On the Optimality of the Kitanidis Filter

Bernard Delyon, Qinghua Zhang

On the Optimality of the Kitanidis Filter

Bernard Delyon* and Qinghua Zhang**

* IRMAR, Univ. Rennes, 35042 Rennes Cedex. bernard.delyon@univ-rennes1.fr
** Univ. Gustave Eiffel, Inria, Cosys-SII, I4S, 35042 Rennes Cedex. qinghua.zhang@inria.fr

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. This optimality does not exclude the possibility that, among the class of unbiased recursive filters, another filter may lead to a lower trace criterion. In this paper, it is shown that the Kitanidis filter is indeed optimal in the sense of the whole gain sequence, thus excluding the aforementioned possibility. Moreover, this gain sequence optimality holds also in a stronger sense, in terms of positive definite matrix inequality, which notably implies that the optimality holds not only in the sense of the trace criterion, but also of the matrix spectral norm criterion.

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

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

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

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, 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

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

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

$$\tilde{x}_k \triangleq x_k - \hat{x}_k, \tag{3}$$

the Kitanidis filter, characterized by a particular gain matrix sequence

$$L^*_{1:(k+1)} = (L^*_1, L^*_2, \dots, L^*_{k+1}), \tag{4}$$

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

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

subject to the unbiaisedness constraint

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

for any unknown input sequence

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

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* $\mathsf{E}(\tilde{x}_{k+1}|\cdot)$.

It is important to remark that 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 $L^*_{1:k}$. This *one-step* optimization does not exclude the possibility that there *may* exist another gain sequence $L_{1:(k+1)}$ leading to a smaller trace of $\mathsf{Cov}(\tilde{x}_{k+1})$.

The purpose of this paper is to formally prove that the Kitanidis filter is not only *one-step optimal*, but also optimal in the sense of the *whole gain sequence* $L_{1:(k+1)}$. Moreover, the optimality holds also in a sense stronger than the trace criterion, in terms of matrix positive definiteness.

The one-step minimization of the trace criterion, like in (5), is also one of the known ways for deriving the classical *Kalman filter*. With white Gaussian noises, the Kalman filter has a much stronger interpretation: the conditioned probability distribution of the state given past observations. Such an interpretation of the Kitanidis filter is not possible, since the state in (1) is, in general, not Gaussian distributed, because of the totally arbitrary unknown input d_k. It is thus important to investigate the optimality of the Kitanidis filter with other interpretations, in addition to its one-step optimality.

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 $\mathrm{E}(w_k w_k^T) = Q_k$, $\mathrm{E}(v_k v_k^T) = R_k$, 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

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

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

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

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

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

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

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

$$\begin{aligned}L^*_{k+1} &= P_{k+1|k}C_{k+1}^T\Sigma_{k+1}^{-1} \\ &\quad + \Gamma_{k+1}\Xi_{k+1}^{-1}E_k^T C_{k+1}^T \Sigma_{k+1}^{-1}. \quad (9f)\end{aligned}$$

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 $P_0 \in \mathbb{R}^{n\times n}$ 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

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

which, like L, satisfies the constraint

$$L_*\Omega = E. \quad (15)$$

□

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

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

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 $L = L_*$. □

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

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

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

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

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

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

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, it is easy to build a sequence $d_{0:k}$ so that $\mathsf{E}(\tilde{x}_{k+1}|L_{1:(k+1)}) \neq 0$. 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

$$\begin{aligned}&\mathsf{Cov}(\tilde{x}_{k+1}|L_{1:(k+1)})\\ &\quad= (I_n - L_{k+1}C_{k+1})A_k\mathsf{Cov}(\tilde{x}_k|L_{1:k})A_k^T\\ &\qquad\cdot (I_n - L_{k+1}C_{k+1})^T\\ &\quad+ (I_n - L_{k+1}C_{k+1})Q_k(I_n - L_{k+1}C_{k+1})^T\\ &\quad+ L_{k+1}R_{k+1}L_{k+1}^T. \end{aligned} \quad (19)$$

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

$$\begin{aligned}\mathsf{Cov}(\tilde{x}_1|L_1) &= (I_n - L_1C_1)A_0P_{0|0}A_0^T(I_n - L_1C_1)^T\\ &\quad+ (I_n - L_1C_1)Q_0(I_n - L_1C_1)^T + L_1R_1L_1^T. \quad (20)\\ &= L_1\Big[C_1\left(A_0P_{0|0}A_0^T + Q_0\right)C_1^T + R_1\Big]L_1^T\\ &\quad- L_1C_1\left(A_0P_{0|0}A_0^T + Q_0\right)\\ &\quad- \left(A_0P_{0|0}A_0^T + Q_0\right)C_1^T L_1^T\\ &\quad+ A_0P_{0|0}A_0^T + Q_0. \end{aligned} \quad (21)$$

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

$$\begin{aligned}L &= L_1\\ \Omega &= C_1E_0\\ E &= E_0\\ \Sigma &= C_1\left(A_0P_{0|0}A_0^T + Q_0\right)C_1^T + R_1\\ \Psi &= C_1\left(A_0P_{0|0}A_0^T + Q_0\right).\end{aligned}$$

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 $k > 0$.

For the proof by induction, assume, for any $j \leq k$,

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

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

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

Combine this inequality with equality (19), then

$$\begin{aligned}
&\mathsf{Cov}(\tilde{x}_{k+1}|L_{1:(k+1)}) \\
&\quad \geq (I_n - L_{k+1}C_{k+1})A_k \mathsf{Cov}(\tilde{x}_k|L^*_{1:k})A_k^T \\
&\quad\quad \cdot (I_n - L_{k+1}C_{k+1})^T \\
&\quad + (I_n - L_{k+1}C_{k+1})Q_k(I_n - L_{k+1}C_{k+1})^T \\
&\quad + L_{k+1}R_{k+1}L^T_{k+1}. \qquad (24)\\
&= L_{k+1}\Big[C_{k+1}\left(A_k \mathsf{Cov}(\tilde{x}_k|L^*_{1:k})A_k^T + Q_k\right)C^T_{k+1} \\
&\qquad\qquad + R_{k+1}\Big]L^T_{k+1} \\
&\quad - L_{k+1}C_{k+1}\left(A_k \mathsf{Cov}(\tilde{x}_k|L^*_{1:k})A_k^T + Q_k\right) \\
&\quad - \left(A_k \mathsf{Cov}(\tilde{x}_k|L^*_{1:k})A_k^T + Q_k\right)C^T_{k+1}L^T_{k+1} \\
&\quad + A_k \mathsf{Cov}(\tilde{x}_k|L^*_{1:k})A_k^T + Q_k. \qquad (25)
\end{aligned}$$

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

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

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 expression in (25) satisfies

$$\begin{aligned}
&L_{k+1}\Big[C_{k+1}\left(A_k \mathsf{Cov}(\tilde{x}_k|L^*_{1:k})A_k^T + Q_k\right)C^T_{k+1} \\
&\qquad\qquad + R_{k+1}\Big]L^T_{k+1} \\
&\quad - L_{k+1}C_{k+1}\left(A_k \mathsf{Cov}(\tilde{x}_k|L^*_{1:k})A_k^T + Q_k\right) \\
&\quad - \left(A_k \mathsf{Cov}(\tilde{x}_k|L^*_{1:k})A_k^T + Q_k\right)C^T_{k+1}L^T_{k+1} \\
&\quad + A_k \mathsf{Cov}(\tilde{x}_k|L^*_{1:k})A_k^T + Q_k \\
&\geq L^*_{k+1}\Big[C_{k+1}\left(A_k \mathsf{Cov}(\tilde{x}_k|L^*_{1:k})A_k^T + Q_k\right)C^T_{k+1} \\
&\qquad\qquad + R_{k+1}\Big]L^{*T}_{k+1} \\
&\quad - L^*_{k+1}C_{k+1}\left(A_k \mathsf{Cov}(\tilde{x}_k|L^*_{1:k})A_k^T + Q_k\right) \\
&\quad - \left(A_k \mathsf{Cov}(\tilde{x}_k|L^*_{1:k})A_k^T + Q_k\right)C^T_{k+1}L^{*T}_{k+1} \\
&\quad + A_k \mathsf{Cov}(\tilde{x}_k|L^*_{1:k})A_k^T + Q_k \qquad (26)\\
&= \mathsf{Cov}(\tilde{x}_{k+1}|L^*_{1:(k+1)}), \qquad (27)
\end{aligned}$$

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

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

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

Compared to the optimality initially established in (Kitanidis, 1987), this new result is stronger, not only for its optimality in the sense of the whole gain sequence, but also due to the matrix inequality (10) instead of trace minimization. For two symmetric matrices $M \in \mathbb{R}^{n\times n}$ and $N \in \mathbb{R}^{n\times n}$, $M \geq N$ implies that $v^TMv \geq v^TNv$ for any vector $v \in \mathbb{R}^n$. In particular, by filling v with zeros except a single component with one, then $v^TMv \geq v^TNv$ means that each diagonal entry of M is larger than or equal to the corresponding diagonal entry of N. This relationship is much stronger than the trace inequality $\mathsf{Trace}(M) \geq \mathsf{Trace}(N)$. It is also known that $M \geq N$ implies $\|M\| \geq \|N\|$ (for the matrix spectral norm). These facts then lead to the following corollary.

Corollary 1 *The trace optimality of the Kitanidis filter holds in the sense of the whole gain sequence, i.e., Theorem 1 remains true if the matrix inequality* (10) *is replaced by*

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

Moreover, the Kitanidis filter minimizes also 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 symmetric matrix inequality, notably implying the optimalities in terms of the state estimation error covariance matrix trace and spectral norm.

References

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.

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

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

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