On the Optimality of the Kitanidis Filter for State Estimation Rejecting Unknown Inputs

Bernard Delyon, Qinghua Zhang

On the Optimality of the Kitanidis Filter for State Estimation Rejecting Unknown Inputs

Bernard Delyon [a], Qinghua Zhang [b],

[a] *Univ Rennes, CNRS, IRMAR - UMR 6625, 35000 Rennes, France*

[b] *Univ Gustave Eiffel, Inria, Cosys-SII, I4S, 35042 Rennes, France*

Abstract

As a natural extension of the Kalman filter to systems subject to arbitrary unknown inputs, the Kitanidis filter has been designed by one-step minimization of the trace of the state estimation error covariance matrix. In this technical communiqué, it is shown that the Kitanidis filter is also optimal for the whole gain sequence in the sense of matrix positive definiteness, which notably implies that the Kitanidis filter minimizes not only the trace criterion, but also the matrix spectral norm criterion.

Key words: optimal state estimation, unknown input observer, Kalman filter, disturbance rejection, time varying system.

1 Introduction

The Kitanidis filter is a natural extension of the Kalman filter to stochastic linear *time varying* (LTV) systems subject to *unknown inputs* in the form of

$$x_{k+1} = A_k x_k + B_k u_k + E_k d_k + w_k \qquad (1a)$$
$$y_k = C_k x_k + v_k, \qquad (1b)$$

where $x_k \in \mathbb{R}^n$ is the state, $y_k \in \mathbb{R}^m$ the output, $u_k \in \mathbb{R}^p$ the (known) input, $d_k \in \mathbb{R}^q$ the *unknown input*, $w_k \in \mathbb{R}^n$ the state noise, $v_k \in \mathbb{R}^m$ the output noise, and A_k, B_k, C_k, E_k are matrices of appropriate sizes at each discrete time instant $k = 0, 1, 2, \dots$.

The *unknown input* (or unknown disturbance) d_k is a *totally arbitrary and unknown* vector sequence.

State estimation while rejecting unknown inputs is usually called the *unknown input observer problem* (Yang and Wilde, 1988; Darouach et al., 1994; Chen and Patton, 1999). In the stochastic framework formulated in (1), the optimal state estimation, in the sense of an error covariance criterion subject to the unbiasedness constraint, is given by the Kitanidis filter (Kitanidis, 1987). Such results are useful for robust prediction (Kitanidis, 1987), for robust control (Ioannou and Sun, 1996), and for fault diagnosis (Chen and Patton, 1999). The Kitanidis filter (Kitanidis, 1987) has been designed

as follows. Among all recursive linear filters of the form

$$\begin{aligned} \hat{x}_{k+1} = {} & A_k \hat{x}_k + B_k u_k \\ & + L_{k+1}(y_{k+1} - C_{k+1} A_k \hat{x}_k - C_{k+1} B_k u_k), \end{aligned} \quad (2)$$

with the state estimate $\hat{x}_k \in \mathbb{R}^n$, the filter gain matrix $L_{k+1} \in \mathbb{R}^{n \times m}$, and the filter error

$$\tilde{x}_k \triangleq x_k - \hat{x}_k, \qquad (3)$$

the Kitanidis filter, characterized by a particular gain matrix sequence

$$L_{1:(k+1)}^* = (L_1^*, L_2^*, \dots, L_{k+1}^*), \qquad (4)$$

is the unbiased minimum variance filter, in the sense that, at instant $k + 1$, L_{k+1}^* is determined by solving the optimization problem

$$L_{k+1}^* = \arg \min_{L_{k+1}} \mathsf{Trace}\, \mathsf{Cov}(\tilde{x}_{k+1} | L_{1:k}^*, L_{k+1}) \qquad (5)$$

subject to the unbiaisedness constraint

$$\mathsf{E}(\tilde{x}_{k+1} | L_{1:(k+1)}^*) = 0 \qquad (6)$$

for *any unknown* input sequence

$$d_{0:k} = (d_0, d_1, d_2, \dots, d_k), \qquad (7)$$

where the dependence of the filter error $\tilde{x}_{k+1}$ on the gain sequence is indicated in the notations of *error covariance* $\mathsf{Cov}(\tilde{x}_{k+1} | \cdot)$ and *error mean* $\mathsf{E}(\tilde{x}_{k+1} | \cdot)$.

_____

Email addresses: `bernard.delyon@univ-rennes1.fr` (Bernard Delyon), `qinghua.zhang@inria.fr` (Qinghua Zhang).

The covariance trace minimization (5) is a *one-step* optimization, in the sense that the trace of $\mathsf{Cov}(\tilde{x}_{k+1})$ is minimized among the filters which were, up to instant k, defined with the previously optimized gain sequence $L_{1:k}^*$. In other words, at instant $k + 1$, the optimization is focused on the current gain matrix L_{k+1} only, without revising the past gain sequence $L_{1:k}^*$. It was later confirmed in (Kerwin and Prince, 2000) that the trace of the error covariance is indeed minimized by the Kitanidis filter among all unbiased linear causal filters.

The purpose of this Technical communiqué is to improve the existing results with a stronger optimality in the sense of matrix positive definiteness. The new result will imply that, for *any* vector $\eta \in \mathbb{R}^n$, the variance of the linear form $\eta^T \tilde{x}_{k+1}$ is minimized. This optimality is not specific to a particular vector η, but holds simultaneously for all $\eta \in \mathbb{R}^n$. In this sense, the Kitanidis filter provides the optimal estimation for *any* linear combination of states, including each individual state component. It will also be shown that the Kitanidis filter minimizes the spectral norm of the error covariance. More details will be given towards the end of Section 2, as corollaries of the main result.

The Kitanidis filter is optimal among the class of unbiased filters satisfying the unbiasedness constraint for *any* unknown input sequence. The error dynamics of such unbiased filters is totally decoupled from unknown inputs. Besides such unbiased filters, there exist filters incorporating different unknown input estimators. See (Hsieh and Majidi, 2014; Majidi et al., 2018) for some examples. Such filters can outperform the Kitanidis filter if the implemented unknown input estimator is appropriate to the actual unknown input. See also (Kong and Sukkarieh, 2019) for a discussion about different filters. It is well known that the Kalman filter has a least squares interpretation (Gillijns et al., 2007; Rawlings et al., 2017). It is also possible to derive the Kitanidis filter following such an approach.

2 New optimality results

In this paper, For a real symmetric matrix M, the inequality $M > 0$ ($M \geq 0$) means that M is (semi)-positive definite. For two real symmetric matrices M, N of the same size, $M > N$ means $M - N > 0$, and $M \geq N$ means $M - N \geq 0$. For a real matrix M, $\|M\|$ denotes its spectral norm, *i.e.*, the matrix norm induced by the Euclidean vector norm, which is equal to the largest singular value of the matrix. I_n is the $n \times n$ identity matrix.

Assumptions

(i) The initial state $x_0 \in \mathbb{R}^n$ is a random vector of mean $\bar{x}_0 \in \mathbb{R}^n$ and of covariance $P_0 > 0$.
(ii) The noises w_k and v_k are white, of zero mean, independent of each other and of x_0, and their covariance matrices $\mathsf{E}(w_k w_k^T) = Q_k$, $\mathsf{E}(v_k v_k^T) = R_k$, with $R_k > 0$, for all $k \geq 0$.
(iii) The matrix product $C_{k+1} E_k$ has a full column rank, for all $k \geq 0$.

Assumptions (i) and (ii) are like in the classical Kalman filter theory, whereas Assumption (iii) ensures the invertibility of some matrices involved in the Kitanidis filter.

The Kitanidis filter (Kitanidis, 1987) is given by

$$\hat{x}_{k+1} = A_k \hat{x}_k + B_k u_k \\ + L_{k+1}^*(y_{k+1} - C_{k+1} A_k \hat{x}_k - C_{k+1} B_k u_k) \quad (8)$$

with the filter gain L_{k+1}^* recursively computed as

$$P_{k+1|k} = A_k P_{k|k} A_k^T + Q_k \quad (9a)$$

$$\Sigma_{k+1} = C_{k+1} P_{k+1|k} C_{k+1}^T + R_{k+1} \quad (9b)$$

$$\Gamma_{k+1} = E_k - P_{k+1|k} C_{k+1}^T \Sigma_{k+1}^{-1} C_{k+1} E_k \quad (9c)$$

$$\Xi_{k+1} = E_k^T C_{k+1}^T \Sigma_{k+1}^{-1} C_{k+1} E_k \quad (9d)$$

$$P_{k+1|k+1} = P_{k+1|k} - P_{k+1|k} C_{k+1}^T \Sigma_{k+1}^{-1} C_{k+1} P_{k+1|k} \\ + \Gamma_{k+1} \Xi_{k+1}^{-1} \Gamma_{k+1}^T \quad (9e)$$

$$L_{k+1}^* = P_{k+1|k} C_{k+1}^T \Sigma_{k+1}^{-1} \\ + \Gamma_{k+1} \Xi_{k+1}^{-1} E_k^T C_{k+1}^T \Sigma_{k+1}^{-1}. \quad (9f)$$

At the initial instant $k = 0$, this filter is initialized as $\hat{x}_0 = \bar{x}_0$, $P_{0|0} = P_0$, with $\bar{x}_0 \in \mathbb{R}^n$ and $P_0 \in \mathbb{R}^{n \times n}$ being respectively the mean and the covariance matrix of the initial state x_0.

The new optimality result of the Kitanidis filter in the sense of the whole gain sequence is stated as follows.

Theorem 1 *The Kitanidis gain sequence $L_{1:(k+1)}^*$ as computed in (9) is optimal among all gain sequences $L_{1:(k+1)}$ leading to an unbiased recursive linear filter (2), in the sense that*

$$\mathsf{Cov}(\tilde{x}_{k+1}|L_{1:(k+1)}) \geq \mathsf{Cov}(\tilde{x}_{k+1}|L_{1:(k+1)}^*), \quad (10)$$

for any gain sequence $L_{1:(k+1)}$ such that

$$\mathsf{E}(\tilde{x}_{k+1}|L_{1:(k+1)}) = 0 \quad (11)$$

for any unknown input sequence $d_{0:k}$. □

The inequality in (10) is in the sense of matrix positive definiteness (see the notice at the beginning of this section).

The proof of this theorem will be based on the following lemma.

Lemma 1 *Given integers q, m, n such that $0 < q \leq m \leq n$ and matrices $\Psi \in \mathbb{R}^{m \times n}$, $\Sigma \in \mathbb{R}^{m \times m}$, $\Omega \in \mathbb{R}^{m \times q}$, $E \in \mathbb{R}^{n \times q}$ such that Σ is symmetric positive definite and Ω has a full column rank. Then any matrix $L \in \mathbb{R}^{n \times m}$ satisfying the constraint*

$$L\Omega = E \quad (12)$$

complies with the inequality

$$L\Sigma L^T - L\Psi - \Psi^T L^T \geq L_* \Sigma L_*^T - L_* \Psi - \Psi^T L_*^T \quad (13)$$

with

$$L_* = \Psi^T \Sigma^{-1} + (E - \Psi^T \Sigma^{-1} \Omega)(\Omega^T \Sigma^{-1} \Omega)^{-1} \Omega^T \Sigma^{-1} \tag{14}$$

which, like L, satisfies the constraint

$$L_* \Omega = E. \tag{15}$$

Proof of Lemma 1.
It is trivial to check that L_* satisfies (15).
Some simple computations lead to

$$(L - L_*)\Sigma(L - L_*)^T = L\Sigma L^T + L_* \Sigma L_*^T - L_* \Sigma L^T - L\Sigma L_*^T$$

and

$$L_* \Sigma L^T = \left(\Psi^T + (E - \Psi^T \Sigma^{-1} \Omega)(\Omega^T \Sigma^{-1} \Omega)^{-1} \Omega^T\right) L^T$$
$$= \Psi^T L^T + (E - \Psi^T \Sigma^{-1} \Omega)(\Omega^T \Sigma^{-1} \Omega)^{-1} E^T.$$

Then

$$(L - L_*)\Sigma(L - L_*)^T = L\Sigma L^T - L\Psi - \Psi^T L^T + \Theta$$

where Θ contains terms independent of L. The matrix Σ is assumed positive definite, then

$$L\Sigma L^T - L\Psi - \Psi^T L^T = (L - L_*)\Sigma(L - L_*)^T - \Theta$$

reaches its minimum $-\Theta$ when $L = L_*$. $\qquad\square$

Proof of Theorem 1.
Let us first show that the unbiasedness constraint (11) holds regardless of the arbitrary unknown input sequence $d_{0:k}$, *if and only if* the gain sequence $L_{1:(k+1)}$ satisfies

$$(I_n - L_{j+1} C_{j+1})E_j = 0 \tag{16}$$

for all $j = 0, 1, \ldots, k$ and for all integer $k \geq 0$.
For the linear filter (2) with any gain matrix $L_k \in \mathbb{R}^{n \times m}$, it is straightforward to check from (1) and (2) that the filter error $\tilde{x}_k$, as defined in (3), satisfies

$$\tilde{x}_{k+1} = (I_n - L_{k+1} C_{k+1})A_k \tilde{x}_k + (I_n - L_{k+1} C_{k+1})E_k d_k$$
$$+ (I_n - L_{k+1} C_{k+1})w_k - L_{k+1} v_{k+1}. \tag{17}$$

If $L_{1:(k+1)}$ satisfies (16), then the term involving d_k disappears from (17), yielding

$$\tilde{x}_{k+1} = (I_n - L_{k+1} C_{k+1})A_k \tilde{x}_k$$
$$+ (I_n - L_{k+1} C_{k+1})w_k - L_{k+1} v_{k+1}. \tag{18}$$

According to Assumptions (i) and (ii), $\mathsf{E}(w_k) = 0$, $\mathsf{E}(v_{k+1}) = 0$ and $\mathsf{E}(\tilde{x}_0) = \mathsf{E}(x_0 - \hat{x}_0) = 0$. It is then recursively shown that $\mathsf{E}(\tilde{x}_{k+1}|L_{1:(k+1)}) = 0$ for all $k \geq 0$. Hence (16) is a sufficient condition for (11).

On the other hand, if (16) is not satisfied, a sequence $d_{0:k}$ such that $\mathsf{E}(\tilde{x}_{k+1}|L_{1:(k+1)}) \neq 0$ can be built as follows. Let $d_j = 0$ for $j = 0, 1, \ldots, (k-1)$, and choose d_k such that $(I_n - L_{k+1} C_{k+1})E_k d_k \neq 0$ (such a d_k exists because in this considered case (16) is not satisfied). It then recursively follows from (17) that $\mathsf{E}(\tilde{x}_{k+1}|L_{1:(k+1)}) = (I_n - L_{k+1} C_{k+1})E_k d_k \neq 0$.
It is then proved that (16) is a *necessary and sufficient* condition for $L_{1:(k+1)}$ to satisfy the unbiasedness constraint (11).
Therefore, in what follows, (16) will replace (11) as a constraint on the gain sequence $L_{1:(k+1)}$.
Notice that $\tilde{x}_k$, w_k and v_{k+1} are pairwise independent, then taking the mathematical expectations of the squares of both sides of (18) yields

$$\mathsf{Cov}(\tilde{x}_{k+1}|L_{1:(k+1)})$$
$$= (I_n - L_{k+1} C_{k+1})A_k \mathsf{Cov}(\tilde{x}_k|L_{1:k})A_k^T (I_n - L_{k+1} C_{k+1})^T$$
$$+ (I_n - L_{k+1} C_{k+1})Q_k(I_n - L_{k+1} C_{k+1})^T$$
$$+ L_{k+1} R_{k+1} L_{k+1}^T. \tag{19}$$

The remaining part of the proof will be made by induction, first for $k = 0$, then recursively for any $k > 0$.
Equality (19) holds for all integer $k \geq 1$ and any gain sequence satisfying (16) with $j \leq k$. The case with $k = 0$ is slightly different, because the initial error $\tilde{x}_0 = x_0 - \hat{x}_0$ does not depend on any gain matrix, then $\mathsf{Cov}(\tilde{x}_k|L_{1:k})$ becomes simply $\mathsf{Cov}(\tilde{x}_0) = P_{0|0}$, and (19) becomes

$$\mathsf{Cov}(\tilde{x}_1|L_1) = (I_n - L_1 C_1)A_0 P_{0|0} A_0^T (I_n - L_1 C_1)^T$$
$$+ (I_n - L_1 C_1)Q_0(I_n - L_1 C_1)^T + L_1 R_1 L_1^T. \tag{20}$$
$$= L_1 \left[C_1 \left(A_0 P_{0|0} A_0^T + Q_0\right) C_1^T + R_1\right] L_1^T$$
$$- L_1 C_1 \left(A_0 P_{0|0} A_0^T + Q_0\right) - \left(A_0 P_{0|0} A_0^T + Q_0\right) C_1^T L_1^T$$
$$+ A_0 P_{0|0} A_0^T + Q_0. \tag{21}$$

It is already shown that $L_{1:(k+1)}$ satisfying (11) satisfies also (16). Apply Lemma 1 to (21) excluding the last two terms independent of L_1, with

$$L = L_1$$
$$\Omega = C_1 E_0$$
$$E = E_0$$
$$\Sigma = C_1 \left(A_0 P_{0|0} A_0^T + Q_0\right) C_1^T + R_1$$
$$\Psi = C_1 \left(A_0 P_{0|0} A_0^T + Q_0\right).$$

The constraint in (12) corresponds to (16) with $j = 0$. The positive definiteness of Σ is due to Assumption (ii). The full column rank of Ω is ensured by Assumption (iii). The conditions of Lemma 1 are thus fulfilled. Then it is straightforward (though tedious) to check that the matrix L_* resulting from the application of (14) in Lemma 1 coincides with the optimal gain L_1^* computed in (9).

According to Lemma 1, under constraint (16) with $j = 0$, the covariance matrix $\mathsf{Cov}(\tilde{x}_1|L_1)$ reaches its minimum when $L_1 = L_1^*$, hence inequality (10) is proved for the initial instant $k = 0$.

Now consider any integer $k > 0$.

<u>For the proof by induction</u>, assume, for any $j \le k$,

$$\mathsf{Cov}(\tilde{x}_j|L_{1:j}) \ge \mathsf{Cov}(\tilde{x}_j|L_{1:j}^*). \tag{22}$$

In particular, for $j = k$, it amounts to

$$\mathsf{Cov}(\tilde{x}_j|L_{1:k}) \ge \mathsf{Cov}(\tilde{x}_k|L_{1:k}^*). \tag{23}$$

Combine this inequality with equality (19), then

$$\mathsf{Cov}(\tilde{x}_{k+1}|L_{1:(k+1)})$$
$$\ge (I_n - L_{k+1}C_{k+1})A_k\mathsf{Cov}(\tilde{x}_k|L_{1:k}^*)A_k^T(I_n - L_{k+1}C_{k+1})^T$$
$$+ (I_n - L_{k+1}C_{k+1})Q_k(I_n - L_{k+1}C_{k+1})^T$$
$$+ L_{k+1}R_{k+1}L_{k+1}^T. \tag{24}$$
$$= L_{k+1}\Big[C_{k+1}\left(A_k\mathsf{Cov}(\tilde{x}_k|L_{1:k}^*)A_k^T + Q_k\right)C_{k+1}^T$$
$$\qquad + R_{k+1}\Big]L_{k+1}^T$$
$$- L_{k+1}C_{k+1}\left(A_k\mathsf{Cov}(\tilde{x}_k|L_{1:k}^*)A_k^T + Q_k\right)$$
$$- \left(A_k\mathsf{Cov}(\tilde{x}_k|L_{1:k}^*)A_k^T + Q_k\right)C_{k+1}^T L_{k+1}^T$$
$$+ A_k\mathsf{Cov}(\tilde{x}_k|L_{1:k}^*)A_k^T + Q_k. \tag{25}$$

It is already shown that $L_{1:(k+1)}$ satisfying (11) satisfies also (16). Apply Lemma 1 to (25) except the last two terms independent of L_{k+1}, with

$$
\begin{aligned}
L &= L_{k+1}\\
\Omega &= C_{k+1}E_k\\
E &= E_k\\
\Sigma &= C_{k+1}\left(A_k\mathsf{Cov}(\tilde{x}_k|L_{1:k}^*)A_k^T + Q_k\right)C_{k+1}^T + R_{k+1}\\
\Psi &= C_{k+1}\left(A_k\mathsf{Cov}(\tilde{x}_k|L_{1:k}^*)A_k^T + Q_k\right).
\end{aligned}
$$

The constraint in (12) corresponds to (16) with $j = k$. The positive definiteness of Σ is due to Assumption (ii). The full column rank of Ω is ensured by Assumption (iii). The conditions of Lemma 1 are thus fulfilled. Moreover, it is straightforward to check that the matrix L_* resulting from the application of (14) coincides with the optimal gain L_{k+1}^* computed in (9) where $P_{k|k} = \mathsf{Cov}(\tilde{x}_k|L_{1:k}^*)$. By applying Lemma 1, under constraint (16), the ex-

pression in (25) satisfies

$$L_{k+1}\Big[C_{k+1}\left(A_k\mathsf{Cov}(\tilde{x}_k|L_{1:k}^*)A_k^T + Q_k\right)C_{k+1}^T$$
$$\qquad + R_{k+1}\Big]L_{k+1}^T$$
$$- L_{k+1}C_{k+1}\left(A_k\mathsf{Cov}(\tilde{x}_k|L_{1:k}^*)A_k^T + Q_k\right)$$
$$- \left(A_k\mathsf{Cov}(\tilde{x}_k|L_{1:k}^*)A_k^T + Q_k\right)C_{k+1}^T L_{k+1}^T$$
$$+ A_k\mathsf{Cov}(\tilde{x}_k|L_{1:k}^*)A_k^T + Q_k$$
$$\ge L_{k+1}^*\Big[C_{k+1}\left(A_k\mathsf{Cov}(\tilde{x}_k|L_{1:k}^*)A_k^T + Q_k\right)C_{k+1}^T$$
$$\qquad + R_{k+1}\Big]L_{k+1}^{*T}$$
$$- L_{k+1}^*C_{k+1}\left(A_k\mathsf{Cov}(\tilde{x}_k|L_{1:k}^*)A_k^T + Q_k\right)$$
$$- \left(A_k\mathsf{Cov}(\tilde{x}_k|L_{1:k}^*)A_k^T + Q_k\right)C_{k+1}^T L_{k+1}^{*T}$$
$$+ A_k\mathsf{Cov}(\tilde{x}_k|L_{1:k}^*)A_k^T + Q_k \tag{26}$$
$$= \mathsf{Cov}(\tilde{x}_{k+1}|L_{1:(k+1)}^*), \tag{27}$$

where the last equality is a particular case of (19), which holds for any gain sequence satisfying (16), including $L_{1:(k+1)}^*$.

Combining (24), (25), (26) and (27) then yields

$$\mathsf{Cov}(\tilde{x}_{k+1}|L_{1:(k+1)}) \ge \mathsf{Cov}(\tilde{x}_{k+1}|L_{1:(k+1)}^*). \tag{28}$$

By induction it is then concluded that (10) holds for all integer $k \ge 0$. □

The following corollaries help to interpret this main result.

Corollary 1 *For any vector $\eta \in \mathbb{R}^n$, the linear form $\eta^T x_{k+1}$ of the state vector x_{k+1} is optimally estimated by the Kitanidis filter in the sense that*

$$\mathsf{E}((\eta^T\tilde{x}_{k+1})^2|L_{1:(k+1)}) \ge \mathsf{E}((\eta^T\tilde{x}_{k+1})^2|L_{1:(k+1)}^*) \tag{29}$$

for any gain sequence $L_{1:(k+1)}$ satisfying (11), and, in particular, for all $i = 1, 2, \ldots, n$,

$$\mathsf{E}((\tilde{x}_{k+1}^{(i)})^2|L_{1:(k+1)}) \ge \mathsf{E}((\tilde{x}_{k+1}^{(i)})^2|L_{1:(k+1)}^*), \tag{30}$$

where $\tilde{x}_{k+1}^{(i)}$ is the i-th component of the error vector $\tilde{x}_{k+1}$. □

In this result, the optimality is not specific to a particular vector η, but holds simultaneously for all $\eta \in \mathbb{R}^n$.

While the trace criterion like in (5) minimizes $\mathsf{E}[(\tilde{x}_{k+1}^{(1)})^2 + \cdots + (\tilde{x}_{k+1}^{(n)})^2]$, the inequality in (30) is a stronger result, by showing that *every* individual term $\mathsf{E}[(\tilde{x}_{k+1}^{(i)})^2]$ is minimized by the Kitanidis filter.

The estimation of linear combinations of states is usually addressed with *functional observers* (Bezzaoucha et al., 2017).

Proof of Corollary 1.
As $\tilde{x}_{k+1}$ has a zero mean value, its covariance

$$\mathsf{Cov}(\tilde{x}_{k+1}|L_{1:(k+1)}) = \mathsf{E}(\tilde{x}_{k+1}\tilde{x}_{k+1}^T|L_{1:(k+1)}).$$

For any vector $\eta \in \mathbb{R}^n$, inequality (10) implies

$$\eta^T\mathsf{Cov}(\tilde{x}_{k+1}|L_{1:(k+1)})\eta \geq \eta^T\mathsf{Cov}(\tilde{x}_{k+1}|L_{1:(k+1)}^*)\eta, \quad (31)$$

hence

$$\eta^T\mathsf{E}(\tilde{x}_{k+1}\tilde{x}_{k+1}^T|L_{1:(k+1)})\eta \geq \eta^T\mathsf{E}(\tilde{x}_{k+1}\tilde{x}_{k+1}^T|L_{1:(k+1)}^*)\eta,$$

which leads to (29) by noticing that $\eta^T\tilde{x}_{k+1}\tilde{x}_{k+1}^T\eta = (\eta^T\tilde{x}_{k+1})^2$. $\qquad\square$

By choosing η to be the eigenvector of $\mathsf{Cov}(\tilde{x}_{k+1}|L_{1:(k+1)}^*)$ associated to its largest eigenvalue, then (31) leads to the following result, where $\|\cdot\|$ denotes the matrix spectral norm.

Corollary 2 *The Kitanidis filter minimizes the spectral norm of the error covariance matrix, i.e., Theorem 1 remains true if* (10) *is replaced by*

$$\|\mathsf{Cov}(\tilde{x}_{k+1}|L_{1:(k+1)})\| \geq \|\mathsf{Cov}(\tilde{x}_{k+1}|L_{1:(k+1)}^*)\|.$$

3 Conclusion

It has been shown in this paper that, though the Kitanidis filter has been designed by one-step minimization of the trace criterion, it is indeed optimal for the whole gain sequence in the sense of matrix positive definiteness.

References

Bezzaoucha, S., Voos, H., and Darouach, M. (2017). A new polytopic approach for the unknown input functional observer design. *International Journal of Control*, 2017.

Chen, J. and Patton, R. (1999). *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Kluwer, Boston.

Darouach, M., Zasadzinski, M., and Xu., S.-J. (1994). Full-order observers for linear systems with unknown inputs. *IEEE Transactions on Automatic Control*, 39(3):606–609.

Gillijns, S., Haverbeke, N., and De Moor, B. (2007). Information, covariance and square-root filtering in the presence of unknown inputs. In *European Control Conference (ECC)*, pages 2213–2217.

Hsieh, C.-S. and Majidi, M. A. (2014). Implementation issues of unbiased minimum-variance state estimation for systems with unknown inputs. In *CACS International Automatic Control Conference*, pages 323–328.

Ioannou, P. and Sun, J. (1996). *Robust Adaptive Control*. Prentice Hall.

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

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

Kong, H. and Sukkarieh, S. (2019). An internal model approach to estimation of systems with arbitrary unknown inputs. *Automatica*, 108:1–11.

Majidi, M. A., Hsieh, C.-S., and Yazdi, H. S. (2018). Kalman filter reinforced by least mean square for systems with unknown inputs. *Circuits, Systems, and Signal Processing*, 37:4955–4972.

Rawlings, J. B., Mayne, D. Q., and Diehl, M. M. (2017). *Model Predictive Control: Theory, Computation, and Design.* Nob Hill Publishing, 2nd edition.

Yang, F. and Wilde, R. (1988). Observers for linear systems with unknown inputs. *IEEE Transactions on Automatic Control*, 33(7):677–681.