Signal Processing 154 (2019) 294–303

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Fractional central difference Kalman filter with unknown prior information

Tianyu Liu, Songsong Cheng, Yiheng Wei, Ang Li, Yong Wang*

Department of Automation, University of Science and Technology of China, Hefei 230027, China

A R T I C L E I N F O

Article history:
Received 17 October 2017
Revised 30 July 2018
Accepted 10 August 2018
Available online 11 September 2018

Keywords:
Fractional calculus
Adaptive filter
Fractional Kalman filter
Central difference Kalman filter
Maximum a posteriori principle

A B S T R A C T

In this paper, a generalized fractional central difference Kalman filter for nonlinear discrete fractional dynamic systems is proposed. Based on the Stirling interpolation formula, the presented algorithm can be implemented as no derivatives are needed. Besides, in order to estimate the state with unknown prior information, a maximum a posteriori principle based adaptive fractional central difference Kalman filter is derived. The adaptive algorithm can estimate the noise statistics and system state simultaneously. The unbiasedness of the proposed algorithm is analyzed. Several numerical examples demonstrate the accuracy and effectiveness of the two Kalman filters.

1. Introduction

The optimum Kalman filter is a recursive state estimation algorithm for integer order linear state space systems. It is widely used in numerous engineering applications, such as aerospace, navigation [1], econometrics, computer vision [2], autopilots [3] and many others where estimation is relevant. The accuracy of the Kalman filter depends largely on certain assumptions, such as noise statistics. The problem is observed that the noise prior knowledge is unknown or time-varying in circumstances. The adaptive Kalman filter is a common tool to deal with this problem.

The classical Kalman filter was applied to the estimation problem for discrete dynamic systems [4]. Then based on the Taylor series approximation, Bucy and Sunahara put forward the extended Kalman filter (EKF) [5,6]. Although the EKF is widely used for various engineering fields, there still exist some theoretical limitations, for example, nonlinear functions must be continuously differentiable and the filter is required to calculate the Jacobian matrix. Following the intuition that "*it is easier to approximate a probability distribution than it is to approximate an arbitrary nonlinear function or transformation*", using the unscented transformation, Julier and Uhlmann et al. presented a new approach to approximate the posterior mean and the posterior error covariance [7]. The corresponding filter is known as the unscented Kalman filter (UKF). The UKF ensures an accuracy of at least the second order Taylor series approximation. But the implementation of a UKF is more computationally expensive than an EKF. Therefore, Biswas et al. proposed a new single propagated unscented Kalman filter and an extrapolated single propagated unscented Kalman filter to reduce computational complexity [8].

For nonlinear Gaussian systems, Ito et al. presented the systematic formulation of Gaussian optimal recursive filters, and obtained a novel central difference filter [9]. At the same time, NøRgaard et al. utilized the Stirling interpolation formula to approximate the posterior mean and the posterior covariance. Then the divided difference filter is developed [10]. Those two filters are essentially identical and can be referred to as the central difference Kalman filter (CDKF) [11].

The performance of the KF depends largely on prior information of noise statistics. The use of imprecise information will result in estimation errors or even filtering divergence. Adaptive filtering is an effective way to solve this problem. Most of the adaptive filtering methods are applied to linear systems. It can be divided into four categories: Bayesian, maximum likelihood, correlation and covariance matching [12]. Based on the maximum a posteriori (MAP) principle, the popular Sage-Husa AKF (SHAKF) [13], which estimates the noise statistics and state recursively , also can be considered as a covariance matching method. Besides, the variational Bayesian based AKF is also an approximation of the Bayesian method [14]. For nonlinear systems, several approaches are investigated.

On the other hand, thanks to that many systems can be described accurately with the introduction of fractional calculus, fractional systems have attracted much attention from engineering

* Corresponding author.
E-mail address: yongwang@ustc.edu.cn (Y. Wang).

https://doi.org/10.1016/j.sigpro.2018.08.006
0165-1684/© 2018 Published by Elsevier B.V.

and physics fields. Besides, the application of fractional calculus in control systems also has rapidly development, especially in stability analysis [15,16], controller design [17,18], adaptive filtering [19,20], etc. An important class of theoretical and practical problems is how to obtain the exact state when state variables cannot be measured directly. Motivated by this, the fractional Kalman filter (FKF) and the fractional extended Kalman filter (FEKF) are proposed [21]. The FKF algorithm is used for state estimation in the systems with ultracapacitor [22], fractional nonlinear systems in a chaotic communication scheme [23] and over networks with packet losses [24], etc. The prime difference between the FKF and the integer Kalman filter is that the integer order dynamic systems can be considered as a Markov process, but fractional dynamic systems can not. Because of the existence of the fractional differential operator, the estimated state $\mathbf{x}_t$ of the FKF depends on all of the previous state, which leads to significant complexity. Meanwhile the defects of the integer order EKF also exist in the FEKF.

Motivated by the previous discussions, a generalized fractional central difference Kalman filter (FCDKF) is presented. Based on the conventional CDKF, the proposed FCDKF is also a derivative-free filtering algorithm. Furthermore, considering that the prior information is hard to obtain, an adaptive fractional central difference Kalman filter (AFCDKF) is addressed, which can evaluate the system state and noise statistics simultaneously. The main contributions are concluded as follows

- A FCDKF and an AFCDKF are addressed to estimate the system state for different prior information conditions;
- The unbiasedness of the AFCDKF algorithm is analyzed and then an unbiased recursive algorithm is developed;
- The approximate accuracy and numerical complexity of proposed algorithms are analyzed.

The rest of this paper is organized as follows. Section 2 reviews the fundamental knowledge of fractional calculus and CDKF. The FCDKF and AFCDKF for fractional discrete nonlinear systems with stochastic perturbation are designed in Section 3. Section 4 provides several illustrative numerical examples. Finally, Section 5 draws some conclusions.

2. Preliminaries

2.1. Problem statement

The fractional discrete nonlinear system with stochastic perturbation can be described as follow

Definition 2.1. The fractional discrete nonlinear system with stochastic perturbation can be described as

$$\begin{cases} \nabla^{\boldsymbol{\alpha}}\mathbf{x}_k = \mathbf{f}_{k-1}(\mathbf{x}_{k-1}) + \boldsymbol{\omega}_{k-1}, \\ \mathbf{x}_k = \nabla^{\boldsymbol{\alpha}}\mathbf{x}_k - \sum_{j=1}^{k}(-1)^j\boldsymbol{\gamma}_j\mathbf{x}_{k-j}, \\ \mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \boldsymbol{\nu}_k, \end{cases} \tag{1}$$

where $\nabla^{\boldsymbol{\alpha}} = [\nabla^{\alpha_1}, \cdots, \nabla^{\alpha_n}]^{\mathrm{T}}$ and $\boldsymbol{\gamma}_j = \mathrm{diag}\left[\binom{\alpha_1}{j}, \cdots, \binom{\alpha_n}{j}\right]$.

Here k denotes the time index, $\mathbf{x}_k \in \mathbb{R}^n$, $\boldsymbol{\alpha} \in \mathbb{R}^n$, and $\mathbf{z}_k \in \mathbb{R}^m$ are the system state, orders of difference and measurement value, respectively. $\mathbf{f}_k : \mathbb{R}^n \to \mathbb{R}^n$ and $\mathbf{h}_k : \mathbb{R}^n \to \mathbb{R}^m$ are the nonlinear state transform function and measurement function. $\boldsymbol{\omega}_k \in \mathbb{R}^n$ and $\boldsymbol{\nu}_k \in \mathbb{R}^m$ mean the system noise and measurement noise. Moreover, $\hat{\mathbf{x}}_{i|j} = \mathrm{E}\{\mathbf{x}_i \mid \mathbf{Z}_j\}$ indicates the state mean conditioned on $\mathbf{Z}_j$, where $\mathbf{Z}_j = [\mathbf{z}_1, \cdots, \mathbf{z}_j]$ is the observed value. ∇ is the nabla operator, and its definition is given by Definition 2.2.

Definition 2.2. The fractional backward difference of the order α is given by

$$\nabla^{\alpha} f(k) = \sum_{j=0}^{k}(-1)^j\binom{\alpha}{j}f(k-j), \tag{2}$$

where $k \in \mathbb{N}_+$ and the corresponding binomial coefficient can be defined as $\binom{\alpha}{j} = \frac{\alpha(\alpha-1)\ \cdots\ (\alpha-j+1)}{j!}$.

The same as the integer order EKF, the FEKF has been proposed to estimate the system state. But the Jacobian matrix of nonlinear functions is also required in FEKF, which is one of the major constraints. Furthermore, the performance of state estimation is positively related to the accuracy of prior noise information. In most situations, those statistics are inexactly known or even completely unknown. This will lead to large estimation errors or even to filtering divergence. Therefore, the objective of this paper is to design a derivative-free FKF algorithm to estimate the system state exactly. In addition, the adaptive FKF with unknown prior information is also investigated, which aims to evaluate the system state and noise statistics concurrently.

To simplify the analysis, the following common assumptions are carried out [25].

Assumption 2.3. The two noise vectors subject to Gaussian distribution

$$\begin{cases} \mathrm{E}\{\boldsymbol{\omega}_k\} = \mathbf{q}_k, \ \mathrm{Cov}(\boldsymbol{\omega}_i, \boldsymbol{\omega}_j) = \mathbf{Q}_i\delta_{ij}, \\ \mathrm{E}\{\boldsymbol{\nu}_k\} = \mathbf{r}_k, \ \mathrm{Cov}(\boldsymbol{\nu}_i, \boldsymbol{\nu}_j) = \mathbf{R}_i\delta_{ij}, \quad \forall\ i, j, k, \\ \mathrm{Cov}(\boldsymbol{\omega}_i, \boldsymbol{\nu}_j) = \mathbf{0}, \end{cases} \tag{3}$$

where δ_{ij} is the *Kronecher-δ* function, $\mathbf{R}$ is a positive definite matrix and $\mathbf{Q}$ is a positive semidefinite matrix.

Assumption 2.4. The initial state $\mathbf{x}_0$ obeys Gaussian distribution, and is uncorrelated with both the system and measurement noises.

Assumption 2.5. $\mathrm{E}\{\mathbf{x}_i \mid \mathbf{Z}_j\} = \mathrm{E}\{\mathbf{x}_i \mid \mathbf{Z}_i\} = \hat{\mathbf{x}}_i, \ \forall\ i \le j.$

Assumption 2.6. $\mathrm{E}\{(\mathbf{x}_i - \hat{\mathbf{x}}_i)(\mathbf{x}_j - \hat{\mathbf{x}}_j)^{\mathrm{T}}\} = \mathbf{0}, \ \forall\ i \ne j.$

2.2. Fundamental knowledge

First, the Stirling interpolation formula is introduced.

Definition 2.7. Assuming that $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{z} = \mathbf{f}(\mathbf{x})$ is a multidimensional differentiable function, applying the Stirling interpolation formula around the point $\mathbf{x} = \bar{\mathbf{x}}$ yields

$$\mathbf{z} = \mathbf{f}(\mathbf{x}) = \mathbf{f}(\bar{\mathbf{x}} + \Delta\mathbf{x}) = \mathbf{f}(\bar{\mathbf{x}}) + \widetilde{\mathbf{D}}_{\Delta\mathbf{x}}\mathbf{f} + \cdots, \tag{4}$$

where $\widetilde{\mathbf{D}}_{\Delta\mathbf{x}}\mathbf{f} = \frac{1}{\hbar}\left(\sum_{i=1}^{n}\Delta x_i\mu_i\delta_i\right)\mathbf{f}(\bar{\mathbf{x}})$ and $\Delta\mathbf{x} = \mathbf{x} - \bar{\mathbf{x}}$.

Here, $\hbar$ denotes a selected interval length, and μ_i and δ_i are the locally difference operators (see [10]).

Next, the so-called Cholesky factorization is introduced. Considering the function $\mathbf{z} = \mathbf{f}(\mathbf{x})$, the stochastic state $\mathbf{x}$ takes on a Gaussian distribution, denoted as $\mathbf{x} \sim \mathcal{N}(\bar{\mathbf{x}}, \mathbf{P}_\mathbf{x})$. Based on the Stirling interpolation formula, the probability distribution of $\mathbf{z} \sim \mathcal{N}(\bar{\mathbf{z}}, \mathbf{P}_\mathbf{z})$ can be deduced. Based on the Cholesky factorization, we derive $\mathbf{P}_\mathbf{x} = \mathbf{S}_\mathbf{x}\mathbf{S}_\mathbf{x}^{\mathrm{T}}$. Next, the following transformation of $\mathbf{x}$ is introduced:

$$\mathbf{y} = \mathbf{S}_\mathbf{x}^{-1}\mathbf{x}, \tag{5}$$

$$\widetilde{\mathbf{f}}(\bar{\mathbf{y}}) = \mathbf{f}(\mathbf{S}_\mathbf{x}\bar{\mathbf{y}}) = \mathbf{f}(\bar{\mathbf{x}}). \tag{6}$$

The following results can be derived [10]

$$\bar{\mathbf{y}} = \mathrm{E}\{\mathbf{y}\} = \mathbf{S}_\mathbf{x}^{-1}\bar{\mathbf{x}}, \tag{7}$$

$$\mathrm{E}\{(\mathbf{y} - \bar{\mathbf{y}})(\mathbf{y} - \bar{\mathbf{y}})^{\mathrm{T}}\} = \mathbf{I}, \tag{8}$$