On Stability of the Kalman Filter for Discrete Time Output Error Systems

Qinghua Zhang

On Stability of the Kalman Filter for Discrete Time Output Error Systems

Qinghua Zhang[a]

[a]*Inria, Campus de Beaulieu, 35042 Rennes Cedex, France (email: qinghua.zhang@inria.fr)*

Abstract

The stability of the Kalman filter is classically ensured by the uniform complete controllability regarding the process noise and the uniform complete observability of linear time varying systems. This paper studies the case of discrete time output error (OE) systems, in which the process noise is totally absent. The classical stability analysis assuming the controllability regarding the process noise is thus not applicable. It is shown in this paper that the uniform complete observability is sufficient to ensure the stability of the Kalman filter applied to time varying OE systems, regardless of the stability of the OE systems. Though the continuous time case has been studied recently, the results on continuous time systems cannot be directly transposed to discrete time systems, because of a difficulty related to the observability of the discrete time filter error dynamics system.

Keywords: Kalman filter, discrete time output error system, time varying system, stability.

1. Introduction

The well known Kalman filter has been extensively studied and is being applied in many different fields (Anderson and Moore, 1979; Jazwinski, 1970; Zarchan and Musof, 2005; Kim, 2011; Grewal and Andrews, 2015). The purpose of the present paper is to study the stability of the Kalman filter in a particular case not yet covered in the literature: the absence of process noise in the state equation of a *discrete time* linear time varying (LTV) system. Such systems are known as *output error* (OE) systems in the literature on system identification. The recent studies on *continuous time* OE systems in (Ni and Zhang, 2015, 2016) have been mainly motivated by applications where state equations originate from physical laws that are believed sufficiently accurate. For *discrete time* systems considered in this paper, the motivation is mainly for OE system identification (Goodman and Dudley, 1987; Forssell and Ljung, 2000; Wang et al., 2015). In control applications, the use of OE models has the advantage of focusing system identification

on the dynamics of the controlled plant, rather than on noise properties (Forssell and Ljung, 2000). The result presented in this paper ensures the stability of the Kalman filter applied to LTV OE systems. This result is particularly useful for linear parameter varying (LPV) system identification based on prediction error minimization, as it ensures stable predictions, regardless of the stability of the estimated LPV models during the iterations of prediction error minimization.

While the optimal properties of the Kalman filter are frequently recalled, its stability properties are less often mentioned in the recent literature. The classical stability analysis is based on both the uniform complete controllability *regarding the process noise* and the uniform complete observability of LTV systems (Kalman, 1963; Jazwinski, 1970). In the case of OE systems, there is no process noise at all in the state equation, hence the controllability regarding the process noise cannot be fulfilled.

The stability of the Kalman filter for *continuous time* OE systems has been recently studied in (Ni and Zhang, 2015, 2016). It is often straightforward to transpose theoretic results from continuous time systems to discrete time systems, and vice versa, but there are exceptions. For the stability problem studied in this paper, there are two main extra difficulties for discrete time systems.

First, in the continuous time case, the observability of an OE system induces the observability of its Kalman filter error dynamics system, and this observability plays an important role in the stability analysis of the error dynamics. In the discrete time case, however, the observability of the Kalman filter error dynamics system cannot be induced in a similar way, as explained in Section 5.

Second, in the continuous time case, in the first step of the proof of the Kalman filter asymptotic stability, the Lyapunov stability is naturally proved within a few lines in (Ni and Zhang, 2016). In the discrete time case, however, by indirectly analyzing the Kalman filter error dynamics, due to the complexity related to the separation between the prediction step and the update step (no such separation exists in the continuous time case), the proof of the Lyapunov stability takes more than one page (see the proof of Theorem 1 in this paper), with non trivial choices of appropriate equalities involved in the discrete time case only.

The classical *optimality* results of the Kalman filter are also valid in the case of OE systems (Jazwinski, 1970, chapter 7). However, it is necessary to complete the stability analysis, as the classical results are not applicable in this case.

As the *main contribution* of this paper, it will be shown that the uniform complete observability is sufficient to guarantee the stability of the dynamics of the Kalman filter applied to a discrete time LTV OE system, *regardless of the stability of the OE system itself*. The boundedness of the state estimate covariance, as well as the boundedness of the Kalman gain, will also be proved under the same condition. These results complete the classical results (Kalman,

1963; Jazwinski, 1970), which do not cover the case of OE systems.

Preliminary results of this study have been submitted to the upcoming IFAC Word Congress (Zhang, 2017). The present paper enriches these preliminary results with technical details, notably the relationship between the Kalman filter error dynamics and the Kalman predictor error dynamics, and also with numerical examples.

The rest of this paper is organized as follows. Some preliminary elements are introduced in Section 2. The problem considered in this paper is formulated in Section 3. The boundedness of the Kalman filter for OE systems is analyzed in Section 4, and the asymptotic stability of the Kalman filter is established in Section 5. Numerical examples are presented in Sections 6. Finally, concluding remarks are drawn in Section 7.

2. Definitions and basic facts

Let m and n be any two positive integers. For a vector $x \in \mathbb{R}^n$, $\|x\|$ denotes its Euclidean norm. For a matrix $A \in \mathbb{R}^{m \times n}$, $\|A\|$ denotes the matrix norm induced by the Euclidean vector norm, which is equal to the largest singular value of A. Then $\|Ax\| \leq \|A\|\|x\|$ for all $A \in \mathbb{R}^{m \times n}$ and all $x \in \mathbb{R}^n$. For two real square symmetric positive definite matrices A and B, $A > B$ means $A - B$ is positive definite.

Let $A(k) \in \mathbb{R}^{m \times n}$ be a sequence of matrices for $k = 0, 1, 2, \ldots$. It is said (upper) bounded if $\|A(k)\|$ is bounded.

Consider the homogeneous discrete time LTV system

$$x(k) = A(k)x(k–1) \tag{1}$$

with $x(k) \in \mathbb{R}^n$ and $A(k) \in \mathbb{R}^{n \times n}$, and with the associated state transition matrix defined as

$$\Phi(k, k) = I_n \tag{2}$$
$$\Phi(k, l) = A(k)A(k–1) \cdots A(l + 1) \tag{3}$$

with I_n denoting the $n \times n$ identity matrix. Then $x(k) = \Phi(k, l)x(l)$.

Definition 1. *System* (1) *is* Lyapunov stable *if there exists a positive constant γ such that, for all integers k, k_0 satisfying $k \geq k_0$, the following inequality holds*

$$\|\Phi(k, k_0)\| \leq \gamma. \tag{4}$$

$\square$

3

Definition 2. *System* (1) *is* asymptotically stable *if it is Lyapunov stable and if the following limiting behavior holds*

$$\lim_{k \to +\infty} \|x(k)\| = 0 \tag{5}$$

for any initial state $x(0) \in \mathbb{R}^n$. $\square$

The following uniform complete observability[1] definition for LTV systems follows (Kalman, 1963).

Definition 3. *The matrix pair $\{A(k), C(k)\}$ with $A(k) \in \mathbb{R}^{n \times n}$ and $C(k) \in \mathbb{R}^{m \times n}$ is* uniformly completely observable *if there exist positive constants ρ_1, ρ_2 and a positive integer h such that, for all $k \geq h$, the following inequalities hold*

$$\rho_1 I_n \leq \sum_{s=k-h}^{k} \Phi^T(s,k) C^T(s) R^{-1}(s) C(s) \Phi(s,k) \tag{6}$$

$$\leq \rho_2 I_n \tag{7}$$

with some bounded symmetric positive definite matrix $R(s) \in \mathbb{R}^{m \times m}$ (typically the covariance matrix of the output noise in a stochastic state space system). $\square$

3. Problem formulation and assumptions

In this section the considered OE system and its Kalman filter are first formulated, before the statement of the assumptions for stability analysis.

3.1. Output error system and Kalman filter

The discrete time output error (OE) systems considered in this paper are in the form of

$$x(k) = A(k)x(k{-}1) + B(k)u(k) \tag{8a}$$

$$y(k) = C(k)x(k) + R^{\frac{1}{2}}(k)v(k) \tag{8b}$$

where $k = 0, 1, 2, \ldots$ represents the discrete time index, $x(k) \in \mathbb{R}^n$ is the state vector, $u(k) \in \mathbb{R}^l$ the input, $y(k) \in \mathbb{R}^m$ the output, $v(k) \in \mathbb{R}^m$ a white Gaussian noise with identity covariance matrix, $A(k), B(k), C(k), R(k)$ are real matrices of appropriate sizes. The noise covariance matrix $R(k)$ is symmetric positive definite. The notation $R^{\frac{1}{2}}(k)$ denotes the symmetric positive definite matrix

---

[1]Some variants of the definition of the uniform complete observability exist in the literature. The definition recalled here follows (Kalman, 1963).

square root of $R(k)$. The initial state $x(0) \in \mathbb{R}^n$ is a random vector following the Gaussian distribution $x(0) \sim \mathcal{N}(x_0, P_0)$ with $x_0 \in \mathbb{R}^n$ and $P_0 \in \mathbb{R}^{n \times n}$.

Compared to the general discrete time state equation

$$x(k) = A(k)x(k{-}1) + B(k)u(k) + Q^{\frac{1}{2}}(k)w(k) \tag{9}$$

with the process noise $w(k)$ and the covariance matrix $Q(k)$, an OE system corresponds to the particular case of $Q(k) \equiv 0$.

After the initialization with $P(0|0) = P_0$ and $\hat{x}(0|0) = x_0$, the Kalman filter for the OE system (8) consists of the following recursions for $k = 1, 2, \ldots,$

$$P(k|k{-}1) = A(k)P(k{-}1|k{-}1)A^T(k) \tag{10a}$$

$$\Sigma(k) = C(k)P(k|k{-}1)C^T(k) + R(k) \tag{10b}$$

$$K(k) = P(k|k{-}1)C^T(k)\Sigma^{-1}(k) \tag{10c}$$

$$P(k|k) = [I_n - K(k)C(k)]P(k|k{-}1) \tag{10d}$$

$$\hat{x}(k|k{-}1) = A(k)\hat{x}(k{-}1|k{-}1) + B(k)u(k) \tag{10e}$$

$$\tilde{y}(k) = y(k) - C(k)\hat{x}(k|k{-}1) \tag{10f}$$

$$\hat{x}(k|k) = \hat{x}(k|k{-}1) + K(k)\tilde{y}(k). \tag{10g}$$

For general LTV systems with a process noise as in the state equation (9), the first equation of the Kalman filter would be

$$P(k|k{-}1) = A(k)P(k{-}1|k{-}1)A^T(k) + Q(k) \tag{11}$$

with an extra term $Q(k)$ representing the process noise covariance matrix. Therefore, the OE system Kalman filter corresponds to the particular case with $Q(k) \equiv 0$ of the general LTV system Kalman filter.

It is known that the dynamics of the general LTV system Kalman filter is stable, provided the matrix pair $\{A(k), Q^{\frac{1}{2}}(k)\}$ is uniformly completely controllable and the matrix pair $\{A(k), C(k)\}$ is uniformly completely observable (Kalman, 1963; Jazwinski, 1970). As OE systems correspond to the case with $Q(k) \equiv 0$, the controllability condition cannot be satisfied. Consequently, the classical results on the *stability* of the Kalman filter cannot be applied here. The main purpose of the present paper is to study the Kalman filter stability in this particular case.

It will be shown that the error dynamics of the Kalman filter (10) is asymptotically stable, and that its iteratively computed variables are all bounded.

Note that the classical *optimality* results of the Kalman filter remain valid in the case of OE systems (Jazwinski, 1970, chapter 7).

3.2. Assumptions

The assumptions stated here are required in *all the following sections* of this paper.

The considered OE system (8) is defined with bounded real matrices $A(k), B(k), C(k), R(k)$ of appropriate sizes. For all $k \geq 0$, the square matrix $A(k)$ is assumed invertible, so that the system dynamics is reversible, *i.e.*, the equality $x(k) = \Phi(k,l)x(l)$ holds (with an appropriate state transition matrix $\Phi(k,l)$) for all integers $k, l \geq 0$, including the case $k < l$. The noise covariance matrix $R(k)$ is symmetric positive definite and has a strictly positive lower bound.

The initial state $x(0) \in \mathbb{R}^n$ is a random vector following the Gaussian distribution

$$x(0) \sim \mathcal{N}(x_0, P_0) \tag{12}$$

with some mean vector $x_0 \in \mathbb{R}^n$ and a symmetric positive definite covariance matrix $P_0 \in \mathbb{R}^{n \times n}$.

It is further assumed that the matrix pair $\{A(k), C(k)\}$ is uniformly completely observable (see Definition 3).

4. Boundedness of the covariance and gain matrices

The purpose of this section is to show that the OE system Kalman filter covariance matrices $P(k|k)$, $P(k|k-1)$ and the Kalman gain $K(k)$ are all bounded under the assumptions stated in Section 3.2.

It will be shown that the covariance of the state estimate $P(k|k)$ is closely related to the matrix sequence $\Omega(k)$ recursively defined by

$$\Omega(0) = P^{-1}(0|0) \tag{13a}$$

$$\Omega(k) = \Phi^T(k{-}1, k)\Omega(k{-}1)\Phi(k{-}1, k) + \Lambda(k), \tag{13b}$$

where

$$\Lambda(k) \triangleq C^T(k)R^{-1}(k)C(k), \tag{14}$$

so that the properties of $P(k|k)$ can be analyzed through those of $\Omega(k)$.

Proposition 1. *Under the assumptions stated in Section 3.2, the matrix sequence $\Omega(k) \in \mathbb{R}^{n \times n}$ defined in (13) is symmetric positive definite for all $k \geq 0$ and has a strictly positive lower bound.* $\qquad\square$

Proof. Repeatedly applying the recursion (13b) yields

$$\Omega(k) = \Phi^T(0, k)\Omega(0)\Phi(0, k) + \sum_{s=1}^{k} \Phi^T(s, k)\Lambda(s)\Phi(s, k). \tag{15}$$

It is assumed that $A(k)$ is invertible for all $k \geq 0$, hence $\Phi(0, k)$ is also invertible. The initial condition $\Omega(0) = P^{-1}(0|0) > 0$ then implies that $\Omega(k) > 0$ for all $k \geq 0$.

When $k < h$ (h is the integer introduced in (6)),

$$\Omega(k) \geq \min_{0 \leq k < h} \Phi^T(0, k)\Omega(0)\Phi(0, k) \geq \rho_0 I_n \tag{16}$$

where $\rho_0 > 0$ is the smallest eigenvalue of $\Phi^T(0, k)\Omega(0)\Phi(0, k)$ for $0 \leq k < h$.

When $k \geq h$, because of the assumed uniforme complete observability (see inequality (6) in Definition 3) and the definition of $\Lambda(k)$ in equation (14)),

$$\Omega(k) \geq \sum_{s=k-h}^{k} \Phi^T(s, k)\Lambda(s)\Phi(s, k) \geq \rho_1 I_n > 0 \tag{17}$$

where ρ_1 is the constant in inequality (6).

Therefore, $\min(\rho_0, \rho_1)I_n > 0$ is a lower bound of $\Omega(k)$ for all $k \geq 0$. $\qquad \square$

Proposition 2. *Under the assumptions stated in Section 3.2, for all $k \geq 0$, the Kalman filter covariance matrix $P(k|k)$, iteratively computed through (10a)-(10d) with the initial condition $P(0|0) = P_0 > 0$, is symmetric positive definite and is upper bounded.* $\qquad \square$

Proof. Recall that $\Phi(k, k\text{–}1) = A(k)$ and $\Phi(k\text{–}1, k) = A^{-1}(k)$. Rewrite (13b) as

$$\Omega(k) = A^{-T}(k)\Omega(k\text{–}1)A^{-1}(k) + \Lambda(k). \tag{18}$$

Take the matrix inverses of both sides of (18) while applying the matrix equality

$$(G + V^T H V)^{-1} = G^{-1} - G^{-1}V^T(H^{-1} + VG^{-1}V^T)^{-1}VG^{-1}$$

with $G = A^{-T}(k)\Omega(k\text{–}1)A^{-1}(k)$, $H = R^{-1}(k)$ and $V = C(k)$, then

$$\begin{aligned}
\Omega^{-1}(k) = {} & A(k)\Omega^{-1}(k\text{–}1)A^T(k) \\
& - A(k)\Omega^{-1}(k\text{–}1)A^T(k)C^T(k)\big[R(k) \\
& + C(k)A(k)\Omega^{-1}(k\text{–}1)A^T(k)C^T(k)\big]^{-1} \\
& \cdot C(k)A(k)\Omega^{-1}(k\text{–}1)A^T(k)
\end{aligned} \tag{19}$$

Define the *new notations*

$$\bar{P}(k|k) \triangleq \Omega^{-1}(k) \tag{20}$$

$$\bar{P}(k|k\text{–}1) \triangleq A(k)\bar{P}(k\text{–}1|k\text{–}1)A^T(k) \tag{21}$$

$$\bar{\Sigma}(k) \triangleq C(k)\bar{P}(k|k\text{–}1)C^T(k) + R(k) \tag{22}$$

$$\bar{K}(k) \triangleq \bar{P}(k|k\text{–}1)C^T(k)\bar{\Sigma}^{-1}(k), \tag{23}$$

7

then the recursion of $\bar{P}(k|k) = \Omega^{-1}(k)$ in (19) is rewritten as

$$
\begin{aligned}
\bar{P}(k|k) &= \bar{P}(k|k{-}1) - \bar{P}(k|k{-}1)C^T(k)\big[R(k) \\
&\quad + C(k)\bar{P}(k|k{-}1)C^T(k)\big]^{-1} \\
&\quad \cdot C(k)\bar{P}(k|k{-}1) \\
&= \bar{P}(k|k{-}1) - \bar{P}(k|k{-}1)C^T(k)\bar{\Sigma}^{-1}(k) \\
&\quad \cdot C(k)\bar{P}(k|k{-}1),
\end{aligned}
$$

hence

$$
\bar{P}(k|k) = \big[I_n - \bar{K}(k)C(k)\big]\bar{P}(k|k{-}1). \tag{24}
$$

It turns out that, the recursions (21)-(24) coincide exactly with those of (10a)-(10d). Moreover, $\bar{P}(0|0) = \Omega^{-1}(0) = P(0|0)$, therefore, $\bar{P}(k|k) = \Omega^{-1}(k) = P(k|k)$ for all $k \geq 0$.

The results of Proposition 1 then imply that $P(k|k) = \Omega^{-1}(k)$ is positive definite and upper bounded. $\qquad\square$

Corollary 1. *Under the assumptions stated in Section 3.2, for all $k \geq 0$, the prediction error covariance matrix $P(k|k-1)$ computed with (10a) is symmetric positive definite and upper bounded, and the gain matrix $K(k)$ computed with (10c) is bounded.* $\qquad\square$

The proof of this corollary is trivial: based on Proposition 2 and on the assumption that $A(k)$ is bounded and invertible, the positive definiteness and the boundedness of $P(k|k-1)$ follow from (10a); while the boundedness of $K(k)$ is due to the boundedness of $P(k|k-1)$ and of $C(k)$ in (10c), and also to the assumed strictly positive lower bound of $R(k)$ in (10b) (see Section 3.2 for the assumptions).

5. Asymptotic stability of the Kalman filter

It has been shown in Proposition 2 and Corollary 1 that the covariance matrices $P(k|k)$, $P(k|k{-}1)$ and the Kalman gain $K(k)$ are all bounded. These are obviously important "stability" properties of a recursive algorithm. The purpose of this section is to study the asymptotic stability of the Kalman filter error dynamics, in the sense of Definition 2.

5.1. Kalman filter error dynamics

The computations in the steps (10e)-(10g) can be combined into a single step for state estimation:

$$
\begin{aligned}
\hat{x}(k|k) &= A(k)\hat{x}(k{-}1|k{-}1) + B(k)u(k) \\
&\quad + K(k)\big\{y(k) - C(k)\big[A(k)\hat{x}(k{-}1|k{-}1) + B(k)u(k)\big]\big\}.
\end{aligned} \tag{25}
$$

The homogeneous part of this *iterative* equation (from $\hat{x}(k{-}1|k{-}1)$ to $\hat{x}(k|k)$) is

$$z(k|k) = [I_n - K(k)C(k)]A(k)z(k{-}1|k{-}1), \qquad (26)$$

which corresponds to the dynamics of the mathematical expectation of the state estimation error, *i.e.*,

$$z(k|k) = \mathrm{E}[x(k) - \hat{x}(k|k)] \qquad (27)$$

with the initial condition $z(0|0) = \mathrm{E}x(0) - \hat{x}(0|0)$. For this reason, equation (26) will be referred to as the *Kalman filter error dynamics* equation.

Like in (Kalman, 1963; Jazwinski, 1970), the Kalman filter stability analysis consists in analyzing the stability of the error dynamics (26), which reflects the intrinsic stability property of the Kalman filter dynamics as expressed in (25).

For this stability analysis, it seems natural to follow the approach adopted in the continuous time case as studied in (Ni and Zhang, 2015). In this approach, an important step would be to establish the *uniform complete observability* of the error dynamics equation (26) adjoined with the "output equation"[2]

$$\bar{y}(k) = C(k)z(k|k). \qquad (28)$$

In (Ni and Zhang, 2015), the continuous time counterpart of (26) and (28) is

$$dz(t)/dt = [A(t) - K(t)C(t)]z(t) \qquad (29)$$

$$\bar{y}(t) = C(t)z(t). \qquad (30)$$

In (29), the term $K(t)C(t)z(t) = K(t)\bar{y}(t)$ is known from the "output" $\bar{y}(t)$ and can be viewed as an "input" term, hence intuitively the pair $\{[A(t) - K(t)C(t)], C(t)\}$ is observable if the pair $\{A(t), C(t)\}$ is observable. Unfortunately, in the discrete time case, such a reasoning cannot be applied to (26) and (28), since the corresponding term in (26), namely $K(k)C(k)A(k)z(k{-}1|k{-}1)$, cannot be written as $K(k)\bar{y}(k)$ with $\bar{y}(k)$ as in (28), because of the matrix $A(k)$ between $C(k)$ and $z(k{-}1|k{-}1)$, and also because of the different indexes k and $k{-}1$.

In order to bypass this difficulty, let us first analyze the stability of the Kalman predictor dynamics. The stability of the filter dynamics will then be established through a relationship between the filter dynamics and the predictor dynamics.

---

[2]The fictive "output" $\bar{y}(k)$ is introduced for stability analysis only. There is no need to know or to measure it in a real Kalman filter implementation.

5.2. Stability of the Kalman predictor

By combining the computations in the steps (10e)-(10g), the Kalman predictor dynamics is summarized as

$$\hat{x}(k|k–1) = A(k)\hat{x}(k–1|k–2)$$
$$+ A(k)K(k–1)[y(k–1) − C(k–1)\hat{x}(k–1|k–2)]$$
$$+ B(k)u(k).$$

The homogeneous part of this *iterative* equation (from $\hat{x}(k–1|k–2)$ to $\hat{x}(k|k–1)$) is

$$z(k|k–1) = A(k)[I_n − K(k–1)C(k–1)]z(k–1|k–2), \tag{31}$$

which corresponds to the dynamics of the mathematical expectation of the state prediction error, *i.e.*,

$$z(k|k–1) = \mathrm{E}[x(k) − \hat{x}(k|k–1)] \tag{32}$$

with the initial condition $z(1|0) = \mathrm{E}x(1) − \hat{x}(1|0)$. For this reason, equation (31) will be referred to as the *Kalman predictor error dynamics* equation.

The stability of predictor error dynamics (31) will be analyzed in this subsection.

In what follows, the *shorter notation*

$$z(k) \triangleq z(k|k–1) \tag{33}$$

will be adopted, thus the predictor error dynamics equation (31) is rewritten as

$$z(k) = A(k)[I_n − K(k–1)C(k–1)]z(k–1). \tag{34}$$

The observability of this "state equation", when adjoined with the "output equation"[3]

$$\bar{y}(k) = C(k)z(k), \tag{35}$$

will play a key role in the stability analysis of (34). The uniform complete observability of the pair $\{A(k), C(k)\}$ is assumed in this paper (see Section 3.2). Notice that the two equations (34) and (35) can be rewritten in the *equivalent* form

$$z(k) = A(k)z(k–1) − A(k)K(k–1)\bar{y}(k–1) \tag{36a}$$
$$\bar{y}(k) = C(k)z(k). \tag{36b}$$

---

[3]Again the fictive "output" $\bar{y}(k)$ is introduced for stability analysis only. There is no need to know or to measure it in a real Kalman filter implementation.

By viewing $A(k)K(k{-}1)\bar{y}(k{-}1)$ as a (known) input term, this system is clearly observable, since the pair $\{A(k), C(k)\}$ is assumed observable. Equivalently, the system defined in (34) and (35) is also observable. Nevertheless, this intuitive reasoning is not sufficient for the following stability analysis, which requires the *uniform complete observability* of the pair $\{A(k)[I_n - K(k{-}1)C(k{-}1)], C(k)\}$ (see Definition 3), that is established in the following lemma.

Lemma 1. *Let $\bar{\Phi}(k, l)$ denote the state transition matrix of the homogeneous iterative equation* (34)*, such that $z(k) = \bar{\Phi}(k, l)z(l)$. Under the assumptions stated in Section 3.2, in particular the uniform complete observability of the pair $\{A(k), C(k)\}$ as defined in Definition 3, there exists a constant $\rho_3 > 0$ such that*

$$\rho_3 I_n \leq \sum_{s=k-h}^{k} \bar{\Phi}^T(s, k)\Lambda(s)\bar{\Phi}(s, k) \tag{37}$$

for all $k \geq h$, with the same positive integer h as in (6)*, and $\Lambda(s) = C^T(s)R^{-1}(s)C(s)$ as defined in* (14)*.* □

A proof of this lemma is given in Appendix 1. The *continuous time* counterpart of this result is known as part of the lemma on observability conservation by output feedback (Anderson et al., 1986; Sastry and Bodson, 1989; Ioannou and Sun, 1996; Zhang and Zhang, 2015). However, for discrete time systems, this result seems not reported in the literature.

Theorem 1. *Under the assumptions stated in Section 3.2, the Kalman predictor error dynamics equation* (31)*, which is also expressed in the shorter form* (34)*, is asymptotically stable.* □

Proof. Define

$$V(z(k), k) \triangleq z^T(k)P^{-1}(k|k{-}1)z(k), \tag{38}$$

which cannot be a Lyapunov function in the classical sense, because $P(k|k{-}1)$ does not have a strictly positive lower bound. The stability analysis based on $V(z(k), k)$ will not exactly follow the classical Lyapunov stability theory.

Substitute $z(k)$ with (34), then

$$V(z(k), k) = z^T(k{-}1)\Xi(k)z(k{-}1) \tag{39}$$

with

$$\Xi(k) \triangleq [I_n - K(k{-}1)C(k{-}1)]^T A^T(k)P^{-1}(k|k{-}1)$$
$$\cdot A(k)[I_n - K(k{-}1)C(k{-}1)]. \tag{40}$$

11

Rewrite (10a) as

$$A^T(k)P^{-1}(k|k{-}1)A(k) = P^{-1}(k{-}1|k{-}1), \tag{41}$$

then

$$\Xi(k) =[I_n - K(k{-}1)C(k{-}1)]^T P^{-1}(k{-}1|k{-}1) \\ \cdot [I_n - K(k{-}1)C(k{-}1)]. \tag{42}$$

It follows from (10d) that

$$[I_n - K(k{-}1)C(k{-}1)] = P(k{-}1|k{-}1)P^{-1}(k{-}1|k{-}2), \tag{43}$$

then

$$\Xi(k) =[I_n - K(k{-}1)C(k{-}1)]^T P^{-1}(k{-}1|k{-}1) \\ \cdot P(k{-}1|k{-}1)P^{-1}(k{-}1|k{-}2) \tag{44}$$
$$=[I_n - K(k{-}1)C(k{-}1)]^T P^{-1}(k{-}1|k{-}2) \tag{45}$$
$$=P^{-1}(k{-}1|k{-}2) \\ - C^T(k{-}1)K^T(k{-}1)P^{-1}(k{-}1|k{-}2). \tag{46}$$

Rearrange (10c) as

$$K^T(k{-}1)P^{-1}(k{-}1|k{-}2) = \Sigma^{-1}(k{-}1)C(k{-}1), \tag{47}$$

then

$$\Xi(k) =P^{-1}(k{-}1|k{-}2) \\ - C^T(k{-}1)\Sigma^{-1}(k{-}1)C(k{-}1) \tag{48}$$

It is then derived from (39) and (38) that

$$V(z(k),k) - V(z(k{-}1),k{-}1) \\ = z^T(k{-}1)\Xi(k)z(k{-}1) \\ - z^T(k{-}1)P^{-1}(k{-}1|k{-}2)z(k{-}1). \tag{49}$$

Applying (48) yields

$$V(z(k),k) - V(z(k{-}1),k{-}1) \\ = -z^T(k{-}1)C^T(k{-}1)\Sigma^{-1}(k{-}1)C(k{-}1)z(k{-}1). \tag{50}$$

It has been shown that $P(k|k{-}1)$ is an upper bounded positive definite matrix (see Corollary 1), therefore $C(k)P(k|k{-}1)C^T(k)$ is positive semidefinite and upper

12

bounded. It is assumed that the noise covariance matrix $R(k)$ has a strictly positive lower bound (See Section 3.2). Let $\sigma_1 > 0$ be an upper bound of the largest eigenvalue of $C(k)P(k|k{-}1)C^T(k)$ for all $k \geq 0$, and $\sigma_2 > 0$ be a lower bound of the smallest eigenvalue of $R(k)$ for all $k \geq 0$. It then follows from (10b) that

$$\Sigma(k) = C(k)P(k|k{-}1)C^T(k) + R(k) \tag{51}$$

$$\leq \frac{\sigma_1}{\sigma_2}R(k) + R(k) \tag{52}$$

$$= \sigma_3 R(k) \tag{53}$$

with the positive constant $\sigma_3 \triangleq \sigma_1/\sigma_2 + 1$.

It is then derived from (50) that

$$V(z(k),k) - V(z(k{-}1),k{-}1)$$

$$\leq -\frac{1}{\sigma_3}z^T(k{-}1)C^T(k{-}1)R^{-1}(k{-}1)C(k{-}1)z(k{-}1) \tag{54}$$

$$= -\frac{1}{\sigma_3}z^T(k{-}1)\Lambda(k{-}1)z(k{-}1) \tag{55}$$

$$\leq 0, \tag{56}$$

where $\Lambda(k{-}1)$ is as defined in (14).

This result implies that $V(z(k),k)$ cannot increase with k, hence $V(z(k),k) = z^T(k)P^{-1}(k|k{-}1)z(k)$ is upper bounded. Moreover, $P(k|k{-}1)$ is upper bounded (see Corollary 1), therefore $\|z(k)\|$ is also upper bounded. The iterative equation (34) is thus Lyapunov stable.

It still remains to show that $\|z(k)\| \to 0$ for asymptotic stability.

Repeatedly applying inequality (55) yields

$$V(z(k{+}1),k{+}1) - V(z(k{-}h),k{-}h)$$

$$\leq -\frac{1}{\sigma_3}\sum_{s=k-h}^{k} z^T(s)\Lambda(s)z(s) \tag{57}$$

$$= -\frac{1}{\sigma_3}z^T(k)\sum_{s=k-h}^{k} \bar{\Phi}^T(s,k)\Lambda(s)\bar{\Phi}(s,k)z(k). \tag{58}$$

where $\bar{\Phi}(s,k)$ is the state transition matrix of (34) such that $z(s) = \bar{\Phi}(s,k)z(k)$, as introduced in Lemma 1.

Applying Lemma 1 then yields

$$V(z(k{+}1),k{+}1) - V(z(k{-}h),k{-}h)$$

$$\leq -\frac{\rho_3}{\sigma_3}\|z(k)\|^2. \tag{59}$$

13

In order to prove, *by contradiction*, the asymptotic stability of the iterative equation (34), assume that $\|z(k)\|$ *does not* tend to zero when $k \to \infty$. This assumption implies that there exists a constant $\varepsilon > 0$ such that for any (arbitrarily large) integer κ, there exists $k > \kappa$ such that $\|z(k)\| > \varepsilon$, hence

$$V(z(k{+}1), k{+}1) - V(z(k{-}h), k{-}h)$$
$$\leq -\frac{\rho_3}{\sigma_3}\varepsilon^2. \tag{60}$$

Because κ can be arbitrarily large, there exist infinitely many such values of $k > \kappa$, Moreover, it was already shown that the value of $V(z(k), k)$ cannot increase with the time index k. Inequality (60) then says that $V(z(k), k)$ is repeatedly decreased by $\rho_3\varepsilon^2/\sigma_3 > 0$ for larger and larger values of k, each being distant of at least h from the previous k. Consequently, $V(z(k), k)$ will become negative for sufficiently large values of k. This is in *contradiction* with the definition of $V(z(k), k)$ in (38) which is positive definite. Therefore, it is proved that $\|z(k)\|$ tends to zero when $k \to +\infty$.

The asymptotic stability of the Kalman predictor error dynamics (34), or equivalently (31), is then proved. □

5.3. Back to the Kalman filter error dynamics

Now let us consider the stability of the Kalman filter error dynamics as expressed in (26). Given the sequence $z(k|k)$ satisfying the recursive equation (26), let $\zeta(k)$ be another sequence such that

$$z(k|k) = [I_n - K(k)C(k)]\zeta(k). \tag{61}$$

According to Proposition 2 and Corollary 1, both $P(k|k)$ and $P(k|k-1)$ are positive definite, it then follows from (10d) that the matrix $[I_n - K(k)C(k)]$ is invertible.

Insert (61) into (26), then

$$[I_n - K(k)C(k)]\zeta(k) = [I_n - K(k)C(k)]A(k)[I_n - K(k{-}1)C(k{-}1)]\zeta(k{-}1). \tag{62}$$

Remove the invertible factor $[I_n - K(k)C(k)]$ from both sides, then

$$\zeta(k) = A(k)[I_n - K(k{-}1)C(k{-}1)]\zeta(k{-}1). \tag{63}$$

This recursive equation is exactly the same as the predictor error dynamics equation (34), whose asymptotic stability has been shown in Theorem 1. Therefore, the recursive equation (63) characterizing $\zeta(k)$ is asymptotically stable, *i.e.*, $\zeta(k)$ is bounded and tends to zero when $k \to \infty$.

The asymptotic stability of the recursive equation (26) satisfied by $z(k|k)$ is then established through $z(k|k) = [I_n - K(k)C(k)]\zeta(k)$, with bounded $K(k)$ (Corollary 1) and bounded $C(k)$. This result is summarized as follows.

Corollary 2. *Under the assumptions stated in Section 3.2, the Kalman filter error dynamics equation* (26) *is asymptotically stable.* $\qquad\square$

6. Numerical examples

Two numerical examples are presented in this section to illustrate the asymptotic stability of the Kalman filter error dynamics and the boundedness of the covariance matrix.

Example 1

Consider a piecewise constant system with

$$A(k) = \begin{bmatrix} 0.3 & 0.4 \\ 0 & 1.2 \end{bmatrix}, \ C(k) = \begin{bmatrix} -0.5 & 1 \end{bmatrix}, \ R(k) = 0.01, \ \text{for } k < 10 \qquad (64)$$

and

$$A(k) = \begin{bmatrix} -0.15 & 1 \\ -0.2 & -0.15 \end{bmatrix}, \ C(k) = \begin{bmatrix} -0.6 & 1 \end{bmatrix}, \ R(k) = 0.01, \ \text{for } k \geq 10. \quad (65)$$

At $k = 10$, the system switches from an unstable model to a stable model. The error dynamics of $z(k|k)$ as defined in (26) and the covariance matrix $P(k|k)$ as computed through (10d) are fully characterized by $A(k), C(k), R(k)$ and the initial conditions $z(0|0) = [1, \ 1]^T$, $P(0|0) = I_2$. The trajectories of $z(k|k)$ and of the two eigenvalues of $P(k|k)$ at every instant k are respectively illustrated in Figure 1 and Figure 2. These results confirm that the error $z(k|k)$ converge to zero and the covariance $P(k|k)$ is bounded.

Example 2

Instead of the piecewise constant system in the previous example, a system varying at every time instant is considered in this example, with

$$A(k) = \begin{bmatrix} 1.2 & 1 + 0.5\cos(k) \\ 0 & 0.5 + 0.2*\sin(k) \end{bmatrix}, \ C(k) = \begin{bmatrix} 1 & 0 \end{bmatrix}, \ R(k) = 0.01. \qquad (66)$$

The error dynamics of $z(k|k)$ as defined in (26) and the covariance matrix $P(k|k)$ as computed through (10d) are fully characterized by $A(k), C(k), R(k)$ and the initial conditions $z(0|0) = [1, \ 1]^T$, $P(0|0) = I_2$. The trajectories of $z(k|k)$ and of the two eigenvalues of $P(k|k)$ at every instant k are respectively illustrated in Figure 3 and Figure 4. Again these results confirm that the error $z(k|k)$ converge to zero and the covariance $P(k|k)$ is bounded.
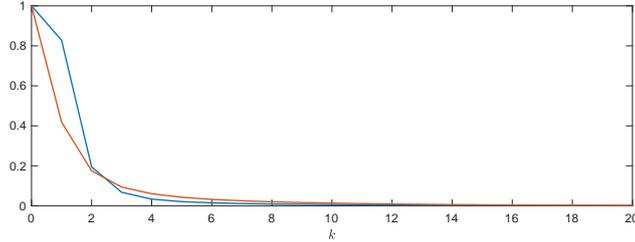
Figure 1: Trajectory of the two components of the filter error mean $z(k|k) = \mathrm{E}[x(k) - \hat{x}(k|k)]$.
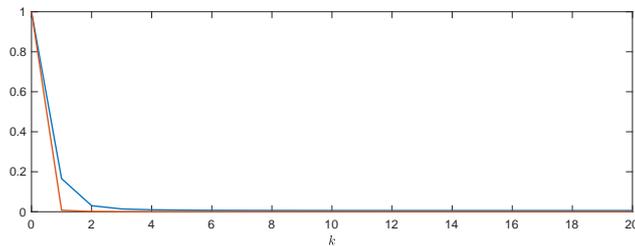

Figure 2: Trajectory of the two eigenvalues of the covariance matrix $P(k|k)$.

7. Conclusion

For any recursive algorithm running in real time, the boundedness of all the involved variables is obviously an important property. It is established in this paper that, when the Kalman filter is applied to discrete time OE systems, all the iteratively computed variables are bounded, essentially under an observability condition. It is further shown that the Kalman filter error dynamics is asymptotically stable. Further analysis on convergence rates will be reported elsewhere, due to space limitation of the present paper.

Since the zero covariance matrix $Q(k) = 0$ of the process noise may be seen as a singularity, in practice it is possible to "regularize" by replacing the zero matrix with some chosen (small) matrix. The classical stability results are then applicable to the "regularized" Kalman filter. However, because a wrong covariance matrix $Q(k)$ is used, the resulting state estimation is no longer optimal in the sense of minimum variance, and the innovation sequence of the Kalman filter is no longer a white noise. The results of this paper show that there is no need to "regularize" the zero covariance matrix, and the optimal Kalman filter for OE systems is indeed stable under the uniform complete observability condition.
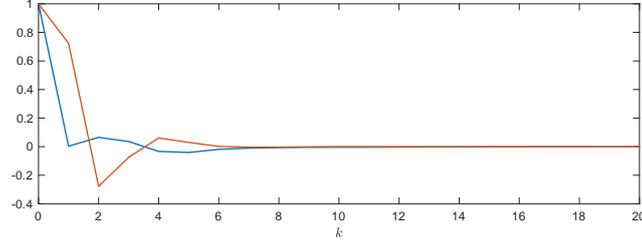
16

Figure 3: Trajectory of the two components of the filter error mean $z(k|k) = \mathrm{E}[x(k) - \hat{x}(k|k)]$.
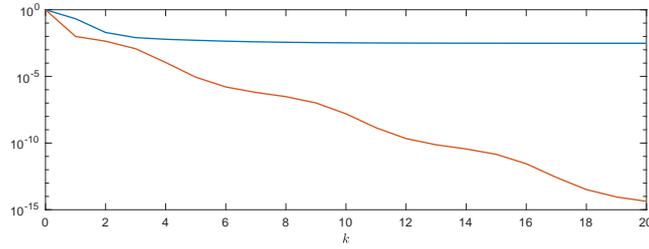

Figure 4: Trajectory of the two eigenvalues of the covariance matrix $P(k|k)$. The logarithmic scale reveals different behaviors of the two eigenvalues.

Appendix A. Proof of Lemma 1

Consider two discrete time state equations

$$x(s) = A(s)x(s{-}1) \tag{A.1}$$
$$z(s) = A(s)[I_n - K(s{-}1)C(s{-}1)]z(s{-}1) \tag{A.2}$$

with their states synchronized at instant k, *i.e.*,

$$x(k) = z(k) \in \mathbb{R}^n. \tag{A.3}$$

Let $\Phi(s,k)$ and $\bar{\Phi}(s,k)$ denote respectively the state transition matrices of these two systems, such that

$$x(s) = \Phi(s,k)x(k) \tag{A.4}$$
$$z(s) = \bar{\Phi}(s,k)z(k). \tag{A.5}$$

Rewrite (A.2) as

$$z(s) = A(s)z(s{-}1) - A(s)K(s{-}1)C(s{-}1)z(s{-}1) \tag{A.6}$$

17

and treat $-A(s)K(s{-}1)C(s{-}1)z(s{-}1)$ as an input term inserted into (A.1), then

$$z(s) = \Phi(s,k)z(k)$$
$$- \sum_{p=k+1}^{s} \Phi(s,p)A(p)K(p{-}1)C(p{-}1)z(p{-}1). \tag{A.7}$$

In this equality, the usual notation convention of the sum for $p = k{+}1, k{+}1, \ldots, s$ assumes that $k < s$. The same equality holds also when $k \geq s$, but the sum is for $p = s, s{+}1, \ldots, k{+}1$, and the equality is accordingly written as

$$z(s) = \Phi(s,k)z(k)$$
$$- \sum_{p=s}^{k+1} \Phi(s,p)A(p)K(p{-}1)C(p{-}1)z(p{-}1). \tag{A.8}$$

This form will be used in what follows, by assuming $s \leq k$.

Because of (A.3) and (A.4), $\Phi(s,k)z(k) = x(s)$, it then follows from (A.8) that

$$x(s) = z(s) + \sum_{p=s}^{k+1} \Phi(s,p)A(p)K(p{-}1)C(p{-}1)z(p{-}1). \tag{A.9}$$

In what follows, the indexes s and p are restricted such that $|p - s| \leq h$, with the integer h as in (6), then $\Phi(s,p)$ with such restricted indexes is bounded. Left multiply both sides of (A.9) by $R^{-\frac{1}{2}}(s)C(s)$ and simplify the writing of the result as

$$R^{-\frac{1}{2}}(s)C(s)x(s) = \sum_{p=s}^{k} \Psi(p)C(p)z(p) \tag{A.10}$$

for appropriate matrices $\Psi(p) \in \mathbb{R}^{m \times m}$, which are bounded due to the boundedness of the involved matrices $R^{-\frac{1}{2}}(s), C(s), \Phi(s,p), A(p)$ and $K(p{-}1)$.

Apply the triangular inequality

$$\|R^{-\frac{1}{2}}(s)C(s)x(s)\| \leq \sum_{p=s}^{k} \|\Psi(p)C(p)z(p)\| \tag{A.11}$$

and take the squares of both sides

$$\|R^{-\frac{1}{2}}(s)C(s)x(s)\|^2 \leq \left[\sum_{p=s}^{k} \|\Psi(p)C(p)z(p)\| \right]^2. \tag{A.12}$$

18

Apply the Schwarz inequality (for $u(p) \in \mathbb{R}$, $v(p) \in \mathbb{R}$)

$$\left[\sum_{p=s}^{k} u(p)v(p) \right]^2 \leq \sum_{p=s}^{k} u^2(p) \sum_{p=s}^{k} v^2(p)$$

with $u(p) = 1$ and $v(p) = \|\Psi(p)C(p)z(p)\|$, then

$$\|R^{-\frac{1}{2}}(s)C(s)x(s)\|^2 \leq (k-s+1) \sum_{p=s}^{k} \|\Psi(p)C(p)z(p)\|^2$$

$$\leq (k-s+1) \sum_{p=s}^{k} \|\Psi(p)\|^2 \cdot \|C(p)z(p)\|^2$$

$$\leq (k-s+1)c_1 \sum_{p=s}^{k} \|C(p)z(p)\|^2$$

where c_1 is the upper bound of $\|\Psi(p)\|^2$. Then

$$\sum_{s=k-h}^{k} \|R^{-\frac{1}{2}}(s)C(s)x(s)\|^2 \tag{A.13}$$

$$\leq (h+1)c_1 \sum_{s=k-h}^{k} \sum_{p=s}^{k} \|C(p)z(p)\|^2 \tag{A.14}$$

$$\leq (h+1)c_1 c_2 \sum_{p=k-h}^{k} \|C(p)z(p)\|^2 \tag{A.15}$$

where $c_2 = h+1$ is to account for the fact that in the double sum each term $\|C(p)z(p)\|^2$ appears not more than $h+1$ times.

By inserting *upper bounded* $R(s)$ in the right hand side, there exists a constant $c_3 > 0$ such that

$$\sum_{s=k-h}^{k} \|R^{-\frac{1}{2}}(s)C(s)x(s)\|^2 \leq c_3 \sum_{p=k-h}^{k} \|R^{-\frac{1}{2}}(p)C(p)z(p)\|^2.$$

Then, by substituting with $x(s) = \Phi(s,k)x(k)$ and $z(p) = \bar{\Phi}(p,k)z(k)$

$$x^T(k) \sum_{s=k-h}^{k} \Phi^T(s,k)C^T(s)R^{-1}(s)C(s)\Phi(s,k)x(k)$$

$$\leq c_3 z^T(k) \sum_{p=k-h}^{k} \bar{\Phi}^T(p,k)C^T(p)R^{-1}(p)C(p)\bar{\Phi}(p,k)z(k)$$

19

This last inequality holds for *any* $x(k) = z(k) \in \mathbb{R}^n$, hence by recalling inequality (6),

$$\rho_1 I_n \leq \sum_{s=k-h}^{k} \Phi^T(s,k)C^T(s)R^{-1}(s)C(s)\Phi(s,k) \tag{A.16}$$

$$\leq c_3 \sum_{p=k-h}^{k} \bar{\Phi}^T(p,k)C^T(p)R^{-1}(p)C(p)\bar{\Phi}(p,k). \tag{A.17}$$

Lemma 1 is then established with $\rho_3 = \rho_1/c_3$. $\qquad\square$

References

Anderson, B. D., Bitmead, R. R., Johnson, C. R. J., Kokotovic, P. V., Kosut, R. L., Mareels, I. M., Praly, L., Riedle, B. D., 1986. Stability of adaptive systems: passivity and averaging analysis. Signal processing, optimization, and control. MIT Press, Cambridge, Massachusetts.

Anderson, B. D. O., Moore, J. B., 1979. Optimal Filtering. Prentice-Hall, Inc., New Jersey.

Forssell, U., Ljung, L., 2000. Identification of unstable systems using output error and box–jenkins model structures. IEEE Transactions on Automatic Control 45 (1), 137–141.

Goodman, D. M., Dudley, D. G., 1987. An output error model and algorithm for electromegnetic system identification. Circuits Systems Signal Process 6 (4), 471–505.

Grewal, M. S., Andrews, A. P., 2015. Kalman Filtering: Theory and Practice with Matlab, 4th Edition. Wiley-IEEE Press.

Ioannou, P., Sun, J., 1996. Robust Adaptive Control. Prentice Hall, electronic copy http://www-rcf.usc.edu/∼ioannou/Robust_Adaptive_Control.htm.

Jazwinski, A. H., 1970. Stochastic Processes and Filtering Theory. Vol. 64 of Mathematics in Science and Engineering. Academic Press, New York.

Kalman, R. E., 1963. New methods in Wiener filtering theory. In: Bogdanoff, J. L., Kozin, F. (Eds.), Proceedings of the First Symposium on Engineering Applications of Random Function Theory and Probability. John Wiley & Sons, New York.

Kim, P., 2011. Kalman Filter for Beginners: with Matlab Examples. CreateSpace.

Ni, B., Zhang, Q., October 2015. Stability of the Kalman filter for output error systems. In: 17th IFAC Symposium on System Identification (SYSID). Beijing.

Ni, B., Zhang, Q., August 2016. Stability of the Kalman filter for continuous time output error systems. Systems & Control Letters 94, 172–180.

Sastry, S., Bodson, M., 1989. Adaptive Control: Stability, Convergence and Robustness. Prentice-Hall.

Wang, L., Ortega, R., Su, H., Liu, Z., Liu, X., 2015. A robust output error identifier for continuous-time systems. International Journal of Adaptive Control and Signal Processing 29 (4), 443–456.

Zarchan, P., Musof, H., 2005. Fundamentals of Kalman filtering: a practical approach. AIAA.

Zhang, L., Zhang, Q., 2015. Observability conservation by output feedback and observability Gramian bounds. Automatica 60, 38–42.

Zhang, Q., 2017. On stability of the Kalman filter for discrete time output error systems. In: Submitted to the 20th IFAC World Congress. Toulouse, France.