

Stability of the adaptive fading extended Kalman filter with the matrix forgetting factor

Cenker BİÇER*, Esin KÖKSAL BABACAN, Levent ÖZBEK

*Modeling and Simulation Laboratory, Department of Statistics, Faculty of Science,
Ankara University, 06100 Tandoğan, Ankara-TURKEY
e-mails: {cbicer, ekoksal, ozbek}@science.ankara.edu.tr*

Received: 11.08.2010

Abstract

The extended Kalman filter is extensively used in nonlinear state estimation problems. As long as the system characteristics are correctly known, the extended Kalman filter gives the best performance. However, when the system information is partially known or incorrect, the extended Kalman filter may diverge or give biased estimates. An extensive number of works has been published to improve the performance of the extended Kalman filter. Many researchers have proposed the introduction of a forgetting factor, both into the Kalman filter and the extended Kalman filter, to improve the performance. However, there are 2 fundamental problems with this approach: the incorporation of the optimal forgetting factor into the (extended) Kalman filter and the selection of the optimal forgetting factor. These problems have not yet been fully resolved and are still open problems in the field. In this study, we propose a new adaptive fading extended Kalman filter with a matrix forgetting factor, and 2 methods are analyzed for the selection of the optimal forgetting factor. The stability properties of the proposed filter are also investigated. Results of the stability analysis show that the proposed filter is an exponential observer for nonlinear deterministic systems. Additionally, the convergence speed of the filter is simulated.

Key Words: *Adaptive fading Kalman filter, extended Kalman filter, adaptive fading extended Kalman filter, forgetting factor, stability analysis*

1. Introduction

The Kalman filter (KF) and the extended Kalman filter (EKF) are the most popular estimation techniques for solving state estimation problems of linear and nonlinear systems, respectively [1]. The EKF was derived from the KF and is used for state estimation of both nonlinear stochastic and nonlinear deterministic systems [2-4]. As long as the characteristics of a system are correctly known, the KF and the EKF will run with the best estimation performance. That is, both dynamic and statistical model parameters should be exactly known so that the KF and EKF give the best estimation performance [5]. However, in many practical cases, system

*Corresponding author: Modeling and Simulation Laboratory, Department of Statistics, Faculty of Science, Ankara University, 06100 Tandoğan, Ankara-TURKEY

characteristics are either unknown or partially known. This lack of information may also seriously reduce the performance of the filter or even cause a divergence. To overcome this problem, in linear systems, several adaptive filtering techniques have been proposed [6-10]. The adaptive fading Kalman filter (AFKF) is one of these proposed filters; it is based on the weighting of the error covariance equation with a scalar forgetting factor [8]. The degree of the lack of information in the model parameters may be different for each parameter; thus, it is not appropriate for the error covariance of complex multiple systems that are weighted with a scalar forgetting factor, but may be appropriate for univariate systems. To address this issue, an AFKF with a symmetrical matrix forgetting factor was proposed by Özbek et al. [7], and a method for the adaptive estimation of multiple forgetting factors in the KF was proposed by Geng and Wang [10]. In nonlinear systems, several adaptive EKF's have been developed. One of these filters is the adaptive fading extended Kalman filter (AFEKF), first proposed by Ozbek and Efe [11]. The AFEKF was designed as an adaptation of the AFKF into nonlinear filter form [11]. In addition, a different AFEKF was proposed to improve the performance of the filter by Kim et al. [12] for when the information of the dynamics or the measurement equation is incomplete. Another AFEKF, for which the authors proposed an adaptive tracking technique with a diagonal matrix forgetting factor to identify time-varying parameters of linear and nonlinear structures, was proposed by Yang et al. [13].

As reported in previous studies [14,15] and confirmed by simulation results, the EKF is more sensitive to the initial values and the selection of appropriate values of the arbitrary matrices in the model. In this study, we propose a new AFEKF with a diagonal matrix forgetting factor to reduce the sensitivity to the initial values and to solve the problem of divergence that arises for a variety of reasons. We call it the matrix adaptive fading extended Kalman filter (MAFEKF). Section 2 briefly introduces the MAFEKF, Section 3 analyzes the stability of the MAFEKF, Section 4 gives some numerical results, and Section 5 summarizes and concludes the work.

2. Adaptive fading extended Kalman filter with the matrix forgetting factor

Consider the nonlinear discrete-time deterministic system given by:

$$x_{n+1} = f(x_n, u_n), \tag{1}$$

$$y_n = h(x_n), \tag{2}$$

where n is the discrete time index, x_n is the $q \times 1$ state vector, u_n is the $p \times 1$ known input vector, and y_n is the $m \times 1$ measurement output vector. Assume that $f(.,.)$ and $h(.)$ are continuously differentiable with respect to x_n , i.e. C^1 functions.

$$\hat{x}_{n+1}^- = f(\hat{x}_n^+, u_n), \tag{3}$$

$$\hat{x}_n^+ = f(\hat{x}_n^-, u_n) + K_n(y_n - h(\hat{x}_n^-)). \tag{4}$$

Eqs. (3) and (4) introduce observers for the given system in Eqs. (1) and (2), where K_n is the time-varying $q \times m$ observer gain and \hat{x}_n^- and \hat{x}_n^+ are a priori and a posteriori estimates, respectively [4]. Since $f(.,.)$ and $h(.)$ are C^1 , they can be expanded by Taylor series expansion as:

$$f(x_n, u_n) - f(\hat{x}_n^+, u_n) = A_n(x_n - \hat{x}_n^+) + \phi(x_n, \hat{x}_n^+, u_n), \tag{5}$$

$$h(x_n) - h_n(\hat{x}_n^-) = C_n(x_n - \hat{x}_n^-) + \chi(x_n, \hat{x}_n^-), \quad (6)$$

where ϕ and χ are higher-order terms in functions $f(\cdot, \cdot)$ and $h(\cdot)$, where

$$A_n = \frac{\partial f}{\partial x}(\hat{x}_n^+, u_n), \quad (7)$$

$$C_n = \frac{\partial h}{\partial x}(\hat{x}_n^-). \quad (8)$$

The general discrete-time extended Kalman filter for the system given by Eqs. (1) and (2) is introduced by the following coupled difference equations:

$$\hat{x}_{n+1}^- = f(\hat{x}_n^+, u_n), \quad (9)$$

$$P_{n+1}^- = A_n P_n^+ A_n' + Q_n, \quad (10)$$

$$K_n = P_n^- C_n' (C_n P_n^- C_n + R_n)^{-1}, \quad (11)$$

$$\hat{x}_n^+ = f(\hat{x}_n^-, u_n) + K_n(y_n - h(\hat{x}_n^-)), \quad (12)$$

$$P_n^+ = (I - K_n C_n) P_n^-, \quad (13)$$

where \hat{x}_{n+1}^- is the state prediction, \hat{x}_n^+ is the state estimation, P_{n+1}^- is the error covariance of state prediction, P_n^+ is the error covariance of state estimation, K_n is the Kalman gain, and Q and R are symmetric positive definite matrices with dimensions $q \times q$ and $m \times m$, respectively [4]. The AFKF is based on the weighting of the error covariance with a forgetting factor. As mentioned above, the main problem is how to incorporate the forgetting factor into the error covariance and how to find the optimal forgetting factor.

In this study, we propose a new AFEKF with the following diagonal matrix forgetting factor:

$$\Lambda_n = \begin{bmatrix} \lambda_{1,n} & 0 & \cdots & 0 \\ 0 & \lambda_{2,n} & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & & \lambda_{q,n} \end{bmatrix}, \quad (14)$$

and weighting of the error covariance matrix by the following equation:

$$P_{n+1}^- = A_n \Lambda_n P_n^+ \Lambda_n' A_n' + \Lambda_n Q_n \Lambda_n'. \quad (15)$$

In this case, the MAFEKF for the system given by Eqs. (1) and (2) is defined by Definition 1. The selection of the optimal matrix forgetting factor is given in the Appendix.

Definition 1. A discrete-time MAFEKF is given by the following coupled difference equations:

$$\hat{x}_{n+1}^- = f(\hat{x}_n^+, u_n), \quad (16)$$

$$P_{n+1}^- = A_n \Lambda_n P_n^+ \Lambda_n' A_n' + \Lambda_n Q_n \Lambda_n', \quad (17)$$

$$K_n = P_n^- C_n' (C_n P_n^- C_n + R_n)^{-1}, \quad (18)$$

$$\hat{x}_n^+ = f(\hat{x}_n^-, u_n) + K_n(y_n - h(\hat{x}_n^-)), \tag{19}$$

$$P_n^+ = (I - K_n C_n) P_n^-, \tag{20}$$

where Λ_n is a time-varying $q \times q$ diagonal matrix forgetting factor.

3. Stability of the adaptive fading extended Kalman filter with the matrix forgetting factor

Under the assumption that the nonlinear system information is perfectly known, Reif and Unbehauen [15] analyzed the behavior of the EKF as a state estimator for nonlinear deterministic systems. This section shows the analysis of the error behavior of the MAFEKF; results of this section are based on [4], [15], and [16].

When the nonlinear system information is partially known or incorrect, we prove, using the second method of Lyapunov, that under certain conditions the proposed MAFEKF is an exponential observer; in other words, the dynamics of the estimation error are exponentially stable. We define the estimation error as:

$$\varsigma_n = x_n - \hat{x}_n^-, \tag{21}$$

and the recursive expression for the estimation error [4] as:

$$\varsigma_{n+1} = A_n (I - K_n C_n) \varsigma_n + r_n \tag{22}$$

and

$$r_n = \phi(x_n, \hat{x}_n^+, u_n) - A_n K_n \chi(x_n, \hat{x}_n^-). \tag{23}$$

Two definitions are presented for the sake of completeness following [17].

Definition 2. The difference equation in Eq. (22) has an exponentially stable equilibrium point at 0 if there are real numbers $\varepsilon, \eta > 0$ and $\theta > 1$ such that:

$$\|\varsigma_n\| \leq \eta \|\varsigma_0\| \theta^n \tag{24}$$

holds for every $n \geq 0$ and for every solution ς_n of Eq. (17) with $\varsigma_0 \in B_\varepsilon$, where $B_\varepsilon = \{v \in R^q \mid \|v\| < \varepsilon\}$.

Definition 3. The observer given by Eqs. (3) and (4) is an exponential observer if the difference equation in Eq. (22) has an exponentially stable equilibrium at 0.

We employ the following lemmas from Reif and Unbehauen [15] in order to prove that the proposed discrete-time MAFEKF is an exponential observer.

The first lemma bounds r_n , which is given by Eq. (23).

Lemma 1. Consider that the following assumptions hold for the given real vectors $x, \hat{x}^-, \hat{x}^+ \in R^q$ and $u \in R^p$, $A_{q \times q}, C_{m \times q}$, and $K_{q \times m}$ matrices, and the functions $\phi(., ., .)$ and $\chi(., .)$.

A1. There are real numbers $\bar{a}, \bar{c}, \bar{k} > 0$ that satisfy the bounds given below:

$$\|A_n\| \leq \bar{a}, \tag{25a}$$

$$\|C_n\| \leq \bar{c}, \tag{25b}$$

$$\|K_n\| \leq \bar{k}. \quad (25c)$$

A2. Positive real numbers $\varepsilon_\phi, \varepsilon_\chi, \kappa_\phi, \kappa_\chi > 0$ exist such that the inequalities:

$$\|\phi(x, \hat{x}^+, u)\| \leq \kappa_\phi \|x - \hat{x}^+\|^2, \quad (26a)$$

$$\|\chi(x, \hat{x}^+)\| \leq \kappa_\chi \|x - \hat{x}^+\|^2, \quad (26b)$$

hold for $\|x - \hat{x}^+\| \leq \varepsilon_\phi$ and $\|x - \hat{x}^-\| \leq \varepsilon_\chi$, respectively.

A3. \hat{x}^+ satisfies:

$$\hat{x}^+ = \hat{x}^- + KC(x - \hat{x}^-) + K\chi(x, \hat{x}^-). \quad (27)$$

We drop the time index “ n ” for the ease of notation. Let r be given by Eq. (23); then positive real numbers $\varepsilon, \kappa > 0$ exist such that:

$$\|r\| \leq \kappa \|x - \hat{x}^-\|^2 \quad (28)$$

holds for $\|x - \hat{x}^-\| \leq \varepsilon$.

Proof. See [15] for proof.

The second lemma is a well known matrix inversion lemma.

Lemma 2. Consider 2 nonsingular matrices, $\Gamma_{q \times q}$ and $\Delta_{q \times q}$, and assume that $\Gamma^{-1} + \Delta$ is also nonsingular. Then:

$$(\Gamma^{-1} + \Delta)^{-1} = \Gamma - \Gamma(\Gamma + \Delta^{-1})^{-1}\Gamma. \quad (29)$$

Proof. See [2, p.139] for proof.

The last lemma is the matrix inequality lemma for the solution of the state and prediction covariances, P_n^+ and P_{n+1}^- , respectively.

Lemma 3. Consider the symmetric positive definite solutions P_n^+ and P_{n+1}^- , $n \geq 0$, of Eqs. (17) and (20), respectively. Define:

$$\Pi_n^- = (P_n^-)^{-1}, \quad (30a)$$

$$\Pi_n^+ = (P_n^+)^{-1}, \quad (30b)$$

and assume that for $n \geq 0$, the inverse of A_n exists. Then:

$$\begin{aligned} \Pi_{n+1}^- &\leq A_n^{-T} \Lambda_n^{-T} (I - K_n C_n)^{-T} \\ &\times \left[\Pi_n^- - \Pi_n^- (\Pi_n^+ + \Lambda_n' A_n' \Lambda_n^{-T} Q^{-1} \Lambda_n^{-1} A_n \Lambda_n)^{-1} \Pi_n^- \right] \\ &\times (I - K_n C_n)^{-1} \Lambda_n^{-1} A_n^{-1} \end{aligned} \quad (31)$$

Proof. Rewrite Eq. (20) as in [3, p.108]:

$$P_n^+ = (I - K_n C_n) P_n^- (I - K_n C_n)' + K_n R_n K_n'. \quad (32)$$

From Eq. (32), we have:

$$P_n^+ \geq (I - K_n C_n) P_n^- (I - K_n C_n)' \tag{33}$$

Inverting Eq. (33), we obtain:

$$(P_n^+)^{-1} \leq (I - K_n C_n)^{-T} (P_n^-)^{-1} (I - K_n C_n)^{-1}. \tag{34}$$

If we rearrange Eq. (17) and take the inverse, we have:

$$\Pi_{n+1}^- = A_n^{-T} \Lambda_n^{-T} (P_n^+ + \Lambda_n^{-1} A_n^{-1} \Lambda_n Q_n \Lambda_n' A_n^{-T} \Lambda_n^{-T})^{-1} \Lambda_n^{-1} A_n^{-1}. \tag{35}$$

Eq. (35) and Lemma 2 now imply:

$$\Pi_{n+1}^- = A_n^{-T} \Lambda_n^{-T} \left[\Pi_n^+ - \Pi_n^+ (\Pi_n^+ + \Lambda_n' A_n' \Lambda_n^{-T} Q^{-1} \Lambda_n^{-1} A_n \Lambda_n)^{-1} \Pi_n^+ \right] \Lambda_n^{-1} A_n^{-1}. \tag{36}$$

From Eq. (34) and since $\Pi_n^+ = \Pi_n^- (I - K_n C_n)^{-1} = (I - K_n C_n)^{-1} \Pi_n^- = (I - K_n C_n)^{-T} \Pi_n^-$, then:

$$\begin{aligned} \Pi_{n+1}^- &\leq A_n^{-T} \Lambda_n^{-T} (I - K_n C_n)^{-T} \\ &\quad \times \left[\Pi_n^- - \Pi_n^- (\Pi_n^+ + \Lambda_n' A_n' \Lambda_n^{-T} Q^{-1} \Lambda_n^{-1} A_n \Lambda_n)^{-1} \Pi_n^- \right] \\ &\quad \times (I - K_n C_n)^{-1} \Lambda_n^{-1} A_n^{-1}. \end{aligned} \tag{37}$$

Theorem 1. Consider the nonlinear deterministic system described by Eqs. (1) and (2) and a MAFEKF as given by Definition 1. Let the following assumptions hold:

A4. There are positive real numbers $\bar{a}, \bar{c}, \underline{p}, \bar{p} > 0$ and $\lambda_{\min}, \lambda_{\max} \geq 1$ such that the following bounds on various matrices are satisfied for every $n \geq 0$.

$$\|A_n\| \leq \bar{a} \tag{38a}$$

$$\|C_n\| \leq \bar{c} \tag{38b}$$

$$\underline{p}I \leq P_n^- \leq \bar{p}I \tag{38c}$$

$$\underline{p}I \leq P_n^+ \leq \bar{p}I \tag{38d}$$

$$\lambda_{\min} I \leq \Lambda_n \leq \lambda_{\max} I \tag{38e}$$

Here, λ_{\min} is the minimum of the λ_i values and λ_{\max} is the maximum of the λ_i values ($i = 1, 2, \dots, q$).

A5. A_n is nonsingular for every $n \geq 0$.

A6. There are positive real numbers $\epsilon_\varphi, \epsilon_\chi, \kappa_\varphi, \kappa_\chi > 0$ such that the nonlinear functions φ, χ in Eq. (23) are bounded via:

$$\|\varphi(x, \hat{x}, u)\| \leq \kappa_\varphi \|x - \hat{x}\|^2, \tag{39a}$$

$$\|\chi(x, \hat{x})\| \leq \kappa_\chi \|x - \hat{x}\|^2. \tag{39b}$$

The MAFEKF is then an exponential observer.

Proof. Proof of the theorem is similar to that of theorem 3.1 in [15]. The Lyapunov function given by Eq. (40) will be used to prove the exponential stability of ς_n . Consider the Lyapunov function:

$$V_n = \varsigma_n' \Pi_n^- \varsigma_n, \quad (40)$$

where Π_n^- is the $(P_n^-)^{-1}$. Eqs. (38c) and (40) imply:

$$\frac{1}{\underline{p}} \|\varsigma_n\|^2 \leq V_n(\varsigma_n) \leq \frac{1}{\underline{p}} \|\varsigma_n\|^2. \quad (41)$$

Eqs. (38c) and (38d) ensure the nonsingularity of P_n^- and P_n^+ and the existence of the inverse of $(I - K_n C_n)$, which is $P_n^- \Pi_n^+$. Note that along with A2, the requirements of Lemma 3 are satisfied. Estimating $V_{n+1}(\varsigma_{n+1})$ with Eqs. (22), (37), and (38e), we obtain:

$$\begin{aligned} V_{n+1}(\varsigma_{n+1}) &= \varsigma_{n+1}' \Pi_{n+1}^- \varsigma_{n+1} \\ &\leq \underline{\lambda}_{\min}^{-2} \varsigma_n' \left[\Pi_n^- - \Pi_n^- (\Pi_n^+ + \Lambda_n' A_n' \Lambda_n^{-T} Q^{-1} \Lambda_n^{-1} A_n \Lambda_n)^{-1} \Pi_n^- \right] \varsigma_n \\ &\quad + 2r_n' \Pi_{n+1}^- A_n (I - K_n C_n) \varsigma_n + r_n' \Pi_{n+1}^- r_n. \end{aligned} \quad (42)$$

Let \underline{r} be the smallest eigenvalue of the R and rewrite the Kalman gain as $K_n = P_n^+ C_n' R_n^{-1}$ as in [3, p.112]. With the bound on C_n given by A1 and P_n^+ , we then have:

$$\|K_n\| \leq \|P_n^+\| \cdot \|C_n\| \cdot \|R^{-1}\| \leq \bar{k}, \quad (43)$$

where $\bar{k} = \bar{p}\bar{c}/\underline{r}$. Eqs. (38a), (38b), (39a), (39b), and (43) make it possible to apply Lemma 1 to Eq. (42). Let $\underline{q} > 0$ be the smallest eigenvalue of Q ; with all the bounds stated in A1, for $\|\varsigma_n\| \leq \varepsilon$, we then have:

$$\begin{aligned} V_{n+1}(\varsigma_{n+1}) &\leq \lambda_{\min}^{-2} \varsigma_n' \Pi_n^- \varsigma_n - \frac{1}{\lambda_{\min}^2 \bar{p}^2 \left(\frac{1}{\underline{p}} + \frac{\bar{\alpha}^2 \lambda_{\max}^2}{\underline{q} \lambda_{\min}^2} \right)} \|\varsigma_n\|^2 + 2\kappa \|\varsigma_n\|^2 \frac{\bar{\alpha}(1+\bar{k}\bar{c})}{\underline{p}} \|\varsigma_n\| \\ &\quad + \kappa \|\varsigma_n\|^2 \frac{1}{\underline{p}} \kappa \varepsilon \|\varsigma_n\|. \end{aligned} \quad (44)$$

Define κ' as follows:

$$\kappa' = \frac{\kappa}{\underline{p}} (2\bar{\alpha}(1 + \bar{k}\bar{c}) + \kappa \varepsilon). \quad (45)$$

Using Eq. (40), rearranging the inequality in Eq. (44) yields the following inequality:

$$V_{n+1}(\varsigma_{n+1}) \leq \lambda_{\min}^{-2} V_n(\varsigma_n) - \left(\frac{1}{\lambda_{\min}^2 \bar{p}^2 \left(\frac{1}{\underline{p}} + \frac{\bar{\alpha}^2 \lambda_{\max}^2}{\underline{q} \lambda_{\min}^2} \right)} - \kappa' \|\varsigma_n\| \right) \|\varsigma_n\|^2. \quad (46)$$

Define:

$$\varepsilon' = \min \left(\varepsilon, \frac{1}{2\lambda_{\min}^2 \bar{p}^2 \left(\frac{1}{\underline{p}} + \frac{\bar{\alpha}^2 \lambda_{\max}^2}{\underline{q} \lambda_{\min}^2} \right)} \right). \quad (47)$$

We will then obtain the following inequalities:

$$V_{n+1}(\varsigma_{n+1}) \leq \lambda_{\min}^{-2} V_n(\varsigma_n) - \left(\frac{1}{2\lambda_{\min}^2 \bar{p}^2 \left(\frac{1}{\underline{p}} + \frac{\bar{\alpha}^2 \lambda_{\max}^2}{\underline{q} \lambda_{\min}^2} \right)} \|\varsigma_n\|^2 \right),$$

$$V_{n+1}(\varsigma_{n+1}) - V_n(\varsigma_n) \leq -\frac{1}{2\lambda_{\min}^2 \bar{p}^2 \left(\frac{1}{\underline{p}} + \frac{\bar{\sigma}^2 \lambda_{\max}^2}{q\lambda_{\min}^2}\right)} \|\varsigma_n\|^2 + (\lambda_{\min}^{-2} - 1) V_n(\varsigma_n), \tag{48}$$

for $\|\varsigma_n\| \leq \varepsilon'$. The error dynamics of the discrete-time MAFEKF are locally negative definite. To see this, use the fact that the right-hand side of the second inequality in Eq. (48) is negative semidefinite due to the bounds of the Lyapunov function and $\underline{\lambda} \geq 1$. If the standard results of Lyapunov functions are applied as in [17, p.108], then it is concluded that Eq. (22) has an asymptotically stable equilibrium point at 0; in other words, the error is bounded and the MAFEKF is an exponential observer.

If the analysis is continued to quantify the degree of stability of the MAFEKF, utilizing Eqs. (41) and (48), we have:

$$V_{n+1}(\varsigma_{n+1}) \leq V_n(\varsigma_n) \left(\lambda_{\min}^{-2} - \frac{\underline{p}}{2\lambda_{\min}^2 \bar{p}^2 \left(\frac{1}{\underline{p}} + \frac{\bar{\sigma}^2 \lambda_{\max}^2}{q\lambda_{\min}^2}\right)} \right),$$

$$V_n(\varsigma_n) \leq V_0(\varsigma_0) \left(\lambda_{\min}^{-2} - \frac{\underline{p}}{2\lambda_{\min}^2 \bar{p}^2 \left(\frac{1}{\underline{p}} + \frac{\bar{\sigma}^2 \lambda_{\max}^2}{q\lambda_{\min}^2}\right)} \right)^n. \tag{49}$$

Without loss of generality, let $\bar{p} > 1$. We then have:

$$1 - \frac{\underline{p}}{2\lambda_{\min}^2 \bar{p}^2 \left(\frac{1}{\underline{p}} + \frac{\bar{\sigma}^2 \lambda_{\max}^2}{q\lambda_{\min}^2}\right)} > 0.$$

Using the bounds on the Lyapunov function and Eq. (49), we obtain:

$$\|\varsigma_n\| \leq \sqrt{\bar{p}/\underline{p}} \|\varsigma_0\| \left(\frac{\lambda_{\min}}{1 - \frac{\underline{p}}{2\lambda_{\min}^2 \bar{p}^2 \left(\frac{1}{\underline{p}} + \frac{\bar{\sigma}^2 \lambda_{\max}^2}{q\lambda_{\min}^2}\right)}} \right)^{-n}.$$

Recall Eq. (26) and define:

$$\eta = \sqrt{\bar{p}/\underline{p}} > 0;$$

then:

$$\theta = \frac{\lambda_{\min}}{\sqrt{1 - \frac{\underline{p}}{2\lambda_{\min}^2 \bar{p}^2 \left(\frac{1}{\underline{p}} + \frac{\bar{\sigma}^2 \lambda_{\max}^2}{q\lambda_{\min}^2}\right)}}} \tag{50}$$

is obtained.

Remark. Recall that for $\Lambda = I$, the filter becomes the standard EKF, and choosing $\lambda_i > 1$, ($i = 1, 2, \dots, q$) provides exponential weighting to the estimates and defines how much emphasis is put on the latest estimate.

4. Simulation study

In this section, we present some simulation results to demonstrate the behavior of the MAFEKF introduced in the previous section. Let us consider the state-space model given below.

$$\begin{aligned}
 x_{n+1} &= \begin{bmatrix} x_{1,n+1} \\ x_{2,n+1} \end{bmatrix} = \begin{bmatrix} 1 - k_1\Delta_t & 0 \\ k_1\Delta_t & 1 - k_2\Delta_t \end{bmatrix} \begin{bmatrix} x_{1,n} \\ x_{2,n} \end{bmatrix} \\
 y_n &= [01] x_n
 \end{aligned}
 \tag{51}$$

The model is a compartmental model that is used to characterize the ingestion, distribution, and metabolism of a drug in an individual. In the model, Δ_t is the integration time interval subdivider; $x_{1,n}$ and $x_{2,n}$ denote the drug mass in the first compartment, the gastrointestinal tract of the individual, and in the second compartment, the bloodstream of the individual, respectively. At the same time, k_1 , the positive constant characterizing the gastrointestinal tract of the given individual, and k_2 , the positive constant characterizing the metabolic and excretory processes of the individual, are defined as unknown parameters that can be constant or time-varying, to be estimated along with the states [11,19]. Assume that $\Phi_n(\psi)$ is a known vector that is a function of some unknown vector given as $\psi = [k_1 \ k_2]'$. Now ψ can be thought of as a random walk process; that is:

$$\psi_{n+1} = \psi_n + \delta_n,
 \tag{52}$$

where δ_n is any zero-mean white noise sequence uncorrelated with the measurement noise variance, v_n , and with preassigned positive definite variance $Var(\delta_n) = S_n$. In applications, S_n may be chosen as $S_n = S > 0$ for all n . The system described by Eqs. (51) and (52) can be reformulated as a nonlinear model as follows:

$$\begin{aligned}
 \begin{bmatrix} x_{n+1} \\ \psi_{n+1} \end{bmatrix} &= \begin{bmatrix} \Phi_n(\psi_n)x_n \\ \psi_n \end{bmatrix} + \begin{bmatrix} w_n \\ \delta_n \end{bmatrix}, \\
 y_n &= [C_n \ 0] \begin{bmatrix} x_n \\ \psi_n \end{bmatrix} + v_n,
 \end{aligned}
 \tag{53}$$

where C_n is the known measurement matrix and w_n is the process noise, which is uncorrelated with v_n . Considering its nonlinear nature, the EKF can be applied to the problem at hand in order to estimate the state vector that contains ψ_n as one of its components. Note that k_1 and k_2 are parts of the state vector through Eq. (53), i.e. the state vector is defined as $x = [x_1 \ x_2 \ k_1 \ k_2]'$, where x_1 and x_2 are the states at the output, and k_1 and k_2 are constant or time-varying parameters. The values of parameters k_1 and k_2 in simulation are set as follows:

$$\begin{aligned}
 k_1 &= \begin{cases} 0.9, & n \leq 37 \\ 0.55, & n > 37 \end{cases}, \\
 k_2 &= \begin{cases} 0.1, & n \leq 37 \\ 0.4, & n > 37 \end{cases}.
 \end{aligned}$$

Table 1. Initials values and noise terms.

	Small initial estimation error	Large initial estimation error
Initial states	$[10 \ 10 \ 0.7 \ 0.3]^T$	$[5 \ 5 \ 0.5 \ 0.5]^T$
Process noise	$1 \times 10^{-6} I_4$	$1 \times 10^{-2} I_4$
Measurement noise	1×10^{-5}	1×10^{-3}

The simulation was run with the initial values and noise terms given in Table 1. The aim was to compare the performance of the EKF and the performance of the MAFEKF. Results of the simulations with 250 replications are given in Figures 1-5. Figure 1 shows the estimation of the k_1 and k_2 parameters under a small initial estimation error. Both the EKF and MAFEKF display similar performances until the change in the values of the parameters. After the change of the value of the parameters, the filters detect the change with some delay and try to adjust their estimations accordingly. However, the filter employing the MAFEKF displays a faster adaptation to the change and a better convergence to the true value of the parameter. Figure 2 shows the elements of the matrix forgetting factor. Figure 3 shows the mean square errors (MSEs) of all state and parameter estimations for the EKF and MAFEKF. The MAFEKF has a smaller MSE than the EKF under the small initial estimation error. Figures 4 and 5 show the estimations of the k_1 and k_2 parameters and the MSEs of the filters, respectively, for a large initial estimation error. Thus, we can claim that the MAFEKF has better performance than the EKF for both small initial estimation errors and large initial estimation errors.

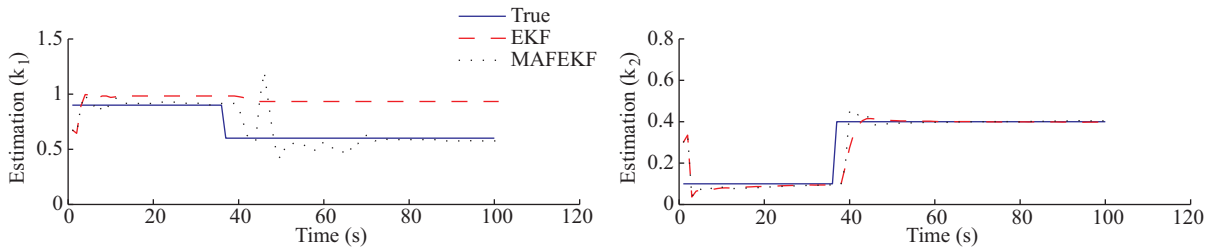


Figure 1. Estimation of parameters k_1 and k_2 (small initial error and process noise).

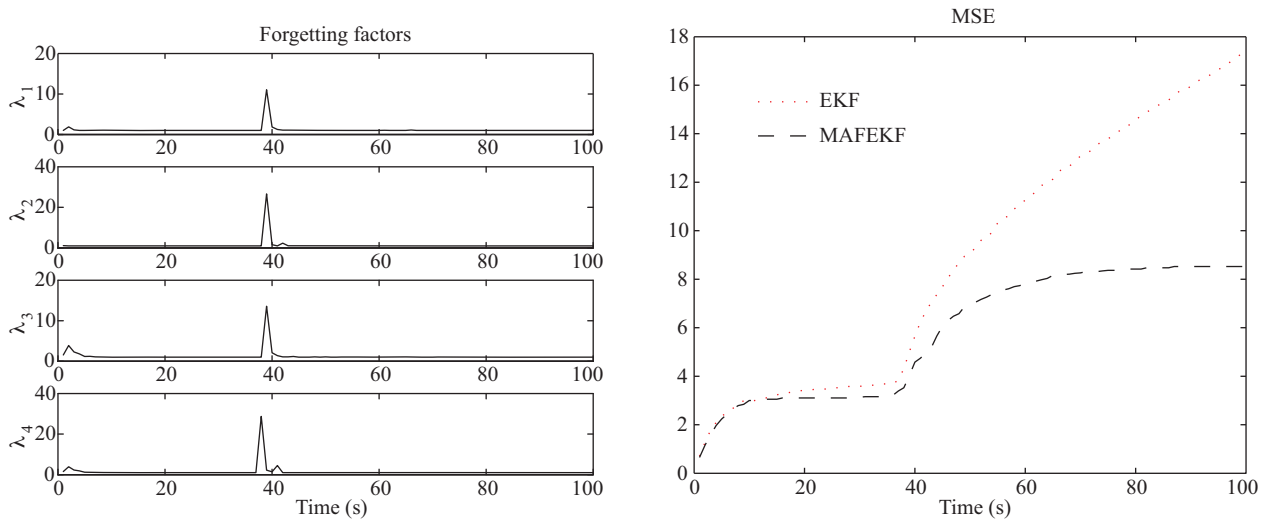


Figure 2. Forgetting factors computed by Algorithm 1.

Figure 3. Mean square errors (small initial error and process noise).

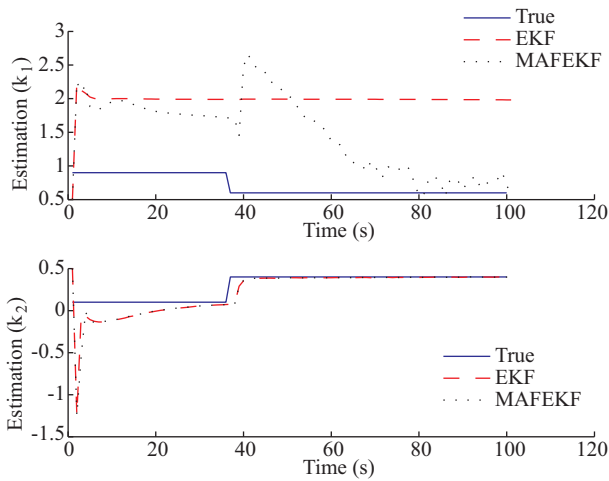


Figure 4. Estimation of parameters k_1 and k_2 (large initial error and process noise).

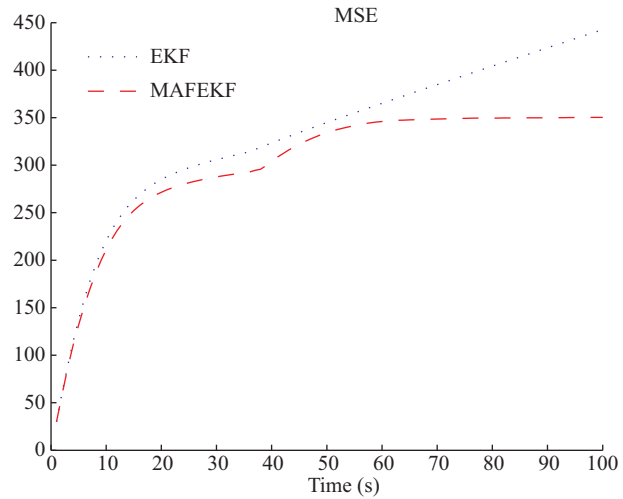


Figure 5. Mean square error (large initial error and process noise).

In the second simulation study, to compare the performance of the proposed algorithms for the selection of the optimal matrix forgetting factor, the Lotka-Volterra model was considered, which is a reproduction model of 2 interactive species. The model is described by the following system of ordinary differential equations:

$$\begin{aligned} \frac{dx_1(t)}{dt} &= ax_1(t) - bx_1(t)x_2(t), \\ \frac{dx_2(t)}{dt} &= -mx_2(t) + rx_1(t)x_2(t). \end{aligned} \tag{54}$$

The state-space notation of the model in Eq. (54) is given below.

$$x_{n+1} = \begin{bmatrix} x_{1,n+1} \\ x_{2,n+1} \end{bmatrix} = \begin{bmatrix} 1 + a - bx_{2,n} & 0 \\ 0 & 1 - m + rx_{1,n} \end{bmatrix} \begin{bmatrix} x_{1,n} \\ x_{2,n} \end{bmatrix} \tag{55}$$

$$y_n = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x_n \tag{56}$$

Here, x_1 is the number of prey in time n , x_2 is the number of predators in time n , y_n is the measurements, a is the reproduction rate of the prey, m is the death rate of the predators, and b and r are the interaction rates between prey and predators [20]. For the simulation study, the true values of parameters a , b , m , and r were set as in Table 2. Simulation for the initial values given in Table 3 was also conducted, the results of which are displayed in Figures 6-8. Figure 6 displays the MSE of the filters and Figure 7 displays the forgetting

Table 2. Values of the parameters used in simulation.

Parameter	True value	Initial estimation used in simulation
a	0.20	0.20
b	0.16	0.16
m	0.30	0.10
r	0.01	0.03

factors computed by Algorithm 1. Figure 8 displays the forgetting factors computed by Algorithm 2. The results of the simulation study reveal that the forgetting factors computed by Algorithm 1 and Algorithm 2 are different. However, the moment of change for the forgetting factors is similar. According to the results of the simulation study, we can claim that the performance of the MAFEKF, as established by using either Algorithm 1 or Algorithm 2, is better than that of the EKF.

Table 3. Initial values.

	Initial estimation	
Initial states	$[2 \ 0.2]^T$	
Arbitrary matrix Q	0.4	0
	0	0.1
Arbitrary matrix R	0.1	0
	0	0.1

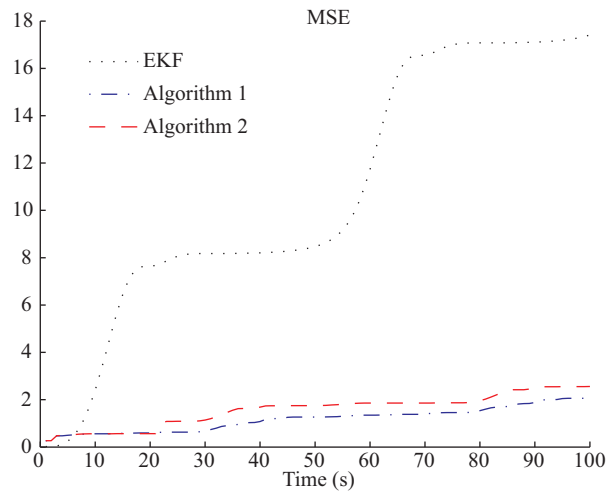


Figure 6. Mean square error of filters.

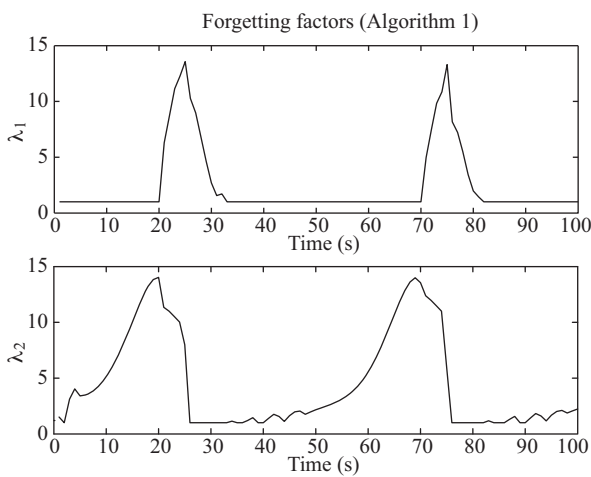


Figure 7. Forgetting factors computed by Algorithm 1.

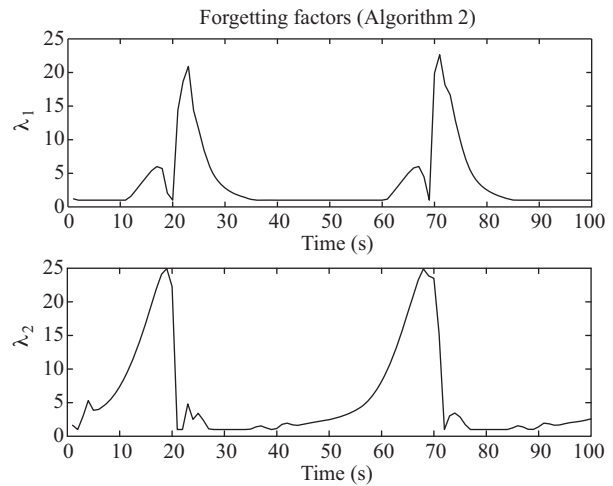


Figure 8. Forgetting factors computed by Algorithm 2.

5. Conclusion

When the system information is partially known or incorrect, the EKF may diverge or give biased estimates. To solve this problem and to improve the estimation performance of the EKF, we have presented MAFEKF and how to make a selection of the optimal matrix forgetting factor. Moreover, the stability properties of the proposed MAFEKF were investigated in a way similar way to that described in [15] for the EKF. It has been proven that, like the EKF, under certain conditions the proposed MAFEKF is an exponential observer for deterministic systems. Finally, the performance of the MAFEKF was demonstrated with simulation studies, including parameter estimation and state estimation. Results of the simulation studies showed that the MAFEKF provides performance gain.

Appendix

Consider the nonlinear discrete time deterministic system given by Eqs. (1) and (2) and the MAFEKF defined by Eqs. (16)-(20). Innovation is defined by:

$$z_n = y_n - h(\widehat{x}_n^-), \quad (\text{A1})$$

the covariance of the innovation is:

$$V_{0,n} = E[z_n z_n'] = C_n P_n^- C_n' + R_n, \quad (\text{A2})$$

and the autocovariance of the innovation is:

$$V_{j,n} = E[z_{n+j} z_n'] = C_{n+j} A_{n+j-1} (I - K_{n+j-1} C_{n+j-1}') \cdots A_{n+1} (I - K_{n+1} C_{n+1}') A_n (P_n^- C_n' - K_n V_{0,n}). \quad (\text{A3})$$

When the information of the nonlinear system is complete, the innovation sequence is a white noise sequence. In order to obtain the optimal forgetting factor, Xia et al. [6] used the fact that the autocovariance of the innovation is zero when the conventional Kalman filter is optimal. Özbek [7] used the same property to obtain the optimal matrix forgetting factor for the AFKF. We also employ the same property to obtain the optimal matrix forgetting factor.

Define:

$$S_n = P_n^- C_n' - K_n V_{0,n}. \quad (\text{A4})$$

The necessary condition for the optimality of the EKF is $S_n = 0$. Using this property, we propose 2 algorithms to obtain the optimal matrix forgetting factor.

Algorithm 1. The optimal diagonal matrix forgetting factor can be obtained in a manner similar to the second method in [7]. We assume that C_n is full rank. In this situation, the optimal matrix forgetting factor can be obtained as a solution of the following nonlinear equation:

$$A_{n-1} \Lambda_{n-1} P_{n-1}^+ \Lambda_{n-1}' A_{n-1} + \Lambda_{n-1} Q_{n-1} \Lambda_{n-1}' = (C_n' C_n)^{-1} C_n' (V_{0,n} - R_n) C_n (C_n' C_n)^{-1}. \quad (\text{A5})$$

Proof. Substituting Eq. (18) into Eq. (Ap4), we obtain the results below.

$$\begin{aligned}
 P_n^- C_n' - (P_n^- C_n' (C_n P_n^- C_n' + R_n)^{-1}) V_{0,n} &= 0 \\
 P_n^- C_n' (I - (C_n P_n^- C_n' + R_n)^{-1} V_{0,n}) &= 0 \\
 (I - (C_n P_n^- C_n' + R_n)^{-1} V_{0,n}) &= 0 \\
 (C_n P_n^- C_n' + R_n)^{-1} V_{0,n} &= I \\
 (C_n P_n^- C_n' + R_n)^{-1} &= (V_{0,n})^{-1} \\
 C_n P_n^- C_n' + R_n &= V_{0,n} \\
 C_n P_n^- C_n' &= V_{0,n} - R_n
 \end{aligned}$$

Hence:

$$P_n^- = (C_n' C_n)^{-1} C_n' (V_{0,n} - R_n) C_n (C_n' C_n)^{-1}. \tag{A6}$$

Using Eqs. (17) and (Ap6), we have:

$$A_{n-1} \Lambda_{n-1} P_{n-1}^+ \Lambda_{n-1}' A_{n-1}' + \Lambda_{n-1} Q_{n-1} \Lambda_{n-1}' = (C_n' C_n)^{-1} C_n' (V_{0,n} - R_n) C_n (C_n' C_n)^{-1}. \tag{A7}$$

In [7], $V_{0,n}$ was estimated from observed data using recursive equations.

$$V_{0,n} = D_{1,n} D_{2,n}^{-1} \tag{A8}$$

$$D_{1,n} = D_{1,n-1} \Lambda_{n-1}^{-1} + z_n z_n' \tag{A9}$$

$$D_{2,n} = D_{2,n-1} \Lambda_{n-1}^{-1} + I \tag{A10}$$

The initial values were:

$$D_{1,0} = 0, \quad D_{2,0} = 0.$$

Hence, Eq. (Ap7) is a nonlinear equation of Λ and it can be solved using a numerical method (for example, the Newton-Raphson method).

Algorithm 2. The performance of the EKF depends on a function defined by:

$$f(\lambda_1, \lambda_2, \dots, \lambda_q, n) = \sum_{i=1} \sum_{j=1} S_{ij,n}^2, \tag{A11}$$

where $S_{ij,n}$ is the (i, j) th element of S_n . The smaller the value of $f(\lambda_1, \lambda_2, \dots, \lambda_q, n)$ is, the higher the performance of the filter is. Moreover, the absolute minimum of $f(\lambda_1, \lambda_2, \dots, \lambda_q, n)$ gives the optimal estimate. Thus, the matrix forgetting factor should be chosen to minimize $f(\lambda_1, \lambda_2, \dots, \lambda_q, n)$. This minimization problem can be numerically solved using nonlinear minimization algorithms (for example, the gradient descent method).

References

[1] Y. Bar-Shalom, X.R. Li, T. Kirubarajan, Estimation with Applications to Tracking and Navigation, New York, Wiley Interscience, 2001.

- [2] B.D.O Anderson, J.B. Moore, *Optimal Filtering*, New Jersey, Prentice Hall, 1979.
- [3] A. Gelb, *Applied Optimal Estimation*, Cambridge, MIT Press, 1984.
- [4] E.K. Babacan, L. Ozbek, M. Efe, "Stability of the extended Kalman filter when the states are constrained", *IEEE Transactions on Automatic Control*, Vol. 53, pp. 2707-2711, 2008.
- [5] K.H. Kim, J.G. Lee, C.G. Park, G.I. Jee, "The stability analysis of the adaptive fading extended Kalman filter", *16th IEEE International Conference on Control*, pp. 982-987, 2007.
- [6] Q. Xia, M. Rao, Y. Ying, X. Shen, "Adaptive fading Kalman filter with an application", *Automatica*, Vol. 30, pp. 1333-1338, 1994.
- [7] L. Özbek, F. Öztürk, F. Aliev, "Kalman filtresinde kayıpları önlemek için bir yöntem", *Otomatik Kontrol Ulusal Toplantısı, İstanbul*, Yayın No: 588, pp. 31-38, 1996.
- [8] L. Özbek, F.A. Aliev, "Comments on 'Adaptive fading Kalman filter with an applications'", *Automatica*, Vol. 34, pp. 1163-1164, 1998.
- [9] C. Hide, T. Moore, M. Smith, "Adaptive Kalman filtering for low-cost INS/GPS", *Journal of Navigation*, Vol. 56, pp. 143-152, 2003.
- [10] Y. Geng, J. Wang, "Adaptive estimation of multiple fading factors in Kalman filter for navigation applications", *GPS Solutions*, Vol. 12, pp. 273-279, 2008.
- [11] L. Ozbek, M. Efe, "An adaptive extended Kalman filter with application to compartment models", *Communication in Statistics - Simulation and Computation*, Vol. 3, pp. 145-158, 2003.
- [12] K.H. Kim, G.I. Jee, C.G. Park, J.G. Lee, "The stability analysis of the adaptive fading extended Kalman filter using the innovation covariance", *International Journal of Control, Automation, and Systems*, Vol. 7, pp. 49-56, 2009.
- [13] J.N. Yang, S. Lin, H. Huang, L. Zhou, "An adaptive extended Kalman filter for structural damage identification", *Structural Control and Health Monitoring*, Vol. 13, pp. 849-867, 2006.
- [14] M. Boutayeb, H. Rafaralahy, M. Darouach, "Convergence analysis of the extended Kalman filter used as an observer for nonlinear deterministic discrete-time systems", *IEEE Transactions on Automatic Control*, Vol. 42, pp. 581-586, 1997.
- [15] K. Reif, R. Unbehauen, "The extended Kalman filter as an exponential observer for nonlinear systems", *IEEE Transactions on Signal Processing*, Vol. 47, pp. 2324-2328, 1999.
- [16] L. Özbek, E.K. Babacan, M. Efe, "Stochastic stability of the discrete-time constrained extended Kalman filter", *Turkish Journal of Electrical Engineering & Computer Sciences*, Vol. 18, pp. 211-223, 2010.
- [17] V. Lakshmikantham, D. Trigiante, *Theory of Difference Equations: Numerical Methods and Applications*, London, Academic Press, 1988.
- [18] F.L. Lewis, *Optimal Estimation*, New York, Wiley, 1986.
- [19] N.M. McClamroch, *State Models of Dynamic Systems*, New York, Springer-Verlag, 1980.
- [20] F. Öztürk, L. Özbek, *Matematiksel Modelleme ve Simülasyon*, Ankara, Gazi Kitabevi, 2004.