Unbiased Minimum-Variance Input and State Estimation for Linear Discrete-Time Systems ⋆

Steven Gillijns [a,1], Bart De Moor [a],

[a] *K.U.Leuven, Department of Electrical Engineering (ESAT), Research group SCD-SISTA, Kasteelpark Arenberg 10, 3001 Leuven, Belgium*

Abstract

This paper addresses the problem of simultaneously estimating the state and the input of a linear discrete-time system. A recursive filter, optimal in the minimum-variance unbiased sense, is developed where the estimation of the state and the input are interconnected. The input estimate is obtained from the innovation by least-squares estimation and the state estimation problem is transformed into a standard Kalman filtering problem. Necessary and sufficient conditions for the existence of the filter are given and relations to earlier results are discussed.

Key words: Kalman filtering; Recursive state estimation; Unknown input estimation; Minimum-variance estimation

1 Introduction

Thanks to its applications in fault detection and in state estimation for geophysical processes with unknown disturbances, the problem of state estimation for linear systems with unknown inputs has received considerable attention during the last decades.

For continuous-time systems, necessary and sufficient conditions for the existence of an optimal state estimator are well-established [KVR80,HP98b]. Furthermore, design procedures for the reconstruction of unknown inputs have received considerable attention [HP98a,XS03].

For discrete-time systems, earliest attempts were based on augmenting the state vector with an unknown input vector, where a prescribed model for the unknown input is assumed. To reduce computation costs of the augmented state filter, Friedland [Fri69] proposed the two-stage Kalman filter where the estimation of the state and the unknown input are decoupled. Although successfully used in many applications, both methods are limited to the case where a model for the dynamical evolution of the unknown input is available.

Kitanidis [Kit87], on the other hand, developed an optimal recursive state filter which is based on the assumption that no prior information about the unknown input is available. His result was extended by Darouach and Zasadzinski [DZ97] who established stability and convergence conditions and developed a new design method for the optimal state filter.

Hsieh [Hsi00] established a connection between the two-stage filter and the Kitanidis filter by showing that Kitanidis' result can be derived by making the two-stage filter independent of the underlying input model. Furthermore, his method yields an estimate of the unknown input. However, the optimality of the input estimate has not been proved.

This paper is an extension of [Kit87] and [DZ97] to combined MVU input and state estimation. We propose a recursive filter where the estimation of the unknown input and the state are interconnected. We prove that this approach yields the same state update as in [Kit87] and [DZ97] and the same input estimate as in [Hsi00], thereby also showing that the latter input estimate is indeed optimal.

This paper is organized as follows. In section 2, the problem is formulated and the structure of the recursive filter is presented. Section 3 deals with optimal reconstruction of the unknown input. Next, the state estimation problem is solved in section 4. Finally, a proof of global optimality is provided in section 5.

⋆ This paper was not presented at any IFAC meeting.

Email addresses: steven.gillijns@esat.kuleuven.be (Steven Gillijns), bart.demoor@esat.kuleuven.be (Bart De Moor).

[1] Corresponding author. Tel. +32-16-32-17-09. Fax +32-16-32-19-70.

2 Problem formulation

Consider the linear discrete-time system

$$x_{k+1} = A_k x_k + G_k d_k + w_k, \tag{1}$$
$$y_k = C_k x_k + v_k, \tag{2}$$

where $x_k \in \mathbb{R}^n$ is the state vector, $d_k \in \mathbb{R}^m$ is an unknown input vector, and $y_k \in \mathbb{R}^p$ is the measurement. The process noise $w_k \in \mathbb{R}^n$ and the measurement noise $v_k \in \mathbb{R}^p$ are assumed to be mutually uncorrelated, zero-mean, white random signals with known covariance matrices, $Q_k = \mathbb{E}[w_k w_k^\mathsf{T}]$ and $R_k = \mathbb{E}[v_k v_k^\mathsf{T}]$, respectively. Results are easily generalized to the case where w_k and v_k are correlated by transforming (1)-(2) into an equivalent system where process and measurement noise are uncorrelated [AM79, Chap. 5.5].

Throughout the paper, we assume that (C_k, A_k) is observable and that x_0 is independent of v_k and w_k for all k. Also, we assume that the following well-established condition for the existence of an unbiased state estimator is satisfied;

Assumption 1 ([Kit87,DZ97])
rank $C_k G_{k-1}$ = *rank* $G_{k-1} = m$, *for all* k.

Notice that Assumption 1 implies $n \geq m$ and $p \geq m$.

The objective of this paper is to make MVU estimates of the system state x_k and the unknown input d_{k-1}, given the sequence of measurements $Y_k = \{y_0, y_1, \ldots, y_k\}$. No prior knowledge about d_{k-1} is assumed to be available and no prior assumption is made. In fact, the unknown input d_{k-1} can be any type of signal and may for example be an unknown nonlinear function of x_{k-1} and d_{k-2}.

We consider a recursive filter of the form

$$\hat{x}_{k|k-1} = A_{k-1}\hat{x}_{k-1|k-1}, \tag{3}$$
$$\hat{d}_{k-1} = M_k(y_k - C_k\hat{x}_{k|k-1}), \tag{4}$$
$$\hat{x}^\star_{k|k} = \hat{x}_{k|k-1} + G_{k-1}\hat{d}_{k-1}, \tag{5}$$
$$\hat{x}_{k|k} = \hat{x}^\star_{k|k} + K_k(y_k - C_k\hat{x}^\star_{k|k}), \tag{6}$$

where $M_k \in \mathbb{R}^{m\times p}$ and $K_k \in \mathbb{R}^{n\times p}$ still have to be determined. Let $\hat{x}_{k-1|k-1}$ be an unbiased estimate of x_{k-1}, then $\hat{x}_{k|k-1}$ is biased due to the unknown input in the true system. Therefore, an unbiased estimate of the unknown input is calculated from the measurement in (4) and used to obtain an unbiased state estimate $\hat{x}^\star_{k|k}$ in (5). In the final step, the variance of $\hat{x}^\star_{k|k}$ is minimized by using an update similar to the Kalman filter.

Conditions on the matrix M_k to obtain unbiased and MVU estimates of the unknown input, are derived in section 3. The gain matrix K_k minimizing the variance of $\hat{x}_{k|k}$, is computed in section 4.

3 Input estimation

In this section, we consider the estimation of the unknown input. In 3.1, we determine the matrix M_k such that (4) is an unbiased estimator of d_{k-1}. In 3.2, we extend to MVU input estimation.

3.1 Unbiased input estimation

Defining the *innovation* $\tilde{y}_k$ by

$$\tilde{y}_k \triangleq y_k - C_k\hat{x}_{k|k-1}, \tag{7}$$

it follows from (1)-(3) that

$$\tilde{y}_k = C_k G_{k-1} d_{k-1} + e_k, \tag{8}$$

where e_k is given by

$$e_k = C_k(A_{k-1}\tilde{x}_{k-1} + w_{k-1}) + v_k, \tag{9}$$

with $\tilde{x}_k \triangleq x_k - \hat{x}_{k|k}$.

Let $\hat{x}_{k-1|k-1}$ be unbiased, then it follows from (9) that $\mathbb{E}[e_k] = 0$ and consequently from (8) that

$$\mathbb{E}[\tilde{y}_k] = C_k G_{k-1} d_{k-1}. \tag{10}$$

Equation (10) indicates that an unbiased estimate of the unknown input d_{k-1} can be obtained from the innovation.

Theorem 1 *Let* $\hat{x}_{k-1|k-1}$ *be unbiased, then (3)-(4) is an unbiased estimator of* d_{k-1} *if and only if* M_k *satisfies*

$$M_k C_k G_{k-1} = I_m. \tag{11}$$

PROOF. Substituting (8) in (4), yields

$$\hat{d}_{k-1} = M_k C_k G_{k-1} d_{k-1} + M_k e_k.$$

Noting that $\hat{d}_{k-1}$ is unbiased if and only if M_k satisfies $M_k C_k G_{k-1} = I_m$, concludes the proof.

The matrix M_k corresponding to the least-squares (LS) solution of (8), satisfies (11). The LS solution is thus unbiased. However, it does not have minimum-variance because e_k does not have unit variance and thus (8) does not satisfy the assumptions of the Gauss-Markov theorem [KSH00, Chap. 3.4.2]. Nevertheless, the variance of e_k can be computed from the covariance matrices of the state estimator. An MVU estimator of d_{k-1} is then obtained by weighted least-squares (WLS) estimation with weighting matrix $(\mathbb{E}[e_k e_k^\mathsf{T}])^{-1}$, as will be shown in the next section.

3.2 Unbiased minimum-variance input estimation

Denoting the variance of e_k by $\tilde{R}_k$, a straightforward calculation yields

$$\begin{aligned}\tilde{R}_k &= \mathbb{E}[e_k e_k^\mathsf{T}], \\ &= C_k(A_{k-1}P_{k-1|k-1}A_{k-1}^\mathsf{T} + Q_k)C_k^\mathsf{T} + R_k,\end{aligned} \tag{12}$$

where $P_{k|k} \triangleq \mathbb{E}[\tilde{x}_k \tilde{x}_k^\mathsf{T}]$. Furthermore, defining

$$P_{k|k-1} \triangleq A_{k-1}P_{k-1|k-1}A_{k-1}^\mathsf{T} + Q_{k-1},$$

it follows that $\tilde{R}_k$ can be rewritten as

$$\tilde{R}_k = C_k P_{k|k-1} C_k^\mathsf{T} + R_k.$$

An MVU input estimate is then obtained as follows.

Theorem 2 *Let Assumption 1 hold, let $\hat{x}_{k-1|k-1}$ be unbiased, let $\tilde{R}_k$ be positive definite and let M_k be given by*

$$M_k = \left(F_k^\mathsf{T}\tilde{R}_k^{-1}F_k\right)^{-1} F_k^\mathsf{T}\tilde{R}_k^{-1}, \tag{13}$$

where $F_k \triangleq C_k G_{k-1}$, then (3)-(4) is the MVU estimator of d_{k-1} given the innovation $\tilde{y}_k$. The variance of the corresponding input estimate, is given by $(F_k^\mathsf{T}\tilde{R}_k^{-1}F_k)^{-1}$.

PROOF. Under the assumption that $\tilde{R}_k$ is positive definite, an invertible matrix $\tilde{S}_k \in \mathbb{R}^{p\times p}$ satisfying $\tilde{S}_k\tilde{S}_k^\mathsf{T} = \tilde{R}_k$, can always be found, for example by a Cholesky factorization. We now transform (8) to

$$\tilde{S}_k^{-1}\tilde{y}_k = \tilde{S}_k^{-1}C_k G_{k-1}d_{k-1} + \tilde{S}_k^{-1}e_k. \tag{14}$$

Under the assumption that $\tilde{S}_k^{-1}C_k G_{k-1}$ has full column rank, the LS solution $\hat{d}_{k-1}$ of (14) equals

$$\hat{d}_{k-1} = (F_k^\mathsf{T}\tilde{R}_k^{-1}F_k)^{-1}F_k^\mathsf{T}\tilde{R}_k^{-1}\tilde{y}_k, \tag{15}$$

where $F_k = C_k G_{k-1}$. Note that solving (14) by LS estimation is equivalent to solving (8) by WLS estimation with weighting matrix $\tilde{R}_k^{-1}$. In addition, since the weighting matrix is chosen such that $\tilde{S}_k^{-1}e_k$ has unit variance, equation (14) satisfies the assumptions of the Gauss-Markov theorem. Hence, (15) is the MVU estimate of d_{k-1} given $\tilde{y}_k$ [KSH00, Chap. 2.2.3]. The variance of the WLS solution (15) is given by $(F_k^\mathsf{T}\tilde{R}_k^{-1}F_k)^{-1}$. □

This input estimator has a strong connection to the filter designed in [Hsi00].

Theorem 3 *Let M_k be given by (13), then we obtain the same input estimate as in [Hsi00, Sec. III].*

In [Hsi00, Sec. III], the input estimate follows by making the two-stage Kalman filter independent of the underlying input model. However, the optimality of the input estimate has not been shown. Here, we obtain the same estimate from the innovation in an optimal way, showing that the input estimate of [Hsi00] is indeed optimal.

4 State estimation

Consider a state estimator for the system (1)-(2) which takes the recursive form (3)-(6). In 4.1, we search for a condition on the gain matrix K_k such that (6) is an unbiased estimator of x_k. In 4.2, we extend to MVU state estimation.

4.1 Unbiased state estimation

Defining $\tilde{x}_k^\star \triangleq x_k - \hat{x}_{k|k}^\star$, it follows from (1)-(3) and (5) that

$$\tilde{x}_k^\star = A_{k-1}\tilde{x}_{k-1} + G_{k-1}\tilde{d}_{k-1} + w_{k-1}, \tag{16}$$

where $\tilde{d}_k \triangleq d_k - \hat{d}_k$. The following theorem is a direct consequence of (16).

Theorem 4 *Let $\hat{x}_{k-1|k-1}$ and $\hat{d}_{k-1}$ be unbiased, then (5)-(6) are unbiased estimators of x_k for any value of K_k.*

This unbiased state estimator has a strong connection to the filter designed in [Kit87]. Substituting (4) and (5) in (6), yields

$$\begin{aligned}\hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k\tilde{y}_k + (I_n - K_kC_k)G_{k-1}\hat{d}_{k-1}, &(17)\\ &= \hat{x}_{k|k-1} + K_k\tilde{y}_k + (I_n - K_kC_k)G_{k-1}M_k\tilde{y}_k. &(18)\end{aligned}$$

Defining

$$L_k \triangleq K_k + (I_n - K_kC_k)G_{k-1}M_k,$$

(18) is rewritten as

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + L_k(y_k - C_k\hat{x}_{k|k-1}), \tag{19}$$

which is the kind of update considered in [Kit87].

Theorem 5 *Let M_k be given by (13) and K_k by*

$$K_k = P_{k|k-1}C_k\tilde{R}_k^{-1}, \tag{20}$$

then we obtain the state update of [Kit87].

In [Kit87] only state estimation is considered. However, we conclude from the equivalence of (17) and (19) that Kitanidis' filter implicitly estimates the unknown input from the innovation by WLS estimation.

4.2 Unbiased minimum-variance state estimation

In this section, we calculate the optimal gain matrix K_k as function of M_k. The derivation holds for any M_k satisfying (11) and yields the MVU estimate $\hat{x}_{k|k}$ of x_k given the value of M_k used in (4).

First, we search for an expression of $\tilde{d}_{k-1}$. It follows from (4) and (8)-(9) that

$$\tilde{d}_{k-1} = (I_m - M_k C_k G_{k-1}) d_{k-1} - M_k e_k = -M_k e_k, \quad (21)$$

where the last step follows from the unbiasedness of the input estimator. Substituting (21) in (16), yields

$$\tilde{x}_k^\star = A_{k-1}^\star \tilde{x}_{k-1} + w_{k-1}^\star, \quad (22)$$

where

$$A_{k-1}^\star = (I_n - G_{k-1} M_k C_k) A_{k-1}, \quad (23)$$
$$w_{k-1}^\star = (I_n - G_{k-1} M_k C_k) w_{k-1} - G_{k-1} M_k v_k. \quad (24)$$

An expression for the error covariance matrix $P_{k|k}^\star \triangleq \mathbb{E}[\tilde{x}_k^\star \tilde{x}_k^{\star\mathsf{T}}]$ follows from (22)-(24),

$$\begin{aligned} P_{k|k}^\star &= A_{k-1}^\star P_{k-1|k-1} A_{k-1}^{\star\mathsf{T}} + Q_{k-1}^\star, \\ &= (I_n - G_{k-1} M_k C_k) P_{k|k-1} (I_n - G_{k-1} M_k C_k)^\mathsf{T} + \\ &\quad G_{k-1} M_k R_k M_k^\mathsf{T} G_{k-1}^\mathsf{T}, \end{aligned} \quad (25)$$

where $Q_k^\star \triangleq \mathbb{E}[w_k^\star w_k^{\star\mathsf{T}}]$.

Next, we search for an expression of the error covariance matrix $P_{k|k}$. It follows from (6) that

$$\tilde{x}_k = (I_n - K_k C_k) \tilde{x}_k^\star - K_k v_k. \quad (26)$$

Substituting (22) in (26), yields

$$\tilde{x}_k = (I_n - K_k C_k)(A_{k-1}^\star \tilde{x}_{k-1} + w_{k-1}^\star) - K_k v_k, \quad (27)$$

where $\mathbb{E}[w_{k-1}^\star v_k^\mathsf{T}] = -G_{k-1} M_k R_k$. Notice that (27) has a close connection to the Kalman filter. This expression represents the dynamical evolution of the error in the state estimate of a Kalman filter with gain matrix K_k for the system $(A_k^\star, C_k)$, where the process noise $w_{k-1}^\star$ is correlated with the measurement noise v_k. The calculation of the optimal gain matrix K_k has thus been reduced to a standard Kalman filtering problem.

It follows from (26) and (25) that the error covariance matrix $P_{k|k}$ is given by

$$P_{k|k} = K_k \tilde{R}_k^\star K_k^\mathsf{T} - V_k^\star K_k^\mathsf{T} - K_k V_k^{\star\mathsf{T}} + P_{k|k}^\star, \quad (28)$$

where

$$\begin{aligned} \tilde{R}_k^\star &= C_k P_{k|k}^\star C_k^\mathsf{T} + R_k + C_k S_k^\star + S_k^{\star\mathsf{T}} C_k^\mathsf{T}, \quad (29) \\ V_k^\star &= P_{k|k}^\star C_k^\mathsf{T} + S_k^\star, \\ S_k^\star &= \mathbb{E}[\tilde{x}_k^\star v_k^\mathsf{T}] = -G_{k-1} M_k R_k. \end{aligned}$$

Notice that $\tilde{R}_k^\star$ equals the variance of the zero-mean signal $\tilde{y}_k^\star$, $\tilde{R}_k^\star = \mathbb{E}[\tilde{y}_k^\star \tilde{y}_k^{\star\mathsf{T}}]$, where

$$\tilde{y}_k^\star \triangleq y_k - C_k \hat{x}_{k|k}^\star = (I_p - C_k G_{k-1} M_k) e_k. \quad (30)$$

Using (30) and (12), (29) can be rewritten as

$$\tilde{R}_k^\star = (I_p - C_k G_{k-1} M_k) \tilde{R}_k (I_p - C_k G_{k-1} M_k)^\mathsf{T}.$$

From Kalman filtering theory, we know that uniqueness of the optimal gain matrix K_k requires invertibility of $\tilde{R}_k^\star$. However, we now show that $\tilde{R}_k^\star$ is singular by proving that $I_p - C_k G_{k-1} M_k$ is not of full rank.

Lemma 6 *Let M_k satisfy (11), then $I_p - C_k G_{k-1} M_k$ has rank $p - m$.*

PROOF. Because M_k satisfies (11), it is a left inverse of $C_k G_{k-1}$. Consequently, $C_k G_{k-1} M_k$ and $I_p - C_k G_{k-1} M_k$ are idempotent [Ber05, Fact 3.8.7 and Fact 3.8.9]. The rank of $I_p - C_k G_{k-1} M_k$ is then given by

$$\begin{aligned} \text{rank } I_p - C_k G_{k-1} M_k &= p - \text{rank } C_k G_{k-1} M_k, \\ &= p - m, \end{aligned}$$

where the first equality follows from [Ber05, Fact 3.8.6] and the second equality from [Ber05, Prop. 2.6.2]. □

Consequently, the optimal gain matrix K_k is not unique. Let r be the rank of $\tilde{R}_k^\star$, we then propose a gain matrix K_k of the form

$$K_k = \bar{K}_k \alpha_k, \quad (31)$$

where $\alpha_k \in \mathbb{R}^{r \times p}$ is an arbitrary matrix which has to be chosen such that $\alpha_k \tilde{R}_k^\star \alpha_k^\mathsf{T}$ has full rank. The optimal gain matrix K_k is then given in the following theorem.

Theorem 7 *Let M_k satisfy (11) and let $\alpha_k \in \mathbb{R}^{r \times p}$, with $r = \text{rank } \tilde{R}_k^\star$, be an arbitrary matrix, chosen such that $\alpha_k \tilde{R}_k^\star \alpha_k^\mathsf{T}$ has full rank, then the gain matrix K_k of the form (31) minimizing the variance of $\hat{x}_{k|k}$, is given by*

$$K_k = (P_{k|k}^\star C_k^\mathsf{T} + S_k^\star) \alpha_k^\mathsf{T} (\alpha_k \tilde{R}_k^\star \alpha_k^\mathsf{T})^{-1} \alpha_k. \quad (32)$$

PROOF. Substituting (31) in (28) and minimizing the trace of $P_{k|k}$ over $\bar{K}_k$, yields (32). □

Substituting (32) in (28), yields the following update for the error covariance matrix,

$$P_{k|k} = P^{\star}_{k|k} - K_k(P^{\star}_{k|k}C_k^{\mathsf{T}} + S^{\star}_k)^{\mathsf{T}}.$$

We now give the relation to [DZ97].

Theorem 8 *Let M_k satisfy (11) and let K_k be given by (32) with $r = p - m$, then we obtain the same state update as in [DZ97]. Furthermore, as shown in [DZ97], for M_k given by (13) and $\alpha_k = [0\ I_r]U_k^{\mathsf{T}}\tilde{S}_k^{-1}$, where U_k is an orthogonal matrix containing the left singular vectors of $\tilde{S}_k^{-1}C_kG_{k-1}$ in its columns, the Kitanidis filter is obtained.*

By parameterizing the unbiasedness conditions in [Kit87], Darouach and Zasadzinski [DZ97] showed that the gain matrix is not unique. Here, the same result is obtained by a procedure which has a closer connection to the Kalman filter.

Notice that the expression (32) implicitly depends on the choice of M_k. *Given the value of M_k used in (4)*, (32) yields the gain matrix K_k for which the variance of $\hat{x}_{k|k}$ is minimal. Our result does not allow to conclude which value(s) of M_k should optimally be used in (4) to minimize the variance of $\hat{x}_{k|k}$.

5 Proof of optimality

In [KP00], it is proved that a recursive MVU state estimator which can be written in the form (3),(19), minimizes the mean square error of $\hat{x}_{k|k}$ over the class of all linear unbiased state estimates based on Y_k. By a similar derivation, we now prove that the estimate of d_{k-1} minimizing the mean square error over the class of all linear unbiased estimates based on Y_k, can be written in the form (4). The proof is inspired by the optimality proof in [KP00].

We relax the recursivity assumption and consider $\hat{d}_{k-1}$ to be the most general linear combination of $\hat{x}_{0|0}$ and Y_k. As pointed out in [KP00], because the innovation $\tilde{y}_k$ is itself a linear combination of $\hat{x}_{0|0}$ and Y_k, the most general estimate of d_{k-1} can be written in the form

$$\hat{d}_{k-1} = M_k\tilde{y}_k + \sum_{i=0}^{k-1} H_i\tilde{y}_i + N\hat{x}_{0|0}, \tag{33}$$

where we dropped the dependence of H_i and N on k for notational simplicity. A necessary and sufficient condition for (33) to be an unbiased estimator of d_{k-1}, is given in the following lemma.

Lemma 9 *The estimator (33) is unbiased if and only if $N = 0$, M_k satisfies (11) and $H_iC_iG_{i-1} = 0$ for every $i < k$.*

PROOF. Sufficiency – It follows from (10) that if $H_iC_iG_{i-1} = 0$ for every $i < k$, then $\sum_{i=0}^{k-1} H_i\mathbb{E}[\tilde{y}_i] = 0$. Furthermore, for M_k satisfying (11), $M_k\tilde{y}_k$ and consequently also (33), with $N = 0$, are unbiased estimators of d_{k-1}.
Necessity – Assume that (33) is an unbiased estimator of d_{k-1}. Since no prior information about d_{k-1} is available and since y_k is the first measurement containing information about d_{k-1}, we conclude that $\mathbb{E}[M_k\tilde{y}_k] = d_{k-1}$ and that consequently also (11) must hold. Furthermore, the expected value of the sum of the last two terms in (33) is zero for any unknown input sequence $d_0, d_1, \ldots, d_{k-1}$ if and only if $H_iC_iG_{i-1} = 0$ for every $i < k$ and $N = 0$. □

In the remainder of this section, we only consider unbiased input estimators of the form (33). We now prove that the mean square error

$$\sigma^2_{k-1} \triangleq \mathbb{E}[\|d_{k-1} - \hat{d}_{k-1}\|_2^2] \tag{34}$$

achieves a minimum when $H_0 = H_1 = \ldots = H_{k-1} = 0$.

Theorem 10 *Let $\hat{d}_{k-1}$ given by (33) be unbiased, then the mean square error (34) achieves a minimum when $H_0 = H_1 = \ldots = H_{k-1} = 0$.*

In the proof of Theorem 10, we make use of the following lemma, which provides an orthogonality relationship.

Lemma 11 (See [KP00], Lemma 2) *Let $\tilde{y}_i$ be defined by (7), then for every $i < k$ and every H_i satisfying $H_iC_iG_{i-1} = 0$, $\mathbb{E}[\tilde{y}_k(H_i\tilde{y}_i)^{\mathsf{T}}] = 0$ and $\mathbb{E}[d_{k-1}(H_i\tilde{y}_i)^{\mathsf{T}}] = 0$.*

The proof of Theorem 10 is then given as follows.

PROOF. Inspired by the proof of Theorem 3 in [KP00], we write $d_{k-1} - \hat{d}_{k-1} = f_M - g_H$, where $f_M \triangleq d_{k-1} - M_k\tilde{y}_k$ and $g_H \triangleq \sum_{i=0}^{k-1} H_i\tilde{y}_i$. It follows from Lemma 11 that $\mathbb{E}[f_Mg_H^{\mathsf{T}}] = 0$, so that

$$\begin{aligned}\sigma^2_{k-1} &= \mathrm{trace}\{(f_M + g_H)(f_M + g_H)^{\mathsf{T}}\},\\ &= \mathbb{E}[\|f_M\|_2^2] + \mathbb{E}[\|g_H\|_2^2].\end{aligned} \tag{35}$$

The second term in (35) is minimized when $g_H = 0$, which occurs for $H_0 = H_1 = \ldots = H_{k-1} = 0$. That solution also satisfies $H_iC_iG_{i-1} = 0$, which completes the proof. □

It follows from Theorem 10 and (33) that the globally optimal linear estimate of d_{k-1} based on Y_k can be written in the recursive form (4). Furthermore, because the matrix M_k given by (13) minimizes $\mathbb{E}[\|f_M\|_2^2]$, it follows that (4) yields the globally optimal linear estimate of d_{k-1} for this value of M_k. Combining this result with Theorem 5 and the global optimality of the Kitanidis filter proved in [KP00], yields the following theorem.

Theorem 12 *Consider a simultaneous input and state estimator of the recursive form (3)-(6). Let M_k be given by (13) and let K_k be given by (20), then (4) and (6) are unbiased estimators of d_{k-1} and x_k minimizing the mean square error over the class of all linear unbiased estimates based on $\hat{x}_{0|0}$ and $Y_k = \{y_0, y_1, \ldots, y_k\}$.*

6 Conclusion

An optimal filter is developed which simultaneously estimates the input and the state of a linear discrete-time system. The estimate of the input is obtained from the innovation by least-squares estimation. The state estimation problem is transformed into a standard Kalman filtering problem for a system with correlated process and measurement noise. We prove that this approach yields the same state update as in [Kit87] and [DZ97], and the same input estimate as in [Hsi00]. Finally, a proof is included, showing that the optimal input estimate over the class of all linear unbiased estimates may be written in the proposed recursive form.

Acknowledgements

Our research is supported by Research Council KULeuven: GOA AMBioRICS, several PhD/postdoc & fellow grants; Flemish Government: FWO: PhD/postdoc grants, projects, G.0407.02 (support vector machines), G.0197.02 (power islands), G.0141.03 (Identification and cryptography), G.0491.03 (control for intensive care glycemia), G.0120.03 (QIT), G.0452.04 (new quantum algorithms), G.0499.04 (Statistics), G.0211.05 (Nonlinear), research communities (ICCoS, ANMMM, MLDM); IWT: PhD Grants, GBOU (McKnow); Belgian Federal Science Policy Office: IUAP P5/22 ('Dynamical Systems and Control: Computation, Identification and Modelling', 2002-2006); PODO-II (CP/40: TMS and Sustainability); EU: FP5-Quprodis; ERNSI; Contract Research/agreements: ISMC/IPCOS, Data4s, TML, Elia, LMS, Mastercard

References

[AM79] B.D.O. Anderson and J.B. Moore. *Optimal filtering*. Prentice-Hall, 1979.

[Ber05] D.S. Bernstein. *Matrix Mathematics: Theory, Facts, and Formulas with Application to Linear Systems Theory*. Princeton University Press, Princeton, New Jersey, 2005.

[DZ97] M. Darouach and M. Zasadzinski. Unbiased minimum variance estimation for systems with unknown exogenous inputs. *Automatica*, 33(4):717–719, 1997.

[Fri69] B. Friedland. Treatment of bias in recursive filtering. *IEEE Trans. Autom. Control*, 14:359–367, 1969.

[HP98a] M. Hou and R.J. Patton. Input observability and input reconstruction. *Automatica*, 34(6):789–794, 1998.

[HP98b] M. Hou and R.J. Patton. Optimal filtering for systems with unknown inputs. *IEEE Trans. Autom. Control*, 43(3):445–449, 1998.

[Hsi00] C.S. Hsieh. Robust two-stage Kalman filters for systems with unknown inputs. *IEEE Trans. Autom. Control*, 45(12):2374–2378, 2000.

[Kit87] P.K. Kitanidis. Unbiased minimum-variance linear state estimation. *Automatica*, 23(6):775–778, 1987.

[KP00] W.S. Kerwin and J.L. Prince. On the optimality of recursive unbiased state estimation with unknown inputs. *Automatica*, 36:1381–1383, 2000.

[KSH00] T. Kailath, A.H. Sayed, and B. Hassibi. *Linear Estimation*. Prentice Hall, Upper Saddle River, New Jersey, 2000.

[KVR80] P. Kudva, N. Viswanadham, and A. Ramakrishna. Observers for linear systems with unknown inputs. *IEEE Trans. Autom. Control*, 25(1):113–115, 1980.

[XS03] Y. Xiong and M. Saif. Unknown disturbance inputs estimation based on a state functional observer design. *Automatica*, 39:1389–1398, 2003.