1} \xi + \sum_{j=0}^{2k} \bar{u}_j^* Q_j^{-1} \bar{u}_j + \sum_{j=0}^{2k} v_j^* R_j^{-1} v_j \\ &= \xi_0^* \Pi_0^{-1} \xi + \sum_{j=0}^{k-1} \bar{u}_{2j+1}^* \bar{u}_{2j+1} + \sum_{j=0}^{k-1} v_{2j+1}^* R_{2j+1}^{-1} v_{2j+1} + \sum_{j=0}^{k} v_{2j}^* R_{2j}^{-1} v_{2j} \\ &= \xi_0^* \Pi_0^{-1} \xi + \sum_{j=0}^{k-1} \bar{u}_{2j+1}^* \bar{u}_{2j+1} + \sum_{j=0}^{k-1} |y_j - H_j \xi_{2j+1}|^2 - \gamma_p^{-2} \sum_{j=0}^{k} |\check{s}_j - L_j \xi_{2j}|^2. \end{aligned}$$

From (3.3.32) we see that $\xi_{2j} = \xi_{2j+1} = x_j$. Using this fact, and defining $\bar{u}_{2j+1} = u_j$, we readily see that $J_{\xi,2k} = J_{p,k}$.

Note also that the Riccati recursion for the model (3.3.32) is

$$\left\{ \begin{array}{lcl} \Sigma_{2j+1} &=& \Sigma_{2j} - \Sigma_{2j} L_j^* (-\gamma_p^2 I + L_j \Sigma_{2j} L_j^*)^{-1} L_j \Sigma_{2j} \\ \Sigma_{2j+2} &=& F_j \Sigma_{2j+1} F_j^* + G_j G_j^* - F_j \Sigma_{2j+1} H_j^* (I + H_j \Sigma_{2j+1} H_j^*)^{-1} H_j \Sigma_{2j+1} F_j^* \end{array} \right. \quad \Sigma_0 = \Pi_0. \tag{3.3.33}$$

Existence Conditions

Using Lemma 2.7.3, the condition for a minimum is that $R_{e,j}$ and R_j should have the same inertia for all $j = 0, 1, \ldots, 2i$ (since each two time steps in (3.3.32) correspond to one time step in $J_{p,k}$). Thus the condition for a minimum is

$$-\gamma_p^2 I + L_j \Sigma_{2j} L_j^* < 0 \quad \text{and} \quad I + H_j \Sigma_{2j+1} H_j^* > 0. \tag{3.3.34}$$

The second of the above conditions is obvious since when we have a minimum Σ_j is positive definite. If the $\left[\begin{array}{cc} F_j & G_j \end{array} \right]$ have full rank, then using Lemma 2.7.4, the condition for a minimum is

$$\Sigma_{2j}^{-1} - \gamma_p^{-2} L_j^* L_j > 0 \quad \text{and} \quad \Sigma_{2j+1}^{-1} + H_j^* H_j > 0 \tag{3.3.35}$$

where once more the second of the above conditions is redundant.

In order to connect with the results of Theorem 3.2.2, we may note that by defining $P_j = \Sigma_{2j}$ and combining the coupled pair of Riccati recursions in (3.3.33), we can write the following Riccati recursion for P_j

$$P_{j+1} = F_jP_jF_j^* + G_jG_j^* - F_jP_j \begin{bmatrix} L_j^* & H_j^* \end{bmatrix} R_{e,j}^{-1} \begin{bmatrix} L_j \\ H_j \end{bmatrix} P_jF_j^*, \quad P_0 = \Pi_0 \qquad (3.3.36)$$

with

$$R_{e,j} = \begin{bmatrix} -\gamma_p^2 I & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} L_j \\ H_j \end{bmatrix} P_j \begin{bmatrix} L_j^* & H_j^* \end{bmatrix}. \qquad (3.3.37)$$

But this is the same Riccati as (3.2.9) in Theorem 3.2.2. Thus the condition for a minimum (3.3.35) becomes

$$P_j^{-1} - \gamma_p^{-2} L_j^* L_j > 0, \qquad (3.3.38)$$

which is the condition (3.2.13) of Theorem 3.2.2.

To express the condition (3.3.34) in a form that is more similar to that of Lemma 3.3.3, we introduce the Krein space state-space model

$$\left\{ \begin{array}{rcl} \mathbf{x}_{j+1} & = & F_j\mathbf{x}_j + G_j\mathbf{u}_j \\ \begin{bmatrix} \check{\mathbf{s}}_{j|j} \\ \mathbf{y}_j \end{bmatrix} & = & \begin{bmatrix} L_j \\ H_j \end{bmatrix} \mathbf{x}_j + \mathbf{v}_j \end{array} \right. \qquad (3.3.39)$$

where $\Pi_0 > 0$, $Q_j = I$, $S_j = 0$ and

$$R_j = \begin{bmatrix} -\gamma_p^2 I & 0 \\ 0 & I \end{bmatrix}.$$

Note that the only difference between the state-space models (3.3.6) and (3.3.39) is that the order of the output equations has been reversed.

We can now use the state-space model (3.3.39) to express the condition (3.3.34) in the form of the following Lemma.

Lemma 3.3.8 (Alternative Test for Existence) *The condition (3.2.13) can be replaced by the condition that all leading submatrices of*

$$R_j = \begin{bmatrix} -\gamma_p^2 I & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix} \quad \textit{and} \quad R_{e,j} = \begin{bmatrix} -\gamma_p^2 I & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix} + \begin{bmatrix} L_j \\ H_j \end{bmatrix} P_j \begin{bmatrix} L_j^* & H_j^* \end{bmatrix}$$

have the same inertia for all $0 \le j \le i$. In other words,

$$-\gamma_p^2 I + L_j P_j L_j^* < 0 \quad \textit{and} \quad I + H_j \tilde{P}_j H_j^* > 0$$

where $\tilde{P}_j^{-1} = P_j^{-1} - \gamma_p^{-2} L_j^* L_j$. We no longer require that $\begin{bmatrix} F_j & G_j \end{bmatrix}$ have full rank, and the size of the matrices involved is generally smaller than in (3.2.8).

Note that compared to Lemma 3.3.3, the condition in Lemma 3.3.8 is more stringent since it requires that all leading submatrices of R_j and $R_{e,j}$ have the same inertia. This distinction is especially important in square-root implementations of the H^∞ filters (see [HSK94c] and Chapter 5).

Construction of the H^∞ a priori filters

To complete the proof of Theorem 3.2.2 we still need to show that if a minimum over $\{x_0, u_0, \ldots, u_{k-1}\}$ exists for all $0 \le k \le i$, then we can find the estimates $\{\check{s}_k\}_{k=0}^{i}$ such that the value of $J_{p,k}(x_0, u_0, \ldots, u_{k-1}, y_0, \ldots, y_{k-1})$ at its minimum is positive.

According to Theorem 2.7.4, the minimum value of $J_{p,k}(x_0, u_0, \ldots, u_{k-1}, y_0, \ldots, y_{k-1})$ is

$$\sum_{j=0}^{k-1} \begin{bmatrix} e_{s,j}^* & e_{y,j}^* \end{bmatrix} R_{e,j}^{-1} \begin{bmatrix} e_{s,j} \\ e_{y,j} \end{bmatrix} + e_{s,k}^*(-\gamma_p^2 I + L_k P_k L_k^*)^{-1} e_{s,k} =$$

$$\sum_{j=0}^{k-1} \begin{bmatrix} \check{s}_j - \hat{s}_{j|j-1} \\ y_j - \hat{y}_{j|j-1} \end{bmatrix}^* \begin{bmatrix} -\gamma_p^2 I + L_j P_j L_j^* & L_j P_j H_j^* \\ H_j P_j L_j^* & I + H_j P_j H_j^* \end{bmatrix}^{-1} \begin{bmatrix} \check{s}_j - \hat{s}_{j|j-1} \\ y_j - \hat{y}_{j|j-1} \end{bmatrix}$$
$$+(\check{s}_k - \hat{s}_{k|k-1})^*(-\gamma_p^2 I + L_k P_k L_k^*)^{-1}(\check{s}_k - \hat{s}_{k|k-1}) > 0$$

where $\hat{s}_{j|j-1}$ and $\hat{y}_{j|j-1}$ are obtained from the Krein space projections of $\check{\mathbf{s}}_j$ and $\mathbf{y}_j$ onto $\mathcal{L}\left\{\{\check{\mathbf{s}}_l\}_{l=0}^{j-1}, \{\mathbf{y}_l\}_{l=0}^{j-1}\right\}$, respectively. Thus $\hat{s}_{j|j-1}$ is a linear function of $\{y_l\}_{l=0}^{j-1}$. Using the block triangular factorization of the $R_{e,j}$ we may rewrite the above as

$$\begin{aligned} &\textstyle\sum_{j=0}^{k}(\check{s}_j - \hat{s}_{j|j-1})^*(-\gamma_p^2 I + L_j P_j L_j^*)^{-1}(\check{s}_j - \hat{s}_{j|j-1}) + \\ &\textstyle\sum_{j=0}^{k-1}(y_j - \bar{y}_{j|j-1})^*(I + H_j \tilde{P}_j H_j^*)^{-1}(y_j - \bar{y}_{j|j-1}) > 0 \end{aligned} \tag{3.3.40}$$

where

$$\begin{bmatrix} \check{s}_j - \hat{s}_{j|j-1} \\ y_j - \bar{y}_{j|j-1} \end{bmatrix} = \begin{bmatrix} I & 0 \\ -H_j P_j L_j^*(-\gamma_p^2 I + L_j P_j L_j^*)^{-1} & I \end{bmatrix} \begin{bmatrix} \check{s}_j - \hat{s}_{j|j-1} \\ y_j - \hat{y}_{j|j-1} \end{bmatrix}. \tag{3.3.41}$$

Or, in other words,

$$\bar{y}_{j|j-1} = \hat{y}_{j|j-1} + H_j P_j L_j^*(-\gamma_p^2 I + L_j P_j L_j^*)^{-1}(\check{s}_j - \hat{s}_{j|j-1}).$$

Note that $\bar{y}_{j|j-1}$ is given by the Krein space projection of $\mathbf{y}_j$ onto $\left\{\{\check{\mathbf{s}}_l\}_{l=0}^{j}, \{\mathbf{y}_l\}_{l=0}^{j-1}\right\}$. Recall from Lemma 3.3.8 that

$$-\gamma_p^2 I + L_j P_j L_j^* < 0 \quad , \quad I + H_j \tilde{P}_j H_j^* > 0.$$

Any choice of $\check{s}_{j|j-1}$ that renders (3.3.40) positive will do, and the simplest choice is $\check{s}_{j|j-1} = \hat{s}_{j|j-1} = L_j \hat{x}_{j|j-1}$, where $\hat{x}_{j|j-1}$ is given by the Krein space projection of $\mathbf{x}_j$ onto $\left\{\{\check{\mathbf{s}}_l\}_{l=0}^{j-1}, \{\mathbf{y}_j\}_{l=0}^{j-1}\right\}$. We may now utilize the Krein space Kalman filter corresponding to the state-space model (3.3.39) to recursively compute $\hat{x}_{j|j-1}$, *viz.*,

$$\hat{x}_{j+1|j} = F_j \hat{x}_{j|j-1} + F_j P_j \begin{bmatrix} L_j^* & H_j^* \end{bmatrix} R_{e,j}^{-1} \begin{bmatrix} \check{s}_j - L_j \hat{x}_{j|j-1} \\ y_j - H_j \hat{x}_{j|j-1} \end{bmatrix} \tag{3.3.42}$$

Setting $\check{s}_j - L_j \hat{x}_{j|j-1} = 0$ and simplifying we get the desired recursion for $\hat{x}_{j+1|j}$.

■

We should also mention here that in the a priori case the $\mathcal{L}_{ij}$ of the canonical spectral factorization are given by the following state-space representations,

$$\begin{bmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} \\ \mathcal{L}_{21} & \mathcal{L}_{22} \end{bmatrix} : \quad \left\{ \begin{array}{rcl} X_{i+1} & = & A_i X_i + B_i U_i \\ Y_i & = & C_i X_i + D_i U_i \end{array} \right. \tag{3.3.43}$$

where

$$\left\{ \begin{array}{rcl} A_i & = & F_i \\ G_i & = & K_{p,i} \begin{bmatrix} (I_p + H_i \tilde{P}_i H_i^*)^{1/2} & -H_i P_i L_i^*(-\gamma^2 I_q + L_i P_i L_i^*)^{-*/2} \\ 0 & (\gamma^2 I_q - L_i P_i L_i^*)^{1/2} \end{bmatrix} \\ C_i & = & \begin{bmatrix} H_i \\ -L_i \end{bmatrix} \\ D_i & = & \begin{bmatrix} I_p & 0 \\ 0 & I_q \end{bmatrix} \end{array} \right. . \tag{3.3.44}$$

Note now that the transfer operator, $\mathcal{L}_{21}$ is strictly causal. Using the above state-space representation, it is straightforward to show that the central filter, $\mathcal{K}_{cen} = -\mathcal{L}_{21}\mathcal{L}_{11}^{-1}$, has state-space model given by the recursions of Theorem 3.2.2. It is also possible to use the $\mathcal{L}_{ij}$ (obtained above) to parametrize all possible H^∞ filters of level γ, however, we shall take a route similar to the one taken in Sec. 3.3.4.

3.3.6 All H^∞ A Priori Filters

The positivity condition (3.3.40) gives a full parametrization of all H^∞ a priori estimators. We thus have the following result.

Lemma 3.3.9 (All H^∞ A Priori Estimators) *All H^∞ a priori estimators that achieve a level γ_p (assuming they exist) are given by any $\check{s}_j = \mathcal{F}_{p,j}(y_0, \ldots y_{j-1})$ that satisfy*

$$\begin{array}{c} \sum_{j=0}^{k-1} \begin{bmatrix} y_j - H_j\hat{x}_j \\ \check{s}_j - L_j\hat{x}_j \end{bmatrix}^* R_{e,j}^{-1} \begin{bmatrix} y_j - H_j\hat{x}_j \\ \check{s}_j - L_j\hat{x}_j \end{bmatrix} \\ -(\check{s}_k - L_k\hat{x}_k)^*(\gamma_p^2 I_q - L_k P_k L_k^*)^{-1}(\check{s}_k - L_k\hat{x}_k) \geq 0, \end{array} \qquad 0 \leq k \leq i \tag{3.3.45}$$

where $\hat{x}_j$ satisfies the recursion,

$$\hat{x}_{j+1} = F_j\hat{x}_j + K_{p,j}\begin{bmatrix} y_j - H_j\hat{x}_j \\ \check{s}_j - L_j\hat{x}_j \end{bmatrix}, \quad \hat{x}_0 \tag{3.3.46}$$

with

$$K_{p,j} = F_j P_j \begin{bmatrix} H_j^* L_j^* \end{bmatrix} R_{e,j}^{-1} \quad \textit{and} \quad R_{e,j} = \begin{bmatrix} I_p & 0 \\ 0 & -\gamma_p^2 I_q \end{bmatrix} + \begin{bmatrix} H_j \\ L_j \end{bmatrix} P_j \begin{bmatrix} H_j^* & L_j^* \end{bmatrix} \tag{3.3.47}$$

and

$$P_{j+1} = F_j P_j F_j^* + G_j Q_j G_j^* - K_{p,j} R_{e,j} K_{p,j}^*, \quad P_0 = \Pi_0. \tag{3.3.48}$$

We can also give a more explicit parametrization as follows.

Theorem 3.3.2 (All H^∞ A Priori Estimators) *All H^∞ a priori estimators that achieve a level γ_p (assuming they exist) are given by*

$$\begin{aligned} \check{s}_j \;=\; & L_j\hat{x}_j + (\gamma_p^2 I - L_j P_j L_j^*)^{\frac{1}{2}} \\ & \mathcal{S}_j\left((I + H_{j-1}\tilde{P}_{j-1}H_{j-1}^*)^{-\frac{1}{2}}(y_{j-1} - H_{j-1}\bar{x}_{j-1}), \ldots, (I + H_0\tilde{P}_0 H_0^*)^{-\frac{1}{2}}(y_0 - H_0\bar{x}_0)\right) \end{aligned} \tag{3.3.49}$$

where

$$\bar{x}_k = \hat{x}_k + P_k L_k^*(-\gamma_p^2 I + L_k L_k^*)^{-1}(\check{s}_k - L_k \hat{x}_k), \tag{3.3.50}$$

$\hat{x}_j$ satisfies the recursion

$$\hat{x}_{j+1|j} = F_j \hat{x}_{j|j-1} + F_j P_j \left[\begin{array}{cc} L_j^* & H_j^* \end{array} \right] R_{e,j}^{-1} \left[\begin{array}{c} \check{s}_j - L_j \hat{x}_{j|j-1} \\ y_j - H_j \hat{x}_{j|j-1} \end{array} \right], \tag{3.3.51}$$

with P_j, $\tilde{P}_j$ and $R_{e,j}$ given by Theorem 3.2.2, and $\mathcal{S}$ is any (possibly nonlinear) contractive causal mapping.

Proof: Referring to (3.3.41), we see that the $\bar{y}_{j|j-1} = H_j \bar{x}_j$ differ from $\hat{y}_{j|j-1} = H_j \hat{x}_j$, via the additional projection onto $\check{s}_{j|j-1}$. Thus we can write

$$\bar{x}_{j|j-1} = \hat{x}_{j|j-1} + P_j L_j^*(-\gamma_p^2 I + L_j P_j L_j^*)^{-1}(\check{s}_{j|j-1} - \hat{s}_{j|j-1}),$$

which proves (3.3.50). Moreover from the proof of Theorem 3.2.2 (see (3.3.42)), the recursion for $\hat{x}_j$ is given by (3.3.51). Condition (3.3.40) can now be rewritten as

$$\sum_{j=0}^{k}(\check{s}_{j|j-1} - \hat{s}_{j|j-1})^*(-\gamma_p^2 I + L_j P_j L_j^*)^{-1}(\check{s}_{j|j-1} - \hat{s}_{j|j-1}) + \\ \sum_{j=0}^{k-1}(y_j - \bar{y}_{j|j-1})^*(I + H_j \tilde{P}_j H_j^*)^{-1}(y_j - \bar{y}_{j|j-1}) > 0 \tag{3.3.52}$$

and an argument similar to the one given in the proof of Theorem 3.3.1 will yield the desired result.

∎

3.4 The H^∞ Smoother

If instead of $e_{f,k}$ and $e_{p,k}$, which correspond to the a posteriori and a priori filters, respectively, we consider the *smoothed* error

$$e_{s,k} = \check{s}_{k|i} - L_k x_k, \quad k \leq i$$

where $\check{s}_{k|i} = \mathcal{F}_s(y_0, y_1, \ldots, y_i)$ is the estimate of s_k given all observations $\{y_j\}$ from time 0 until time i, we are led to the so-called H^∞ smoothers. Such estimators guarantee that the maximum energy gain from the disturbances $\left\{\Pi_0^{-1/2}(x_0 - \check{x}_0), \{u_j, v_j\}_{j=0}^i\right\}$ to the smoothing errors $\{e_{s,j}\}_{j=0}^i$ is bounded by γ_s, *i.e.*,

$$\sup_{x_0, u \in h^2, v \in h^2} \frac{\sum_{j=0}^i e_{s,j}^* e_{s,j}}{(x_0 - \check{x}_0)^* {\Pi_0}^{-1}(x_0 - \check{x}_0) + \sum_{j=0}^i u_j^* u_j + \sum_{j=0}^i v_j^* v_j} < \gamma_s^2. \tag{3.4.1}$$

Using an argument similar to the ones given before, we are led to the following quadratic form

$$\begin{gathered} J_{s,i}(x_0, u_0, \ldots, u_i, y_0, \ldots, y_i) = x_0 \Pi_0^{-1} x_0^* + \sum_{k=0}^i u_k^* u_k + \\ \sum_{k=0}^i (y_k - H_k x_k)^*(y_k - H_k x_k) - \gamma^{-2} \sum_{k=0}^i (\check{s}_{k|i} - L_k x_k)^*(\check{s}_{k|i} - L_k x_k). \end{gathered} \tag{3.4.2}$$

Note that the only difference between $J_{s,i}$ and $J_{f,i}$ is that $\check{s}_{k|k}$ has been replaced by $\check{s}_{k|i}$ (*i.e.*, filtered estimates have been replaced by smoothed estimates). Once more it can be shown that an H^∞ smoother of level γ_s will exist if, and only if, there exists some $\check{s}_{k|i}$ such that $J_{s,i} \geq 0$. The rather interesting result shown below, and which has already been pointed out in the literature (see *e.g.*, [KN91, ST92, SFB92]), is that one H^∞ smoother is given by the conventional H^2 smoother (which does not even depend on the value of γ_s).

Theorem 3.4.1 (H^∞ Smoother) *For a given $\gamma_s > 0$, an H^∞ smoother that achieves level γ_s exists, if, and only if, the block diagonal matrix*

$$R_e = R_{e,0} \oplus R_{e,1} \oplus \ldots \oplus R_{e,i},$$

where

$$R_{e,j} = \begin{bmatrix} I & 0 \\ 0 & -\gamma_s^2 \end{bmatrix} + \begin{bmatrix} H_j \\ L_j \end{bmatrix} P_j \begin{bmatrix} H_j^* & L_j^* \end{bmatrix}$$

and P_j is the same as in Theorem 3.2.1, has $(i+1)p$ positive eigenvalues and $(i+1)q$ negative eigenvalues. In other words, if, and only if,

$$In[R_e] = \begin{bmatrix} (i+1)p & 0 & (i+1)q \end{bmatrix}.$$

If this is the case, one possible H_∞ smoother is given by the H^2 smoother.

Proof: The condition for $J_{s,i}(x_0, u_0, \ldots, u_i, y_0, \ldots, y_i)$ to have a minimum is slightly different than the earlier cases, since we do *not* require that $J_{s,k}(x_0, u_0, \ldots, u_k, y_0, \ldots, y_k)$ have a minimum over the disturbances for all past values $k < i$. Thus using Lemma 2.6.4, the condition for a minimum over $\{x_0, u_0, \ldots, u_i\}$ is that the matrices

$$R_e \quad \text{and} \quad R = R_0 \oplus R_1 \oplus \ldots \oplus R_i$$

have the same inertia, where $R_j = I_p \oplus (-\gamma_s^2 I_q)$. But this is precisely the inertia condition given in the statement of the Theorem.

The value of $J_{s,i}$ at its minimum is (see Chapter 2)

$$\begin{bmatrix} y^* & \check{s}_{|i}^* \end{bmatrix} \begin{bmatrix} R_y & R_{y\check{s}} \\ R_{\check{s}y} & R_{\check{s}} \end{bmatrix}^{-1} \begin{bmatrix} y \\ \check{s}_{|i} \end{bmatrix}$$

where we have defined

$$\begin{bmatrix} R_y & R_{y\check{s}} \\ R_{\check{s}y} & R_{\check{s}} \end{bmatrix} = \langle \begin{bmatrix} \mathbf{y} \\ \check{\mathbf{s}}_{|i} \end{bmatrix}, \begin{bmatrix} \mathbf{y} \\ \check{\mathbf{s}}_{|i} \end{bmatrix} \rangle \tag{3.4.3}$$

with

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_0 \\ \vdots \\ \mathbf{y}_i \end{bmatrix}, \quad \check{\mathbf{s}}_{|i} = \begin{bmatrix} \check{\mathbf{s}}_{0|i} \\ \vdots \\ \check{\mathbf{s}}_{i|i} \end{bmatrix}, \quad y = \begin{bmatrix} y_0 \\ \vdots \\ y_i \end{bmatrix}, \quad \check{s}_{|i} = \begin{bmatrix} \check{s}_{0|i} \\ \vdots \\ \check{s}_{i|i} \end{bmatrix},$$

and where the $\{\mathbf{y}_j\}$ and $\{\check{\mathbf{s}}_{j|i}\}$ satisfy the Krein state-space model (3.3.6). In this case all the entries in $\check{s}_{|i}$ are unknown and there is no causal dependence between the $\{\check{s}_{j|k}\}$ and the $\{y_j\}$. Using a block triangular factorization, or a completion of squares argument, the value at the minimum can be rewritten as

$$y^* R_y^{-1} y + (\check{s}_{|i} - R_{\check{s}y} R_y^{-1} y)^* (R_{\check{s}} - R_{\check{s}y} R_y^{-1} R_{y\check{s}})^{-1} (\check{s}_{|i} - R_{\check{s}y} R_y^{-1} y).$$

But $R_y > 0$ (since it is the covariance of a Hilbert space state-space model), and hence one possible choice of $\check{s}_{|i}$ to guarantee $J_{s,i} > 0$ is to choose $\check{s}_{|i} = R_{\check{s}y} R_y^{-1} y = \hat{s}_{|i}$, which is clearly the H^2 smoothed estimate of $\check{s}$.

■

The following result is now straightforward.

Theorem 3.4.2 (All H^∞ Smoothers) *All H^∞ smoothers that achieve a level γ_s (assuming they exist) are given by*

$$\check{s}_{|i} = \hat{s}_{|i} + (-R_{\check{s}} + R_{\check{s}y}R_y^{-1}R_{y\check{s}})^{1/2}\mathcal{S}(R_y^{-1/2}y) \tag{3.4.4}$$

where $\mathcal{S}$ is any (not necessarily causal) contractive mapping, $\hat{s}_{|i}$ is the usual H^2 smoothed estimate, and R_y and $R_{\check{s}} - R_{\check{s}y}R_y^{-1}yR_{y\check{s}}$ are defined in (3.4.3).

It is clear from the discussions so far in this chapter that the Krein space estimation formalism provide simple derivations of H^∞ estimators. These estimators turn out to be certain Krein space Kalman filters, and show that Krein space estimation yields a unified approach to H^2 and H^∞ problems. To derive such filters, and to solve other related problems as discussed ahead, all one essentially needs is to identify an indefinite quadratic form, and to construct a convenient auxiliary state-space model with the appropriate Gramians. Before considering further applications in the next chapter, it will be useful to consider the problem of l-step H^∞ prediction.

3.5 l-step H^∞ Prediction

Consider once more the standard state-space model,

$$\begin{cases} x_{i+1} &= F_i x_i + G_i u_i, \qquad x_0 \\ y_i &= H_i x_i + v_i, \qquad i \geq 0 \\ s_i &= L_i x_i \end{cases} \tag{3.5.1}$$

where, as before, the initial condition x_0 and the disturbances $\{u_i, v_i\}$ are unknown quantities, y_i is the observed output, and s_i is the signal we intend to estimate. Suppose now that we would like to find l-step predictions of s_i, *i.e.*, estimates

$$\check{s}_{i|i-l} = \mathcal{F}_{l,i}(y_0 . \ldots, y_{i-l}), \tag{3.5.2}$$

that can only make use of the observations from time 0 to time $i-l$, where $l > 0$ is some given value. [Obviously, $l = 1$ corresponds to the a priori filtering problem just considered.] In this case, our estimation errors will be the l-step prediction errors,

$$e_{l,i} = L_i x_i - \check{s}_{i|i-l}. \tag{3.5.3}$$

We can now define the problem of finding l-step H^∞ predictors. Such predictors will have the property that the maximum energy gain from the disturbances

$$\left\{\Pi_0^{-1/2}(x_0-\check{x}_0),\{u_j\}_{j=0}^{i},\{v_j\}_{j=0}^{i-l}\right\},$$

to the prediction errors

$$\{e_{l,j}\}_{j=0}^{i},$$

is bounded by some given γ_l, *i.e.*,

$$\sup_{x_0,u\in h^2,v\in h^2}\frac{\sum_{j=0}^{i}e_{l,j}^*e_{l,j}}{(x_0-\check{x}_0)^*\Pi_0^{-1}(x_0-\check{x}_0)+\sum_{j=0}^{i}u_j^*u_j+\sum_{j=0}^{i-l}v_j^*v_j}<\gamma_l^2. \tag{3.5.4}$$

Note that in the above definition and inequality we have only included the measurement disturbances v_j until time $i-l$ since measurement disturbances after time $i-l$ only affect the observations, $\{y_j, j>i-l\}$, which the estimator is not allowed to use.

Before attempting to solve this H^∞ prediction problem, let us see what happens in the H^2 case. There to obtain, $\hat{x}_{i|i-l}$, one needs to project x_i onto $\mathcal{L}\{y_j\}_{j=0}^{i-l}$, *i.e.*, the linear space spanned by the random variables, $\{y_0,\ldots,y_{i-l}\}$. Using the state equation in (3.5.1) this leads to

$$\hat{x}_{j+1|i-l}=F_j\hat{x}_{j|i-l}+G_j\hat{u}_{j|i-l}. \tag{3.5.5}$$

Now if we assume $j>i-l$, then $\hat{u}_{j|i-l}=0$ (since u_j is independent of the observations, $\{y_k\}_{k=0}^{i-l}$). This implies that

$$\hat{x}_{j+1|i-l}=F_j\hat{x}_{j|i-l},\quad j>i-l. \tag{3.5.6}$$

In other words, the predictions $\hat{x}_{j|i-l}$ satisfy an "unforced" version of the same recursion that the states, x_j satisfy. Therefore to find $\hat{x}_{i|i-l}$, all one needs to do is propagate the above recursion starting with the predicted estimate, $\hat{x}_{i-l+1}$, as the initial condition. Thus, we may write

$$\hat{x}_{i|i-l}=F_{i-1}\ldots F_{i-l+1}\hat{x}_{i-l+1}. \tag{3.5.7}$$

We shall presently see that the above expression is no longer true for l-step H^∞ predictions. It turns out that the recursion that needs to be propagated is slightly more complicated.

To obtain the required H^∞ estimates let us first note that (3.5.4) implies the inequalities,

$$\sup_{x_0, u\in h^2, v\in h^2} \frac{\sum_{j=0}^{k} e_{l,j}^* e_{l,j}}{(x_0-\check{x}_0)^*{\Pi_0}^{-1}(x_0-\check{x}_0)+\sum_{j=0}^{k} u_j^* u_j + \sum_{j=0}^{k-l} v_j^* v_j} < \gamma_l^2, \tag{3.5.8}$$

for all $k = 0, \ldots, i$. This in turn implies that for all x_0 and for all $\{u_j\}_{j=0}^{i}$ (that are not simultaneously zero), we must be able to choose the $\check{s}_{j|j-l} = \mathcal{F}_{l,i}(y_0, \ldots, y_{j-l})$ such that

$$J_{l,k} = (x_0-\check{x}_0)^*{\Pi_0}^{-1}(x_0-\check{x}_0) + \sum_{j=0}^{k} u_j^* u_j + \sum_{j=0}^{k-l} v_j^* v_j - \gamma_l^{-2}\sum_{j=0}^{k} e_{l,j}^* e_{l,j} > 0, \tag{3.5.9}$$

for all $k = 0, \ldots, i$. Of course, the above is true if, and only if,

(i) $J_{l,k}$ has a minimum over the free variables $\{x_0, \{u_j\}_{j=0}k\}$, for all $k = 0, \ldots, i$.

(ii) The $\check{s}_{j|j-l} = \mathcal{F}_{l,i}(y_0, \ldots, y_{j-l})$ can be chosen such that the value of $J_{l,k}$ at this minimum is positive (for all $k = 0, \ldots, i$).

Let us now proceed to check and construct the above two conditions recursively. At time $i = 0$, we have need to guarantee

$$J_{l,k} = (x_0-\check{x}_0)^*{\Pi_0}^{-1}(x_0-\check{x}_0) + u_0^* u_0 - {\gamma_l}^{-2}(\check{s}_{0|-l} - L_0 x_0)^*(\check{s}_{0|-l} - L_0 x_0) > 0. \tag{3.5.10}$$

Now since $\Pi_0 > 0$, it is straightforward to see that we will have a minimum over the free variables $\{x_0, u_0\}$ if, and only if,

$$-\gamma_l^2 I_q + L_0 \Pi_0 L_0^* < 0. \tag{3.5.11}$$

Moreover, the value at the minimum is given by

$$(\check{s}_{0|-l} - L_0\check{x}_0)^* \left[-\gamma_l^2 I_q + L_0\Pi_0 L_0^*\right]^{-1} (\check{s}_{0|-l} - L_0\check{x}_0), \tag{3.5.12}$$

from which we readily infer that the (here) *unique* prediction is

$$\check{s}_{0|-l} = L_0\check{x}_0. \tag{3.5.13}$$

Assume now that we have solved the problem until time $k-1$, *i.e.*, we have checked the condition for a minimum for all $J_{l,j}$ from time $j = 0$ to time $j = k-1$, and have chosen the $\{\check{s}_{j|j-l}\}_{j=0}^{k-1}$ such that the values of the $J_{l,j}$ at their minima are positive. We would now like to solve the problem for time k.

To this end, let us write $J_{l,k}$ as

$$J_{l,k} = J_{l,k}^1 + J_{l,k}^2, \tag{3.5.14}$$

where

$$\begin{gathered} J_{l,k}^1 = (x_0 - \check{x}_0)^*\Pi_0{}^{-1}(x_0 - \check{x}_0) + \textstyle\sum_{j=0}^{k-l} u_j^* u_j + \\ \textstyle\sum_{j=0}^{k-l} \begin{bmatrix} y_j - H_j x_j \\ \check{s}_{j|j-l} - L_j x_j \end{bmatrix}^* \begin{bmatrix} I_p & 0 \\ 0 & -\gamma_l^{-2} I_q \end{bmatrix} \begin{bmatrix} y_j - H_j x_j \\ \check{s}_{j|j-l} - L_j x_j \end{bmatrix} \end{gathered} \tag{3.5.15}$$

and

$$J_{l,k}^2 = \sum_{j=k+1-l}^{k} u_j^* u_j - \sum_{j=k+1-l}^{k} \gamma_l^{-2} (\check{s}_{j|j-l} - L_j x_j)^* (\check{s}_{j|j-l} - L_j x_j). \tag{3.5.16}$$

The reason why we have decomposed $J_{l,k}$ in this fashion is that $J_{l,k}^1$ is a function of $\{x_0, u_0, \ldots, u_{k-l}\}$, or equivalently, $\{u_0, \ldots, u_{k-l}, x_{k-l+1}\}$, *i.e.*,

$$J_{l,k}^1 = J_{l,k}^1(u_0, \ldots, u_{k-l}, x_{k-l+1}), \tag{3.5.17}$$

and $J_{l,k}^2$ is a function of $\{x_{k-l+1}, u_{k-l+1}, \ldots, u_k\}$, *i.e.*,

$$J_{l,k}^2 = J_{l,k}^2(x_{k-l+1}, u_{k-l+1}, \ldots, u_k). \tag{3.5.18}$$

Thus, in effect, the above decomposition decouples the effects of $\{u_j, j \leq k-l\}$ and $\{u_j, j > k-l\}$ and allows us to minimize over the two sets separately.

Now the condition that $J_{l,k}^1$ have a minimum over $\{u_j, j \leq k-l\}$ is readily seen to be that the matrices

$$R_{e,j} = \begin{bmatrix} I_p & 0 \\ 0 & -\gamma_l^2 I_q \end{bmatrix} + \begin{bmatrix} H_j \\ L_j \end{bmatrix} P_j \begin{bmatrix} H_j^* & L_j^* \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix} \tag{3.5.19}$$

have the same inertia for all $j = 0, \ldots k-l$, where P_j satisfies the Riccati recursion

$$P_{j+1} = F_j P_j F_j^* + G_j G_j^* - F_j P_j \begin{bmatrix} H_j^* & L_j^* \end{bmatrix} R_{e,j}^{-1} \begin{bmatrix} H_j \\ L_j \end{bmatrix}, \quad P_0 = \Pi_0. \tag{3.5.20}$$

Moreover the minimum value is

$$J^1_{l,k}(\min) = \sum_{j=0}^{k-l} \begin{bmatrix} y_j - H_j\hat{x}_j \\ \check{s}_{j|j-l} - L_j\hat{x}_j \end{bmatrix}^* R_{e,j}^{-1} \begin{bmatrix} y_j - H_j\hat{x}_j \\ \check{s}_{j|j-l} - L_j\hat{x}_j \end{bmatrix} + \tilde{x}^*_{k+1-l} P^{-1}_{k+1-l} \tilde{x}^*_{k+1-l}, \tag{3.5.21}$$

where $\hat{x}_j$ satisfies the Krein space Kalman filter recursions,

$$\hat{x}_{j+1} = F_j\hat{x}_j + K_{p,j} \begin{bmatrix} y_j - H_j\hat{x}_j \\ \check{s}_{j|j-l} - L_j\hat{x}_j \end{bmatrix}, \quad \hat{x}_0 = \check{x}_0 \tag{3.5.22}$$

with

$$K_{p,j} = F_j P_j \begin{bmatrix} H_j^* & L_j^* \end{bmatrix} R_{e,j}^{-1}. \tag{3.5.23}$$

Since the first term (*i.e.*, the summation) in $J^1_{l,k}(\min)$ is independent of x_{k+1-l}, to minimize $J_{l,k}$ over the remaining variables $\{x_{k-l+1}, u_{k-l+1}, \ldots, u_k\}$ it suffices to minimize

$$\begin{gathered} J^2_{l,k} + \tilde{x}^*_{k+1-l} P^{-1}_{k+1-l} \tilde{x}^*_{k+1-l} = \\ \tilde{x}^*_{k+1-l} P^{-1}_{k+1-l} \tilde{x}^*_{k+1-l} + \textstyle\sum_{j=k+1-l}^{k} u_j^* u_j - \sum_{j=k+1-l}^{k} \gamma_l^{-2} (\check{s}_{j|j-l} - L_j x_j)^* (\check{s}_{j|j-l} - L_j x_j) \end{gathered} \tag{3.5.24}$$

over those variables. But the condition for this minimum is readily seen to be

$$R_{e,j}^{k+1-l} \triangleq -\gamma_l^2 I_q + L_j P_j^{k+1-l} L_j^* < 0, \quad j = k+1-l, \ldots, k \tag{3.5.25}$$

where $P_{j,k+1-l}$ satisfies the Riccati recursion

$$P_{j+1}^{k+1-l} = F_j P_j^{k+1-l} F_j^* + G_j G_j^* - F_j P_j^{k+1-l} L_j^* \left[R_{e,j}^{k+1-l}\right]^{-1} L_j P_j^{k+1-l} F_j^*, \quad P_{k+1-l}^{k+1-l} = P_{k+1-l}. \tag{3.5.26}$$

Moreover, in this case the value at the minimum is given by

$$\sum_{j=k+1-l}^{k} (\check{s}_{j|j-l} - L_j\hat{x}_j^{k+1-l})^* \left[R_{e,j}^{k+1-l}\right]^{-1} (\check{s}_{j|j-l} - L_j\hat{x}_j^{k+1-l}), \tag{3.5.27}$$

where $\hat{x}_j^{k+1-l}$ satisfies the Krein space Kalman filter recursions,[7]

$$\hat{x}_{j+1}^{k+1-l} = F_j\hat{x}_j^{k+1-l} + K_{p,j}^{k+1-l}(\check{s}_{j|j-l} - L_j\hat{x}_j^{k+1-l}), \quad j = k+1-l, \ldots k \tag{3.5.28}$$

[7]Note that at time k all the $\check{s}_{j|j-l}$ (for $j = k+1-l, \ldots k$) in the given recursion are known.

where

$$K_{p,j}^{k+1-l} = F_j P_j^{k+1-l} L_j^* \left[R_{e,j}^{k+1-l}\right]^{-1}. \tag{3.5.29}$$

Therefore the value of $J_{l,k}$ at it its minimum is given by

$$\begin{array}{c} J_{l,k}(\min) = \sum_{j=0}^{k-l} \begin{bmatrix} y_j - H_j\hat{x}_j \\ \check{s}_{j|j-l} - L_j\hat{x}_j \end{bmatrix}^* R_{e,j}^{-1} \begin{bmatrix} y_j - H_j\hat{x}_j \\ \check{s}_{j|j-l} - L_j\hat{x}_j \end{bmatrix} + \\ \sum_{j=k+1-l}^{k} (\check{s}_{j|j-l} - L_j\hat{x}_j^{k+1-l})^* \left[R_{e,j}^{k+1-l}\right]^{-1} (\check{s}_{j|j-l} - L_j\hat{x}_j^{k+1-l}). \end{array} \tag{3.5.30}$$

Now if we define

$$A \triangleq J_{l,k}(\min) - (\check{s}_{k|k-l} - L_j\hat{x}_k^{k+1-l})^* \left[R_{e,k}^{k+1-l}\right]^{-1} (\check{s}_{k|k-l} - L_j\hat{x}_k^{k+1-l}), \tag{3.5.31}$$

then it possible to show that $A > 0$. Indeed,

$$\begin{aligned} A &= \sum_{j=0}^{k-l} \begin{bmatrix} y_j - H_j\hat{x}_j \\ \check{s}_{j|j-l} - L_j\hat{x}_j \end{bmatrix}^* R_{e,j}^{-1} \begin{bmatrix} y_j - H_j\hat{x}_j \\ \check{s}_{j|j-l} - L_j\hat{x}_j \end{bmatrix} + \\ &\quad \sum_{j=k-l}^{k-1} (\check{s}_{j|j-l} - L_j\hat{x}_j^{k+1-l})^* \left[R_{e,j}^{k+1-l}\right]^{-1} (\check{s}_{j|j-l} - L_j\hat{x}_j^{k+1-l}) \\ &= \sum_{j=0}^{k-l-1} \begin{bmatrix} y_j - H_j\hat{x}_j \\ \check{s}_{j|j-l} - L_j\hat{x}_j \end{bmatrix}^* R_{e,j}^{-1} \begin{bmatrix} y_j - H_j\hat{x}_j \\ \check{s}_{j|j-l} - L_j\hat{x}_j \end{bmatrix} + \\ &\quad \min_{x_{k-l},u_{k-l},\ldots,u_{k-1}} \left[\tilde{x}_{k-l}^* P_{k-l}^{-1}\tilde{x}_{k-l} + v_{k-l}^* v_{k-l} + \sum_{j=k-l}^{k-1} u_j^* u_j - \sum_{j=k-l}^{k-1} \gamma_l^{-2} e_{l,j}^* e_{l,j}\right] \\ &\geq \sum_{j=0}^{k-l-1} \begin{bmatrix} y_j - H_j\hat{x}_j \\ \check{s}_{j|j-l} - L_j\hat{x}_j \end{bmatrix}^* R_{e,j}^{-1} \begin{bmatrix} y_j - H_j\hat{x}_j \\ \check{s}_{j|j-l} - L_j\hat{x}_j \end{bmatrix} + \\ &\quad \min_{x_{k-l},u_{k-l},\ldots,u_{k-1}} \left[\tilde{x}_{k-l}^* P_{k-l}^{-1}\tilde{x}_{k-l} + \sum_{j=k-l}^{k-1} u_j^* u_j - \sum_{j=k-l}^{k-1} \gamma_l^{-2} e_{l,j}^* e_{l,j}\right] \\ &= J_{f,k-1}(\min) > 0. \end{aligned}$$

[Note that in the last step we have made use of the fact that the $\{\check{s}_{j|j-l}\}_{j=0}^{k-1}$ were chosen such that $J_{f,k-1}(\min) > 0$.

Thus writing,

$$J_{l,k}(\min) = A + (\check{s}_{k|k-l} - L_j\hat{x}_k^{k+1-l})^* \left[R_{e,k}^{k+1-l}\right]^{-1} (\check{s}_{k|k-l} - L_j\hat{x}_k^{k+1-l}),$$

we note that any choice of $\check{s}_{k|k-l}$ that renders $J_{l,k}(\min) > 0$ is an acceptable estimate. Since $A > 0$ and $R_{e,k}^{k+1-l} < 0$ there are obviously many choices of which the choice

$$\check{s}_{k|k-l} = L_j \hat{x}_k^{k+1-l}, \tag{3.5.32}$$

seems the most natural (and is referred to as the central filter).

We can summarize the results obtained so far in the following theorem.

Theorem 3.5.1 (l-Step H^∞ Predictor) *Given $\gamma_l > 0$, an l-step H^∞ predictor, $\check{s}_{j|j-l} = \mathcal{F}_{l,j}(y_0, \ldots y_{j-l})$ that achieves*

$$\sup_{x_0, u\in h^2, v\in h^2} \frac{\sum_{j=0}^{i}(s_j - \check{s}_{j|j-l})^*(s_j - \check{s}_{j|j-l})}{(x_0 - \check{x}_0)^*{\Pi_0}^{-1}(x_0 - \check{x}_0) + \sum_{j=0}^{i} u_j^* u_j + \sum_{j=0}^{i-l} v_j^* v_j} < \gamma_l^2, \tag{3.5.33}$$

exists if, and only if, for each $j = 0, \ldots, i-l$,

(i) the two matrices

$$R_{e,j} = \begin{bmatrix} I_p & 0 \\ 0 & -\gamma_l^2 I_q \end{bmatrix} + \begin{bmatrix} H_j \\ L_j \end{bmatrix} P_j \begin{bmatrix} H_j^* & L_j^* \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}$$

have the same inertia, where P_j satisfies the Riccati recursion

$$P_{j+1} = F_j P_j F_j^* + G_j G_j^* - F_j P_j \begin{bmatrix} H_j^* & L_j^* \end{bmatrix} R_{e,j}^{-1} \begin{bmatrix} H_j \\ L_j \end{bmatrix}, \quad P_0 = \Pi_0. \tag{3.5.34}$$

(ii) the sequence of matrices

$$R_{e,j+m}^{j+1} \triangleq -\gamma_l^2 I_q + L_{j+m} P_{j+m}^{j+1} L_{j+m}^*, \quad m = 1, \ldots l$$

are negative definite, where P_{j+m}^{j+1} satisfies the Riccati recursion

$$\begin{aligned} P_{j+m+1}^{j+1} \;=\;& F_{j+m} P_{j+m}^{j+1} F_{j+m}^* + G_{j+m} G_{j+m}^* - \\ & F_{j+m} P_{j+m}^{j+1} L_{j+m}^* \left[R_{e,j+m}^{j+1}\right]^{-1} L_{j+m} P_{j+m}^{j+1} F_{j+m}^* \end{aligned}, \quad P_{j+1}^{j+1} = P_{j+1}. \tag{3.5.35}$$

If this is the case then one possible l-step H^∞ predictor is given by[8]

$$\check{s}_{j+l|j} = L_{j+l}\hat{x}^{j+1}_{j+l}, \tag{3.5.36}$$

where $\hat{x}^{j+1}_{j+m}$ satisfies the recursion,

$$\hat{x}^{j+1}_{j+m+1} = F_{j+m}\hat{x}^{j+1}_{j+m} + K^{j+1}_{p,j+m}(\check{s}_{j+m|j+m-l} - L_{j+m}\hat{x}^{j+1}_{j+m}), \quad \hat{x}^{j+1}_{j+1} = \hat{x}_{j+1} \tag{3.5.37}$$

with

$$K^{j+1}_{p,j+m} = F_{j+m}P^{j+1}_{j+m}L^*_{j+m}\left[R^{j+1}_{e,j+m}\right]^{-1}, \tag{3.5.38}$$

and where $\hat{x}_j$ satisfies the recursion,

$$\hat{x}_{j+1} = F_j\hat{x}_j + K_{p,j}\begin{bmatrix} y_j - H_j\hat{x}_j \\ \check{s}_{j|j-l} - L_j\hat{x}_j \end{bmatrix}, \quad \hat{x}_0 = \check{x}_0 \tag{3.5.39}$$

with

$$K_{p,j} = F_jP_j\begin{bmatrix} H_j^* & L_j^* \end{bmatrix}R^{-1}_{e,j}. \tag{3.5.40}$$

Remark: It will be useful to further explain the structure of the l-step H^∞ filter given in the above theorem.

1. Set $\check{s}_{k|k-l} = L_kF_{k-1}\ldots F_0\check{x}_0$ for $k = 0, \ldots, l-1$.
2. Set $j = 0$. (j is the observation index that runs from 0 to $i - l$.)
3. Check inertia condition (i). Compute P_{j+1} (using (3.5.34)) and $\hat{x}_{j+1}$ (using (3.5.39)).
4. Perform the l-step propagation of the Riccati variable, P^{j+1}_{j+m}, (using (3.5.35)) and check the negativity conditions (ii).
5. Perform the l-step propagation of $\hat{x}^{j+1}_{j+m}$ (using (3.5.37)) to find the estimate, $\check{s}_{j+l|j} = L_{j+l}\hat{x}^{j+1}_{j+l}$.

[8] We also have $\check{s}_{k|k-l} = L_kF_{k-1}\ldots F_0\check{x}_0$ for $k = 0, \ldots, l-1$.

6. Set $j = j+1$ and go to 3.

The important feature of the l-step H^∞ predictor is that at each iteration (*i.e.*, after each observation) in addition to propagating one step of the usual Riccati recursion (3.5.34) and state estimate recursion (3.5.39)) (which are the same recursions as those of the a posteriori filter) we need to propagate l steps of the auxiliary Riccati recursion for P_{i+m}^{i+1} and l steps of the auxiliary state estimate, $\hat{x}_{i+m}^{i+1}$. This is clearly different from the case of l-step H^2 prediction.

It is also straightforward to parametrize all possible l-step H^∞ predictors of a given level γ_l.

Lemma 3.5.1 (All l-Step H^∞ Predictors) *All l-step H^∞ predictors that achieve a level γ_l (assuming they exist) are given by any $\check{s}_{j|j-l} = \mathcal{F}_{l,j}(y_0, \ldots y_{j-l})$ that satisfy*

$$\begin{array}{c}\sum_{j=0}^{k-l}\begin{bmatrix} y_j - H_j\hat{x}_j \\ \check{s}_{j|j-l} - L_j\hat{x}_j\end{bmatrix}^* R_{e,j}^{-1}\begin{bmatrix} y_j - H_j\hat{x}_j \\ \check{s}_{j|j-l} - L_j\hat{x}_j\end{bmatrix} + \\ \sum_{j=k+1-l}^{k}(\check{s}_{j|j-l} - L_j\hat{x}_j^{k+1-l})^*\left[R_{e,j}^{k+1-l}\right]^{-1}(\check{s}_{j|j-l} - L_j\hat{x}_j^{k+1-l}) > 0\end{array} \qquad k = 0, \ldots, i \tag{3.5.41}$$

where $\hat{x}_{j+m}^{j+1}$ and $\hat{x}_j$ satisfy the recursions

$$\hat{x}_{j+m+1}^{j+1} = F_{j+m}\hat{x}_{j+m}^{j+1} + K_{p,j+m}^{j+1}(\check{s}_{j+m|j+m-l} - L_{j+m}\hat{x}_{j+m}^{j+1}), \quad \hat{x}_{j+1}^{j+1} = \hat{x}_{j+1} \tag{3.5.42}$$

and

$$\hat{x}_{j+1} = F_j\hat{x}_j + K_{p,j}\begin{bmatrix} y_j - H_j\hat{x}_j \\ \check{s}_{j|j-l} - L_j\hat{x}_j\end{bmatrix}, \quad \hat{x}_0 = \check{x}_0 \tag{3.5.43}$$

respectively, and where $R_{e,j}^{k+1-l}$, $R_{e,j}$, $K_{p,j+m}^{j+1}$ and $K_{p,j}$ as in Theorem 3.5.1.

3.6 Conclusion

Certain studies in least-squares estimation, adaptive filtering and H^∞ filtering motivated us to develop a theory for linear estimation in certain indefinite metric spaces, called Krein spaces. The main difference from the conventional Hilbert space framework for Kalman filtering and LQG control are that projections in Krein spaces may not necessarily exist or be unique, and that quadratic forms may have stationary points that are not necessarily extreme points (*i.e.*, minima or maxima). We showed that these simple but fundamental differences explain both the unexpected similarities and differences between the well-known Kalman filter solution for stochastic state-space systems and the solution for the completely nonstochastic H^∞ filtering problem.

The main points are the following. There are many problems whose solution can be reduced to the recursive minimization of some indefinite quadratic form. A stationary point, when it exists, of the quadratic form can be computed as follows: set up a (partially equivalent) problem of projecting a vector in a Krein space onto a certain subspace. The advantage is that when there is state-space structure, this projection can be recursively computed by using the innovations approach to derive a Krein space Kalman filter. The equivalence is only partial because the Krein space projection only defines the stationary point of the quadratic form and further conditions need to be checked to determine if this point is also a minimum. It turns out that this checking can also be done recursively using quantities arising in the Kalman filtering algorithms.

Apart from quite straightforward derivations of known results in H^2, H^∞ and risk-sensitive estimation and control, the above approach allows us to extend to the H^∞ setting some of the huge body of results and insights developed over the last three decades in the field of Kalman filtering (and LQG control). A first bonus is the derivation (see [HSK94c] and Chapter 5) of square-root and (fast) Chandrasekhar algorithms for H^∞ estimation and control, a possibility that is much less obvious in current approaches. These square-root algorithms, which are now increasingly standard in H^2 Kalman filtering, have two advantages over the earlier H^∞ algorithms: they eliminate the need for explicitly checking the existence conditions of the filters, and have various potential numerical and implementational advantages.

Application of the Krein space formulation to adaptive filtering arises from the approach in [SK94b] where it was shown how to recast adaptive filtering problems as state-space estimation problems. If we further allow the elements of the state-space model to belong to a Krein space, then we can solve finite memory and H^∞ adaptive filtering problems. In the finite memory case, this allows us to consider general sliding patterns with windows of varying lengths. In the H^∞ adaptive case, this has allowed us to establish that the famed LMS (or stochastic gradient) algorithm is an optimal H^∞ filter [HSK96a] (see also Chapter 9).

We also remark that, although not pursued here, it is also possible to construct *dual* (rather than partially equivalent) Krein state-space models (via the concept of a dual basis), which can be used to extend the methods to the solution of H^2 and H^∞ control problems. This will be done in Chapter 6.