Proceedings of the 40th IEEE Conference on Decision and Control Orlando, Florida USA, December 2001

ThA13-4

A Unified Approach to Linear Estimation for Discrete-Time Systems-Part I: H_2 Estimation

Huanshui Zhang[1] Lihua Xie[2] Yeng Chai Soh

BLK S2, School of Electrical and Electronic Engineering
Nanyang Technological University, Singapore 639798
Email: elhxie@ntu.edu.sg
Tel: (65) 7904524; Fax: (65) 7912687

Abstract

The paper addresses the H_2 estimation problem for linear discrete systems with current and delayed measurements. By re-organizing the measurements and introducing a re-organized innovation sequence, a unified solution to the H_2 filtering, prediction and fixed-lag smoothing is derived based on an innovation analysis method together with projection in Hilbert space. Our solution does not require system augmentation and the estimator is given in terms of a Riccati difference equation of the same order as that of the system state. The proposed results will find important applications in sensor and data fusion. Furthermore, it will be shown in a companion paper that our solution to the H_2 estimation forms the basis of solving the more complicated problem of H_∞ fixed-lag smoothing without resorting to system augmentation.

1 Introduction and Problem Statement

Linear estimation has been a well known subject since the 1960s. The early work of optimal estimation includes the Wiener filtering [5] and the celebrated Kalman filtering [3]. The Kalman filtering has found significant applications in engineering such as in space technology, target tracking and telecommunications, etc. In this paper we shall revisit this problem by considering systems with measurements from multiple sensors of different delays. That is, the current measurements consist of information of state at different time instants. The solution to this optimal estimation has important applications in data and sensor fusion. It also turns out that the proposed approach can be extended to solve the more complicated H_∞ fixed-lag smoothing problem, which will be addressed in a companion paper.

We'll consider the linear stochastic system described by

$$\begin{aligned} \mathbf{x}(t+1) &= \Phi_t \mathbf{x}(t) + \Gamma_t \mathbf{u}(t) & (1)\\ \mathbf{y}(t) &= H_t \mathbf{x}(t) + \mathbf{v}(t) & (2)\\ \mathbf{z}_{t-d}(t) &= L_{t-d}\mathbf{x}(t-d) + \mathbf{v}_z(t) & (3)\end{aligned}$$

where $\mathbf{x}(t) \in R^n$ is the state, $\mathbf{u}(t) \in R^r$ is the input noise, $\mathbf{y}(t) \in R^m$ and $\mathbf{z}_{t-d}(t) \in R^{m_0}$ are respectively the current and delayed measurements, $\mathbf{v}(t) \in R^m$ and $\mathbf{v}_z(t) \in R^{m_0}$ are the measurements noises. The initial state $\mathbf{x}(0)$ and $\mathbf{u}(t)$, $\mathbf{v}(t)$ and $\mathbf{v}_z(t)$ are uncorrelated white noises with zero means and known covariance matrices $\mathcal{E}\left[\mathbf{x}(0)\mathbf{x}^T(0)\right] = P_0$, $\mathcal{E}\left[\mathbf{u}(i)\mathbf{u}^T(j)\right] = Q_u(i)\delta_{ij}$, $\mathcal{E}\left[\mathbf{v}(i)\mathbf{v}^T(j)\right] = Q_v(i)\delta_{ij}$ and $\mathcal{E}\left[\mathbf{v}_z(i)\mathbf{v}_z^T(j)\right] = Q_{v_z}(i)\delta_{ij}$, respectively.

In equation (3), $\mathbf{z}_{t-d}(t)$ means the observation of the state $\mathbf{x}(t-d)$ at time t, with delay d. So the system (1)-(3) is not in a standard form to which the Kalman filtering is applicable. The measurement equations can be rewritten as

$$\mathbf{y}_s(t) = \begin{cases} H_t\mathbf{x}(t) + \mathbf{v}_s(t), & 0 \le t < d \\ \begin{bmatrix} H_t & 0 \\ 0 & L_{t-d}\end{bmatrix}\begin{bmatrix}\mathbf{x}(t) \\ \mathbf{x}(t-d)\end{bmatrix} + \mathbf{v}_s(t), & t \ge d \end{cases} \tag{4}$$

where

$$\mathbf{y}_s(t) = \begin{cases} \mathbf{y}(t), & 0 \le t < d \\ \begin{bmatrix}\mathbf{y}(t) \\ \mathbf{z}_{t-d}(t)\end{bmatrix}, & t \ge d \end{cases},$$
$$\mathbf{v}_s(t) = \begin{cases} \mathbf{v}(t), & 0 \le t < d \\ \begin{bmatrix}\mathbf{v}(t) \\ \mathbf{v}_z(t)\end{bmatrix}, & t \ge d \end{cases} \tag{5}$$

and $\mathbf{v}_s(t)$ is a white noise with zero mean and covari-

[1]On leave from Research Center of Information and Control, Dalian University of Technology, P.R. China.

[2]Author to be contacted.

ance matrix $Q_{v_s}(t)$, where

$$Q_{v_s}(t) = \begin{cases} Q_v(t), & 0 \le t < d \\ \begin{bmatrix} Q_v(t) & 0 \\ 0 & Q_{v_z}(t) \end{bmatrix}, & t \ge d \end{cases} \quad (6)$$

The H_2 optimal estimation problem can be stated as: *Given the observation $\{\{\mathbf{y}_s(i)\}_{i=0}^t\}$ and an integer l, find a linear mean square error estimator $\hat{\mathbf{x}}(t-l \mid t)$ of $\mathbf{x}(t-l)$.*

The above estimation with current measurement $\mathbf{y}(t)$ and delayed measurement $\mathbf{z}_{t-d}(t)$ can find important application in many engineering problems such as communications and sensor fusion ([2]). In the companion paper, we'll show that the H_∞ estimation, including filtering, multi-step prediction and fixed-lag smoothing, is in fact an estimation problem of this kind in Krein space.

It should be pointed out that the above estimation problem has been well studied for the case when no delayed measurement ($d = 0$) is involved. In particular, the smoothing problem can be approached by a filtering problem of an augmented system ([4]) or a procedure involving backward and forward filtering ([6]). For the case when $d \neq 0$, the above estimation problem is different from the classical Kalman filtering although it may be transformed into the latter through system augmentation.

2 Optimal Estimator

In this section, we shall present a solution to the H_2 estimation for the system (1)-(3) involving delayed measurements using the projection in Hilbert space. The key to our discussion in this section is to organize the current and delayed measurements and introduce an associated innovation sequence. As is well known, given the measurement sequence $\{\mathbf{y}_s(i)\}_{i=0}^t$, the optimal state estimator $\hat{\mathbf{x}}(t-l \mid t)$ is the projection of $\mathbf{x}(t-l)$ on the linear space spanned by the measurement sequence, denoted by $\mathcal{L}\{\{\mathbf{y}_s(i)\}_{i=0}^t\}$ [1].

First, observe from (??) that for $d > t \ge 0$,

$$\mathcal{L}\{\{\mathbf{y}_s(i)\}_{i=0}^t\} = \mathcal{L}\{\{\mathbf{y}(i)\}_{i=0}^t\} \quad (7)$$

and the estimator $\hat{\mathbf{x}}(t-l \mid t)$ is a standard H_2 estimator (filter($l = 0$), smoother ($l > 0$) or predictor ($l < 0$)). For the simplicity of discussion, we'll assume that $t \ge d$ in this section. In the latter case, it is easy to know that the linear space $\mathcal{L}\{\{\mathbf{y}_s(i)\}_{i=0}^t\}$ can be re-organized as

$$\mathcal{L}_f(t, t-d) \triangleq \mathcal{L}\{\{\mathbf{y}_f(i)\}_{i=0}^{t-d}, \mathbf{y}(t-d+1), \cdots \mathbf{y}(t)\} \quad (8)$$

where for $i = 0, 1, \cdots, t-d$

$$\mathbf{y}_f(i) \triangleq \begin{bmatrix} \mathbf{y}(i) \\ \mathbf{z}_i(i+d) \end{bmatrix} = \begin{bmatrix} H_i \\ L_i \end{bmatrix} \mathbf{x}(i) + \mathbf{v}_f(i) \quad (9)$$

with

$$\mathbf{v}_f(i) = \begin{bmatrix} \mathbf{v}(i) \\ \mathbf{v}_z(i+d) \end{bmatrix} \quad (10)$$

being a white noise of zero mean and covariance matrix $Q_{v_f}(i) = \begin{bmatrix} Q_v(i) & 0 \\ 0 & Q_{v_z}(i+d) \end{bmatrix}$. It should be noted that $\mathbf{y}_f(i)$ contains measurements of the state $\mathbf{x}$ at time instants i and $i+d$. Observe that (1) and (9) give a standard state space representation.

The following notations will be used throughout the paper. $t_l \triangleq t - l$ and similarly $t_d \triangleq t - d$. $\hat{\xi}(j \mid t)$ stands for the optimal estimate of $\xi(j)$ based on the observations $\{\mathbf{y}_s(0), \cdots \mathbf{y}_s(t)\}$ whereas $\hat{\xi}(j \mid t+i, t)$ the optimal estimate of $\xi(j)$ based on the observations $\{\mathbf{y}_f(0), \cdots \mathbf{y}_f(t), \mathbf{y}(t+1), \cdots \mathbf{y}(t+i), i \ge 0\}$.

It is obvious that $\hat{\xi}(j \mid t, t)$ is the standard Kalman estimator and the estimator $\hat{\mathbf{x}}(t_l \mid t)$ to be sought for the system (1)-(3) can be re-denoted as $\hat{\mathbf{x}}(t_l \mid t, t_d)$. In other words, the optimal estimation problem is equivalent to finding the optimal estimate $\hat{\mathbf{x}}(t_l \mid t, t_d)$ of $x(t-l)$.

Before proceeding to compute the state estimate using the projection, we shall introduce the concepts of innovation sequence and Riccati equation.

2.1 Innovation sequence and Riccati equation

First, we introduce the following stochastic sequence:

$$\mathbf{w}(t+i, t) = \mathbf{y}(t+i) - \hat{\mathbf{y}}(t+i \mid t+i-1, t), \; i > 0 \quad (11)$$

$$\mathbf{w}(t, t) = \mathbf{y}_f(t) - \hat{\mathbf{y}}_f(t \mid t-1, t-1) \quad (12)$$

where $\hat{\mathbf{y}}_f(0 \mid -1, -1) = \begin{bmatrix} H_0 \\ L_0 \end{bmatrix} \hat{\mathbf{x}}(0 \mid -1, -1) = 0$ and $\hat{\mathbf{y}}_f(t+i \mid t+i-1, t)$ is the projection of $\mathbf{y}_f(t+i)$ on the linear space of $\{\mathbf{y}_f(0), \cdots \mathbf{y}_f(t), \mathbf{y}(t+1), \cdots \mathbf{y}(t+i-1)\}$ and $\hat{\mathbf{y}}(t \mid t-1, t-1)$ is the projection of $\mathbf{y}(t)$ on the linear space of $\{\mathbf{y}_f(0), \cdots \mathbf{y}_f(t-1)\}$. Then we have the following relationships

$$\mathbf{w}(t+i, t) = H_{t+i}\mathbf{e}(t+i, t) + \mathbf{v}(t+i), \quad i > 0 \quad (13)$$

$$\mathbf{w}(t, t) = \begin{bmatrix} H_t \\ L_t \end{bmatrix} \mathbf{e}(t, t) + \mathbf{v}_f(t) \quad (14)$$

where

$$\mathbf{e}(t+i, t) = \mathbf{x}(t+i) - \hat{\mathbf{x}}(t+i \mid t+i-1, t), \; i > 0 \quad (15)$$

$$\mathbf{e}(t, t) = \mathbf{x}(t) - \hat{\mathbf{x}}(t \mid t-1, t-1) \quad (16)$$

It is clear that $\mathbf{e}(t+1, t) = \mathbf{e}(t+1, t+1)$.

The following lemma shows that $\{\mathbf{w}(\cdot, \cdot)\}$ is in fact the innovation sequence.

Lemma 1

$$\{\mathbf{w}(0,0),\cdots,\mathbf{w}(t_d,t_d),\mathbf{w}(t_d+1,t_d),\cdots\mathbf{w}(t,t_d)\}$$

is the innovation sequence which is an uncorrelated white noise and spans the same linear space as

$$\mathcal{L}\{\mathbf{y}_f(0),\cdots,\mathbf{y}_f(t_d),\mathbf{y}(t_d+1),\cdots\mathbf{y}(t)\}$$

or equivalently $\mathcal{L}\{\mathbf{y}_s(0),\cdots\mathbf{y}_s(t)\}$.

From (13) and (14), the innovation covariance matrix

$$Q_w(t+i,t) \triangleq \mathcal{E}\left[\mathbf{w}(t+i,t)\mathbf{w}^T(t+i,t)\right],\ i>0$$

is given by

$$Q_w(t+i,t) = \begin{cases} H_{t+i}P(t+i,t)H_{t+i}^T + Q_v(t+i), & i \geq 0 \\ \begin{bmatrix} H_t \\ L_t \end{bmatrix} P(t,t) \begin{bmatrix} H_t \\ L_t \end{bmatrix}^T + Q_{v_f}(t), & i=0 \end{cases} \tag{17}$$

where

$$P(t+i,t) \triangleq \mathcal{E}\left[\mathbf{e}(t+i,t)\mathbf{e}^T(t+i,t)\right],\ i=0,1,\cdots$$

is the covariance matrix of the one-step ahead state estimation error, which can be computed using the Lemma below.

Lemma 2 The covariance matrix $P(j+1,t)$ for $j = t+1, t+2, \cdots$ can be calculated recursively as

$$\begin{aligned} P(j+1,t) &= \Phi_j P(j,t)\Phi_j^T + \Gamma_j Q_u(j)\Gamma_j^T, \\ &-\Phi_j P(j,t)H_j^T Q_w^{-1}(j,t)H_j P(j,t)\Phi_j^T \end{aligned} \tag{18}$$

where $P(t+1,t) = P(t+1,t+1)$ and $P(t+1,t+1)$ is the solution of the following standard Riccati equation

$$\begin{aligned} P(t+1,t+1) &= \Phi_t P(t,t)\Phi_t^T + \Gamma_t Q_u(t)\Gamma_t^T \\ &-\Phi_t P(t,t)\begin{bmatrix} H_t \\ L_t \end{bmatrix}^T Q_w^{-1}(t,t)\begin{bmatrix} H_t \\ L_t \end{bmatrix} P(t,t)\Phi_t^T \end{aligned} \tag{19}$$

where $P(0,0) = P_0$ and $Q_w(j,t)$ is as in (17).

In the sequel, for the convenience of discussion we give the following definition.

Definition 1

$$R_{t+i,t}^{t+j} \triangleq \mathcal{E}\left[\mathbf{x}(t+j)\mathbf{e}^T(t+i,t)\right] \tag{20}$$

is the cross-covariance matrix of the state $\mathbf{x}(t+j)$ and the state estimation error $\mathbf{e}(t+i,t)$.

$$K_{t+i,t}^{t+j} \triangleq \mathcal{E}\left[\mathbf{x}(t+j)\mathbf{w}^T(t+i,t)\right]Q_w^{-1}(t+i,t) \tag{21}$$

is the gain matrix of the innovation $\mathbf{w}(t+i,t)$ in the estimation of the state $\mathbf{x}(t+j)$.

For $i>0$, it follows easily from the above definition that

$$K_{t+i,t}^{t+j} = R_{t+i,t}^{t+j}H_{t+i}^T Q_w^{-1}(t+i,t) \tag{22}$$

For $i=0$, we have

$$K_{t,t}^{t+j} = R_{t,t}^{t+j}\begin{bmatrix} H_t \\ L_t \end{bmatrix}^T Q_w^{-1}(t,t) \tag{23}$$

By applying the projection formula and taking into consideration the above definition, it becomes clear that for $i>0$

$$\begin{aligned} \hat{\mathbf{x}}(t+j \mid t+i,t) &= \hat{\mathbf{x}}(t+j \mid t+i-1,t) \\ &+K_{t+i,t}^{t+j}\mathbf{w}(t+i,t) \end{aligned} \tag{24}$$

and

$$\hat{\mathbf{x}}(t+j \mid t,t) = \hat{\mathbf{x}}(t+j \mid t-1,t-1) + K_{t,t}^{t+j}\mathbf{w}(t,t) \tag{25}$$

Therefore, the key to the above update is the computation of the cross-covariance matrix $R_{t+i,t}^{t+j}$ which is given in the following lemma.

Lemma 3 The cross-covariance matrix $R_{t+i,t}^{t+j}$ for $i \geq 0$ can be computed as

$$R_{t+i,t}^{t+j} = \begin{cases} P(t+j,t)A^T(t+j,t)\cdots A^T(t+i-1,t), & i \geq j \\ \Phi_{t+j-1}\cdots\Phi_{t+i}P(t+i,t), & i<j \end{cases} \tag{26}$$

where $A(t+k,t)$, $k>0$ is given by

$$\begin{aligned} &A(t+k,t) = \\ &\Phi_{t+k}[I_n - P(t+k,t)H_{t+k}^T Q_w^{-1}(t+k,t)H_{t+k}] \end{aligned} \tag{27}$$

For $k \leq 0$, we let $P(t+k,t) = P(t+k,t+k)$ and $A(t+k,t) = A(t+k,t+k)$ in (26), where the matrix $A(t+k,t+k)$ is given by

$$\begin{aligned} A(t+k,t+k) &= \Phi_{t+k}\left(I_n - P(t+k,t+k)\begin{bmatrix} H_{t+k} \\ L_{t+k} \end{bmatrix}\right)^T \\ &\times Q_w^{-1}(t+k,t+k)\begin{bmatrix} H_{t+k} \\ L_{t+k} \end{bmatrix} \end{aligned} \tag{28}$$

Remark 1 It is readily seen from (26) that for $i \geq 0$ and $i \geq j$, $R_{t+i,t}^{t+j}$ can be calculated through the recursion

$$R_{t+i,t}^{t+j} = R_{t+i-1,t}^{t+j}A^T(t+i-1,t),\ R_{t+j,t}^{t+j} = P(t+j,t) \tag{29}$$

whereas for $i \geq 0$ and $j>i$,

$$R_{t+i,t}^{t+j} = \Phi_{t+j-1}R_{t+i,t}^{t+j-1},\quad R_{t+i,t}^{t+j} = P(t+i,t) \tag{30}$$

2.2 Optimal estimate $\hat{\mathbf{x}}(t_l \mid t,t_d)$ with $d \geq 0$

Based on the discussion in the previous subsection, it is clear that for a given integer $l \geq 0$, the optimal smoothing solution is in fact the optimal estimate $\hat{\mathbf{x}}(t_l \mid t,t_d)$ with $d \geq 0$, which is given in the following theorem.

Theorem 1 Consider the system (1)-(3). Given any integer l satisfying $d > l \geq 0$, the optimal estimator is given by

$$\hat{\mathbf{x}}(t_l \mid t) = \hat{\mathbf{x}}(t_l \mid t, t_d) \tag{31}$$

where $\hat{\mathbf{x}}(t_l \mid t, t_d)$ $(t \geq t_l \geq t_d)$ is given by

$$\hat{\mathbf{x}}(t_l \mid t, t_d) = \hat{\mathbf{x}}(t_l \mid t_l - 1, t_d) + \sum_{i=0}^{l} R^{t_l}_{t_l+i,t_d} H^T_{t_l+i}$$

$$Q_w^{-1}(t_l + i, t_d)\left[y(t_l + i) - H_{t_l+i}\hat{\mathbf{x}}(t_l + i \mid t_l + i - 1, t_d)\right] \tag{32}$$

while $\hat{\mathbf{x}}(t_l + i \mid t_l + i - 1, t_d)$, $i = 0, 1, \cdots, l$ in (32) is calculated recursively as

$$\begin{aligned}
&\hat{\mathbf{x}}(t_l + i + 1 \mid t_l + i, t_d)\\
= \;&\Phi_{t_l+i}\hat{\mathbf{x}}(t_l + i \mid t_l + i - 1, t_d)\\
&+\Phi_{t_l+i}P(t_l + i, t_d)H^T_{t_l+i}Q_w^{-1}(t_l + i, t_d)\\
&\times\left[\mathbf{y}(t_l + i) - H_{t_l+i}\hat{\mathbf{x}}(t_l + i \mid t_l + i - 1, t_d)\right]
\end{aligned} \tag{33}$$

with the initial value $\hat{\mathbf{x}}(t_d + 1 \mid t_d, t_d)$ given by the Kalman filter, which is computed recursively as

$$\begin{aligned}
&\hat{\mathbf{x}}(t_d + 1 \mid t_d, t_d) = \Phi_{t_d}\hat{\mathbf{x}}(t_d \mid t_d - 1, t_d - 1)\\
&+\Phi_{t_d}P(t_d, t_d)\begin{bmatrix} H_{t_d} \\ L_{t_d} \end{bmatrix}^T Q_w^{-1}(t_d, t_d)\\
&\times\left[\mathbf{y}_f(t_d) - \begin{bmatrix} H_{t_d} \\ L_{t_d} \end{bmatrix}\hat{\mathbf{x}}(t_d \mid t_d - 1, t_d - 1)\right].
\end{aligned} \tag{34}$$

where $\hat{\mathbf{x}}(0 \mid -1, -1) = 0$.

Remark 2 In the theorem we consider the case of $d > l \geq 0$. In the case of $l < 0$ (prediction) or $l \geq d$, the estimator $\hat{\mathbf{x}}(t_l \mid t, t_d)$ can be obtained using a similar method.

Remark 3 A procedure for computing the estimator given in the above theorem can be as follows:
1. Set $t = d$; $\hat{\mathbf{x}}(0 \mid -1) = 0$; $P(0, 0) = P_0$.
2. Compute $P(t_d + 1, t_d + 1)$ and the Kalman filter $\hat{\mathbf{x}}(t_d + 1 \mid t_d, t_d)$ using (19) and (34).
3. Set $i = l - d + 1$.
4. Compute $P(t_l + i, t_d)$ and $\hat{\mathbf{x}}(t_l + i + 1 \mid t_l + i, t_d)$ using (18) and (33).
5. If $i > 0$, compute $A(t_l + i - 1, t_d)$ using (27), and $R^{t_l}_{t_l+i,t_d}$ using (30).
6. Set $i = i + 1$. If $i < l + 1$ then go to Step 4.
7. Compute $\hat{\mathbf{x}}(t_l \mid t, t_d)$ using (32).
8. Set $t = t + 1$, goto Step 2.

2.3 Optimal estimate $\hat{\mathbf{x}}(t_l \mid t, t_d)$ for $d < 0$

In the last subsection, we have considered an optimal estimation problem for the case when $d \geq 0$. We shall now discuss the case of $d < 0$. It should be pointed out that the latter case implies that we have future measurements at the current time instant, which certainly does not make practical sense. However, we shall show in the companion paper that the H_∞ multi-step prediction problem shares a remarkable similarity to the H_2 estimation of this type.

First, note that the optimal estimate $\hat{\mathbf{x}}(t_l \mid t, t_d)$ is the projection of $\mathbf{x}(t_l)$ on the linear space of $\mathcal{L}\{\mathbf{y}_f(0), \cdots \mathbf{y}_f(t), \mathbf{z}_{t+1}(t + d + 1), \cdots, \mathbf{z}_{t_d}(t)\}$. It will become clear later that this problem is dual to the problem which has been investigated in the previous subsection. For the sake of simplicity, we adopt the similar notations as in the last subsection. Let $\hat{\xi}(j \mid t, t + i)$ be the optimal estimate of $\xi(j)$ based on the measurement sequence $\{\mathbf{y}_f(0), \cdots \mathbf{y}_f(t), \mathbf{z}_{t+1}(t + d + 1), \cdots, \mathbf{z}_{t+i}(t + d + i); i \geq 0\}$ and denote for $i > 0$

$$\begin{aligned}
\mathbf{e}(t, t + i) &\triangleq \mathbf{x}(t + i) - \hat{\mathbf{x}}(t + i \mid t, t + i - 1)\\
\mathbf{w}(t, t + i) &\triangleq \mathbf{z}_{t+i}(t + i + d)\\
&\quad\; -\hat{\mathbf{z}}_{t+i}(t + i + d \mid t, t + i - 1),\\
P(t, t + i) &\triangleq \mathcal{E}\left[\mathbf{e}(t, t + i)\mathbf{e}^T(t, t + i)\right]\\
Q_w(t, t + i) &\triangleq \mathcal{E}\left[\mathbf{w}(t, t + i)\mathbf{w}^T(t, t + i)\right]\\
R^{t+j}_{t,t+i} &\triangleq \mathcal{E}\left[\mathbf{x}(t + j)\mathbf{e}^T(t, t + i)\right]\\
K^{t+j}_{t,t+i} &\triangleq \mathcal{E}\left[\mathbf{x}(t + j)\mathbf{w}^T(t, t + i)\right]Q_w^{-1}(t, t + i)
\end{aligned}$$

For $i > 0$, it follows easily from the projection formula and the above definition that

$$K^{t+j}_{t,t+i} = R^{t+j}_{t,t+i}L^T_{t+i}Q_w^{-1}(t, t + i) \tag{35}$$

Similarly, $P(t, t + i)$ and $R^{t+j}_{t,t+i}$ can be calculated as in the following lemma.

Lemma 4 The cross-covariance matrix $R^{t+j}_{t,t+i}$ of $x(t + j)$ and $e(t, t + i)$ for $i \geq 0$ and $i \geq j$ can be computed as

$$R^{t+j}_{t,t+i} = \begin{cases} P(t, t + j)A^T(t, t + j)\cdots A^T(t, t + i - 1), & i \geq j\\ \Phi_{t+j-1}\cdots\Phi_{t+i}P(t, t + i), & i < j \end{cases} \tag{36}$$

where $A(t, t + k)$, $k > 0$ is given by

$$A(t, t + k) = \Phi_{t+k}\left[I_n - P(t, t + k)L^T_{t+k}Q_w^{-1}(t, t + k)L_{t+k}\right] \tag{37}$$

and for $k \leq 0$, we let $P(t, t + k) = P(t + k, t + k)$ and $A(t, t + k) = A(t + k, t + k)$ in (28). $P(t, t + i)$ and $Q_w(t, t + i)$ for $i > 0$ satisfy

$$P(t, t + i + 1) = \Phi_{t+i}P(t, t + i)\Phi^T_{t+i} + \Gamma_{t+i}Q_u(t + i)\Gamma^T_{t+i}$$

$$-\Phi_{t+i}P(t,t+i)L_{t+i}^TQ_w^{-1}(t,t+i)L_{t+i}^TP(t,t+i)\Phi_{t+i}^T \tag{38}$$

where $P(t,t+1)=P(t,t)$ and

$$Q_w(t,t+i)=L_{t+i}P(t,t+i)L_{t+i}^T+Q_{v_z}(t+i) \tag{39}$$

Remark 4 The calculation of $R_{t,t+i}^{t+j}$ can be carried out recursively as

$$\begin{aligned} R_{t,t+i}^{t+j} &= R_{t,t+i-1}^{t+j}A^T(t,t+i-1), \\ R_{t,t+j}^{t+j} &= P(t,t+j) \end{aligned} \tag{40}$$

For the case of $j>i$, it is easy to verify that

$$R_{t,t+i}^{t+j}=\Phi_{t+j-1}R_{t,t+i}^{t+j-1},\quad R_{t,t+i}^{t+i}=P(t,t+i) \tag{41}$$

Theorem 2 The optimal estimator $\hat{\mathbf{x}}(t_l \mid t,t_d)$ for $0>d>l$ ($t_l>t_d>t$) is given by

$$\hat{\mathbf{x}}(t_l \mid t,t_d)=\Phi_{t_l-1}\cdots\Phi_{t_d+1}\mathbf{x}(t_d+1 \mid t,t_d) \tag{42}$$

where $\hat{\mathbf{x}}(t_d+1 \mid t,t_d)$ is computed recursively as

$$\begin{aligned} &\hat{\mathbf{x}}(t+i+1 \mid t,t+i) \\ =\ &\Phi_{t+i}\hat{\mathbf{x}}(t+i \mid t,t+i-1) \\ &+\Phi_{t+i}P(t,t+i)L_{t+i}^TQ_w^{-1}(t,t+i) \\ &\times[\mathbf{z}_{t+i}(t+i+d)-L_{t+i}\hat{\mathbf{x}}(t+i \mid t,t+i-1)] \end{aligned} \tag{43}$$

for $i=1,2,\cdots-d$ with the initial value $\hat{\mathbf{x}}(t+1 \mid t,t)$ given by the Kalman filter of (34).

Remark 5 Theorem 2 presents an H_2 prediction result based on the measurements at time instants t and t_d, which is rather straightforward. For the case of $0>l\geq d$ or $l\geq 0$, the estimator $\hat{\mathbf{x}}(t_l \mid t,t_d)$ can be considered similarly.

3 Steady-state Estimation

In this section, we assume that the system to be considered is time-invariant, i.e.

$$\Phi_t=\Phi,\ H_t=H,\ L_t=L,\ \Gamma_t=\Gamma$$

and the white noises in (1)-(3) have the time-invariant covariance matrices, i.e.

$$Q_u(t)=Q_u,\ Q_v(t)=Q_v,\ Q_{v_z}(t)=Q_{v_z}$$

Assumption 1 Given the initial value P_0, the Riccati equation (19) has a time-invariant solution P as $t\longrightarrow\infty$.

Under the assumption, as t is large enough, $P(t+i,t)$, $Q_w(t+i,t)$, $A(t+i,t)$, $R_{t+i,t}^{t+j}$ and $K_{t+i,t}^{t+j}$ as well as their duals $P(t,t+i)$, $Q_w(t,t+i)$, $A(t,t+i)$, $R_{t,t+i}^{t+j}$, $K_{t,t+i}^{t+j}$ will be independent of the time t, and are rewritten as $P(i,0)$, $Q_w(i,0)$, $A(i,0)$, $R_{i,0}^j$, $K_{i,0}^j$, $P(0,i)$, $Q_w(0,i)$, $A(0,i)$, $R_{0,i}^j$, $K_{0,i}^j$, respectively. Now we consider the calculation of these matrices. First, the matrix $P(i,0)$, $i>0$ is given from (18) as

$$\begin{aligned} P(i+1,0) &= \Phi P(i,0)\Phi^T+\Gamma Q_u\Gamma^T \\ &\quad -\Phi P(i,0)H^TQ_w^{-1}(i,0)HP(i,0)\Phi^T \end{aligned} \tag{44}$$

with $P(1,0)=P$ and

$$Q_w(i,0)=HP(i,0)H^T+Q_v \tag{45}$$

It is obvious from (27), (26) and (22) that the matrices $A(i,0)$, $R_{i,0}^j$, $K_{i,0}^j$, $i>0$ are given as

$$A(i,0)=\Phi\left[I_n-P(i,0)H^TQ_w^{-1}(i,0)H\right] \tag{46}$$

$$R_{i,0}^j = \begin{cases} P(j,0)A^T(j,0)\cdots A^T(i-1,0), & i\geq j \\ \Phi^{j-i}P(i,0), & i<j \end{cases} \tag{47}$$

and

$$K_{i,0}^j=R_{i,0}^jH^TQ_w^{-1}(i,0) \tag{48}$$

Similarly, the matrices $P(0,i)$, $Q_w(0,i)$, $A(0,i)$, $R_{0,i}^j$ and $K_{0,i}^j$, $i>0$ can be computed by (44)-(48) with the matrices H and Q_v being replaced by L and Q_{v_z}, respectively. In the case of $i=0$, $Q_w(0,0)$, $A(0,0)$ and $K_{0,0}^j$ are calculated by (45)-(46) and (48) with the matrix H and Q_v being replaced by $H_L=\begin{bmatrix} H \\ L \end{bmatrix}$ and Q_{v_f}, respectively. The matrix $R_{0,0}^j$ is computed by

$$R_{0,0}^j = \begin{cases} P[A^T(0,0)]^{-j}, & j\leq 0 \\ \Phi^jP, & j>0 \end{cases} \tag{49}$$

3.1 Steady-state estimation with delay $d=0$

The case for filter ($l=0$) and multi step predictor ($l<0$) has been considered in the previous work [4]. For convenience of discussion below, we repeat the results here. The steady state estimator $\hat{\mathbf{x}}(t-l \mid t,t)$ for $l=-1$ is the Kalman filter, which is given as

$$\hat{\mathbf{x}}(t+1 \mid t,t)=\mathcal{F}_{-1}(q^{-1})\mathbf{y}_f(t) \tag{50}$$

where $\mathbf{y}_f(t)=\begin{bmatrix}\mathbf{y}^T(t) & \mathbf{z}_t^T(t)\end{bmatrix}^T$, and

$$\mathcal{F}_{-1}(q^{-1})=[I_n-q^{-1}(\Phi-K_{0,0}^1H_L)]^{-1}K_{0,0}^1 \tag{51}$$

For multi-step prediction ($l<-1$), the steady state multi-step predictor $\hat{\mathbf{x}}(t-l \mid t,t)$ is directly given as

$$\hat{\mathbf{x}}(t-l \mid t,t)=\mathcal{F}_l(q^{-1})\mathbf{y}_f(t) \tag{52}$$

where

$$\mathcal{F}_l(q^{-1}) = \Phi^{-l-1}\mathcal{F}_{-1}(q^{-1}) \tag{53}$$

Now we give the result for fixed lag smoothing problem ($l > 0$) based on the results for one-step ahead predictor.

Lemma 5 The steady state smoother $\hat{\mathbf{x}}(t_l \mid t, t),\ l > 0$ is computed by

$$\hat{\mathbf{x}}(t_l \mid t, t) = \mathcal{F}_l(q^{-1})\mathbf{y}_f(t) \tag{54}$$

where

$$\begin{aligned}\mathcal{F}_l(q^{-1}) \;\; = \;\; & q^{-1}[I_n q^{-l} - K_{0,0}(q^{-1})H_L]\mathcal{F}_{-1}(q^{-1}) \\ & + K_{0,0}(q^{-1})\end{aligned} \tag{55}$$

with

$$K_{0,0}(q^{-1}) = K_{0,0}^{0}q^{-l} + K_{0,0}^{-1}q^{-l+1} + \cdots + K_{0,0}^{-l} \tag{56}$$

3.2 Steady-state estimation with delay $d \neq 0$

Now we consider a more complicated state estimation with $d \neq 0$. The result will be applied to design a steady state H_∞ estimator.

Theorem 3 1) The steady state estimator $\hat{\mathbf{x}}(t_l \mid t, t_d)$ with $d \geq 0$ is computed by

$$\hat{\mathbf{x}}(t_l \mid t, t_d) = \mathcal{T}_{l,d}(q^{-1})\mathbf{y}_f(t_d) + T_{l,d}(q^{-1})\mathbf{y}(t) \tag{57}$$

where $\mathbf{y}_f(i) = [\mathbf{y}^T(i) \quad \mathbf{z}_i^T(i+d)]^T$, and

$$\mathcal{T}_{l,d}(q^{-1}) = \mathcal{F}_{l-d}(q^{-1}) - C_{l,d}(0)\mathcal{F}_{-1}(q^{-1}),$$

$$T_{l,d}(q^{-1}) = \sum_{i=1}^{d} K_{i,0}^{d-l}q^{-d+i} - \sum_{i=1}^{d-1} C_{l,d}(i)K_{i,0}^{i+1}q^{-d+i}$$

while $\mathcal{F}_{-1}(q^{-1})$ is as in (51) d $\mathcal{F}_{l-d}(q^{-1})$ is as in (53) for $l - d < 0$ or (55) for $l - d > 0$, and $C_{l,d}(i),\ i = 0, 1, \cdots, d-1$ is given by

$$C_{l,d}(i) = \sum_{j=i+1}^{d} K_{j,0}^{d-l}HA_{i+1}^{j-1} \tag{58}$$

with

$$A_{i+1}^{j-1} \;\; = \;\; \begin{cases} A(j-1,0)\cdots A(i+1,0), & i < j-1 \\ I_n, & i = j-1 \end{cases} \tag{59}$$

2) The steady state estimator $\hat{\mathbf{x}}(t_l \mid t, t_d)$ with $d < 0$ is computed by

$$\hat{\mathbf{x}}(t_l \mid t, t_d) = \mathcal{T}_{l,d}(q^{-1})\mathbf{y}_f(t) + T_{l,d}(q^{-1})\mathbf{z}_{t_d}(t) \tag{60}$$

where

$$\begin{aligned}\mathcal{T}_{l,d}(q^{-1}) &= \mathcal{F}_l(q^{-1}) - C_{l,d}(0)\mathcal{F}_{-1}(q^{-1}), \\ T_{l,d}(q^{-1}) &= \sum_{i=1}^{-d} K_{0,i}^{-l}q^{d+i} - \sum_{i=1}^{-d-1} C_{l,d}(i)K_{0,i}^{i+1}q^{d+i}\end{aligned} \tag{61}$$

while $C_{l,d}(i),\ i = 0, 1, \cdots, -d-1$ is given by

$$C_{l,d}(i) = \sum_{j=i+1}^{-d} K_{0,j}^{-l}L\bar{A}_{i+1}^{j-1} \tag{62}$$

with

$$\bar{A}_{i+1}^{j-1} \;\; = \;\; \begin{cases} A(0,j-1)\cdots A(0,i+1), & i < j+1 \\ I_n, & i = j-1 \end{cases}$$

4 Conclusion

In this paper, we have investigated the H_2 estimation problems for discrete time system with current and delayed mesurements based on the innovation analysis method and projection formula in Hilbert space. The key to our development is the re-organized innovation sequence which is different from the Kalman filtering innovation sequence ([1]). As will be shown in the companion paper, the unified approach for filtering, prediction and smoothing can be extended to the H_∞ estimation.

References

[1] Z. L. Deng, H. S. Zhang, S. J. Liu and L. Zhou, "Optimal and self-tuning white noise estimators with applications to deconvolution and filtering problems," *Automatica,* vol. 32, pp.199-216, 1996.

[2] L.A. Klein, *Sensor and Data Fusion Concepts and Applications,* SPIE Press, 1999.

[3] R.E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering, Trans. ASME-D,* vol. 82, no. 1, pp. 35-45, 1960.

[4] B. D. O. Anderson and J. B. Moore, *Optimal filtering.* Prentice-Hall, Englewood. Cliffs, N.J, 1979.

[5] N. Wiener, *Extrapolation interpolation, and smoothing of stationary time series.* The Technology Press and Wiley, New York, 1950.

[6] J. M. Mendel, *Lessons in estimation thoery for signal processing, communications, and control,* Prentice-Hall, In., 1995.