Chapter 6

Duality and Control

This chapter deals with duality in (definite and indefinite metric) linear spaces. Duality here is introduced through the geometrical notion of dual bases for linear spaces spanned by a set of nonorthogonal basis vectors. We shall see that, apart from having conceptual value, duality is a useful tool in the study of various problems of interest. In particular, it allows for a dual approach to (definite and indefinite) quadratic problems. A large part of this chapter is therefore devoted to the use of duality in (H^2, H^∞, game-theoretic and risk-sensitive) control, where (within the framework presented here) the notion of duality is indispensable for obtaining the solution.

6.1 Introduction

In this chapter we return to the geometric viewpoint of Chapter 2. We have already seen the power of such a geometrical approach in unifying the treatment of (stochastic) H^2 estimation and (deterministic) H^∞ estimation problems, (where the solutions are obtained by projecting, respectively in a Hilbert and Krein space, one vector onto the linear span of another set of vectors), in the introduction of the innovations process (via Gram-Schmidt orthogonalization) for performing the canonical factorization of (possibly indefinite) Gramian matrices, and in the extension of conventional square-root arrays and Chandrasekhar recursions to the H^∞ setting. Here we shall introduce a further natural consequence of adopting a geometric point of view; more specifically, the study of dual bases and dual linear models. The concept of a dual basis is one that has considerable (albeit, often conceptual) value in studying linear spaces spanned by a set of nonorthogonal basis vectors. It also has several ramifications; in particular, it shows how deterministic and stochastic least-squares problems can be solved using a dual approach, it has certain implications to smoothing problems, and, most importantly, it is very useful in the study of control problems (where the duality between estimation and control has long been recognized).[1] We shall, in fact, spend considerable time on a general linear-quadratic-regulator (LQR) control problem with indefinite weighting matrices, and its special cases in full information and measurement feedback H^∞ control.

The remainder of this chapter is organized as follows. In Sec. 6.2 we begin by considering a different deterministic quadratic form to the one that had been studied in Sec. 2.5.2, and whose stationary point was given by a Krein space projection. These two deterministic quadratic forms are related by the fact that their coefficient matrices are inverses of one another. In Sec. 6.3, by considering the concept of a dual basis, we obtain a geometric interpretation for the stationary point of the quadratic form of Sec. 6.2, and in particular, show that it is given by a dual Krein space projection. Moreover, we obtain various useful relationships (via certain dualities and equivalences) that shed further light on the interplay between deterministic and stochastic least-squares problems (see Table 6.4.1).[2] In particular, in Sec. 6.4, we show that, given any quadratic optimization problem, it can be solved in either one of two ways; by constructing an equivalent Krein space model, for which the solution is given by a Krein space projection, or, by constructing a dual Krein space model, for which the solution is given by the negative conjugate transpose of the operator that yields the dual Krein space projection. Which one of these approaches to use depends upon the application; for estimation and adaptive filtering problems it is more natural to use the equivalent Krein space model, whereas for control problems it more natural to use the dual Krein space model.

Now the dual bases span the orthogonal complement space, or the space of the

[1] See also the comments at the end of Sec. 1.5.1.
[2] For different approaches to duality consult [Wal89] and [Lue92].

smoothing errors, of the original bases. In Sec. 6.5, we show that when the original basis has state-space structure, then the dual basis also has state-space structure. [The dual state-space model is often referred to as the adjoint or complementary state-space model.] In particular, we develop backwards-time, forwards-time and mixed dual state-space models, and in Sec. 6.6 use the mixed dual state-space model to obtain the so-called two filter formulae for smoothing.

Sec. 6.7 considers the LQR (linear-quadratic-regulator) control problem with (possibly) indefinite weighting matrices. This problem is of interest since it subsumes, as special cases, the problems of H^2, H^∞, game-theoretic and risk-sensitive control. We solve the LQR problem using duality, *i.e.*, we first construct the dual Krein space model associated with the indefinite LQR quadratic form (which is a Krein space backwards-time dual state-space model) and then use the (negative conjugate transpose of the operator that performs the) dual Krein space projection to find its stationary point (and hence its solution). The dual projection is performed by the Krein space Kalman filter associated with the backwards-time dual state-space model, and the condition for the stationary point to be a minimum is given by certain inertia requirements on Gramians readily available from the Krein space Kalman filter. The results of Sec. 6.7 are then used in Secs. 6.8 and 6.9 to solve the full information and measurement feedback H^∞ control problems. The solutions for both causal and strictly causal controllers are given, and the measurement feedback problem's solution requires a separation principle that was earlier introduced in Sec. 1.7.2.

6.2 An Alternative Scalar Quadratic Form

In Sec. 2.5.2 we saw that if we form a scalar quadratic form

$$J(z,y) \equiv \begin{bmatrix} z^* & y^* \end{bmatrix} \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}^{-1} \begin{bmatrix} z \\ y \end{bmatrix}$$

where $\begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}$ may be identified as the Gramian of the the Krein space variables $\{\mathbf{z}, \mathbf{y}\}$, then the stationary point of $J(z,y)$ over z can be obtained via computing the projection of $\mathbf{z}$ onto $\mathcal{L}(\mathbf{y})$. However, there is another scalar quadratic form that can be constructed from the Gramian of $\{\mathbf{z}, \mathbf{y}\}$ and that we have not yet considered. This is what is done in the following Lemma.

Lemma 6.2.1 (Matrix to Scalar Second Order Form) *The stationary point y_o^d (over z^d) of*

$$I(z^d, y^d) = \begin{bmatrix} z^{d*} & y^{d*} \end{bmatrix} \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix} \begin{bmatrix} z^d \\ y^d \end{bmatrix}$$

is given by

$$y_o^d = -R_y^{-1} R_{yz} z^d,$$

and

$$I(z^d, y_o^d) = z^{d*}(R_z - R_{zy} R_y^{-1} R_{yz}) z^d.$$

Moreover, this stationary point is a minimum if, and only if,

$$R_y > 0.$$

Proof: The proof involves the (lower-upper) triangular factorization of $\begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}$ and is straightforward.

■

Remark: Note that the coefficient matrix appearing in the solution to the deterministic problem of Lemma 6.2.1 is given by the negative conjugate transpose of the coefficient matrix appearing in the projection of $\mathbf{z}$ onto $\mathbf{y}$. However, the condition for a minimum, $R_y > 0$, is the exact same as that of the stochastic problem of Theorem 2.5.1.

In view of the above result, it is now natural to speculate whether the stationary point of Lemma 6.2.1 is related to a stochastic problem in the same way that the the stationary point of Theorem 2.5.2 was related to the stochastic problem of Theorem 2.5.1. In the next section we shall see that this is indeed the case. The development involves the introduction of so-called dual bases, and will bring forth various dualities that will be useful for the study of both estimation and control problems in sections 6.6 and 6.7.

6.3 Dual Bases

Recall from Chapter 2 the definition of a Krein space $\mathcal{K}$ over a given ring of scalars (such as the ring of complex numbers), $\mathcal{S}$. Consider a set of linearly independent vectors $\{\mathbf{z}_0, \ldots, \mathbf{z}_m, \mathbf{y}_0, \ldots \mathbf{y}_n\}$, (with $\mathbf{z}_i \in \mathcal{K}$ $\mathbf{y}_j \in \mathcal{K}$, $i = 0, \ldots m$, $j = 0, \ldots n$), which we shall denote by $\{\mathbf{z}, \mathbf{y}\}$ where[3]

$$\mathbf{z} = \operatorname{col}\{\mathbf{z}_0, \ldots, \mathbf{z}_m\} \quad \text{and} \quad \mathbf{y} = \operatorname{col}\{\mathbf{y}_0, \ldots, \mathbf{y}_n\}.$$

The corresponding Gramian for this set of vectors is denoted by

$$\langle \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix}, \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix} \rangle = \begin{bmatrix} \langle \mathbf{z}, \mathbf{z} \rangle & \langle \mathbf{z}, \mathbf{y} \rangle \\ \langle \mathbf{y}, \mathbf{z} \rangle & \langle \mathbf{y}, \mathbf{y} \rangle \end{bmatrix} = \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}, \tag{6.3.1}$$

which due to the linear independence is nonsingular. Moreover, the linear independence implies that $\{\mathbf{z}, \mathbf{y}\}$ forms a *basis* for $\mathcal{L}\{\mathbf{z}, \mathbf{y}\}$ (*i.e.*, the linear space of all vectors $a_0\mathbf{z}_0 + \ldots + a_m\mathbf{z}_m + b_0\mathbf{y}_0 + \ldots + b_n\mathbf{y}_n$, for $a_i, b_j \in \mathcal{S}$).

Definition 6.3.1 (Dual Basis) *Given a basis, $\{\mathbf{z}, \mathbf{y}\}$, the dual basis is defined as the pair $\{\mathbf{z}^d, \mathbf{y}^d\}$ with the properties,*

$$\mathcal{L}\{\mathbf{z}^d, \mathbf{y}^d\} = \mathcal{L}\{\mathbf{z}, \mathbf{y}\}, \tag{6.3.2}$$

and

$$\begin{aligned} \langle \mathbf{z}^d, \mathbf{z} \rangle = I \quad &, \quad \langle \mathbf{z}^d, \mathbf{y} \rangle = 0 \;, \\ \langle \mathbf{y}^d, \mathbf{z} \rangle = 0 \quad &, \quad \langle \mathbf{y}^d, \mathbf{y} \rangle = I \;. \end{aligned} \tag{6.3.3}$$

Note that if $\{\mathbf{z}, \mathbf{y}\}$ were an *orthonormal* basis then the dual basis would simply coincide with our original basis. In general, however, the dual basis will be different and (6.3.3) is referred to as the *bi-orthogonality condition.*

We now describe two ways - algebraic and geometric - of finding the dual basis.

[3]As a matter of fact we do not need to partition the set of independent vectors into a set of $\{\mathbf{z}_i\}$ and $\{\mathbf{y}_i\}$ in order to introduce the concept of dual bases. However, since our major interest is estimation, and the set $\{\mathbf{y}_i\}$ will typically designate the observations, while the set $\{\mathbf{z}_i\}$ will designate the quantities we want to estimate, we shall find this partitioning quite convenient.

6.3.1 Algebraic Specification

Clearly since $\{\mathbf{z}, \mathbf{y}\}$ and $\{\mathbf{z}^d, \mathbf{y}^d\}$ span the same space we must have

$$\begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix}, \tag{6.3.4}$$

for some nonsingular block matrix $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$. Now the defining requirements (6.3.3) are equivalent to

$$\left\langle \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix}, \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix} \right\rangle = \begin{bmatrix} \langle \mathbf{z}^d, \mathbf{z} \rangle & \langle \mathbf{z}^d, \mathbf{y} \rangle \\ \langle \mathbf{y}^d, \mathbf{z} \rangle & \langle \mathbf{y}^d, \mathbf{y} \rangle \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}. \tag{6.3.5}$$

But

$$\left\langle \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix}, \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix} \right\rangle = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \left\langle \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix}, \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix} \right\rangle = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}. \tag{6.3.6}$$

Combining the above two relations we have

$$\begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix} = \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix}. \tag{6.3.7}$$

Moreover, the Gramian of $\{\mathbf{z}^d, \mathbf{y}^d\}$ is seen to be

$$\begin{aligned} \left\langle \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix}, \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix} \right\rangle &= \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}^{-1} \left\langle \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix}, \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix} \right\rangle \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}^{-*} \\ &= \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}^{-*} = \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}^{-1}, \end{aligned}$$

since Gramians are Hermitian.

Lemma 6.3.1 (Dual Basis) *Consider the pair $\{\mathbf{z}, \mathbf{y}\}$ with nonsingular Gramian given by (6.3.1). Then the the dual basis $\{\mathbf{z}^d, \mathbf{y}^d\}$ satisfying the bi-orthogonality condition (6.3.3) is given by*

$$\begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix} = \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix}.$$

Moreover, the Gramian of the dual basis $\{\mathbf{z}^d, \mathbf{y}^d\}$ *is*

$$\left\langle \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix}, \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix} \right\rangle = \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}^{-1}.$$

The above arguments and result are quite specific, but may not be very intuitive. Let us now consider the geometric description.

6.3.2 Geometric Specification

Let us first introduce the notations

$$\hat{\mathbf{z}}_{|\mathbf{y}} = \text{the projection of } \mathbf{z} \text{ onto } \mathcal{L}\{\mathbf{y}\}\ , \tag{6.3.8}$$

and the error

$$\tilde{\mathbf{z}} \triangleq \tilde{\mathbf{z}}_{|\mathbf{y}} = \mathbf{z} - \hat{\mathbf{z}}_{|\mathbf{y}}. \tag{6.3.9}$$

From the orthogonality principle we have $\langle \tilde{\mathbf{z}}, \mathbf{y} \rangle = 0$. On the other hand, due to the property that $\langle \mathbf{z}^d, \mathbf{y} \rangle = 0$, and since $\{\tilde{\mathbf{z}}, \mathbf{y}\}$ and $\{\mathbf{z}^d, \mathbf{y}\}$ span the same space, we clearly see that $\{\tilde{\mathbf{z}}\}$ and $\{\mathbf{z}^d\}$ must span the same space as well. Thus we must have

$$\mathbf{z}^d = M\tilde{\mathbf{z}}, \tag{6.3.10}$$

for some nonsingular M. However, note that

$$I = \langle \mathbf{z}^d, \mathbf{z} \rangle = M\langle \tilde{\mathbf{z}}, \mathbf{z} \rangle = M\langle \tilde{\mathbf{z}}, \tilde{\mathbf{z}} + \hat{\mathbf{z}} \rangle = M\langle \tilde{\mathbf{z}}, \tilde{\mathbf{z}} \rangle = MR_{\tilde{z}}\ , \tag{6.3.11}$$

so that

$$\mathbf{z}^d = R_{\tilde{z}}^{-1}\tilde{\mathbf{z}}_{|\mathbf{y}}. \tag{6.3.12}$$

Similarly, of course we have

$$\mathbf{y}^d = R_{\tilde{y}}^{-1}\tilde{\mathbf{y}}_{|\mathbf{z}}. \tag{6.3.13}$$

The reader may want to verify that the invertibility of the matrices R_x, R_y and the Gramian (6.3.1) guarantees the invertibility of $R_{\tilde{z}}$ and $R_{\tilde{y}}$.

Lemma 6.3.2 (Dual Basis - Revisited) *Consider the pair $\{\mathbf{z}, \mathbf{y}\}$ with nonsingular Gramian given by (6.3.1). Then the the dual basis $\{\mathbf{z}^d, \mathbf{y}^d\}$ satisfying the bi-orthogonality condition (6.3.3) is given by*

$$\mathbf{z}^d = R_{\tilde{z}}^{-1}\tilde{\mathbf{z}}_{|\mathbf{y}} \quad \textit{and} \quad \mathbf{y}^d = R_{\tilde{y}}^{-1}\tilde{\mathbf{y}}_{|\mathbf{z}} \ ,$$

where $\tilde{\mathbf{z}}_{|\mathbf{y}}$ and $\tilde{\mathbf{y}}_{|\mathbf{z}}$ are the errors in projecting $\mathbf{z}$ and $\mathbf{y}$ onto $\mathcal{L}\{\mathbf{y}\}$ and $\mathcal{L}\{\mathbf{z}\}$, respectively, and $R_{\tilde{z}}$ and $R_{\tilde{y}}$ are their corresponding Gramians.

The vectors $\{\mathbf{z}^d, \mathbf{y}^d\}$ can be readily sketched as in Fig. 6.1. [In simple cases, determining the dual basis from the geometry can be simpler than inverting the Gramian matrix.]

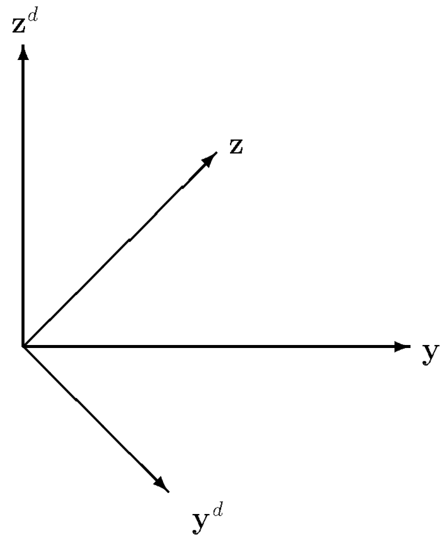

Figure 6.1: Dual bases.

But why introduce dual bases? There are several reasons. One is the fact that to compute the projection of a vector $\mathbf{x}$ onto $\mathcal{L}\{\mathbf{y}, \mathbf{z}\}$, say

$$\hat{\mathbf{x}}_{|\mathbf{y},\mathbf{z}} = A\mathbf{z} + B\mathbf{y} \ , \tag{6.3.14}$$

requires solving a system of linear equations to determine the coefficient matrices $\{A, B\}$, except, of course, when $\mathbf{y}$ and $\mathbf{z}$ are orthogonal. In the orthogonal case, the coefficients are obtained by projecting $\mathbf{x}$ separately on $\mathbf{z}$ and $\mathbf{y}$, so that

$$\hat{\mathbf{x}}_{|\mathbf{y},\mathbf{z}} = \langle \mathbf{x}, \mathbf{z} \rangle \|\mathbf{z}\|^{-2}\mathbf{z} + \langle \mathbf{x}, \mathbf{y} \rangle \|\mathbf{y}\|^{-2}\mathbf{y}, \quad \text{when } \langle \mathbf{y}, \mathbf{z} \rangle = 0. \tag{6.3.15}$$

In the general case, we can obtain a somewhat similar expression by using the dual basis.

Lemma 6.3.3 (Projection via the Dual Basis) $\hat{\mathbf{x}}_{|\mathbf{y},\mathbf{z}}$, *the projection of* $\mathbf{x}$ *onto* $\mathcal{L}\{\mathbf{y}, \mathbf{z}\}$, *can be written as*

$$\hat{\mathbf{x}}_{|\mathbf{y},\mathbf{z}} = \langle \mathbf{x}, \mathbf{z}^d \rangle \mathbf{z} + \langle \mathbf{x}, \mathbf{y}^d \rangle \mathbf{y}, \tag{6.3.16}$$

where $\{\mathbf{y}^d, \mathbf{z}^d\}$ *is the dual basis.*

Proof: We need to find $\{A, B\}$ such that

$$\mathbf{x} - A\mathbf{z} - B\mathbf{y} \ \perp \ \mathcal{L}\{\mathbf{y}, \mathbf{z}\} = \mathcal{L}\{\mathbf{y}^d, \mathbf{z}^d\}.$$

But

$$0 = \langle \mathbf{x} - A\mathbf{z} - B\mathbf{y}, \mathbf{z}^d \rangle = \langle \mathbf{x}, \mathbf{z}^d \rangle - A\langle \mathbf{z}, \mathbf{z}^d \rangle - B\langle \mathbf{y}, \mathbf{z}^d \rangle = \langle \mathbf{x}, \mathbf{z}^d \rangle - A.$$

Similarly,

$$0 = \langle \mathbf{x}, \mathbf{y}^d \rangle - B.$$

■

Of course, the apparent simplicity of the formula (6.3.16) is, in general, only conceptual rather than computational. The algebraic equations for $\{A, B\}$ are

$$\begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix} = \begin{bmatrix} \langle \mathbf{x}, \mathbf{z} \rangle & \langle \mathbf{x}, \mathbf{y} \rangle \end{bmatrix}, \tag{6.3.17}$$

and inverting the (Gramian) coefficient matrix is equivalent to determining the dual basis $\{\mathbf{z}^d, \mathbf{y}^d\}$! Nevertheless, the conceptual simplification makes the introduction of dual bases, and their geometric interpretation, quite useful.

We can also obtain an interesting result by combining the algebraic and geometric characterizations just given. Recall that the geometric characterization was

$$\begin{bmatrix} \tilde{\mathbf{z}}_{|\mathbf{y}} \\ \tilde{\mathbf{y}}_{|\mathbf{z}} \end{bmatrix} = \begin{bmatrix} R_{\tilde{z}} & 0 \\ 0 & R_{\tilde{y}} \end{bmatrix} \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix}. \tag{6.3.18}$$

Therefore we have

$$\begin{aligned} \left\langle \begin{bmatrix} \tilde{\mathbf{z}}_{|\mathbf{y}} \\ \tilde{\mathbf{y}}_{|\mathbf{z}} \end{bmatrix}, \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix} \right\rangle &= \begin{bmatrix} R_{\tilde{z}} & 0 \\ 0 & R_{\tilde{y}} \end{bmatrix} \left\langle \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix}, \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix} \right\rangle \\ &= \begin{bmatrix} R_{\tilde{z}} & 0 \\ 0 & R_{\tilde{y}} \end{bmatrix} \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}^{-1}. \end{aligned} \tag{6.3.19}$$

But note that

$$\langle \tilde{\mathbf{z}}_{|\mathbf{y}}, \mathbf{z}^d \rangle = \langle \mathbf{z} - R_{zy} R_y^{-1} \mathbf{y}, \mathbf{z}^d \rangle = I - 0 = I \ ,$$

while

$$\langle \tilde{\mathbf{z}}_{|\mathbf{y}}, \mathbf{y}^d \rangle = \langle \mathbf{z} - R_{zy} R_y^{-1} \mathbf{y}, \mathbf{y}^d \rangle = 0 - R_{zy} R_y^{-1} I.$$

With similar results for $\langle \tilde{\mathbf{y}}_{|\mathbf{z}}, \mathbf{z}^d \rangle$ and $\langle \tilde{\mathbf{y}}_{|\mathbf{z}}, \mathbf{y}^d \rangle$, we see that

$$\left\langle \begin{bmatrix} \tilde{\mathbf{z}}_{|\mathbf{y}} \\ \tilde{\mathbf{y}}_{|\mathbf{z}} \end{bmatrix}, \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix} \right\rangle = \begin{bmatrix} I & -R_{zy} R_y^{-1} \\ -R_{yz} R_z^{-1} & I \end{bmatrix}.$$

But this leads immediately to the identity

$$\begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}^{-1} = \begin{bmatrix} R_{\tilde{z}}^{-1} & 0 \\ 0 & R_{\tilde{y}}^{-1} \end{bmatrix} \begin{bmatrix} I & -R_{zy} R_y^{-1} \\ -R_{yz} R_z^{-1} & I \end{bmatrix}, \tag{6.3.20}$$

whose origin is otherwise not so evident. We can also rewrite (6.3.20) as

$$\begin{bmatrix} R_{z^d} & R_{z^d y^d} \\ R_{y^d z^d} & R_{y^d} \end{bmatrix} = \begin{bmatrix} R_{\tilde{z}}^{-1} & -R_{\tilde{z}}^{-1} R_{zy} R_y^{-1} \\ -R_{\tilde{y}}^{-1} R_{yz} R_z^{-1} & R_{\tilde{y}}^{-1} \end{bmatrix}. \tag{6.3.21}$$

In particular, this implies that

$$R_{z^d} = R_{\tilde{z}}^{-1} \quad , \quad R_{y^d} = R_{\tilde{y}}^{-1} \tag{6.3.22}$$

and

$$R_{zy} R_y^{-1} = -R_{\tilde{z}} R_{z^d y^d} = -R_{z^d}^{-1} R_{z^d y^d} = -(R_{y^d z^d} R_{z^d}^{-1})^*. \tag{6.3.23}$$

But recalling that

$$\hat{\mathbf{z}}_{|\mathbf{y}} = R_{zy}R_y^{-1}\mathbf{y} \quad \text{and} \quad \hat{\mathbf{y}}^d_{|\mathbf{z}^d} = R_{y^dz^d}R_{z^d}^{-1}\mathbf{z}^d \tag{6.3.24}$$

we see that the gain matrix for estimating $\mathbf{z}$ from $\mathbf{y}$ is the negative conjugate transpose of the gain matrix for estimating $\mathbf{y}^d$ from $\mathbf{z}^d$!

This is an interesting result, and we therefore collect the above identities into the following Lemma.

Lemma 6.3.4 (Dual Computation of Projection) *The projection of $\mathbf{z}$ onto $\mathcal{L}(\mathbf{y})$, say $\hat{\mathbf{z}}_{|\mathbf{y}}$, can be computed as*

$$\hat{\mathbf{z}}_{|\mathbf{y}} = -R_{z^d}^{-1}R_{z^dy^d}\mathbf{y} \ ,$$

where $\{R_{z^d}, R_{z^dy^d}\}$ are the Gramians and cross-Gramians of the dual basis vectors $\{\mathbf{z}^d, \mathbf{y}^d\}$. In other words, we have the identity

$$R_{zy}R_y^{-1} = -R_{z^d}^{-1}R_{z^dy^d}.$$

Also, the minimum mean-square-error covariance matrix is given by

$$R_{\tilde{z}} = R_{z^d}^{-1}.$$

The above lemma captures some of the duality between $\{\mathbf{z}, \mathbf{y}\}$ and the dual basis $\{\mathbf{z}^d, \mathbf{y}^d\}$. However, before proceeding with this duality, we shall study the consequence of the above results to the case where the $\{\mathbf{z}, \mathbf{y}\}$ are related in a linear fashion.

6.3.3 Linear Models

It is interesting to apply the above results to the simple linear model,[4]

$$\mathbf{y} = H\mathbf{z} + \mathbf{v}, \tag{6.3.25}$$

[4] It can, in fact, be shown that there is no loss of generality in assuming such a linear model (rather than assuming the general case of arbitrarily related Krein space variables, $\mathbf{z}$ and $\mathbf{y}$). However, this representation is useful since it has (and will) frequently occur in this thesis.

where the Gramian matrix of $\mathbf{z}$ and $\mathbf{v}$ is block-diagonal,

$$\left\langle \begin{bmatrix} \mathbf{z} \\ \mathbf{v} \end{bmatrix}, \begin{bmatrix} \mathbf{z} \\ \mathbf{v} \end{bmatrix} \right\rangle = \begin{bmatrix} R_z & 0 \\ 0 & R_v \end{bmatrix}. \tag{6.3.26}$$

Moreover, assume that R_z and R_v are invertible. Then

$$\begin{aligned} \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix} = \left\langle \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix}, \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix} \right\rangle &= \begin{bmatrix} I & 0 \\ H & I \end{bmatrix} \left\langle \begin{bmatrix} \mathbf{z} \\ \mathbf{v} \end{bmatrix}, \begin{bmatrix} \mathbf{z} \\ \mathbf{v} \end{bmatrix} \right\rangle \begin{bmatrix} I & H^* \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ H & I \end{bmatrix} \begin{bmatrix} R_z & 0 \\ 0 & R_v \end{bmatrix} \begin{bmatrix} I & H^* \\ 0 & I \end{bmatrix}. \end{aligned}$$

Therefore

$$\begin{aligned} \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix} &= \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix}, \\ &= \left\{ \begin{bmatrix} I & 0 \\ H & I \end{bmatrix} \begin{bmatrix} R_z & 0 \\ 0 & R_v \end{bmatrix} \begin{bmatrix} I & H^* \\ 0 & I \end{bmatrix} \right\}^{-1} \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix}, \\ &= \begin{bmatrix} I & -H^* \\ 0 & I \end{bmatrix} \begin{bmatrix} R_z^{-1} & 0 \\ 0 & R_v^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -H & I \end{bmatrix} \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix}, \\ &= \begin{bmatrix} I & -H^* \\ 0 & I \end{bmatrix} \begin{bmatrix} R_z^{-1} & 0 \\ 0 & R_v^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{z} \\ \mathbf{v} \end{bmatrix}, \\ &= \begin{bmatrix} R_z^{-1}\mathbf{z} - H^* R_v^{-1}\mathbf{v} \\ R_v^{-1}\mathbf{v} \end{bmatrix}. \end{aligned}$$

We thus see that

$$\mathbf{y}^d = R_v^{-1}\mathbf{v}, \tag{6.3.27}$$

and that $\mathbf{z}^d$ arises from a *dual* linear model,

$$\mathbf{z}^d = -H^* R_v^{-1}\mathbf{v} + R_z^{-1}\mathbf{z} = -H^*\mathbf{y}^d + R_z^{-1}\mathbf{z} = -H^*\mathbf{y}^d + \mathbf{v}^c, \tag{6.3.28}$$

where we have introduced $\mathbf{v}^c = R_z^{-1}\mathbf{z}$ and, hence,

$$\left\langle \begin{bmatrix} \mathbf{y}^d \\ \mathbf{v}^c \end{bmatrix}, \begin{bmatrix} \mathbf{y}^d \\ \mathbf{v}^c \end{bmatrix} \right\rangle = \begin{bmatrix} R_v^{-1} & 0 \\ 0 & R_z^{-1} \end{bmatrix}.$$

The Gramian matrix of the dual basis is readily seen to be

$$\langle \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix}, \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix} \rangle = \begin{bmatrix} R_{z^d d} & R_{z^d d y^d} \\ R_{y^d z^d d} & R_{y^d} \end{bmatrix} = \begin{bmatrix} R_z^{-1} + H^* R_v^{-1} H & -H^* R_v^{-1} \\ -R_v^{-1} H & R_v^{-1} \end{bmatrix}. \quad (6.3.29)$$

We can therefore compute the projection of $\mathbf{y}^d$ onto $\mathbf{z}^d$ as follows

$$\hat{\mathbf{y}}^d_{|\mathbf{z}^d} = R_{y^d z^d d} R_{y^d}^{-1} \mathbf{z}^d = -R_v^{-1} H (R_z^{-1} + H^* R_v^{-1} H)^{-1} \mathbf{z}^d. \quad (6.3.30)$$

On the other hand,

$$\hat{\mathbf{z}}_{|\mathbf{y}} = R_{zy} R_y^{-1} = R_z H^* (R_v + H R_z H^*)^{-1} \mathbf{y}. \quad (6.3.31)$$

But as claimed in Lemma 6.3.4, the coefficient matrices for these two problems must be the negative conjugate transpose of each other, *i.e.*, we must have

$$R_z H^* (R_v + H R_z H^*)^{-1} = -(-R_v^{-1} H (R_z^{-1} + H^* R_v^{-1} H)^{-1})^* = (R_z^{-1} + H^* R_v^{-1} H)^{-1} H^* R_v^{-1} \quad (6.3.32)$$

which is an identity that can also be verified algebraically.

We note that these so-called "information form" expressions for the estimates actually correspond to an estimation problem for certain dual variables. To summarize the results, we have the following Lemma.

Lemma 6.3.5 (Linear Models and Dual Bases) *Suppose* $\{\mathbf{z}, \mathbf{y}\}$ *satisfy the linear model*

$$\mathbf{y} = H\mathbf{z} + \mathbf{v},$$

where

$$\langle \begin{bmatrix} \mathbf{z} \\ \mathbf{v} \end{bmatrix}, \begin{bmatrix} \mathbf{z} \\ \mathbf{v} \end{bmatrix} \rangle = \begin{bmatrix} R_z & 0 \\ 0 & R_v \end{bmatrix},$$

and both R_z *and* R_v *are assumed nonsingular. Then the dual basis* $\{\mathbf{z}^d, \mathbf{y}^d\}$ *will also satisfy a linear model, namely,*

$$\mathbf{z}^d = -H^* \mathbf{y}^d + \mathbf{v}^c,$$

where

$$\mathbf{y}^d = R_v^{-1} \mathbf{v} \quad \textit{and} \quad \mathbf{v}^c = R_z^{-1} \mathbf{z}.$$

We therefore have the identities

$$R_{zy}R_y^{-1} = R_z H^*(R_v + HR_z H^*)^{-1} = (R_z^{-1} + H^* R_v^{-1} H)^{-1} H^* R_v^{-1} = -R_{z^d d} R_{z^d d y^d} \ ,$$

and

$$\langle \tilde{\mathbf{z}}_{|y}, \tilde{\mathbf{z}}_{|y} \rangle = R_{\tilde{z}} = R_{z^d}^{-1} = (R_z^{-1} + H^* R_v^{-1} H)^{-1}.$$

6.4 A Pair of Duality and Equivalence Relationships

In the previous section we considered the problem of projecting the dual vector $\mathbf{y}^d$ onto the dual vector $\mathbf{z}^d$, and saw that the problem was dual to the problem (considered so far) of projecting $\mathbf{z}$ onto $\mathbf{y}$. In this section we shall study the consequences of this observation to the solution of deterministic least-squares problems. The main conclusion is, that given a deterministic quadratic minimization problem, one can solve it in two-ways: either by constructing an equivalent Krein space model (which is what has been done so far, say in the H^∞ filtering problems of Chapter 3), where the solution is the same as the original problem, or by constructing a dual Krein space model where the solution is given by the negative transpose of the solution of the original problem. Although, not done here, it is possible to solve the general estimation problem using this dual approach to obtain the (so-called) Information-form Kalman filter [AM79, Kai81]. However, the major benefit of using this dual approach is in the solution of LQR (linear quadratic regulator) control problems that will be studied in Sec. 6.7.

6.4.1 General Equivalence and Duality Relationships

Lemma 6.3.4 captures the duality between the original basis $\{\mathbf{z}, \mathbf{y}\}$ and the dual basis $\{\mathbf{z}^d, \mathbf{y}^d\}$. Most noteworthy it points out that the inverse of the matrix that appears in $I(z^d, y^d)$ of Lemma 6.2.1, is the Gramian of the dual basis $\{\mathbf{z}^d, \mathbf{y}^d\}$. We shall presently continue to pursue this connection.

Note now that the scalar quadratic form appearing in Lemma 6.2.1 can be written as

$$I(z,y) = \begin{bmatrix} z^{d*} & y^{d*} \end{bmatrix} \langle \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix}, \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix} \rangle^{-1} \begin{bmatrix} z^d \\ y^d \end{bmatrix}. \tag{6.4.1}$$

Using Theorems 2.5.1 and 2.5.2 this implies that the stationary point of $I(z^d, y^d)$ over y^d is given by the same coefficient matrix that computes the projection of $\mathbf{y}^d$ onto $\mathbf{z}^d$. Thus

$$\hat{\mathbf{y}}^d_{|z^d} = R_{y^d z^d} R_{z^d}^{-1} \mathbf{z}^d \quad \text{and} \quad y^d_o = R_{y^d z^d} R_{z^d}^{-1} z^d. \tag{6.4.2}$$

Now from Lemma 6.3.4 we have $R_{y^d z^d} R_{z^d}^{-1} = -R_y^{-1} R_{yz}$, so that

$$y^d_o = -R_y^{-1} R_{yz} z, \tag{6.4.3}$$

as in Lemma 6.2.1. The condition for a minimum is (using Theorem 2.5.2 and Lemma 6.3.4)

$$R_{y^d}^{-1} - R_{y^d z^d} R_{z^d}^{-1} R_{z^d y^d} = R_y^{-1} > 0, \tag{6.4.4}$$

which is the same condition as in Lemma 6.2.1. Thus the deterministic problem of Lemma 6.2.1 is related to the stochastic problem of projecting $\mathbf{y}^d$ onto $\mathbf{z}^d$.

The results of this section are summarized in Table 6.4.1. It collects the relationships between the two dual stochastic problems and their corresponding deterministic quadratic forms. Vertical transitions from (i) to (iii) and from (ii) to (iv) correspond to going from the original bases to the dual bases, so that the solutions are the dual (negative conjugate transpose) of each other and the conditions for a minimum are different. Horizontal transitions from (i) to (ii) and (iii) to (iv) correspond to going from matrix-valued to scalar-valued quadratic forms so that these solutions are also the dual of one another, however, now the conditions for a minimum are the same. Diagonal transitions from (i) to (iv) and (ii) to (iii) relate problems with the same solution, as in Theorems 2.5.1 and 2.5.2, but with different conditions for a minimum.

It is important to keep the picture of these four relationships in mind. As we have seen so far, and as we shall see in the remainder of this thesis, with appropriate choice of the variables $\mathbf{z}$ and $\mathbf{y}$, we are able to solve a number of interesting problems in estimation and control, and to study their properties. All these results are special

	Stochastic Problems	Deterministic Problems
Model	(i) $\langle \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix}, \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix} \rangle = \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}$	(ii) $\{z^d, y^d\}$
Problem	$\min_{\hat{\mathbf{z}} \in \mathcal{L}\{\mathbf{y}\}} \langle \mathbf{z} - \hat{\mathbf{z}}, \mathbf{z} - \hat{\mathbf{z}} \rangle$	$\min_{y^d} \begin{bmatrix} y^{d*} & z^{d*} \end{bmatrix} \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix} \begin{bmatrix} y^d \\ z^d \end{bmatrix}$
Solution	$\hat{\mathbf{z}} = K_o \mathbf{y}$ $K_o = R_{zy} R_y^{-1}$	$y_o^d = K_o^d z^d$ $K_o^d = -R_y^{-1} R_{yz}$
Cond. for Min.	$R_y > 0$	$R_y > 0$
Value at Min.	$R_z - R_{zy} R_y^{-1} R_{yz} = R_{z^d}^{-1}$	$z^{d*}(R_z - R_{zy} R_y^{-1} R_{yz}) z^d = z^{d*} R_{z^d}^{-1} z^d$
Model	(iii) $\langle \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix}, \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix} \rangle = \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}^{-1}$	(iv) $\{z, y\}$
Problem	$\min_{\hat{\mathbf{y}}^d \in \mathcal{L}\{\mathbf{z}^d\}} \langle \mathbf{y}^d - \hat{\mathbf{y}}^d, \mathbf{y}^d - \hat{\mathbf{y}}^d \rangle$	$\min_z \begin{bmatrix} z^* & y^* \end{bmatrix} \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}^{-1} \begin{bmatrix} z \\ y \end{bmatrix}$
Solution	$\hat{\mathbf{y}}^d = K_o^d \mathbf{z}^d$ $K_o^d = R_{y^d z^d} R_{z^d}^{-1} = -R_y^{-1} R_{yz}$	$z_o = K_o y$ $K_o = R_{zy} R_y^{-1}$
Cond. for Min.	$R_y^d = (R_y - R_{zy} R_y^{-1} R_{yz})^{-1} > 0$	$R_z - R_{zy} R_y^{-1} R_{yz} > 0$
Value at Min.	$R_{y^d}^{-1} - R_{y^d z^d} R_{z^d}^{-1} R_{z^d y^d} = R_y^{-1}$	$y^* R_y^{-1} y$

Table 6.1: General equivalences and dualities.

cases of these four problems, when we specialize to a particular choice and structure for $\mathbf{z}$ and $\mathbf{y}$.

We can now explain how these four relations may be used to solve various problems. Suppose we are given an optimization problem of the form (iv), *i.e.* ,

$$\min_z \begin{bmatrix} z^* & y^* \end{bmatrix} \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}^{-1} \begin{bmatrix} z \\ y \end{bmatrix}. \tag{6.4.5}$$

Then to obtain the solution we may proceed in either one of two ways. We can construct an *equivalent* Krein space model of the form

$$\{\mathbf{z}, \mathbf{y}\}, \tag{6.4.6}$$

with

$$\langle \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix}, \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix} \rangle = \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}, \tag{6.4.7}$$

where R_z, $R_{zy} = R_{yz}^*$ and R_y are given in (iv). Then the projection of $\mathbf{z}$ onto $\mathbf{y}$ provides the solution to (iv).

Or, we can construct the *dual* Krein space model

$$\{\mathbf{z}^d, \mathbf{y}^d\}, \tag{6.4.8}$$

with

$$\langle \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix}, \begin{bmatrix} \mathbf{z}^d \\ \mathbf{y}^d \end{bmatrix} \rangle = \begin{bmatrix} R_z & R_{zy} \\ R_{yz} & R_y \end{bmatrix}^{-1}, \tag{6.4.9}$$

where, once more, R_z, $R_{zy} = R_{yz}^*$ and R_y are given in (iv). Now if we find the projection of $\mathbf{y}^d$ onto $\mathbf{z}^d$, we can use the negative conjugate transpose of the matrix that performs this projection to solve problem (iv).

Which one of the above two approaches to use depends upon the application at hand. It so turns out that for estimation and adaptive filtering problems it is more natural to use the equivalent Krein space model, whereas for control problems it more natural to use the dual Krein space model.

6.5 Dual State-Space Models

In this section we shall study the structure of the dual bases when the observations $\{\mathbf{y}_i\}$ have an underlying state-space structure. As demonstrated in the previous section, the dual bases $\mathbf{z}^d$ and $\mathbf{y}^d$ span the same space as the *smoothed* estimation errors $\tilde{\mathbf{z}}_{|y}$ and $\tilde{\mathbf{y}}_{|z}$, respectively. Thus, in a certain sense, we will also be obtaining state-space models for the smoothed estimation errors as well. The state-space models obtained here will be useful in the study of smoothing problems, in obtaining the information-form Kalman filter, and especially in the study of LQR control problems.

6.5.1 The Backwards Dual Model

Consider once more the time-varying state-space model

$$\begin{cases} \mathbf{x}_{i+1} &= F_i\mathbf{x}_i + \mathbf{u}_i, \quad 0 \le i \le N \\ \mathbf{y}_i &= H_i\mathbf{x}_i + \mathbf{v}_i \end{cases} \tag{6.5.1}$$

with

$$\left\langle \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{u}_i \\ \mathbf{v}_i \end{bmatrix}, \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{u}_j \\ \mathbf{v}_j \end{bmatrix} \right\rangle = \begin{bmatrix} \Pi_0 & 0 & 0 \\ 0 & Q_i\delta_{ij} & 0 \\ 0 & 0 & R_i\delta_{ij} \end{bmatrix}.$$

Recall the definitions of the observability map,

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

and the impulse response matrix,

$$\Gamma = \begin{bmatrix} 0 & & & \\ H_1G_0 & 0 & & \\ H_2F_1G_0 & H_2G_1 & & \\ \vdots & \vdots & \ddots & \\ H_NF_{N-1}\dots F_1G_0 & H_NF_{N-1}\dots F_2G_1 & \dots & H_NG_{N-1} \end{bmatrix},$$

so that we may write

$$\mathbf{y} = \mathcal{O}\mathbf{x}_0 + \Gamma\mathbf{u} + \mathbf{v} = \begin{bmatrix} \mathcal{O} & \Gamma & I \end{bmatrix} \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \mathcal{O} & \Gamma \end{bmatrix} & I \end{bmatrix} \begin{bmatrix} \mathbf{z} \\ \mathbf{v} \end{bmatrix},$$

where we have defined, $\mathbf{z} \triangleq \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{u} \end{bmatrix}$. We are now interested in $\{\mathbf{z}^d, \mathbf{y}^d\}$, the dual basis to $\{\mathbf{z}, \mathbf{y}\}$. To this end note that

$$\langle \begin{bmatrix} \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{u} \end{bmatrix} \\ \mathbf{v} \end{bmatrix}, \begin{bmatrix} \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{u} \end{bmatrix} \\ \mathbf{v} \end{bmatrix} \rangle = \begin{bmatrix} \begin{bmatrix} \Pi_0 & 0 \\ 0 & Q \end{bmatrix} & 0 \\ 0 & R \end{bmatrix}$$

where

$$Q \triangleq Q_0 \oplus \ldots \oplus Q_N \quad \text{and} \quad R \triangleq R_0 \oplus \ldots \oplus R_N.$$

We now have the following interesting result.

Lemma 6.5.1 (Dual Model for State-Space Structure) *Consider the linear state-space model*

$$\mathbf{y} = \begin{bmatrix} \mathcal{O} & \Gamma \end{bmatrix} \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{u} \end{bmatrix} + \mathbf{v},$$

with

$$\langle \begin{bmatrix} \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{u} \end{bmatrix} \\ \mathbf{v} \end{bmatrix}, \begin{bmatrix} \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{u} \end{bmatrix} \\ \mathbf{v} \end{bmatrix} \rangle = \begin{bmatrix} \begin{bmatrix} \Pi_0 & 0 \\ 0 & Q \end{bmatrix} & 0 \\ 0 & R \end{bmatrix},$$

and Π_0, Q *and* R *all invertible. Then choosing* $\mathbf{z} = \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{u} \end{bmatrix}$, *the dual basis* $\{\mathbf{z}^d, \mathbf{y}^d\}$ *is given by*

$$\mathbf{z}^d = -\begin{bmatrix} \mathcal{O}^* \\ \Gamma^* \end{bmatrix} \mathbf{y}^d + \mathbf{v}^c, \tag{6.5.2}$$

where

$$\mathbf{y}^d = R^{-1}\mathbf{v} \quad \textit{and} \quad \mathbf{v}^c = \begin{bmatrix} \Pi_0^{-1}\mathbf{x}_0 \\ Q^{-1}\mathbf{u} \end{bmatrix}. \tag{6.5.3}$$

In addition we may write

$$\mathbf{z}^d = \begin{bmatrix} -\mathcal{O}^* R^{-1}\mathbf{v} + \Pi_0^{-1}\mathbf{x}_0 \\ -\Gamma^* R^{-1}\mathbf{v} + Q^{-1}\mathbf{u} \end{bmatrix} = \begin{bmatrix} R_{\tilde{x}_{0|y}} & R_{\tilde{x}_{0|y}\tilde{u}_{|y}} \\ R_{\tilde{u}_{|y}\tilde{x}_{0|y}} & R_{\tilde{u}_{|y}} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\mathbf{x}}_{0|y} \\ \tilde{\mathbf{u}}_{|y} \end{bmatrix} = \begin{bmatrix} R^{-1}_{\tilde{x}_{0|y,u}} \tilde{\mathbf{x}}_{0|y,u} \\ R^{-1}_{\tilde{u}_{|y,x_0}} \tilde{\mathbf{u}}_{|y,x_0} \end{bmatrix}. \tag{6.5.4}$$

Proof: Eqs. (6.5.2) and (6.5.3) follow immediately from the dual bases formulae for linear models given in Lemma 6.3.5. The first of the equalities in (6.5.4), *i.e.*,

$$\mathbf{z}^d = \begin{bmatrix} -\mathcal{O}^* R^{-1}\mathbf{v} + \Pi_0^{-1}\mathbf{x}_0 \\ -\Gamma^* R^{-1}\mathbf{v} + Q^{-1}\mathbf{u} \end{bmatrix},$$

follows from (6.5.2-6.5.3), the second equality, *i.e.*,

$$\mathbf{z}^d = \begin{bmatrix} R_{\tilde{x}_{0|y}} & R_{\tilde{x}_{0|y}\tilde{u}_{|y}} \\ R_{\tilde{u}_{|y}\tilde{x}_{0|y}} & R_{\tilde{u}_{|y}} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\mathbf{x}}_{0|y} \\ \tilde{\mathbf{u}}_{|y} \end{bmatrix},$$

follows from the equation $\mathbf{z}^d = R^{-1}_{\tilde{z}_{|y}} \tilde{\mathbf{z}}_{|y}$, given in Lemma 6.3.2, and the third equality, from the fact that

$$\begin{bmatrix} R_{\tilde{x}_{0|y}} & R_{\tilde{x}_{0|y}\tilde{u}_{|y}} \\ R_{\tilde{u}_{|y}\tilde{x}_{0|y}} & R_{\tilde{u}_{|y}} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\mathbf{x}}_{0|y} \\ \tilde{\mathbf{u}}_{|y} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{x}}^d_{0|y} \\ \tilde{\mathbf{u}}^d_{|y} \end{bmatrix} = \begin{bmatrix} R^{-1}_{\tilde{x}_{0|y,u}} \tilde{\mathbf{x}}_{0|y,u} \\ R^{-1}_{\tilde{u}_{|y,x_0}} \tilde{\mathbf{u}}_{|y,x_0} \end{bmatrix}.$$

■

Note that the above lemma gives a dual linear model (viz. (6.5.2-6.5.3)) that describes the dual bases $\{\mathbf{z}^d, \mathbf{y}^d\}$ and also gives a physical interpretation for $\mathbf{z}^d$ in terms of various smoothing errors. It turns out, however, that since we have state-space structure, we can give an even more explicit representation of the dual basis, $\mathbf{z}^d$, and in fact show that it also possesses state-space structure.

Lemma 6.5.2 (Backwards Dual State-Space Model) *The vector* $\mathbf{z}^d$ *in the dual basis* $\{\mathbf{z}^d, \mathbf{y}^d\}$ *of Lemma 6.5.1 is given by*

$$\mathbf{z}^d = \begin{bmatrix} \xi_0^d + \Pi_0^{-1}\mathbf{x}_0 \\ \eta^d \end{bmatrix}, \tag{6.5.5}$$

where $\eta^d = col\{\eta_0^d, \ldots, \eta_N^d\}$ *has the following backwards-time state-space model*

$$\left\{ \begin{array}{lcl} \xi_i^d & = & F_i^* \xi_{i+1}^d - H_i^* R_i^{-1} \mathbf{v}_i \quad \xi_{N+1}^d = 0 \\ \eta_i^d & = & G_i^* \xi_{i+1}^d + Q_i^{-1} \mathbf{u}_i \end{array} \right. , \quad i = N, N-1, \ldots 0. \tag{6.5.6}$$

and where ξ_0^d *is its final state.*

Moreover, the (backwards) estimation errors of the states, $\tilde{\xi}_{i|i}^d \triangleq \xi_i^d - \hat{\xi}_{i|i}^d$*, given the (future) observations* $\{\eta_i^d, \ldots, \eta_N^d\}$ *in (6.5.6) have the following interpretation*

$$\tilde{\xi}_{i|i}^d = P_{i|i}^{-b} \tilde{\mathbf{x}}_{i|i}^b - \Pi_i^{-1} \mathbf{x}_i, \tag{6.5.7}$$

where $\tilde{\mathbf{x}}_{i|i}^b \triangleq \mathbf{x}_i - \hat{\mathbf{x}}_{i|i}^b$ *is the (backwards) error in estimating* $\mathbf{x}_i$ *using the (future) observations* $\{\mathbf{y}_i, \ldots, \mathbf{y}_N\}$*, where* $P_{i|i}^b$ *is the corresponding error Gramian, and where* Π_i *is the Gramian of* $\mathbf{x}_i$*.*

Proof: Note from Lemma 6.5.1 that we need to show

$$\left[\begin{array}{c} -\mathcal{O}^* R^{-1} \mathbf{v} + \Pi_0^{-1} \mathbf{x}_0 \\ -\Gamma^* R^{-1} \mathbf{v} + Q^{-1} \mathbf{u} \end{array} \right] = \left[\begin{array}{c} \xi_0^d + \Pi_0^{-1} \mathbf{x}_0 \\ \eta^d \end{array} \right].$$

Now the state space model (6.5.6) can be written in matrix form as

$$\left[\begin{array}{c} \eta_0^d \\ \eta_1^d \\ \vdots \\ \eta_N^d \end{array} \right] = \mathcal{C}^* \xi_{N+1}^d - \Gamma^* R^{-1} \mathbf{v} + Q^{-1} \mathbf{u}$$

where

$$\mathcal{C} = \left[\begin{array}{cccc} F_N F_{N-1} \ldots F_1 G_0 & F_N F_{N-1} \ldots F_2 G_1 & \ldots & G_N \end{array} \right],$$

is the controllability map of the original state space model. But when $\xi_{N+1}^d = 0$ this is the matrix relation defining η^d.

To prove that the first entry of the dual basis is $\xi_0^d + \Pi_0^{-1} \mathbf{x}_0$, we use the state-space recursion to write

$$\xi_0^d = \Phi_F^* \xi_{N+1}^d - \mathcal{O} R^{-1} \mathbf{v} = -\mathcal{O} R^{-1} \mathbf{v},$$

where $\Phi_F^* = F_0^* \ldots F_N^*$ is the state transition matrix corresponding to (6.5.6) and since $\xi_{N+1}^d = 0$. But if we recall that the first entry of $\mathbf{z}^d$ is

$$-\mathcal{O}^* R^{-1}\mathbf{v} + \Pi_0^{-1}\mathbf{x}_0,$$

we obtain the desired result.

Finally, to prove (6.5.7) let us first recall from (6.5.4) that

$$\mathbf{z}^d = \begin{bmatrix} \xi_0^d + \Pi_0^{-1}\mathbf{x}_0 \\ \eta^d \end{bmatrix} = \begin{bmatrix} R_{\tilde{x}_{0|y}} & R_{\tilde{x}_{0|y}\tilde{u}_{|y}} \\ R_{\tilde{u}_{|y}\tilde{x}_{0|y}} & R_{\tilde{u}_{|y}} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\mathbf{x}}_{0|y} \\ \tilde{\mathbf{u}}_{|y} \end{bmatrix} \tag{6.5.8}$$

and that therefore the Gramian of $\mathbf{z}^d$ is

$$\begin{bmatrix} R_{\tilde{x}_{0|y}} & R_{\tilde{x}_{0|y}\tilde{u}_{|y}} \\ R_{\tilde{u}_{|y}\tilde{x}_{0|y}} & R_{\tilde{u}_{|y}} \end{bmatrix}^{-1}.$$

Let us now consider $\{\xi_0^d + \Pi_0^{-1}\mathbf{x}_0, \eta^d\}$ and try to construct its dual basis $\{(\xi_0^d + \Pi_0^{-1}\mathbf{x}_0)^d, (\eta^d)^d\}$. This is straightforward to find, since

$$\begin{bmatrix} (\xi_0^d + \Pi_0^{-1}\mathbf{x}_0)^d \\ (\eta^d)^d \end{bmatrix} = \begin{bmatrix} R_{\tilde{x}_{0|y}} & R_{\tilde{x}_{0|y}\tilde{u}_{|y}} \\ R_{\tilde{u}_{|y}\tilde{x}_{0|y}} & R_{\tilde{u}_{|y}} \end{bmatrix} \begin{bmatrix} \xi_0^d + \Pi_0^{-1}\mathbf{x}_0 \\ \eta^d \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{x}}_{0|y} \\ \tilde{\mathbf{u}}_{|y} \end{bmatrix},$$

where in the last equality we have used (6.5.8). Therefore we conclude

$$(\xi_0^d + \Pi_0^{-1}\mathbf{x}_0)^d = \tilde{\mathbf{x}}_{0|y}.$$

But on the other hand

$$(\xi_0^d + \Pi_0^{-1}\mathbf{x}_0)^d = M^{-1}(\tilde{\xi}_{0|\eta}^d + \Pi_0^{-1}\tilde{\mathbf{x}}_{0|\eta}),$$

where M is the Gramian of $\tilde{\xi}_{0|\eta}^d + \Pi_0^{-1}\tilde{\mathbf{x}}_{0|\eta}$. But since η is a function of only the $\{\mathbf{u}_i, \mathbf{v}_i\}$, we conclude that $\tilde{\mathbf{x}}_{0|\eta} = \mathbf{x}_0$, and therefore

$$M^{-1}(\tilde{\xi}_{0|\eta}^d + \Pi_0^{-1}\mathbf{x}_0) = \tilde{\mathbf{x}}_{0|y}.$$

Taking the Gramian of both sides of the above equation we conclude that $M = R_{x_{0|y}}^{-1}$. Therefore we may write

$$\tilde{\xi}_{0|\eta}^d = R_{x_{0|y}}^{-1}\tilde{\mathbf{x}}_{0|y} - \Pi_0^{-1}\mathbf{x}_0.$$

But noting that $\tilde{\mathbf{x}}_{0|y} = \tilde{\mathbf{x}}^b_{0|0}$ and $R_{x_{0|y}} = P^b_{0|0}$, we have

$$\tilde{\xi}^d_{0|0} = P^{-b}_{0|0}\tilde{\mathbf{x}}^b_{0|0} - \Pi_0^{-1}\mathbf{x}_0.$$

Since, in principle, we could have taken any other time instant (instead of $i=0$) as the initial point, we also have

$$\tilde{\xi}^d_{i|i} = P^{-b}_{i|i}\tilde{\mathbf{x}}^b_{i|i} - \Pi_i^{-1}\mathbf{x}_i,$$

which is the desired result.

■

The backwards-time state-space model

$$\begin{cases} \xi^d_i &= F_i^*\xi^d_{i+1} - H_i^* R_i^{-1}\mathbf{v}_i \quad \xi^d_{N+1} = 0 \\ \eta^d_i &= G_i^*\xi^d_{i+1} + Q_i^{-1}\mathbf{u}_i \end{cases}, \quad i = N, N-1, \ldots 0. \tag{6.5.9}$$

is well-known in the system theory and is referred to as the *adjoint* state-space model, or as the backwards-time *complementary* model (since its output spans the orthogonal complement space of the output of the original model) [WD81, AK89]. However, we shall continue to call it the backwards dual model since its output describes the dual basis.

The duality with the original model is quite explicit. Indeed, we see that if we perform the following transformations;

$$F_i \to F_i^* \ , \ G_i \to H_i^* \ , \ H_i \to G_i^* \ , \ Q_i \to R_i^{-1} \ , \ R_i \to Q_i^{-1}$$

and reverse forward time to backward time we can obtain the dual state-space model from its original.

One further remark is that, as we shall see in Sec. 6.7 the above backwards dual state-space model will be very useful in the study of LQR control problems.

We close this section with one interesting observation regarding the output of the backwards dual state-space model (6.5.6).

Corollary 6.5.1 (An Orthogonality Result) *Consider the backwards dual model (6.5.6) and the original state-space model (6.5.1). Then we have*

$$\mathbf{x}_i \ \perp \ \eta^d_j \quad j = i, i+1, \ldots, N. \tag{6.5.10}$$

Proof: Follows readily from the orthogonality of the $\{\mathbf{x}_0, \{\mathbf{u}_i\}, \{\mathbf{v}_i\}\}$ and the facts that

$$\mathbf{x}_i \in \mathcal{L}\{\mathbf{u}_0, \mathbf{v}_0, \ldots \mathbf{u}_{i-1}, \mathbf{v}_{i-1}\} \quad \text{and} \quad \eta_j \in \mathcal{L}\{\mathbf{u}_j, \mathbf{v}_j \ldots \mathbf{u}_N, \mathbf{v}_N\}.$$

■

6.5.2 The Forwards Dual Model

In the previous section we obtained the dual basis to $\{\mathbf{z}, \mathbf{y}\}$ where $\mathbf{z} = \text{col}\{\mathbf{x}_0, \mathbf{u}\}$ and $\mathbf{y}$ were related via the standard state-space model (6.5.1). This led us to the basis $\{\mathbf{z}^d, \mathbf{y}^d\}$ of Lemma 6.5.1, where one component of $\mathbf{z}^d$ satisfied the backwards-time state-space model (6.5.6).

As is obvious from the defining relation (6.3.3), the dual basis depends on our choice of the original basis. In the state-space context, $\{\mathbf{u}, \mathbf{x}_0, \mathbf{y}\}$ is just one choice of basis for the underlying space $\mathcal{L}\{\mathbf{u}, \mathbf{x}_0, \mathbf{y}\}$, and alternative dual bases are found when one chooses different bases for $\mathcal{L}\{\mathbf{u}, \mathbf{x}_0, \mathbf{y}\}$. In this and the next section, we shall study the consequences of choosing such different bases.

When the matrices $\{F_i\}_{i=0}^N$ are nonsingular, $\{\mathbf{u}, \mathbf{x}_{N+1}, \mathbf{y}\}$ is a basis for $\mathcal{L}\{\mathbf{u}, \mathbf{x}_0, \mathbf{y}\}$, since the matrix obtained in the transformation

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{x}_{N+1} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ \mathcal{C} & \Phi_F & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{x}_0 \\ \mathbf{y} \end{bmatrix}, \tag{6.5.11}$$

is nonsingular. (Recall that $\mathcal{C}$ and $\Phi_F = F_N \ldots F_0$ are the controllability and state-transition matrices for the state-space model (6.5.1), and that $\mathbf{x}_{N+1} = \Phi_F \mathbf{x}_0 + \mathcal{C}\mathbf{u}$.)

Defining $\mathbf{z}' \triangleq \text{col}\{\mathbf{u}, \mathbf{x}_{N+1}\}$, we can now consider $\{\mathbf{z}', \mathbf{y}\}$ as a basis for $\mathcal{L}\{\mathbf{u}, \mathbf{x}_0, \mathbf{y}\}$. Our goal is to find the dual basis $\{\mathbf{z}'^d, \mathbf{y}^d\}$.

Note that with the above transformation we may write the linear relationship between $\{\mathbf{u}, \mathbf{x}_{N+1}, \mathbf{v}\}$ as follows

$$\begin{aligned} \mathbf{y} &= \mathcal{O}\mathbf{x}_0 + \Gamma\mathbf{u} + \mathbf{v} \\ &= \mathcal{O}(\Phi^{-1}\mathbf{x}_{N+1} - \Phi^{-1}\mathcal{C}) + \Gamma\mathbf{u} + \mathbf{v} \\ &= \mathcal{O}\Phi^{-1}\mathbf{x}_{N+1} + (\Gamma - \mathcal{O}\Phi^{-1}\mathcal{C})\mathbf{u} + \mathbf{v} \end{aligned}$$

$$= \begin{bmatrix} \Gamma - \mathcal{O}\Phi^{-1}\mathcal{C} & \mathcal{O}\Phi^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{x}_{N+1} \end{bmatrix} + \mathbf{v}.$$

The above linear model for $\mathbf{y}$, allows us to apply the results of Lemma 6.3.5 in order to obtain the dual basis $\{\mathbf{z}'^d, \mathbf{y}^d\}$. Thus

$$\begin{aligned}
\mathbf{z}'^d &= -\begin{bmatrix} \Gamma^* - \mathcal{C}^*\Phi^{-*}\mathcal{O}^* \\ \Phi^{-*}\mathcal{O}^* \end{bmatrix} R^{-1}\mathbf{v} + \langle \begin{bmatrix} \mathbf{u} \\ \mathbf{x}_{N+1} \end{bmatrix}, \begin{bmatrix} \mathbf{u} \\ \mathbf{x}_{N+1} \end{bmatrix} \rangle^{-1} \begin{bmatrix} \mathbf{u} \\ \mathbf{x}_{N+1} \end{bmatrix} \\
&= -\begin{bmatrix} \Gamma^* - \mathcal{C}^*\Phi^{-*}\mathcal{O}^* \\ \Phi^{-*}\mathcal{O}^* \end{bmatrix} R^{-1}\mathbf{v} + \langle \begin{bmatrix} \mathbf{u} \\ \mathbf{x}_0 \end{bmatrix}, \begin{bmatrix} \mathbf{u} \\ \mathbf{x}_{N+1} \end{bmatrix} \rangle^{-1} \begin{bmatrix} \mathbf{u} \\ \mathbf{x}_0 \end{bmatrix} \\
&= -\begin{bmatrix} \Gamma^* - \mathcal{C}^*\Phi^{-*}\mathcal{O}^* \\ \Phi^{-*}\mathcal{O}^* \end{bmatrix} R^{-1}\mathbf{v} + \begin{bmatrix} Q & Q\mathcal{C}^* \\ 0 & \Pi_0\Phi^* \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{u} \\ \mathbf{x}_0 \end{bmatrix} \\
&= -\begin{bmatrix} \Gamma^* - \mathcal{C}^*\Phi^{-*}\mathcal{O}^* \\ \Phi^{-*}\mathcal{O}^* \end{bmatrix} R^{-1}\mathbf{v} + \begin{bmatrix} Q^{-1} & -\mathcal{C}^*\Phi^{-*}\Pi_0^{-1} \\ 0 & \Phi^{-*}\Pi_0^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{x}_0 \end{bmatrix}
\end{aligned}$$

from which we finally conclude

$$\mathbf{z}'^d = -\begin{bmatrix} -\mathcal{C}^*\Phi^{-*}\Pi_0^{-1}\mathbf{x}_0 - (\Gamma^* - \mathcal{C}^*\Phi^{-*}\mathcal{O}^*)R^{-1}\mathbf{v} + Q^{-1}\mathbf{u} \\ \Phi^{-*}\Pi_0^{-1}\mathbf{x}_0 - \Phi^{-*}\mathcal{O}^*R^{-1}\mathbf{v} \end{bmatrix}. \tag{6.5.12}$$

The above expression can be used to obtain a state-space model for the dual basis $\mathbf{z}'^d$. Indeed we have the following result.

Lemma 6.5.3 (Forwards Dual State-Space Model) *Consider the standard state-space model (6.5.1). Then the vector* $\mathbf{z}'^d$ *in the dual basis* $\{\mathbf{z}'^d, \mathbf{y}'^d\}$ *of* $\{\mathbf{z}', \mathbf{y}\}$ *(with* $\mathbf{z}' \triangleq col\{\mathbf{u}, \mathbf{x}_{N+1}\}$*) is given by*

$$\mathbf{z}^d = \begin{bmatrix} \eta'^d \\ \xi'^d_{N+1} \end{bmatrix}, \tag{6.5.13}$$

where $\eta'^d = col\{\eta'^d_0, \ldots, \eta'^d_N\}$ *has the following forwards-time state-space model*

$$\left\{ \begin{array}{rcl} \xi'^d_{i+1} &=& F_i^{-*}\xi'^d_i - F_i^{-*}H_i^*R_i^{-1}\mathbf{v}_i, \quad \xi'^d_0 = \Pi_0^{-1}\mathbf{x}_0 \\ \eta'^d_i &=& G_i^*F_i^{-*}\xi'^d_i - G_i^*F_i^{-*}H_i^*R_i^{-1}\mathbf{v}_i + Q_i^{-1}\mathbf{u}_i \end{array} \right. \quad i = 0, 1, \ldots N \tag{6.5.14}$$

and where ξ'^d_{N+1} *is its final state.*

Moreover, the prediction errors of the states, $\tilde{\xi}_i'^d \triangleq \xi_i'^d - \hat{\xi}_i'^d$, *given the (past) observations* $\{\eta_0'^d, \ldots, \eta_{i-1}'^d\}$ *in (6.5.14) have the following interpretation*

$$\tilde{\xi}_i'^d = P_i^{-1}\tilde{\mathbf{x}}_i, \tag{6.5.15}$$

where $\tilde{\mathbf{x}}_i \triangleq \mathbf{x}_i - \hat{\mathbf{x}}_i$ *is the prediction error in estimating* $\mathbf{x}_i$ *using the (past) observations* $\{\mathbf{y}_0, \ldots, \mathbf{y}_{i-1}\}$, *and where* P_i *is the corresponding error Gramian.*

Proof: Note that the state-space equations can be written in global form as

$$\eta'^d = \begin{bmatrix} \eta_0'^d \\ \eta_1'^d \\ \vdots \\ \mathbf{y}_N'^d \end{bmatrix} = \underbrace{\begin{bmatrix} G_0^* F_0^{-*} H_0^{-*} & & & \\ G_1^* F_1^{-*} F_0^{-*} H_0^{-*} & G_1^* F_1^{-*} H_1^{-*} & & \\ \vdots & \vdots & \ddots & \\ G_N^* F_N^{-*} \ldots F_0^{-*} H_0^{-*} & G_N^* F_N^{-*} \ldots F_1^{-*} H_1^{-*} & \ldots & G_N^* F_N^{-*} H_N^{-*} \end{bmatrix}}_{-\mathbf{\Gamma}^* + \mathcal{C}^* \mathcal{F}^{-*} \mathcal{O}^*} \begin{bmatrix} R_0^{-1}\mathbf{v}_0 \\ R_1^{-1}\mathbf{v}_1 \\ \ldots \\ R_N^{-1}\mathbf{v}_N \end{bmatrix}$$

$$- \begin{bmatrix} G_0^* F_0^{-*} \\ G_1^* F_1^{-*} F_0^{-*} \\ \vdots \\ G_N^* \ldots F_0^{-*} \end{bmatrix} \Pi_0^{-1}\mathbf{x}_0 + \begin{bmatrix} Q_0^{-1}\mathbf{u}_0 \\ Q_1^{-1}\mathbf{u}_1 \\ \ldots \\ Q_N^{-1}\mathbf{u}_N \end{bmatrix}. \tag{6.5.16}$$

which is precisely the equation defining η'^d in (6.5.12). Likewise using (6.5.14) we may write

$$\xi_{N+1}'^d = \Phi_F^{-*}\Pi_0^{-1}\mathbf{x}_0 - \Phi_F^{-*}\mathcal{O}^* R^{-1}\mathbf{v}.$$

Comparing with (6.5.12) we obtain the desired result that the econd component of $\mathbf{z}'^d$ is the end-state, $\xi_{N+1}'^d$.

The proof of (6.5.15) is analogous to the proof of (6.5.7) in Lemma 6.5.2, *i.e.*, we show that $(\xi_{N+1}'^d)^d = \tilde{\mathbf{x}}_{N+1}$ and that $(\xi_{N+1}'^d)^d = M^{-1}\tilde{\xi}_{N+1}'^d$ where M is the Gramian of $\tilde{\xi}_{N+1}'^d$. These two equalities then imply that $M = P_{N+1}$, so that we may write

$$\tilde{\xi}_{N+1}'^d = P_{N+1}^{-1}\tilde{\mathbf{x}}_{N+1}.$$

Since the endpoint $i = N+1$ is arbitrary we conclude that for all i,

$$\tilde{\xi}_i'^d = P_i^{-1}\tilde{\mathbf{x}}_i,$$

which is the desired result.

■

The forwards-time state-space model

$$\begin{cases} \xi_{i+1}'^{d} &= F_i^{-*}\xi_i'^{d} - F_i^{-*}H_i^*R_i^{-1}\mathbf{v}_i, \quad \xi_0'^{d} = \Pi_0^{-1}\mathbf{x}_0 \\ \eta_i'^{d} &= G_i^*F_i^{-*}\xi_i'^{d} - G_i^*F_i^{-*}H_i^*R_i^{-1}\mathbf{v}_i + Q_i^{-1}\mathbf{u}_i \end{cases} \qquad i = 0, 1, \ldots N \tag{6.5.17}$$

is less well-known in the system theory and is referred to as the forwards-time complementary state-space model [AK89]. However, as before, we shall continue to call it the backwards dual model since its output describes the dual basis.

The reader at this point may note that (6.5.14) is essentially the backwards dual state-space model of Lemma 6.5.2 whose time has been reversed. Indeed this is one way of obtaining the forwards-time dual model. Another route would be to construct the backwards Markovian model for the original state-space model (6.5.1), and to find its dual using the approach of the previous section. We shall not the provide the details of these derivations here and will leave them as an exercise for the interested reader.

One further remark here is that the above forwards dual state-space model allows one to obtain the information form Kalman filter.

We close this section with one interesting observation regarding the output of the forwards dual state-space model (6.5.14).

Corollary 6.5.2 (An Orthogonality Result) *Consider the forwards dual model (6.5.14) and the original state-space model (6.5.1). Then we have*

$$\mathbf{x}_i \ \perp \ \eta_j'^{d} \quad j = 0, 1, \ldots, i-1. \tag{6.5.18}$$

Proof: Note that $\mathcal{L}\{\eta'^{d}\} = \mathcal{L}\{\tilde{\mathbf{u}}_{|y,x_{N+1}}\}$ (since it is the second component of the dual basis $\mathbf{z}'^{d}$). Therefore we readily see that

$$\mathbf{x}_{N+1} \ \perp \ \eta_j'^{d} \quad j = 0, 1, \ldots, N.$$

But since the endpoint $i = N+1$ is arbitrary, we conclude that for any value of i,

$$\mathbf{x}_i \ \perp \ \eta_j'^{d} \quad j = 0, 1, \ldots, i-1,$$

which is the desired result.

■

6.5.3 The Mixed Dual Model

Note that if we assume that the $\{F_j\}_{j=0}^{i}$ are nonsingular, then, for any $i = 0, \ldots N+1$, we can take $\{\mathbf{u}_0, \ldots, \mathbf{u}_{i-1}, \mathbf{x}_i, \mathbf{u}_i, \ldots, \mathbf{u}_N, \mathbf{y}\}$ as a basis for $\mathcal{L}\{\mathbf{u}, \mathbf{x}_0, \mathbf{y}\}$. For this so-called mixed basis we can obtain the following result using the arguments presented in Secs. 6.5.1 and 6.5.2.

Lemma 6.5.4 (Mixed Dual Model) *Consider the standard state-space model (6.5.1). Then the dual basis to* $\{\mathbf{u}_0, \ldots, \mathbf{u}_{i-1}, \mathbf{x}_i, \mathbf{u}_i, \ldots, \mathbf{u}_N, \mathbf{y}\}$ *is given by*

$$\{\eta_0'^d, \ldots, \eta_{i-1}'^d, \xi_i^d + \xi_i'^d, \eta_i^d, \ldots \eta_N^d, R^{-1}\mathbf{v}\}, \tag{6.5.19}$$

where η_j^d *has the following backwards-time dual state-space model*

$$\left\{ \begin{array}{rcl} \xi_j^d &=& F_j^* \xi_{j+1}^d - H_j^* R_j^{-1} \mathbf{v}_j \quad \xi_{N+1}^d = 0 \\ \eta_j^d &=& G_j^* \xi_{j+1}^d + Q_j^{-1} \mathbf{u}_j \end{array} \right. , \quad j = N, N-1, \ldots i. \tag{6.5.20}$$

with ξ_i^d *its final state, and where* $\eta_j'^d$ *has the following forwards-time state-space model*

$$\left\{ \begin{array}{rcl} \xi_{j+1}'^d &=& F_j^{-*} \xi_j'^d - F_j^{-*} H_j^* R_j^{-1} \mathbf{v}_j, \quad \xi_0'^d = \Pi_0^{-1} \mathbf{x}_0 \\ \eta_j'^d &=& G_j^* F_j^{-*} \xi_j'^d - G_j^* F_j^{-*} H_j^* R_j^{-1} \mathbf{v}_j + Q_j^{-1} \mathbf{u}_j \end{array} \right. \quad i = 0, 1, \ldots i-1 \tag{6.5.21}$$

with $\xi_i'^d$ *its final state.*

Proof: Instead of actually deriving the dual basis using the approach of the previous sections, we shall verify that the expressions given in the Lemma are indeed the dual basis. To show this, we must verify that

$$\left\langle \begin{bmatrix} \eta_0'^d \\ \vdots \\ \eta_{i-1}'^d \\ \xi_i^d + \xi_i'^d \\ \eta_i^d \\ \vdots \\ \eta_N^d \\ R^{-1}\mathbf{v} \end{bmatrix}, \begin{bmatrix} \mathbf{u}_0 \\ \vdots \\ \mathbf{u}_{i-1} \\ \mathbf{x}_i \\ \mathbf{u}_i \\ \vdots \\ \mathbf{u}_N \\ \mathbf{y} \end{bmatrix} \right\rangle = I. \tag{6.5.22}$$

But $\langle R^{-1}\mathbf{v}, \mathbf{y}\rangle = I$ and from our construction of the backwards and forwards dual bases we already know that $\eta_k^d \perp \mathbf{y}_l$ and $\eta_k'^d \perp \mathbf{y}_l$, for all k and l. Thus to verify the inner products with respect to $\mathbf{y}$ in (6.5.22) we need only show

$$\xi_i^d + \xi_i'^d \ \perp \ \mathbf{y}_j \quad j = 0, 1 \ldots, N.$$

But for $j < i$ we have $\xi_i'^d \perp \mathbf{y}_j$, so we need to show

$$\xi_i^d \ \perp \ \mathbf{y}_j \quad j = 0, 1 \ldots, i-1$$

which is immediate since $\xi_i^d \in \mathcal{L}\{\mathbf{v}_i, \ldots, \mathbf{v}_N\}$. On the other hand, for $j \geq i$ we have $\xi_i^d + \Pi_i^{-1}\mathbf{x}_i \perp \mathbf{y}_j$, so we need to show

$$-\Pi_i^{-1}\mathbf{x}_i + \xi_i'^d \ \perp \ \mathbf{y}_j \quad j = i, i+11 \ldots, N$$

which follows from the fact that $\xi_i'^d = \Pi_i^{-1}\mathbf{x}_i + \mu_i$ where $\mu_i \in \mathcal{L}\{\mathbf{v}_0, \ldots, \mathbf{v}_{i-1}\}$.

From Lemmas 6.5.2 and 6.5.3, we already know that

$$\left\langle \begin{bmatrix} \eta_0'^d \\ \vdots \\ \eta_{i-1}'^d \\ \xi_i^d + \xi_i'^d \end{bmatrix}, \begin{bmatrix} \mathbf{u}_0 \\ \vdots \\ \mathbf{u}_{i-1} \end{bmatrix} \right\rangle = I \quad \text{and} \quad \left\langle \begin{bmatrix} \xi_i^d + \xi_i'^d \\ \eta_i^d \\ \vdots \\ \eta_N^d \end{bmatrix}, \begin{bmatrix} \mathbf{u}_i \\ \vdots \\ \mathbf{u}_N \end{bmatrix} \right\rangle = I.$$

Therefore all that remains to be shown is that

$$\left\langle \begin{bmatrix} \eta_0'^d \\ \vdots \\ \eta_{i-1}'^d \end{bmatrix}, \begin{bmatrix} \mathbf{u}_i \\ \vdots \\ \mathbf{u}_N \end{bmatrix} \right\rangle = 0 \quad , \quad \left\langle \begin{bmatrix} \eta_i^d \\ \vdots \\ \eta_N^d \end{bmatrix}, \begin{bmatrix} \mathbf{u}_0 \\ \\ \vdots \\ \\ \mathbf{u}_{i-1} \end{bmatrix} \right\rangle = 0 \;\; \text{and} \;\; \left\langle \begin{bmatrix} \eta_0'^d \\ \vdots \\ \eta_{i-1}'^d \\ \xi_i^d + \xi_i'^d \\ \eta_i^d \\ \vdots \\ \eta_N^d \end{bmatrix}, \mathbf{x}_i \right\rangle = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ I \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

But the first two requirements are obvious since the $\{\eta_j'^d\}_{j=0}^{i-1}$ and the $\{\eta_j^d\}_{j=i}^{N}$ depend only on $\{\mathbf{x}_0, \{\mathbf{u}_j\}_{j=0}^{i-1}, \{\mathbf{v}_j\}_{j=0}^{i-1}\}$ and $\{\{\mathbf{u}_j\}_{j=i}^{N}, \{\mathbf{v}_j\}_{j=i}^{N}\}$. respectively. Also, from our Corollaries 6.5.1 and 6.5.2, $\mathbf{x}_i$ is orthogonal to the $\{\eta_j^d\}_{j=i}^{N}$ and to the $\{\eta_j'^d\}_{j=0}^{i-1}$.

Thus the only requirement to verify is

$$\langle \xi_i^d + \xi_i'^d, \mathbf{x}_i \rangle = I.$$

But

$$\begin{aligned}
\langle \xi_i^d, \mathbf{x}_i \rangle &= \underbrace{\langle \tilde{\xi}_i^d + \hat{\xi}_i^d, \mathbf{x}_i \rangle = \langle \tilde{\xi}_i^d, \mathbf{x}_i \rangle}_{\text{since } \hat{\xi}_i^d \in \mathcal{L}\{\eta_j^d\}_{j=i}^{N}} \\
&= \langle P_{i|i}^{-b} \tilde{\mathbf{x}}_{i|i}^b - \Pi_i^{-1} \mathbf{x}_i, \mathbf{x}_i \rangle \\
&= \langle P_{i|i}^{-b} \tilde{\mathbf{x}}_{i|i}^b, \mathbf{x}_i \rangle - I \\
&= P_{i|i}^{-b} \langle \tilde{\mathbf{x}}_{i|i}^b, \tilde{\mathbf{x}}_{i|i}^b + \hat{\mathbf{x}}_{i|i}^b \rangle - I \\
&= P_{i|i}^{-b} \langle \tilde{\mathbf{x}}_{i|i}^b, \tilde{\mathbf{x}}_{i|i}^b \rangle - I = 0.
\end{aligned}$$

A similar argument shows that

$$\langle \xi_i'^d, \mathbf{x}_i \rangle = I,$$

so that

$$\langle \xi_i^d + \xi_i'^d, \mathbf{x}_i \rangle = 0 + I = I,$$

as desired.

∎

As we shall see in the section, the above mixed dual state-space model will be useful for deriving the so-called two-filter smoothing formulas [May66, Fra67].

6.6 Application to Smoothing

The key observation that allows one to use dual bases for solving smoothing problems is the following. Suppose we are given a Krein space variable $\mathbf{x} \in \mathcal{L}\{\mathbf{z}, \mathbf{y}\}$ and would like to find, $\hat{\mathbf{x}}_{|y}$, the smoothed estimate of $\mathbf{x}$ given $\mathbf{y}$. Then since $\mathbf{z}^d$ is one possible basis for the orthogonal complement space of $\mathcal{L}\{\mathbf{y}\}$ in $\mathcal{L}\{\mathbf{z}, \mathbf{y}\}$ (in other words, $\mathcal{L}\{\mathbf{z}, \mathbf{y}\} = \mathcal{L}\{\mathbf{y}\} \oplus \mathcal{L}\{\mathbf{z}^d\}$), we may uniquely decompose $\mathbf{x}$ as

$$\mathbf{x} = \hat{\mathbf{x}}_{|y} + \hat{\mathbf{x}}_{|z^d},$$

from which it follows

$$\hat{\mathbf{x}}_{|y} = \mathbf{x} - \hat{\mathbf{x}}_{|z^d}. \tag{6.6.1}$$

Thus we can obtain the desired estimate by projecting onto the dual basis, $\mathbf{z}^d$, and subtracting the result from $\mathbf{x}$.

The above approach is a general method for obtaining smoothed estimates. Of course the resulting algorithms will depend on the choice of $\mathbf{z}$ (and hence $\mathbf{z}^d$). By projecting onto the backwards and forwards dual state-space models of Lemmas 6.5.2 and 6.5.3 it is possible to derive the standard Bryson-Frazier and Rauch-Tung-Striebel (RTS) smoothing formulae and their variants [RTS65]. The algebra involved in these derivations is (at times) rather tedious and therefore we shall not present them here. [The interested reader may refer to [AK89] for details.] Instead, we shall illustrate the general method by deriving the so-called two-filter smoothing formulae by projecting onto the mixed dual state-space model of Lemma 6.5.4.

6.6.1 Two-Filter Formulae

Consider the standard state-space model (6.5.1) and suppose that we would like to obtain, $\hat{\mathbf{x}}_{i|N}$, the smoothed estimate of the state $\mathbf{x}_i$ given all the observations $\{\mathbf{y}_0, \ldots, \mathbf{y}_N\}$. To do so, consider the following mixed dual basis that was introduced in Lemma 6.5.4,

$$\{\eta_0'^d, \ldots, \eta_{i-1}'^d, \xi_i^d + \xi_i'^d, \eta_i^d, \ldots \eta_N^d, R^{-1}\mathbf{v}\}. \tag{6.6.2}$$

To find $\hat{\mathbf{x}}_{i|N}$ we will make use of the formula

$$\hat{\mathbf{x}}_{i|N} \triangleq \hat{\mathbf{x}}_{|y} = \mathbf{x}_i - \hat{\mathbf{x}}_{i|z^d}, \tag{6.6.3}$$

where $\mathbf{z}^d$ is the first part of the aforementioned mixed dual basis, *i.e.*,

$$\mathbf{z}^d = \text{col}\{\eta_0'^d, \ldots, \eta_{i-1}'^d, \xi_i^d + \xi_i'^d, \eta_i^d, \ldots \eta_N^d\}. \tag{6.6.4}$$

Now due to Corollaries 6.5.1 and 6.5.2, the state $\mathbf{x}_i$ is orthogonal to both $\{\eta_j^d\}_{j=i}^N$ and $\{\eta_j'^d\}_{j=0}^{i-1}$. Therefore to find $\hat{\mathbf{x}}_{i|z^d}$ we need only project onto $\xi_i^d + \xi_i'^d$.

Moreover, due to the above orthogonalities, projecting $\mathbf{x}_i$ onto $\xi_i^d + \xi_i'^d$ is the same as projecting $\mathbf{x}_i$ onto $\tilde{\xi}_{i|i}^d + \tilde{\xi}_i'^d$, where $\tilde{\xi}_{i|i}^d \triangleq \xi_i^d - \hat{\xi}_{i|i}^d$, is the state estimation error given the (future) observations $\{\eta_i^d, \ldots, \eta_N^d\}$ and $\tilde{\xi}_i'^d \triangleq \xi_i'^d - \hat{\xi}_i'^d$, is the state estimation error given the (past) observations $\{\eta_0'^d, \ldots, \eta_{i-1}'^d\}$. Now in Lemmas 6.5.2 and 6.5.3 we have shown

$$\tilde{\xi}_{i|i}^d = P_{i|i}^{-b}\tilde{\mathbf{x}}_{i|i}^b - \Pi_i^{-1}\mathbf{x}_i \quad \text{and} \quad \tilde{\xi}_i'^d = P_i^{-1}\tilde{\mathbf{x}}_i$$

where $\tilde{\mathbf{x}}_{i|i}^b \triangleq \mathbf{x}_i - \hat{\mathbf{x}}_{i|i}^b$ is the (backwards) error in estimating $\mathbf{x}_i$ using the (future) observations $\{\mathbf{y}_i, \ldots, \mathbf{y}_N\}$, with $P_{i|i}^b$ its corresponding error Gramian, and where $\tilde{\mathbf{x}}_i \triangleq \mathbf{x}_i - \hat{\mathbf{x}}_i$ is the prediction error in estimating $\mathbf{x}_i$ using the (past) observations $\{\mathbf{y}_0, \ldots, \mathbf{y}_{i-1}\}$, with P_i its corresponding error Gramian.

In view of the above arguments, we have

$$\hat{\mathbf{x}}_{i|z^d} = \langle \mathbf{x}_i, P_{i|i}^{-b}\tilde{\mathbf{x}}_{i|i}^b - \Pi_i^{-1}\mathbf{x}_i + P_i^{-1}\tilde{\mathbf{x}}_i\rangle \|P_{i|i}^{-b}\tilde{\mathbf{x}}_{i|i}^b - \Pi_i^{-1}\mathbf{x}_i + P_i^{-1}\tilde{\mathbf{x}}_i\|^{-2}\left(P_{i|i}^{-b}\tilde{\mathbf{x}}_{i|i}^b - \Pi_i^{-1}\mathbf{x}_i + P_i^{-1}\tilde{\mathbf{x}}_i\right). \tag{6.6.5}$$

But

$$\langle \mathbf{x}_i, P_{i|i}^{-b}\tilde{\mathbf{x}}_{i|i}^b - \Pi_i^{-1}\mathbf{x}_i + P_i^{-1}\tilde{\mathbf{x}}_i\rangle = I - I + I = I, \tag{6.6.6}$$

and

$$\|P_{i|i}^{-b}\tilde{\mathbf{x}}_{i|i}^b - \Pi_i^{-1}\mathbf{x}_i + P_i^{-1}\tilde{\mathbf{x}}_i\|^2 = \|\tilde{\xi}_{i|i}^d + \tilde{\xi}_i'^d\|^2 = \|\tilde{\xi}_{i|i}^d\|^2 + \|\tilde{\xi}_i'^d\|^2 = \|P_{i|i}^{-b}\tilde{\mathbf{x}}_{i|i}^b - \Pi_i^{-1}\mathbf{x}_i\|^2 + \|P_i^{-1}\tilde{\mathbf{x}}_i\|^2, \tag{6.6.7}$$

since $\tilde{\xi}_{i|i}^d \in \mathcal{L}\{\mathbf{u}_i, \mathbf{v}_i, \ldots, \mathbf{u}_N, \mathbf{v}_N\}$ and $\tilde{\xi}_i'^d \in \mathcal{L}\{\mathbf{x}_0, \mathbf{u}_0, \mathbf{v}_0, \ldots, \mathbf{u}_{i-1}, \mathbf{v}_{i-1}\}$ (see Eqs. (6.5.6) and (6.5.14)). Moreover,

$$\|P_{i|i}^{-b}\tilde{\mathbf{x}}_{i|i}^b - \Pi_i^{-1}\mathbf{x}_i\|^2 = P_{i|i}^{-b} - \Pi_i^{-1} \quad \text{and} \quad \|P_i^{-1}\tilde{\mathbf{x}}_i\|^2 = P_i^{-1}. \tag{6.6.8}$$

Therefore we may finally write

$$\mathbf{x}_i - \hat{\mathbf{x}}_{i|N} = \hat{\mathbf{x}}_{i|z^d} = I \cdot \left(P_{i|i}^{-b} - \Pi_i^{-1} + P_i^{-1}\right)^{-1}\left(P_{i|i}^{-b}\tilde{\mathbf{x}}_{i|i}^b - \Pi_i^{-1}\mathbf{x}_i + P_i^{-1}\tilde{\mathbf{x}}_i\right), \tag{6.6.9}$$

or

$$\left(P_{i|i}^{-b} - \Pi_i^{-1} + P_i^{-1}\right)\left(\mathbf{x}_i - \hat{\mathbf{x}}_{i|N}\right) = \left(P_{i|i}^{-b}\tilde{\mathbf{x}}_{i|i}^{b} - \Pi_i^{-1}\mathbf{x}_i + P_i^{-1}\tilde{\mathbf{x}}_i\right), \tag{6.6.10}$$

from which we conclude

$$\left(P_{i|i}^{-b} - \Pi_i^{-1} + P_i^{-1}\right)\hat{\mathbf{x}}_{i|N} = P_{i|i}^{-b}\hat{\mathbf{x}}_{i|i}^{b} + P_i^{-1}\hat{\mathbf{x}}_i. \tag{6.6.11}$$

Computing the Gramian of $\tilde{\mathbf{x}}_{i|N}$ from Eq. (6.6.9) we readily see

$$P_{i|N} = \left(P_{i|i}^{-b} - \Pi_i^{-1} + P_i^{-1}\right)^{-1}, \tag{6.6.12}$$

so that (6.6.11) becomes

$$\hat{\mathbf{x}}_{i|N} = P_{i|N}\left(P_{i|i}^{-b}\hat{\mathbf{x}}_{i|i}^{b} + P_i^{-1}\hat{\mathbf{x}}_i\right), \tag{6.6.13}$$

which is the desired two-filter smoother.

It is left as an exercise to the reader to show that if we had instead begun with the dual basis

$$\text{col}\{\eta_0'^d, \ldots, \eta_i'^d, \xi_i^d + \xi_i'^d, \eta_{i+1}^d, \ldots \eta_N^d\} \tag{6.6.14}$$

then we would have obtained the following two-filter formula

$$\hat{\mathbf{x}}_{i|N} = P_{i|N}\left(P_i^{-b}\hat{\mathbf{x}}_i^{b} + P_{i|i}^{-1}\hat{\mathbf{x}}_{i|i}\right), \tag{6.6.15}$$

with

$$P_{i|N}^{-1} = \left(P_i^{-b} + P_{i|i}^{-1} - \Pi_i^{-1}\right)^{-1}, \tag{6.6.16}$$

where $\hat{\mathbf{x}}_i^b$, $\hat{\mathbf{x}}_{i|i}$, P_i^b and $P_{i|i}$ have the obvious meanings.

The results are summarized in the following theorem.

Theorem 6.6.1 (Two-Filter Smoothing Formulae) *Consider the standard state-space model (6.5.1) with the $\{F_i\}$ assumed to be invertible. Then we can write*

$$\hat{\mathbf{x}}_{i|N} = P_{i|N}(P_i^{-1}\hat{\mathbf{x}}_i + P_{i|i}^{-b}\hat{\mathbf{x}}_{i|i}^{b}),$$

with

$$P_{i|N}^{-1} = \left(P_i^{-1} + P_{i|i}^{-b} - \Pi_i^{-1}\right)^{-1},$$

or, alternatively,

$$\hat{\mathbf{x}}_{i|N} = P_{i|N}(P_{i|i}^{-1}\hat{\mathbf{x}}_{i|i} + P_i^{-b}\hat{\mathbf{x}}_i^b),$$

with

$$P_{i|N}^{-1} = \left(P_{i|i}^{-1} + P_i^{-b} - \Pi_i^{-1}\right)^{-1}.$$

Both sets of formulae require forwards and backwards sweeps over the original data: the first formulae require a forwards sweep from $\mathbf{y}_0$ *to* $\mathbf{y}_i$ *and a backwards sweep from* $\mathbf{y}_N$ *down to* $\mathbf{y}_{i+1}$*, while the second formulae require a forwards sweep from* $\mathbf{y}_0$ *to* $\mathbf{y}_{i-1}$ *and a backwards sweep from* $\mathbf{y}_N$ *down to* $\mathbf{y}_i$*.*

6.7 The LQR Control Problem

In this section we shall study the dual problem to Kalman filtering, *i.e.* the linear quadratic regulator (LQR) control problem. The generalization considered here is that the LQR cost function may be indefinite (actually problems in H^∞ control immediately result in such cost functions). Therefore, contrary to the conventional case where the problem always had a minimizing solution, we can at most guarantee a stationarizing solution. Further conditions must be met for the solution to be a minimum.

The usual method for solving the conventional LQR problem is through a dynamic programming argument (see *e.g.* ,[BK65, BH69]). Although the dynamic programming principle can be extended from minimizing solutions to include stationarizing solutions, our approach to solving the LQR problem is similar to the approach we used for solving H^∞ estimation: we identify a Krein state-space model that corresponds to the LQR quadratic form (via simple inspection), and then invoke the Krein space Kalman filter to write down the stationarizing solution and its condition for being a minimum. We should also remark that the Krein space model that most naturally lends itself to the solution of the LQR control problem is the dual model of Sec. 6.4, which is what will be used here.

Finally, we should mention that the results of this section will be used to solve the finite horizon full information and measurement feedback control problems.

6.7.1 Problem Formulation

Consider the (possibly) time-varying state-space model

$$x_{i+1} = F_i x_i + G_i u_i, \qquad 0 \le i \le N \tag{6.7.1}$$

and some given linear combination of the states

$$s_i = L_i x_i, \tag{6.7.2}$$

where the $\{u_i, x_i\}$ are deterministic quantities. One would like to choose the control signal u_i to regulate the s_i in a certain sense. The conventional LQR (linear quadratic regulator) method proposes the following criterion for choosing the u_i

$$\min_{\{u_i\}} \quad x_{N+1}^* P_{N+1}^c x_{N+1} + \sum_{i=0}^{N} u_i^* Q_i^c u_i + \sum_{i=0}^{N} s_i^* R_i^c s_i \tag{6.7.3}$$

subject to the state-space constraints (6.7.1) and (6.7.2), where P_{N+1}^c, Q_i^c and R_i^c are positive-definite matrices that penalize the final state, the inputs and the intermediary states, respectively.

In many applications, such as the H^∞ problems to be studied, one needs to consider the quadratic form in (6.7.3) when the P_{N+1}^c, Q_i^c and R_i^c are indefinite Hermitian matrices. Therefore, let us introduce the notation for the quadratic form in (6.7.3) as

$$J_N^c(x_0, u_0, \ldots, u_N) = x_{N+1}^* P_{N+1}^c x_{N+1} + \sum_{i=0}^{N} u_i^* Q_i^c u_i + \sum_{i=0}^{N} s_i^* R_i^c s_i, \tag{6.7.4}$$

where from now on we shall drop the positivity conditions on the P_{N+1}^c, Q_i^c and R_i^c. Since J_N^c will now be an indefinite quadratic form, it does not necessarily have a minimum over the $\{u_i\}$, and the most that we can guarantee is a stationary point.

With this in mind, we may now state the finite horizon LQR control problem as follows.

Problem 6.7.1 (Finite Horizon LQR Control Problem) *Find the control signals* $\{\hat{u}_i\}_{i=0}^N$ *that stationarize the quadratic form*

$$J_N^c(x_0, u_0, \ldots, u_N) = x_{N+1}^* P_{N+1}^c x_{N+1} + \sum_{i=0}^{N} u_i^* Q_i^c u_i + \sum_{i=0}^{N} s_i^* R_i^c s_i$$

subject to the state-space relations (6.7.1) and (6.7.2), where by a stationary point we mean

$$\frac{\partial J_N^c}{\partial u_i}(x_0, \hat{u}_0, \ldots, \hat{u}_N) = 0, \qquad 0 \leq i \leq N.$$

Moreover, find the necessary and sufficient conditions for this stationary point to be a minimum.

6.7.2 Solution Based on Duality

Let us begin by defining $u = \text{col}\{u_0, \ldots, u_N\}$ and $s = \text{col}\{s_0, \ldots, s_N\}$, so that we may use the state equation (6.7.1) and the output equation (6.7.2) to write

$$x_{N+1} = \Phi_F x_0 + \mathcal{C}u \quad \text{and} \quad y = \mathcal{O}x_0 + \Gamma \tag{6.7.5}$$

where $\Phi_F = F_N F_{N-1} \ldots F_0$ is the state transition matrix,

$$\mathcal{O} = \begin{bmatrix} L_0 \\ L_1 F_0 \\ \vdots \\ L_N F_{N-1} \ldots F_0 \end{bmatrix},$$

is the observability map,

$$\mathcal{C} = \begin{bmatrix} F_N F_{N-1} \ldots F_1 G_0 & F_N F_{N-1} \ldots F_2 G_1 & \ldots & G_0 \end{bmatrix},$$

is the controllability map, and

$$\Gamma = \begin{bmatrix} 0 & & & \\ H_1 G_0 & 0 & & \\ H_2 F_1 G_0 & H_2 G_1 & & \\ \vdots & \vdots & \ddots & \\ H_N F_{N-1} \ldots F_1 G_0 & H_N F_{N-1} \ldots F_2 G_1 & \ldots & H_N G_{N-1} \end{bmatrix},$$

is the impulse response matrix. If we further define,

$$Q^c = \text{diag}\{Q_0^c, Q_1^c, \ldots, Q_N^c\} \quad \text{and} \quad R^c = \text{diag}\{R_0^c, R_1^c, \ldots, R_N^c\},$$

then we can write the LQR cost function, J_N^c, as a quadratic form in terms of x_0 and u. We thus have

$$J_N^c = (x_0^*\Phi_F^* + u^*\mathcal{C}^*)P_{N+1}^c(\Phi_F x_0 + \mathcal{C}u) + (x_0^*\mathcal{O}^* + u^*\Gamma^*)R^c(\mathcal{O}x_0 + \Gamma u) + u^*Q^c u,$$

or, after gathering terms,

$$J_N^c = \begin{bmatrix} x_0^* & u^* \end{bmatrix} \begin{bmatrix} \Phi_F^* P_{N+1}^c \Phi_F + \mathcal{O}^* R^c \mathcal{O} & \Phi_F^* P_{N+1}^c \mathcal{C} + \mathcal{O}^* R^c \Gamma \\ \mathcal{C}^* P_{N+1}^c \Phi_F + \Gamma^* R^c \mathcal{O} & \mathcal{C}^* P_{N+1}^c \mathcal{C} + \Gamma^* R^c \Gamma + Q^c \end{bmatrix} \begin{bmatrix} x_0 \\ u \end{bmatrix}. \tag{6.7.6}$$

Now if we define

$$\begin{cases} \mathbf{x}_0^c &= \Phi_F^* \mathbf{x}_{N+1}^c + \mathcal{O}^* \mathbf{u}^c \\ \mathbf{y}^c &= \mathcal{C}^* \mathbf{x}_{N+1}^c + \Gamma^* \mathbf{u}^c + \mathbf{v}^c \end{cases} \tag{6.7.7}$$

where $\mathbf{x}_{N+1}^c$, $\mathbf{u}^c$ and $\mathbf{v}^c$ are Krein space variables with

$$\left\langle \begin{bmatrix} \mathbf{x}_{N+1}^c \\ \mathbf{u}^c \\ \mathbf{v}^c \end{bmatrix}, \begin{bmatrix} \mathbf{x}_{N+1}^c \\ \mathbf{u}^c \\ \mathbf{v}^c \end{bmatrix} \right\rangle = \begin{bmatrix} P_{N+1}^c & 0 & 0 \\ 0 & R^c & 0 \\ 0 & 0 & Q^c \end{bmatrix},$$

then it is straightforward to see that

$$\left\langle \begin{bmatrix} \mathbf{x}_0^c \\ \mathbf{y}^c \end{bmatrix}, \begin{bmatrix} \mathbf{x}_0^c \\ \mathbf{y}^c \end{bmatrix} \right\rangle = \begin{bmatrix} \Phi_F^* P_{N+1}^c \Phi_F + \mathcal{O}^* R^c \mathcal{O} & \Phi_F^* P_{N+1}^c \mathcal{C} + \mathcal{O}^* R^c \Gamma \\ \mathcal{C}^* P_{N+1}^c \Phi_F + \Gamma^* R^c \mathcal{O} & \mathcal{C}^* P_{N+1}^c \mathcal{C} + \Gamma^* R^c \Gamma + Q^c \end{bmatrix}, \tag{6.7.8}$$

so that the LQR cost function in (6.7.6) can be rewritten as

$$J_N^c = \begin{bmatrix} x_0^* & u^* \end{bmatrix} \begin{bmatrix} R_{x_0^c} & R_{x_0^c y^c} \\ R_{y^c x_0^c} & R_{y^c} \end{bmatrix} \begin{bmatrix} x_0 \\ u \end{bmatrix}. \tag{6.7.9}$$

The minimizing solution over u for J_N^c is now readily seen to be

$$\hat{u} = -R_{y^c}^{-1} R_{y^c x_0^c} x_0, \tag{6.7.10}$$

i.e., its gain matrix $-R_{y^c}^{-1} R_{y^c x_0^c}$ is the negative (conjugate) transpose of the gain matrix in estimating the random variable $\mathbf{x}_0^c$ from $\mathbf{y}^c$. Note that this is the dual stochastic approach to deterministic least-squares problems that was introduced in Sec. 6.4. Therefore to find our desired solution we can project $\mathbf{x}_0^c$ onto $\mathbf{y}^c$ and use the negative (conjugate) transpose of the resulting gain matrix to find $\hat{u}$.

To this end, let us first note that (6.7.7) has the following backwards dual state-space model

$$\begin{cases} \mathbf{x}_i^c = F_i^* \mathbf{x}_{i+1}^c + L_i^* \mathbf{u}_i^c \quad , \quad \mathbf{x}_{N+1}^c \\ \mathbf{y}_i^c = G_i^* \mathbf{x}_{i+1}^c + \mathbf{v}_i^c \end{cases} \tag{6.7.11}$$

with

$$\left\langle \begin{bmatrix} \mathbf{u}_i^c \\ \mathbf{v}_i^c \\ \mathbf{x}_{N+1}^c \end{bmatrix}, \begin{bmatrix} \mathbf{u}_j^c \\ \mathbf{v}_j^c \\ \mathbf{x}_{N+1}^c \end{bmatrix} \right\rangle = \begin{bmatrix} R_i^c \delta_{ij} & 0 & 0 \\ 0 & Q_i^c \delta_{ij} & 0 \\ 0 & 0 & P_{N+1}^c \end{bmatrix},$$

as can be readily checked by defining $\mathbf{y}^c = \text{col}\{\mathbf{y}_0^c, \ldots, \mathbf{y}_N^c\}$, $\mathbf{u}^c = \text{col}\{\mathbf{u}_0^c, \ldots, \mathbf{u}_N^c\}$ and $\mathbf{v}^c = \text{col}\{\mathbf{v}_0^c, \ldots, \mathbf{v}_N^c\}$, noting the relations

$$\mathbf{x}_0^c = \Phi_F^* \mathbf{x}_{N+1}^c + \mathcal{O}^* \mathbf{u}^c \quad , \quad \mathbf{y}^c = \mathcal{C}^* \mathbf{x}_{N+1}^c + \Gamma^* \mathbf{u}^c + \mathbf{v}^c \tag{6.7.12}$$

and computing the respective Gramians.

To find the projection of $\mathbf{x}_0^c$ onto $\mathbf{y}^c$ we can invoke the Krein space Kalman filter corresponding to (6.7.11). Thus

$$\hat{\mathbf{x}}_i^c = F_i^* \hat{\mathbf{x}}_{i+1}^c + K_{p,i}^c \mathbf{e}_i^c, \quad \hat{\mathbf{x}}_{N+1}^c = 0 \tag{6.7.13}$$

where

$$\mathbf{e}_i^c = \mathbf{y}_i^c - G_i^* \hat{\mathbf{x}}_{i+1}^c, \tag{6.7.14}$$

is the innovations,

$$K_{p,i}^c = F_i^* P_{i+1}^c G_i (R_{e,i}^c)^{-1} \quad \text{and} \quad R_{e,i}^c = Q_i^c + G_i^* P_{i+1}^c G_i \tag{6.7.15}$$

are the Kalman gain and innovations variance, respectively, and where P_i^c satisfies the backwards-time Riccati recursion

$$P_i^c = F_i^* P_{i+1}^c F_i + L_i^* R_i^c L_i - K_{p,i}^c R_{e,i}^c K_{p,i}^{c*}, \quad P_{N+1}^c. \tag{6.7.16}$$

Now (6.7.13) may be rewritten as

$$\hat{\mathbf{x}}_i^c = \Phi_i \hat{\mathbf{x}}_{i+1}^c + K_{p,i}^c \mathbf{y}_i^c, \quad \hat{\mathbf{x}}_{N+1}^c = 0 \tag{6.7.17}$$

where

$$\Phi_i = (F_i^* - K_{p,i}^c G_i^*).$$

The recursion (6.7.17) can be solved to yield

$$\hat{\mathbf{x}}_0^c = \left[\begin{array}{cccc} K_{p,0}^c & \Phi_0 K_{p,1}^c & \dots & \Phi_0 \dots \Phi_{N-1} K_{p,N}^c \end{array}\right] \left[\begin{array}{c} \mathbf{y}_0^c \\ \mathbf{y}_1^c \\ \vdots \\ \mathbf{y}_N^c \end{array}\right]. \tag{6.7.18}$$

Now the solution to the LQR problem is given by the negative of the conjugate transpose of the above solution. Thus

$$\left[\begin{array}{c} \hat{u}_0 \\ \hat{u}_1 \\ \vdots \\ \hat{u}_N \end{array}\right] = -\left[\begin{array}{c} K_{p,0}^{c*} \\ K_{p,1}^{c*}\Phi_0^* \\ \vdots \\ K_{p,N}^{c*}\Phi_{N-1}^* \dots \Phi_0^* \end{array}\right] x_0. \tag{6.7.19}$$

But this implies

$$\hat{u}_0 = -K_{p,0}^{c*},$$

and

$$\begin{aligned} \hat{u}_1 = -K_{p,1}^{c*}\Phi_0^* &= -K_{p,1}^{c*}(F_0 - G_0 K_{p,0}^{c*})x_0 \\ &= -K_{p,1}^{c*}(F_0 x_0 + G_0 \hat{u}_0) = -K_{p,1}^{c*} x_1, \end{aligned}$$

and continuing in a similar fashion

$$\hat{u}_i = -K_{p,i}^{c*} x_i, \quad i = 0, \dots, N \tag{6.7.20}$$

which is the wellknown state-feedback law of LQR control.

Equations (6.7.15), (6.7.16) and (6.7.20) constitute the solution to the LQR control problem. Of course, we still need to check the condition for a minimum over, the $\{u_i\}$. But, using Lemma 6.2.1, this is readily seen to be

$$R_{y^c} = \mathcal{C}^* P_{N+1}^c \mathcal{C} + \Gamma^* R^c \Gamma + Q^c > 0. \tag{6.7.21}$$

Moreover, the Krein space Kalman filter corresponding to R_{y^c} allows us to recursively check the above condition via the innovations Gramian, viz.,

$$R_{e,i}^c = Q_i^c + G_i^* P_{i+1}^c G_i > 0, \quad i = 0, \dots, N. \tag{6.7.22}$$

Now the control signal, $\hat{u} = -R_{y^c}^{-1} R_{y^c x_0^c} x_0$, is referred to as the optimal *open-loop* control, since it only depends on the initial state, x_0. The control signal, $\hat{u}_i = -K_{p,i}^{c*} x_i$, on the other hand, is referred to as the optimal *closed-loop* control, since it depends only on the current state, x_i. Note that at any given time, i, the optimal closed-loop control signal, $\hat{u}_i = -K_{p,i}^{c*} x_i$, coincides with the optimal open-loop signal if, and only if, all previous choices of the control signal were also optimal. (Since non-optimal choices of $\{u_j, j < i\}$ affect the value of x_i, and thereby the value of the optimal closed-loop control, $\hat{u}_i = -K_{p,i}^{c*} x_i$.)

The closed-loop control signal, however, has the additional property that it is the optimal control at time i, *irrespective of what the previous control signals were.* To be more specific, it solves the LQR problem,

$$\min_{u_i} \sum_{j=i}^{N} u_j^* Q_j^c u_j + \sum_{j=i}^{N} s_j^* R_j^c s_j + x_{N+1}^* P_{N+1}^c x_{N+1}. \tag{6.7.23}$$

The validity of the above claim is readily seen from the fact that in the above argument, which led to the optimal choice of u_0, *i.e.*, $\hat{u}_0 = -K_{p,0}^{c*} x_0$, the choice of the initial time, $i = 0$, was arbitrary.

The closed-loop control signals, $\hat{u}_i = -K_{p,i}^{c*} x_i$, also allow for a certain decomposition of the LQR cost, which will prove to be extremely useful. Indeed,[5]

$$J_N^c(x_0, u_0, \ldots, u_N) = x_0^* P_0^c x_0 + \sum_{j=0}^{N} (u_j - \hat{u}_j)^* R_{e,j}^c (u_j - \hat{u}_j). \tag{6.7.24}$$

We can now summarize the results obtained so far in the following Theorem, where, to conform with the results of Sec. 1.5.3, we have defined, $K_{c,i} \triangleq K_{p,i}^{c*}$.

---

[5]This identity can be shown in several different ways. One way is to complete the squares in the expression for J_N^c, (6.7.6), to write

$$J_N^c = x_0^* P_0^c x_0 + (u - R_{y^c}^{-1} R_{y^c x_0^c} x_0)^* R_{y^c} (u - R_{y^c}^{-1} R_{y^c x_0^c} x_0).$$

The triangular decomposition, $R_{y^c} = L^* R_{e^c} L$, (obtained from the Krein space Kalman filter for R_{y^c}) can then (after some amount of algebra) be used to show the desired result. A shorter proof is to repeatedly apply the (readily verified via completion of squares) identity,

$$u_j^* Q_j^c u_j + s_j^* R_j^c s_j + x_{j+1}^* P_{j+1}^c x_{j+1} = x_j^* P_j^c x_j + (u_j - \hat{u}_j)^* R_{e,j}^c (u_j - \hat{u}_j),$$

to the quadratic form, J_N^c.

Theorem 6.7.1 (Finite Horizon LQR Solution) *Consider the LQR cost function*

$$J_N^c(x_0, u_0, \ldots, u_N) = x_{N+1}^* P_{N+1}^c x_{N+1} + \sum_{i=0}^{N} u_i^* Q_i^c u_i + \sum_{i=0}^{N} s_i^* R_i^c s_i,$$

with P_{N+1}^c, Q_i^c, and R_i^c, given Hermitian matrices, and suppose that we want to minimize $J_N^c(x_0, u_0, \ldots, u_N)$ over the variables $\{u_i\}_{i=0}^N$ subject to the state-space constraint $x_{i+1} = F_i x_i + G_i u_i$, $i = 0, \ldots, N$. Then the minimizing solution is given by the state-feedback law

$$\hat{u}_i = -K_{c,i} x_i, \quad i = 0, \ldots, N \tag{6.7.25}$$

where

$$K_{c,i} = F_i^* P_{i+1}^c G_i (R_{e,i}^c)^{-1} \tag{6.7.26}$$

and

$$R_{e,i}^c = Q_i^c + G_i^* P_{i+1}^c G_i \tag{6.7.27}$$

and where P_i^c satisfies the backwards-time Riccati recursion

$$P_i^c = F_i^* P_{i+1}^c F_i + H_i^* R_i^c H_i - K_{c,i} R_{e,i}^c K_{c,i}^*, \qquad P_{N+1}^c. \tag{6.7.28}$$

The condition for the above solution to be a minimum is that

$$R_{e,i}^c > 0, \quad i = 0, \ldots, N. \tag{6.7.29}$$

Moreover, we have the identity,

$$J_N^c(x_0, u_0, \ldots, u_N) = x_0^* P_0^c x_0 + \sum_{j=0}^{N} (u_j - \hat{u}_j)^* R_{e,j}^c (u_j - \hat{u}_j). \tag{6.7.30}$$

Remark: Note that we obtained the solution to the LQR problem by appealing to the dual approach for solving deterministic least-squares problems. The solution was obtained by constructing the Kalman filter for the backwards-time dual state-space model. This, therefore, directly suggests the wellknown duality between Kalman filtering and LQR control. To see this duality more explicitly, note that if we apply the transformations given below to the solution of Theorem 6.7.1,

$$\begin{array}{lll} F_i^* \to F_i & P_{i+1}^c \to P_i & K_{c.i} \to K_{p,i}^* \\ H_i^* \to G_i & R_i^c \to Q_i & \text{backward time} \to \text{forward time} \\ G_i^* \to H_i & Q_i^c \to R_i & \end{array}$$

then we recover the Kalman filter solution corresponding to the standard forward-time state-space model.

6.8 Full Information H^∞ Control

In this section we shall study the full information H^∞ control problem. We first introduced, and in fact solved, this problem in Sec. 1.7.1 using an approach based on the canonical factorization of a certain indefinite transfer operator. Here we shall take a slightly different approach, closer to the general spirit of Chapters 2 and 3, that is based upon associating an indefinite quadratic form, rather than an indefinite transfer operator, with the problem.[6] The resulting indefinite quadratic form is a special case of the LQR cost function studied in the previous section and Theorem 6.7.1 will be used to stationarize the quadratic form (thereby obtaining the solution) and to check the conditions for a minimum (thereby yielding the existence conditions for the controller).

For the early motivation of full information H^∞ control and alternative approaches to its solution the reader is referred to [Zam81, Fra87, FD87, BC87, Kim87, DGKF89, GD89, Tad90, Kwa91, GLD$^+$91, IPJ91, LAKG92, BB95, GL95, ZDG96] and the references therein.

[6]For comparison of the two methods, see the remarks at the beginning of Sec. 3.3.

6.8.1 Problem Formulation

Consider the (possibly) time-variant state-space model

$$\begin{aligned} x_{i+1} &= F_i x_i + G_{1,i} w_i + G_{2,i} w_i \\ &= F_i x_i + \left[\begin{array}{cc} G_{1,i} & G_{2,i} \end{array} \right] \left[\begin{array}{c} w_i \\ u_i \end{array} \right] \qquad i = 0, \ldots N, \end{aligned} \tag{6.8.1}$$

where the $\{u_i\}$ are the control inputs, and x_0 and the $\{w_i\}$ are deterministic disturbances. The $\{w_i\}$ may be interpreted as process noise or driving disturbance. Consider a given linear combination of the states,

$$s_i = L_i x_i, \tag{6.8.2}$$

that we intend to regulate using the control signal u_i. In the full information problem, it is assumed that that the control signal, u_i, has access to current and past values of the disturbances, $\{x_0, w_0, \ldots, w_i\}$. Let

$$\check{u}_i = \mathcal{F}_i(x_0, w_0, \ldots, w_i), \tag{6.8.3}$$

denote a such a full information control strategy.

Using this control strategy, let $\mathcal{T}(\mathcal{F})$ denote the transfer operator that maps the disturbances

$$\left\{ \Pi_0^{-\frac{1}{2}} x_0, \{(Q_i^w)^{\frac{1}{2}} w_i\}_{i=0}^N \right\},$$

to the variables

$$\left\{ (P_{N+1}^c)^{\frac{1}{2}} x_{N+1}, \{(Q_i^c)^{\frac{1}{2}} \check{u}_i\}_{i=0}^N, \{(R_i^c)^{\frac{1}{2}} s_i\}_{i=0}^N \right\},$$

where Π_0, P_{N+1}^c, Q_i^w, Q_i^c and R_i^c are positive (semi)definite weighting matrices. The finite horizon full information H^∞ control problem can now be stated as follows.

Problem 6.8.1 (Optimal Full Information H^∞ Control Problem) *Find a finite horizon full information H^∞-optimal control strategy $\check{u}_i = \mathcal{F}_i(x_0, w_0, \ldots, w_i)$ that minimizes $\mathcal{T}(\mathcal{F})$ and obtain the resulting*

$$\gamma_{opt}^2 = \inf_{\mathcal{F}} \| T(\mathcal{F}) \|_\infty^2 = \inf_{\mathcal{F}} \sup_{x_0, w \in h_2} \frac{x_{N+1}^* P_{N+1}^c x_{N+1} + \| (Q_i^c)^{\frac{1}{2}} \check{u}_i \|^2 + \| (R_i^c)^{\frac{1}{2}} s_i \|^2}{x_0 \Pi_0^{-1} x_0 + \| (Q_i^w)^{\frac{1}{2}} w_i \|^2}. \tag{6.8.4}$$

Note that the numerator in (6.8.4) is simply the LQR cost function. Thus the above problem formulation guarantees the smallest LQR cost over all possible disturbances of fixed energy. H^∞ controllers must therefore accommodate for all conceivable disturbances, which reflects in better robust behaviour with respect to disturbance variation.

A closed form solution to the optimal full information H^∞ control problem is available only for some special cases, and a simpler problem results if one relaxes the minimization condition and settles for a suboptimal solution.

Problem 6.8.2 (Suboptimal H^∞ Problem) *Find a suboptimal full information H^∞ control strategy $\check{u}_i = \mathcal{F}_i(x_0, w_0, \ldots, w_i)$ that achieves $\| \mathcal{T}(\mathcal{F}) \|_\infty < \gamma$. This clearly requires checking whether $\gamma \geq \gamma_{opt}$.*

6.8.2 Solution

Before giving the solution to the full information H^∞ control problem, it will be useful to study the structure of the problem in slightly more detail. In view of Prob. 6.8.2, we must find a control strategy $\check{u}_i = \mathcal{F}_i(x_0, w_0, \ldots w_i)$ such that

$$\sup_{x_0, w \in h_2} \frac{x_{N+1}^* P_{N+1}^c x_{N+1} + \| (Q_i^c)^{\frac{1}{2}} \check{u}_i \|^2 + \| (R_i^c)^{\frac{1}{2}} s_i \|^2}{x_0 \Pi_0^{-1} x_0 + \| (Q_i^w)^{\frac{1}{2}} w_i \|^2} < \gamma^2,$$

or, in other words, a control strategy, such that for all nonzero x_0 and $\{w_i\}_{i=0}^N$,

$$\frac{x_{N+1}^* P_{N+1}^c x_{N+1} + \sum_{i=0}^N \check{u}_i^* Q_i^c \check{u}_i + \sum_{i=0}^N s_i^* R_i^c z_i}{x_0 \Pi_0^{-1} x_0 + \sum_{i=0}^N w_i^* Q_i^w w_i} < \gamma^2. \tag{6.8.5}$$

Multiplying (6.8.5) by the nonzero denominator, we obtain the following result.

Lemma 6.8.1 (Full Information H^∞ Control and Quadratic Forms) *Given a scalar $\gamma > 0$, there exists a full information controller that achieves $\| \mathcal{T}(\mathcal{F}) \|_\infty < \gamma$ if, and only if, there exists $\check{u}_i = \mathcal{F}_i(x_0, w_0, \ldots w_i)$ (for all $i = 0, \ldots, N$) such that for all nonzero complex vectors x_0 and all nonzero sequences $\{w_i\}_{i=0}^N$, the scalar quadratic form*

$$J_N^c = x_0^* \Pi_0^{-1} x_0 - \gamma^{-2} (x_{N+1}^* P_{N+1}^c x_{N+1} + \sum_{i=0}^N \check{u}_i^* Q_i^c \check{u}_i - \gamma^2 \sum_{i=0}^N w_i^* Q_i^w w_i + \sum_{i=0}^N s_i^* R_i^c s_i)$$

$$= \quad x_0^*\Pi_0^{-1}x_0 - \gamma^{-2}(x_{N+1}^* P_{N+1}^c x_{N+1} + \sum_{i=0}^{N} \left[\begin{array}{cc} \check{u}_i^* & w_i^* \end{array}\right] \left[\begin{array}{cc} Q_i^c & 0 \\ 0 & -\gamma^2 Q_i^w \end{array}\right] \left[\begin{array}{c} \check{u}_i \\ w_i \end{array}\right] + \sum_{i=0}^{N} s_i^* R_i^c s_i)$$

satisfies

$$J_N^c > 0.$$

The quadratic form appearing in the parentheses of J_N^c is simply the LQR cost function associated with the state-space model (6.8.1), and with weighting matrices P_{N+1}, $\left[\begin{array}{cc} Q_i^c & 0 \\ 0 & -\gamma^2 Q_i^w \end{array}\right]$, and R_i^c, respectively. Note that one of the weighting matrices is indefinite, which is why we considered the Krein space setting in Sec. 6.7. We may thus use Eq. (6.7.30) of Theorem 6.7.1 to rewrite J_N^c as

$$J_N^c = x_0^*\Pi_0^{-1}x_0 - \gamma^{-2}\,(x_0^* P_0^c x_0 +$$

$$\sum_{i=0}^{N} \left[\begin{array}{c} \check{u}_i - \hat{u}_i \\ w_i - \hat{w}_i \end{array}\right]^* \left[\begin{array}{cc} Q_i^c + G_{2,i}^* P_{i+1}^c G_{2,i} & G_{2,i}^* P_{i+1}^c G_{1,i} \\ G_{1,i}^* P_{i+1}^c G_{2,i} & -\gamma^2 Q_i^w + G_{1,i}^* P_{i+1}^c G_{1,i} \end{array}\right] \left[\begin{array}{c} \check{u}_i - \hat{u}_i \\ w_i - \hat{w}_i \end{array}\right]\Bigg). \tag{6.8.6}$$

where $\hat{u}_i$ and $\hat{w}_i$ are given by

$$\left[\begin{array}{c} \hat{u}_i \\ \hat{w}_i \end{array}\right] = -K_{c,i} x_i, \tag{6.8.7}$$

with

$$K_{c,i} = (R_{e,i}^c)^{-1} \left[\begin{array}{c} G_{2,i}^* \\ G_{1,i}^* \end{array}\right] P_{i+1}^c F_i, \tag{6.8.8}$$

where

$$R_{e,i}^c = \left[\begin{array}{cc} Q_i^c + G_{2,i}^* P_{i+1}^c G_{2,i} & G_{2,i}^* P_{i+1}^c G_{1,i} \\ G_{1,i}^* P_{i+1}^c G_{2,i} & -\gamma^2 Q_i^w + G_{1,i}^* P_{i+1}^c G_{1,i} \end{array}\right], \tag{6.8.9}$$

and where P_i^c satisfies the backwards Riccati recursion

$$P_i^c = F_i^* P_{i+1}^c F_i + L_i^* R_i^c L_i - K_{c,i}^* R_{e,i}^c K_{c,i}, \quad P_{N+1}^c. \tag{6.8.10}$$

In order to solve the full information problem, let us introduce the following block LDU factorization of $R_{e,i}^c$,

$$\left[\begin{array}{cc} I_{m_2} & 0 \\ G_{1,i}^* P_{i+1}^c G_{2,i} R_{G^c,i}^{-1} & I_{m_1} \end{array}\right] \left[\begin{array}{cc} R_{G^c,i} & 0 \\ 0 & \Delta_i \end{array}\right] \left[\begin{array}{cc} I_{m_1} & R_{G^c,i}^{-1} G_{2,i}^* P_{i+1}^c G_{1,i} \\ 0 & I_{m_1} \end{array}\right], \tag{6.8.11}$$

where we have defined

$$R_{G^c,i} \triangleq Q_i^c + G_{2,i}^* P_{i+1}^c G_{2,i}, \tag{6.8.12}$$

and the Schur complement,

$$\Delta_i \triangleq -\gamma^2 Q_i^w + G_{1,i}^* P_{i+1}^c G_{1,i} - G_{1,i}^* P_{i+1}^c G_{2,i} R_{G^c,i}^{-1} G_{2,i}^* P_{i+1}^c G_{1,i}. \tag{6.8.13}$$

The factorization in (6.8.11) allows us to write J_N^c as

$$J_N^c = x_0^*(\Pi_0^{-1} - \gamma^{-2} P_0^c)x_0 - \gamma^{-2}\sum_{i=0}^{N}(\check{u}_i - \bar{u}_i)^* R_{G^c,i}(\check{u}_i - \bar{u}_i) - \gamma^{-2}\sum_{i=0}^{N}(w_i - \hat{w}_i)^*\Delta_i(w_i - \hat{w}_i), \tag{6.8.14}$$

where we have defined $\bar{u}_i$ via

$$\begin{bmatrix} \check{u}_i - \bar{u}_i \\ w_i - \hat{w}_i \end{bmatrix} = \begin{bmatrix} I_{m_2} & 0 \\ G_{1,i}^* P_{i+1}^c G_{2,i} R_{G^c,i}^{-1} & I_{m_1} \end{bmatrix} \begin{bmatrix} \check{u}_i - \hat{u}_i \\ w_i - \hat{w}_i \end{bmatrix}, \tag{6.8.15}$$

so that, after some algebraic simplification, we obtain

$$\bar{u}_i = -R_{G^c,i}^{-1} G_{2,i}^* P_{i+1}^c F_i x_i - R_{G^c,i}^{-1} G_{2,i}^* P_{i+1}^c G_{1,i} w_i. \tag{6.8.16}$$

Note that $\bar{u}_i$ is a function of $\{x_0, w_0, \ldots, w_i\}$, whereas $\hat{u}_i$ and $\hat{w}_i$ are functions of $\{x_0, w_0, \ldots, w_{i-1}\}$.

Referring back to Eq. (6.8.14) we see that since $\check{u}_i$ is only allowed to be a function of $\{x_0, w_0, \ldots, w_i\}$ it cannot influence the first and third terms in the summation for J_N^c.[7] Thus a necessary condition for J_N^c to be positive for all nonzero $\{x_0, w_0, \ldots w_N\}$ is that the first and third terms be positive for all such disturbances. In other words, if

$$\Pi_0^{-1} - \gamma^{-2} P_0^c > 0, \tag{6.8.17}$$

and

$$\Delta_i < 0, \quad i = 0, \ldots N. \tag{6.8.18}$$

The above conditions are also sufficient, since we can always choose the control signal to be,

$$\check{u}_i = \bar{u}_i, \tag{6.8.19}$$

[7] To see why, note that $x_0^*(\Pi_0^{-1} - \gamma^{-2} P_0^c)x_0$ is independent of $\check{u}_i$, and, moreover, that the controller cannot influence the value of $(w_i - \hat{w}_i)^*\Delta_i(w_i - \hat{w}_i)$ since $\hat{w}_i$ depends only on previous controls which have no knowledge of w_i.

which yields, incidentally, the central controller.

We have thus established the following result.

Theorem 6.8.1 (Finite-Horizon Full Information H^∞ Control) *Consider the state-space model (6.8.1) and define $\mathcal{T}(\mathcal{F})$ as the transfer operator that maps the disturbances $\{\Pi_0^{-\frac{1}{2}}x_0, \{(Q_i^w)^{\frac{1}{2}}w_i\}_{i=0}^N\}$ to the variables $\{(P_{N+1}^c)^{\frac{1}{2}}x_{N+1}, \{(Q_i^c)^{\frac{1}{2}}\check{u}_i, (R_i^c)^{\frac{1}{2}}s_i\}_{i=0}^N\}$, where Π_0, P_{N+1}^c, Q_i^w, Q_i^c and R_i^c are positive (semi)definite weighting matrices. Then for any given $\gamma > 0$, a full information H^∞ control strategy, $\check{u}_i = \mathcal{F}_i(x_0, w_0, \ldots, w_i)$, that yields*

$$\| \mathcal{T}(\mathcal{F}) \|_\infty < \gamma,$$

exists if, and only if,

(i) $\Pi_0^{-1} - \gamma^{-2}P_0^c > 0$

(ii) $\Delta_i = -\gamma^2 Q_i^w + G_{1,i}^* P_{i+1}^c G_{1,i} - G_{1,i}^* P_{i+1}^c G_{2,i} R_{G^c,i}^{-1} G_{2,i}^* P_{i+1}^c G_{1,i} < 0 \quad i = 0, \ldots, N$

where $R_{G^c,i} = Q_i^c + G_{2,i}^* P_{i+1}^c G_{2,i}$, and P_{i+1}^c satisfies the backwards Riccati recursion

$$P_i^c = F_i^* P_{i+1}^c F_i + L_i^* R_i^c L_i - Kc, i^* R_{e,i}^c K_{c,i}, \quad P_{N+1}^c$$

with

$$K_{c,i} = (R_{e,i}^c)^{-1} \begin{bmatrix} G_{2,i}^* \\ G_{1,i}^* \end{bmatrix} P_{i+1}^c F_i,$$

and

$$R_{e,i}^c = \begin{bmatrix} Q_i^c + G_{2,i}^* P_{i+1}^c G_{2,i} & G_{2,i}^* P_{i+1}^c G_{1,i} \\ G_{1,i}^* P_{i+1}^c G_{2,i} & -\gamma^2 Q_i^w + G_{1,i}^* P_{i+1}^c G_{1,i} \end{bmatrix}.$$

If this is the case, then all possible full information H^∞ control strategies, $\check{u}_i = \mathcal{F}_i(x_0, w_0, \ldots, w_i)$, are given by those that satisfy,

$$x_0^*(\Pi_0^{-1} - \gamma^{-2}P_0^c)x_0 - \gamma^{-2}\sum_{i=0}^N (\check{u}_i - \bar{u}_i)^* R_{G^c,i}(\check{u}_i - \bar{u}_i) - \gamma^{-2}\sum_{i=0}^N (w_i - \hat{w}_i)^* \Delta_i (w_i - \hat{w}_i) > 0,$$

where $\hat{w}_i$ is defined via,

$$\begin{bmatrix} \hat{u}_i \\ \hat{w}_i \end{bmatrix} = -K_{c,i} x_i,$$

and $\bar{u}_i$, *via*

$$\bar{u}_i = -R_{G^c,i}^{-1} G_{2,i}^* P_{i+1}^c F_i x_i - R_{G^c,i}^{-1} G_{2,i}^* P_{i+1}^c G_{1,i} w_i.$$

Finally, the so-called central controller is given by,

$$\check{u}_i = \bar{u}_i. \tag{6.8.20}$$

Remarks:

(i) Note that the above Theorem is simply the finite-horizon time-variant counterpart of Theorem 1.7.3. (See also the comments at the end of Sec. 1.7.3.)

(ii) The solution to the H^∞ control problem with perfect state measurement, as given by Theorem 6.8.1, is a special case of the LQR control problem studied in Sec. 6.7. Note that we needed the Krein space extension of the LQR theory (with indefinite weighting matrices) since the H^∞ control problem leads to such types of cost functions.

(iii) There are two conditions for the existence of H^∞ controllers: one is a certain positivity check for the solution of the backwards Riccati at each time instant, and one is a coupling condition between the end-point solution of the Riccati and the weighting on the initial state x_0. This last condition is reminiscent of the separation principle and coupling condition that occurs in measurement feedback H^∞ control, which we shall study in the next section.

(iv) The central full information controllers given by Eq. (6.8.20) have the same basic properties that the central filters of H^∞ estimation had. In other words, they are risk-sensitive optimal with risk-sensitivity parameter, $\theta = -\gamma^{-2}$, they are the solution to certain quadratic dynamic games, and they are maximum entropy controllers.

(v) We should mention that the central control law (6.8.1) is not exactly a state feedback law (as happens in continuous-time full information H^∞ control) since the second term involves w_i. The reason is, of course, that we have allowed the control signal to be a causal function of the disturbances, *i.e.*, a function of $\{x_0, w_0, \ldots w_i\}$. If we restrict the control signal to be a strictly causal function of the disturbances, *i.e.*, a function of $\{x_0, w_0, \ldots w_{i-1}\}$, then, as shown below, we do obtain a state feedback law for the central controller. (Compare with the remarks following Theorem 1.7.3.)

When the control signal is restricted to be a strictly causal function of the disturbances, *i.e.*, $\check{u}_i = \mathcal{F}_i(x_0, w_0, \ldots w_{i-1})$, then the only difference in the arguments preceding Theorem 6.8.1 is that we need to replace the block LDU factorization (6.8.11) of $R^c_{e,i}$ with its block upper-diagonal-lower (UDL) factorization,

$$\begin{bmatrix} I_{m_2} & G^*_{2,i}P^c_{i+1}G_{1,i}(R'_{G^c,i})^{-1} \\ 0 & I_{m_1} \end{bmatrix} \begin{bmatrix} \Delta'_i & 0 \\ 0 & R'_{G^c,i} \end{bmatrix} \begin{bmatrix} I_{m_1} & 0 \\ (R'_{G^c,i})^{-1}G^*_{1,i}P^c_{i+1}G_{2,i} & I_{m_1} \end{bmatrix}, \tag{6.8.21}$$

where we have defined

$$R'_{G^c,i} \triangleq -\gamma^2 Q^w_i + G^*_{1,i}P^c_{i+1}G_{1,i}, \tag{6.8.22}$$

and the Schur complement,

$$\Delta'_i \triangleq Q^c_i + G^*_{2,i}P^c_{i+1}G_{2,i} - G^*_{2,i}P^c_{i+1}G_{1,i}(R'_{G^c,i})^{-1}G^*_{1,i}P^c_{i+1}G_{2,i}. \tag{6.8.23}$$

The factorization in (6.8.21) allows us to write J^c_N as

$$J^c_N = x^*_0(\Pi_0^{-1} - \gamma^{-2}P^c_0)x_0 - \gamma^{-2}\sum_{i=0}^{N}(\check{u}_i - \hat{u}_i)^*\Delta'_i(\check{u}_i - \hat{u}_i) - \gamma^{-2}\sum_{i=0}^{N}(w_i - \bar{w}_i)^*R'_{G^c,i}(w_i - \bar{w}_i), \tag{6.8.24}$$

where $\bar{w}_i$ is defined via,

$$\begin{bmatrix} \check{u}_i - \hat{u}_i \\ w_i - \bar{w}_i \end{bmatrix} = \begin{bmatrix} I_{m_2} & G^*_{2,i}P^c_{i+1}G_{1,i}(R'_{G^c,i})^{-1} \\ 0 & I_{m_1} \end{bmatrix} \begin{bmatrix} \check{u}_i - \hat{u}_i \\ w_i - \hat{w}_i \end{bmatrix}. \tag{6.8.25}$$

Note that since u_i is only a function of $(x_0, w_0, \ldots w_{i-1})$, it can only effectively influence the second term in the summation for J^c_N in Eq. (6.8.24). Therefore proceeding with an argument similar to the one that led to the proof of Theorem 6.8.1, we obtain the following result.

Theorem 6.8.2 (Strictly Causal Full Information H^∞ Control) *Consider the state-space model (6.8.1) and define $\mathcal{T}(\mathcal{F})$ as the transfer operator that maps the disturbances $\{\Pi_0^{-\frac{1}{2}}x_0, \{(Q_i^w)^{\frac{1}{2}}w_i\}_{i=0}^N\}$ to the variables $\{(P_{N+1}^c)^{\frac{1}{2}}x_{N+1}, \{(Q_i^c)^{\frac{1}{2}}\check{u}_i, (R_i^c)^{\frac{1}{2}}s_i\}_{i=0}^N\}$, where Π_0, P_{N+1}^c, Q_i^w, Q_i^c and R_i^c are positive (semi)definite weighting matrices. Then for any given $\gamma > 0$, a strictly causal full information H^∞ control strategy, $\check{u}_i = \mathcal{F}_i(x_0, w_0, \ldots, w_{i-1})$, that yields*

$$\| \mathcal{T}(\mathcal{F}) \|_\infty < \gamma,$$

exists if, and only if,

(i) $\Pi_0^{-1} - \gamma^{-2}P_0^c > 0$

(ii) $R'_{G^c,i} = -\gamma^2 Q_i^w + G_{1,i}^* P_{i+1}^c G_{1,i} < 0 \quad i = 0, \ldots, N$

where P_{i+1}^c is the same as in Theorem 6.8.1. If this is the case, then all possible strictly causal full information H^∞ control strategies, $\check{u}_i = \mathcal{F}_i(x_0, w_0, \ldots, w_{i-1})$, are given by those that satisfy,

$$x_0^*(\Pi_0^{-1} - \gamma^{-2}P_0^c)x_0 - \gamma^{-2}\sum_{i=0}^N(\check{u}_i - \hat{u}_i)^*\Delta_i'(\check{u}_i - \hat{u}_i) - \gamma^{-2}\sum_{i=0}^N(w_i - \bar{w}_i)^* R'_{G^c,i}(w_i - \bar{w}_i) > 0,$$

where $\Delta_i' = Q_i^c + G_{2,i}^* P_{i+1}^c G_{2,i} - G_{2,i}^* P_{i+1}^c G_{1,i}(R'_{G^c,i})^{-1} G_{1,i}^* P_{i+1}^c G_{2,i}$, and where

$$\bar{w}_i = -(R'_{G^c,i})^{-1} G_{1,i}^* P_{i+1}^c F_i x_i - (R'_{G^c,i})^{-1} G_{1,i}^* P_{i+1}^c G_{2,i}\check{u}_i, \tag{6.8.26}$$

and

$$\hat{u}_i = -(Q_i^c + G_{2,i}^* \tilde{P}_{i+1}^c G_{2,i})^{-1} G_{2,i}^* \tilde{P}_{i+1}^c F_i x_i, \tag{6.8.27}$$

with

$$\tilde{P}_{i+1}^c = P_{i+1}^c - P_{i+1}^c G_{1,i}(R'_{G^c,i})^{-1} G_{1,i}^* P_{i+1}^c. \tag{6.8.28}$$

Finally, the so-called central controller is given by,

$$\check{u}_i = \hat{u}_i. \tag{6.8.29}$$

6.9 Measurement Feedback H^∞ Control

In this section we shall study the finite horizon H^∞ control problem with measurement feedback. We first introduced, and in fact solved, this problem in Sec. 1.7.2. The approach used there was based on a certain separation principle that reduced the measurement feedback problem to two coupled problems: one a full information control problem, and one an estimation problem (essentially for estimating the unobservable full information control signals using the available measurements). The approach taken here is the same, and is a direct application of our earlier results on H^∞ filtering (from Chapter 3) and full information H^∞ control (from Sec. 6.8), coupled with the aforementioned separation principle. We should mention that the separation principle presented here is slightly different from the separation principles first presented in H^∞ control [DGKF89] and risk-sensitive control [Whi90], and is closer to the separation principle given in [GL95]. We should also mention, however, as pointed out earlier in Remark (i) following Theorem 1.7.4, that the two separation principles are closely related, although we believe that the one presented here is more natural since it does not need to appeal to state-space models.

6.9.1 Problem Formulation

Consider the (possibly) time-variant state-space model

$$\left\{\begin{array}{rcl} x_{i+1} & = & F_i x_i + G_{1,i} w_i + G_{2,i} u_i \\ y_i & = & H_i x_i + v_i \\ s_i & = & L_i x_i \end{array}\right. \qquad i = 0, \ldots N, \tag{6.9.1}$$

where the $\{u_i\}$ are the control inputs, the $\{x_0, \{w_i\}, \{v_i\}\}$ are unknown disturbances, the $\{y_i\}$ are the observed outputs, and the $\{s_i\}$ are the signals we intend to regulate. The $\{w_i\}$ may be interpreted as process noise and the $\{v_i\}$ as measurement noise.

Moreover let $\check{u}_i$ be a control strategy that uses current and past observations, *i.e.*

$$\check{u}_i = \mathcal{F}_i(y_0, \ldots, y_i).$$

With this control strategy, let $\mathcal{T}(\mathcal{F})$ denote the transfer operator that maps the unknown disturbances

$$\left\{\Pi_0^{-\frac{*}{2}}x_0, \{(Q_i^w)^{\frac{1}{2}}w_i\}_{i=0}^N, \{(R_i^v)^{-\frac{*}{2}}v_i\}_{i=0}^N\right\},$$

to the variables

$$\left\{(P_{N+1}^c)^{\frac{1}{2}}x_{N+1}, \{(R_i^c)^{\frac{1}{2}}s_i\}_{i=0}^N, \{(Q_i^c)^{\frac{1}{2}}\check{u}_i\}_{i=0}^N\right\},$$

where Π_0, P_{N+1}^c, Q_i^w, Q_i^c R_i^v and R_i^c are positive (semi)definite weighting matrices. The H^∞ control problem with output feedback can now be stated as follows. (See figure 6.2.)

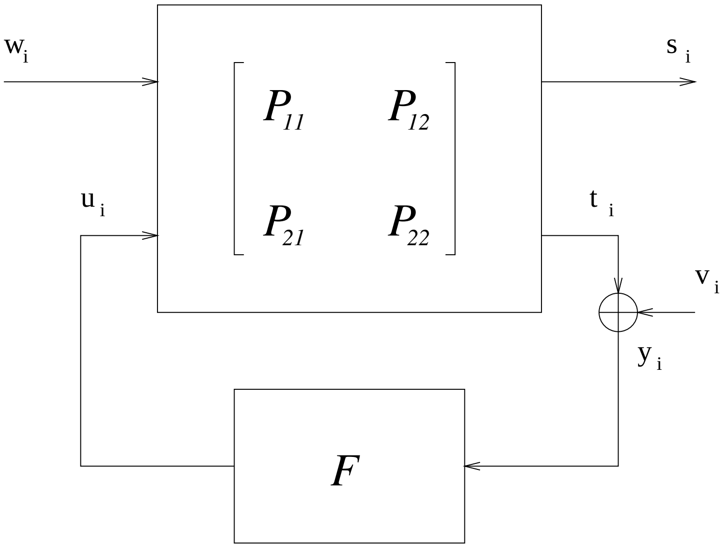

Figure 6.2: H^∞ control with measurement feedback.

Problem 6.9.1 (Optimal Measurement Feedback H^∞ Control Problem) *Find a finite horizon H^∞-optimal measurement feedback control strategy $\check{u}_i = \mathcal{F}_i(y_0, \ldots, y_i)$ that minimizes $\mathcal{T}(\mathcal{F})$ and obtain the resulting*

$$\gamma_{opt}^2 = \inf_{\mathcal{F}} \| \mathcal{T}(\mathcal{F}) \|_\infty^2 = \inf_{\mathcal{F}} \sup_{x_0, w\in h_2, v\in h_2} \frac{x_{N+1}^* P_{N+1}^c x_{N+1} + \| (Q_i^c)^{\frac{1}{2}}\check{u}_i \|^2 + \| (R_i^c)^{\frac{1}{2}} L_i x_i \|^2}{x_0 \Pi_0^{-1} x_0 + \| (Q_i^w)^{\frac{1}{2}} w_i \|^2 + \| (R_i^v)^{-\frac{*}{2}} v_i \|^2}. \tag{6.9.2}$$

Note that the numerator in (6.9.2) is simply the LQR cost function. Thus, as before, the above problem formulation guarantees the smallest LQR cost over all possible disturbances of fixed energy. The resulting H^∞ controllers are thus conservative, which reflects in a better robust behaviour with respect to disturbance variation.

As we have seen earlier, closed form solutions to the optimal H^∞ problem are available for only some special cases. We therefore obtain a simpler problem if we relax the minimization condition and settle for a suboptimal solution.

Problem 6.9.2 (Suboptimal H^∞ Problem) *Find a suboptimal measurement feedback H^∞ control strategy $\check{u}_i = \mathcal{F}_i(y_0, \ldots y_i)$ that achieves $\| \mathcal{T}(\mathcal{F}) \|_\infty < \gamma$. This clearly requires checking whether $\gamma \geq \gamma_{opt}$.*

6.9.2 Solution

We begin by examining the structure of the problem in slightly more detail. In view of Prob. 6.9.2, we must find a control strategy $\check{u}_i = \mathcal{F}_i(y_0, \ldots, y_i)$ such that

$$\sup_{x_0, w \in h_2, v \in h_2} \frac{x^*_{N+1} P^c_{N+1} x_{N+1} + \| (Q^c_i)^{\frac{1}{2}} \check{u}_i \|^2 + \| (R^c_i)^{\frac{1}{2}} L_i x_i \|^2}{x_0 \Pi_0^{-1} x_0 + \| (Q^w_i)^{\frac{1}{2}} w_i \|^2 + \| (R^v_i)^{-\frac{*}{2}} v_i \|^2} < \gamma^2,$$

or, in other words, a control strategy, such that for all nonzero x_0, $\{v_i\}_{i=0}^N$ and $\{w_i\}_{i=0}^N$,

$$\frac{x^*_{N+1} P^c_{N+1} x_{N+1} + \sum_{i=0}^N \check{u}_i^* Q^c_i \check{u}_i + \sum_{i=0}^N x_i^* L_i * R^c_i L_i x_i}{x_0 \Pi_0^{-1} x_0 + \sum_{i=0}^N w_i^* Q^w_i w_i + \sum_{i=0}^N v_i^* (R^v_i)^{-1} v_i} < \gamma^2. \tag{6.9.3}$$

Multiplying (6.9.3) by the nonzero denominator, we obtain the following result.

Lemma 6.9.1 (Measurement Feedback H^∞ Control and Quadratic Forms) *Given a scalar $\gamma > 0$, there exists a measurement feedback controller that achieves $\| \mathcal{T}(\mathcal{F}) \|_\infty < \gamma$ if, and only if, there exists $\check{u}_i = \mathcal{F}_i(y_0, \ldots, y_i)$ (for all $i = 0, \ldots, N$) such that for all nonzero complex vectors x_0, and all nonzero sequences $\{v_i\}_{i=0}^N$ and $\{w_i\}_{i=0}^N$, the scalar quadratic form*

$$\begin{aligned} J_N^m \quad = \quad & x_0^* \Pi_0^{-1} x_0 + \sum_{i=0}^N w_i^* Q^w_i w_i + \sum_{i=0}^N (y_i - H_i x_i)^* (R^v_i)^{-1} (y_i - H_i x_i) - \\ & \gamma^{-2} (\sum_{i=0}^N x_i^* L_i^* R^c_i L_i x_i + \sum_{i=0}^N \check{u}_i^* Q^c_i \check{u}_i + x^*_{N+1} P^c_{N+1} x_{N+1}), \end{aligned}$$

satisfies

$$J_N^m > 0.$$

In other words if, and only if,

(i) $J_N^m(x_0, w_0, \ldots, w_N, \check{u}_0, \ldots, \check{u}_N)$ *has a minimum over* $\{x_0, w_0, \ldots, w_N\}$.

(ii) The $\{\check{u}_0, \ldots, \check{u}_N\}$ *can be chosen such that the value of* J_N^m *at this minimum be positive.*

The above Lemma requires us to minimize J_N^m over the variables, $\{x_0, w_0, \ldots, w_N\}$ and to compute its value at this minimum. Since the expression for J_N^m is quite complicated, this appears to be a formidable task. To somewhat alleviate the problem, let us reorganize J_N^m in the following fashion,

$$\begin{aligned} J_N^m &= x_0^*\Pi_0^{-1}x_0 + \sum_{i=0}^{N}(y_i - H_i x_i)^*(R_i^v)^{-1}(y_i - H_i x_i) \\ &\quad -\gamma^{-2}\left[\sum_{i=0}^{N} x_i^* L_i^* R_i^c L_i x_i + \sum_{i=0}^{N}\begin{bmatrix} w_i \\ \check{u}_i \end{bmatrix}^* \begin{bmatrix} -\gamma^2 Q_i^w & 0 \\ 0 & Q_i^c \end{bmatrix}\begin{bmatrix} w_i \\ \check{u}_i \end{bmatrix} + x_{N+1}^* P_{N+1}^c x_{N+1}\right]. \end{aligned}$$

Note that the expression inside the brackets is simply the indefinite quadratic form associated with the full information problem. (See Lemma 6.8.1). Therefore we can use Eq. (6.8.6) to write it as

$$x_0^* P_0^c x_0 + \sum_{i=0}^{N}\begin{bmatrix} w_i - \hat{w}_i \\ \check{u}_i - \hat{u}_i \end{bmatrix}^* (R_{e,i}^c)^{-1}\begin{bmatrix} w_i - \hat{w}_i \\ \check{u}_i - \hat{u}_i \end{bmatrix}, \qquad (6.9.4)$$

where $\hat{w}_i$ and $\hat{u}_i$ are given by

$$\begin{bmatrix} \hat{w}_i \\ \hat{u}_i \end{bmatrix} = -K_{c,i}x_i \triangleq -\begin{bmatrix} K_{w,i} \\ K_{u,i} \end{bmatrix} x_i, \qquad (6.9.5)$$

with

$$K_{c,i} = (R_{e,i}^c)^{-1}\begin{bmatrix} G_{1,i}^* \\ G_{2,i}^* \end{bmatrix} P_{i+1}^c F_i, \qquad (6.9.6)$$

where

$$R^c_{e,i} = \begin{bmatrix} -\gamma^2 Q^w_i + G^*_{1,i}P^c_{i+1}G_{1,i} & G^*_{1,i}P^c_{i+1}G_{2,i} \\ G^*_{2,i}P^c_{i+1}G_{1,i} & Q^c_i + G^*_{2,i}P^c_{i+1}G_{2,i} \end{bmatrix}, \tag{6.9.7}$$

and where P^c_i satisfies the backwards Riccati recursion

$$P^c_i = F^*_i P^c_{i+1}F_i + L^*_i R^c_i L_i - K^*_{c,i}R^c_{e,i}K_{c,i}, \quad P^c_{N+1}. \tag{6.9.8}$$

Now Eq. (6.9.4) allows us to write J^m_N as follows,

$$\begin{aligned} J^m_N &= x^*_0(\Pi_0^{-1} - \gamma^{-2}P^c_0)x_0 + \\ &\sum_{i=0}^{N} \begin{bmatrix} w_i - \hat{w}_i \\ \check{u}_i - \hat{u}_i \\ y_i - H_i x_i \end{bmatrix}^* \begin{bmatrix} -\gamma^{-2}\begin{bmatrix} R^c_{e,i}(1,1) & R^c_{e,i}(1,2) \\ R^c_{e,i}(2,1) & R^c_{e,i}(2,2) \end{bmatrix} & 0 \\ 0 & (R^v_i)^{-1} \end{bmatrix} \begin{bmatrix} w_i - \hat{w}_i \\ \check{u}_i - \hat{u}_i \\ y_i - H_i x_i \end{bmatrix}, \end{aligned}$$

where the $R^c_{e,i}(i,j)$ denote the (obvious) block entries of $R^c_{e,i}$. Now if we define,

$$\begin{bmatrix} \Delta_i^{-1} & \bar{S}_i \\ \bar{S}^*_i & (\Delta'_i)^{-1} \end{bmatrix} \triangleq \begin{bmatrix} R^c_{e,i}(1,1) & R^c_{e,i}(1,2) \\ R^c_{e,i}(2,1) & R^c_{e,i}(2,2) \end{bmatrix}, \tag{6.9.9}$$

then J^m_N can be written as,

$$J^m_N = x^*_0(\Pi_0^{-1} - \gamma^{-2}P^c_0)x_0 + \sum_{i=0}^{N} \begin{bmatrix} w_i - \hat{w}_i \\ \begin{bmatrix} \check{u}_i \\ y_i \end{bmatrix} - \begin{bmatrix} K_{u,i} \\ H_i \end{bmatrix} \end{bmatrix}^* \begin{bmatrix} Q_i & S_i \\ S^*_i & R_i \end{bmatrix}^{-1} \begin{bmatrix} w_i - \hat{w}_i \\ \begin{bmatrix} \check{u}_i \\ y_i \end{bmatrix} - \begin{bmatrix} K_{u,i} \\ H_i \end{bmatrix} \end{bmatrix}, \tag{6.9.10}$$

where we have defined,

$$Q_i \triangleq -\gamma^2\Delta_i^{-1} \quad , \quad S_i \triangleq -\gamma^2 \begin{bmatrix} \bar{S}_i & 0 \end{bmatrix} \quad , \quad R_i \triangleq \begin{bmatrix} \gamma^2(\Delta'_i)^{-1} & 0 \\ 0 & R^v_i \end{bmatrix}. \tag{6.9.11}$$

But note that J^m_N, as given by (6.9.10), is now the standard estimation quadratic form of Sec. 2.6 (see Eq. (2.6.2)) for the Krein space state-space model,[8]

$$\left\{ \begin{aligned} \mathbf{x}_{i+1} &= (F_i - G_{1,i}K_{w,i})\mathbf{x}_i + G_{1,i}(\mathbf{w}_i - \hat{\mathbf{w}}_i) + G_{2,1}\check{u}_i \\ \begin{bmatrix} \check{\mathbf{u}}_i \\ \mathbf{y}_i \end{bmatrix} &= \begin{bmatrix} -K_{u,i} \\ H_i \end{bmatrix} \mathbf{x}_i + \mathbf{v}_i \end{aligned} \right. , \tag{6.9.12}$$

[8]Note that we have used $\hat{\mathbf{w}}_i = -K_{w,i}\mathbf{x}_i$ to rewrite the state equation as

$$\mathbf{x}_{i+1} = F_i\mathbf{x}_i + G_{1,i}\mathbf{w}_i + G_{2,1}\check{u}_i = (F_i - G_{1,i}K_{w,i})\mathbf{x}_i + G_{1,i}(\mathbf{w}_i - \hat{\mathbf{w}}_i) + G_{2,1}\check{u}_i.$$

with

$$\left\langle \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{w}_i - \hat{\mathbf{w}}_i \\ \mathbf{v}_i \end{bmatrix}, \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{w}_j - \hat{\mathbf{w}}_j \\ \mathbf{v}_j \end{bmatrix} \right\rangle = \begin{bmatrix} (\Pi_0^{-1} - \gamma^{-2} P_0^c)^{-1} & 0 \\ 0 & \begin{bmatrix} Q_i & S_i \\ S_i^* & R_i \end{bmatrix} \delta_{ij} \end{bmatrix}. \tag{6.9.13}$$

We can now use the results of Theorem 2.7.4 and Lemma 2.7.3 to minimize J_N^m over the variables, $\{x_0, w_0, \ldots w_N\}$ and to check the conditions for a minimum. Note that the condition for a minimum is that,

(i) $\Pi_0^{-1} - \gamma^{-2} P_0^c > 0$

(ii) $Q_i = -\gamma^2 \Delta_i^{-1} > 0$, for all $j = 0, \ldots, N$

(iii) $R_i - S^* Q_i^{-1} S$ and $R_{e,i}$ have the same inertia, for all $j = 0, \ldots, N$

where

$$R_{e,i} = R_i + \begin{bmatrix} -K_{u,i} \\ H_i \end{bmatrix} P_i \begin{bmatrix} -K_{u,i}^* & H_i^* \end{bmatrix}, \tag{6.9.14}$$

and P_i satisfies the Riccati recursion,

$$P_{i+1} = (F_i - G_{1,i} K_{w,i}) P_i (F_i - G_{1,i} K_{w,i})^* + -\gamma^2 G_{1,i} \Delta_i^{-1} G_{1,i}^* - K_{p,i} R_{e,i} K_{p,i}^*, \tag{6.9.15}$$

initialized with $P_0 = (\Pi_0^{-1} - \gamma^{-2} P_0^c)^{-1}$, and with

$$K_{p,i} = \left\{ (F_i - G_{1,i} K_{w,i}) P_i \begin{bmatrix} -K_{u,i}^* & H_i^* \end{bmatrix} + G_{1,i} S_i \right\} R_{e,i}^{-1}. \tag{6.9.16}$$

Now some simple algebra shows that

$$R_i - S^* Q_i^{-1} S = \begin{bmatrix} -\gamma^2 \left[Q_i^c + G_{2,i}^* P_{i+1}^c G_{2,i} \right]^{-1} & 0 \\ 0 & R_i^v \end{bmatrix}, \tag{6.9.17}$$

so that using the facts that $R_i^v > 0$ and $Q_i^c + G_{2,i}^* P_{i+1}^c G_{2,i} > 0$,[9] we readily see that condition (iii) can be replaced by the condition that

(iii) $(-I_{m_2}) \oplus I_p$ and $R_{e,i}$ have the same inertia, for all $j = 0, \ldots, N$.

[9] This follows from the full information control problem.

Now from Theorem 2.7.4 the value of J_N^m at its minimum is given by

$$J_N^m(\min) = \sum_{i=0}^{N} \begin{bmatrix} \check{u}_i + K_{u,i}\hat{x}_i \\ y_i - H_i\hat{x}_i \end{bmatrix}^* R_{e,i}^{-1} \begin{bmatrix} \check{u}_i + K_{u,i}\hat{x}_i \\ y_i - H_i\hat{x}_i \end{bmatrix}, \tag{6.9.18}$$

where $\hat{x}_i$ obeys the Krein space Kalman filter recursions,

$$\hat{x}_{i+1} = (F_i - G_{1,i}K_{w,i})\hat{x}_i + K_{p,i} \begin{bmatrix} \check{u}_i + K_{u,i}\hat{x}_i \\ y_i - H_i\hat{x}_i \end{bmatrix} + G_{2,i}\check{u}_i, \quad \hat{x}_0 = 0. \tag{6.9.19}$$

Now any choice of control strategy, $\check{u}_i = \mathcal{F}_i(y_0, \ldots, y_i)$ that renders $J(\min)$ positive is an acceptable control. To see that, under the aforementioned conditions, such a control strategy always exists since, via a UDL factorization of $R_{e,i}$ or a completion of squares argument, we can write $J(\min)$ as

$$J_N^m(\min) = \sum_{i=0}^{N}(\check{u}_i - \bar{u}_i)^*\Delta_{R,i}^{-1}(\check{u}_i - \bar{u}_i) + \sum_{i=0}^{N}(y_i - H_i\hat{x}_i)^*(R_i^v + H_iP_iH_i^*)^{-1}(y_i - H_i\hat{x}_i), \tag{6.9.20}$$

where

$$\bar{u}_i = -K_{u,i}\hat{x}_i - K_{u,i}P_iH_i^*(R_i^v + H_iP_iH_i^*)^{-1}(y_i - H_i\hat{x}_i), \tag{6.9.21}$$

and $\Delta_{R,i}$ is the Schur complement of the $(2,2)$ entry $R_i^v + H_iP_iH_i^*$ in $R_{e,i}$. Now in view of the inertia condition (iii) on the $R_{e,i}$ (and the positivity of P_i and R_i^v) we have

$$R_i^v + H_iP_iH_i^* > 0 \quad \text{and} \quad \Delta_{R,i} < 0. \tag{6.9.22}$$

Thus the choice $\check{u}_i = \bar{u}_i$ always renders J_N^m positive. This control strategy is referred to as the central controller. There are, of course, many other acceptable strategies.

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

Theorem 6.9.1 (Finite Horizon Measurement Feedback H^∞ Control) *Consider the state-space model (6.9.1) and define $\mathcal{T}(\mathcal{F})$ as the transfer operator that maps the unknown disturbances $\{\Pi_0^{-\frac{*}{2}}x_0, \{(Q_i^w)^{\frac{1}{2}}w_i, (R_i^v)^{-\frac{*}{2}}v_i\}_{i=0}^N\}$ to the variables*

$$\left\{(P_{N+1}^c)^{\frac{1}{2}}x_{N+1}, \{(R_i^c)^{\frac{1}{2}}s_i, (Q_i^c)^{\frac{1}{2}}z_i\}_{i=0}^N\right\},$$

where Π_0, P^c_{N+1}, Q^w_i, Q^c_i R^v_i *and* R^c_i *are positive (semi)definite weighting matrices. Then for any given* $\gamma > 0$, *a measurement feedback* H^∞ *control strategy* $\check{u}_i = \mathcal{F}_i(y_0, \ldots, y_i)$ *that yields*

$$\| \mathcal{T}(\mathcal{F}) \|_\infty < \gamma,$$

exists if, and only if,

(i) $\Pi_0^{-1} - \gamma^{-2} P^c_0 > 0$

(ii) $\Delta_i = -\gamma^2 Q^w_i + G^*_{1,i} P^c_{i+1} G_{1,i} - G^*_{1,i} P^c_{i+1} G_{2,i} R^{-1}_{G^c,i} G^*_{2,i} P^c_{i+1} G_{1,i} < 0 \quad i = 0, \ldots, N$

(iii) the matrices

$$\begin{bmatrix} -I_{m_2} & 0 \\ 0 & I_p \end{bmatrix} \quad \text{and} \quad R_{e,i} = R_i + \begin{bmatrix} -K_{u,i} \\ H_i \end{bmatrix} P_i \begin{bmatrix} -K^*_{u,i} & H^*_i \end{bmatrix}$$

have the same inertia, for all $j = 0, \ldots, N$

where P^c_{i+1} *and* P_i *satisfy the Riccati recursions*

$$\begin{cases} P^c_i &= F^*_i P^c_{i+1} F_i + L^*_i R^c_i L_i - K^*_{c,i} R^c_{e,i} K_{c,i}, \\ P_{i+1} &= (F_i - G_{1,i} K_{w,i}) P_i (F_i - G_{1,i} K_{w,i})^* + -\gamma^2 G_{1,i} \Delta_i^{-1} G^*_{1,i} - K_{p,i} R_{e,i} K^*_{p,i}, \end{cases}$$

initialized with P^c_{N+1} *and* $P_0 = (\Pi_0^{-1} - \gamma^{-2} P^c_0)^{-1}$, *and where*

$$\begin{cases} R_{G^c,i} &= Q^c_i + G^*_{2,i} P^c_{i+1} G_{2,i} \\ K_{c,i} &= (R^c_{e,i})^{-1} \begin{bmatrix} G^*_{1,i} \\ G^*_{2,i} \end{bmatrix} P^c_{i+1} F_i \\ R^c_{e,i} &= \begin{bmatrix} Q^c_i + G^*_{1,i} P^c_{i+1} G_{1,i} & G^*_{1,i} P^c_{i+1} G_{2,i} \\ G^*_{2,i} P^c_{i+1} G_{1,i} & -\gamma^2 Q^w_i + G^*_{2,i} P^c_{i+1} G_{2,i} \end{bmatrix} \\ \begin{bmatrix} K_{w,i} \\ K_{u,i} \end{bmatrix} &= K_{c,i} \\ K_{p,i} &= \left[(F_i - G_{1,i} K_{w,i}) P_i \begin{bmatrix} -K^*_{u,i} & H^*_i \end{bmatrix} + G_{1,i} S_i \right] R^{-1}_{e,i} \\ R_i &= \begin{bmatrix} (\Delta'_i)^{-1} & 0 \\ 0 & R^v_i \end{bmatrix} \\ \Delta'_i &= Q^c_i + G^*_{2,i} P^c_{i+1} G_{2,i} - G^*_{2,i} P^c_{i+1} G_{1,i} (R'_{G^c,i})^{-1} G^*_{1,i} P^c_{i+1} G_{2,i} \\ R'_{G^c,i} &= -\gamma^2 Q^w_i + G^*_{1,i} P^c_{i+1} G_{1,i} \end{cases}$$

If this is the case, then all possible measurement feedback H^∞ control strategies, $\check{u}_i = \mathcal{F}_i(y_0, \ldots, y_i)$, are given by those that satisfy

$$\sum_{i=0}^{N}(\check{u}_i - \bar{u}_i)^* \Delta_{R,i}^{-1}(\check{u}_i - \bar{u}_i) + \sum_{i=0}^{N}(y_i - H_i\hat{x}_i)^*(R_i^v + H_iP_iH_i^*)^{-1}(y_i - H_i\hat{x}_i) > 0, \quad (6.9.23)$$

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

$$\hat{x}_{i+1} = (F_i - G_{1,i}K_{w,i})\hat{x}_i + K_{p,i}\begin{bmatrix} \check{u}_i + K_{u,i}\hat{x}_i \\ y_i - H_i\hat{x}_i \end{bmatrix} + G_{2,i}\check{u}_i, \quad \hat{x}_0 = 0$$

where

$$\Delta_{R,i} = -\gamma^2(\Delta_i^{'})^{-1} + K_{u,i}P_iK_{u,i}^* - K_{u,i}P_iH_i^*(R_i^v + H_iP_iH_i^*)^{-1}H_iK_{u,i}^*,$$

and where

$$\bar{u}_i = -K_{u,i}\hat{x}_i - K_{u,i}P_iH_i^*(R_i^v + H_iP_iH_i^*)^{-1}(y_i - H_i\hat{x}_i).$$

The so-called central controller is given by

$$\check{u}_i = \bar{u}_i \stackrel{\Delta}{=} -K_{u,i}\hat{x}_{i|i}, \quad (6.9.24)$$

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

$$\hat{x}_{i+1|i+1} = (F_i - G_{1,i}K_{w,i} - K_{s,i+1}H_{i+1}F_i)\hat{x}_{i|i} + K_{s,i+1}y_{i+1} + G_{2,i}\check{u}_i, \quad \hat{x}_{-1|-1} = 0,$$

with $K_{s,i} = P_iH_i^*(R_i^v + H_iP_iH_i^*)^{-1}$.

Remarks:

(i) Note that the above Theorem is simply the finite-horizon time-variant counterpart of Theorem 1.7.4. (See also the comments at the end of Sec. 1.7.3.)

(ii) There are three conditions for the existence of H^∞ controllers: the first is a certain positivity check for the solution of the backwards Riccati recursion at each time instant, the second is a certain inertia check for the solution of the forwards Riccati recursion at each time instant, and the third is a coupling condition between the end-point solution of the backwards Riccati recursion and the weighting on the initial state x_0.

(iii) Note that, the Riccati recursions for P_i^c and P_i are coupled. Indeed, the recursion for P_i depends on the solution of the recursion for P_i^c (but not vice versa). We should mention that it is possible, through a suitable change of variables (essentially a bilinear transformation involving P_i and P_i^c) to come up with an auxiliary variable, say P_i^d, that satisfies a recursion which is independent of P_i^c. However, the price we have to pay is a certain coupling condition on the spectral radius of $P_i^c P_i^d$ at each time instant. In fact, the first solutions to the measurement feedback H^∞ and risk-sensitive control problems (see [DGKF89] and [Whi90]) presented such Riccati equations. However, we shall not gain much by presenting that algebra here.

(iv) The central full information controllers given by Eq. (6.9.24) have the same basic properties that the central filters of H^∞ estimation and full information control had. They are thus risk-sensitive optimal with risk-sensitivity parameter, $\theta = -\gamma^{-2}$, they are the solution to certain quadratic dynamic games, and they are maximum entropy controllers.

We can also consider the measurement feedback problem with strictly causal controllers,

$$\check{u}_i = \mathcal{F}_i(y_0, \ldots y_{i-1}). \tag{6.9.25}$$

The development is exactly the same as that given prior to the statement of Theorem 6.9.1 with the exception that in the expression for $J_N^m(\min)$,

$$J_N^m(\min) = \sum_{i=0}^{N} \begin{bmatrix} \check{u}_i + K_{u,i}\hat{x}_i \\ y_i - H_i\hat{x}_i \end{bmatrix}^* R_{e,i}^{-1} \begin{bmatrix} \check{u}_i + K_{u,i}\hat{x}_i \\ y_i - H_i\hat{x}_i \end{bmatrix}, \tag{6.9.26}$$

instead of performing the UDL factorization of $R_{e,i}$, we need to perform its LDU factorization. This allows us to write,

$$J_N^m(\min) = \sum_{i=0}^{N}(\check{u}_i - \hat{u}_i)^* \left(R_{e,i}(1,1)\right)^{-1} (\check{u}_i - \hat{u}_i) + \sum_{i=0}^{N}(y_i - H_i\bar{x}_i)^*(\Delta_i^{'})^{-1}(y_i - H_i\bar{x}_i), \tag{6.9.27}$$

where

$$R_{e,i}(1,1) = -\gamma^2 \left[Q_i^c + G_{2,i}^* P_{i+1}^c G_{2,i}\right]^{-1} + K_{u,i} P_i K_{u,i}^*,$$

is the block $(1,1)$ entry of $R_{e,i}$,

$$\Delta_i' = R_i^v + H_i P_i H_i^* - H_i P_i K_{u,i}^* \left(R_{e,i}(1,1)\right)^{-1} K_{u,i} P_i H_i^*, \tag{6.9.28}$$

is its Schur complement, and

$$\bar{x}_i = \hat{x}_i - P_i K_{u,i}^* \left(R_{e,i}(1,1)\right)^{-1} (\check{u}_i - \hat{u}_i). \tag{6.9.29}$$

We thus can show the following result.

Theorem 6.9.2 (Strictly Causal Measurement Feedback H^∞ Control) *Consider the state-space model (6.9.1) and define $\mathcal{T}(\mathcal{F})$ as the transfer operator that maps the unknown disturbances $\{\Pi_0^{-\frac{*}{2}} x_0, \{(Q_i^w)^{\frac{1}{2}} w_i, (R_i^v)^{-\frac{*}{2}} v_i\}_{i=0}^N\}$ to the variables*

$$\left\{(P_{N+1}^c)^{\frac{1}{2}} x_{N+1}, \{(R_i^c)^{\frac{1}{2}} s_i, (Q_i^c)^{\frac{1}{2}} z_i\}_{i=0}^N\right\},$$

where Π_0, P_{N+1}^c, Q_i^w, Q_i^c R_i^v and R_i^c are positive (semi)definite weighting matrices. Then for any given $\gamma > 0$, a strictly causal measurement feedback H^∞ control strategy $\check{u}_i = \mathcal{F}_i(y_0, \ldots, y_{i-1})$ that yields

$$\| \mathcal{T}(\mathcal{F}) \|_\infty < \gamma,$$

exists if, and only if,

(i) $\Pi_0^{-1} - \gamma^{-2} P_0^c > 0$

(ii) $\Delta_i = -\gamma^2 Q_i^w + G_{1,i}^* P_{i+1}^c G_{1,i} - G_{1,i}^* P_{i+1}^c G_{2,i} R_{G^c,i}^{-1} G_{2,i}^* P_{i+1}^c G_{1,i} < 0 \quad i = 0, \ldots, N$

(iii) all leading submatrices of

$$\begin{bmatrix} -I_{m_2} & 0 \\ 0 & I_p \end{bmatrix} \quad \text{and} \quad R_{e,i} = R_i + \begin{bmatrix} -K_{u,i} \\ H_i \end{bmatrix} P_i \begin{bmatrix} -K_{u,i}^* & H_i^* \end{bmatrix}$$

have the same inertia, for all $j = 0, \ldots, N$

where all the variables are as in Theorem 6.9.1. If this is the case, then all possible strictly causal measurement feedback H^∞ control strategies, $\check{u}_i = \mathcal{F}_i(y_0, \ldots, y_{i-1})$, are given by those that satisfy

$$\sum_{i=0}^N (\check{u}_i + K_{u,i}\hat{x}_i)^* \left(R_{e,i}(1,1)\right)^{-1} (\check{u}_i + K_{u,i}\hat{x}_i) + \sum_{i=0}^N (y_i - H_i \bar{x}_i)^* (\Delta_i')^{-1} (y_i - H_i \bar{x}_i) > 0, \tag{6.9.30}$$

where

$$\begin{cases} R_{e,i}(1,1) &= -\gamma^2 \left[Q_i^c + G_{2,i}^* P_{i+1}^c G_{2,i}\right]^{-1} + K_{u,i} P_i K_{u,i}^* \\ \Delta_i' &= R_i^v + H_i P_i H_i^* - H_i P_i K_{u,i}^* \left(R_{e,i}(1,1)\right)^{-1} K_{u,i} P_i H_i^* \end{cases}$$

and

$$\bar{x}_i = \hat{x}_i - P_i K_{u,i}^* \left(R_{e,i}(1,1)\right)^{-1} (\check{u}_i + K_{u,i}\hat{x}_i), \tag{6.9.31}$$

with $\hat{x}_i$ *satisfying the recursion,*

$$\hat{x}_{i+1} = (F_i - G_{1,i}K_{w,i})\hat{x}_i + K_{p,i} \begin{bmatrix} \check{u}_i + K_{u,i}\hat{x}_i \\ y_i - H_i\hat{x}_i \end{bmatrix} + G_{2,i}\check{u}_i, \quad \hat{x}_0 = 0.$$

The so-called central controller is given by

$$\check{u}_i = -K_{u,i}\hat{x}_{i|i}. \tag{6.9.32}$$

6.10 Conclusion

In this chapter we studied duality in linear spaces through the notion of a dual basis. This study led to certain dual approaches for solving (definite and indefinite) quadratic problems with various applications. We, in particular, used duality to solve the LQR control problem with indefinite weighting matrices, and the closely-related problems of full information and measurement feedback H^∞ control.