Applied Mathematics and Computation 221 (2013) 662–671

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

Higher order sigma point filter: A new heuristic for nonlinear time series filtering

Ksenia Ponomareva *, Paresh Date

Department of Mathematical Sciences, Brunel University, Kingston Lane, Uxbridge, Middlesex UB8 3PH, United Kingdom

ARTICLE INFO

Keywords:
State estimation
Sigma point filters
Moment matching
Nonlinear time series

ABSTRACT

In this paper we present some new results related to the higher order sigma point filter (HOSPoF), introduced in [1] for filtering nonlinear multivariate time series. This paper makes two distinct contributions. Firstly, we propose a new algorithm to generate a discrete statistical distribution to match exactly a specified mean vector, a specified covariance matrix, the average of specified marginal skewness and the average of specified marginal kurtosis. Both the sigma points and the probability weights are given in closed-form and no numerical optimization is required. Combined with HOSPoF, this random sigma point generation algorithm provides a new method for generating proposal density which propagates the information about higher order moments. A numerical example on nonlinear, multivariate time series involving real financial market data demonstrates the utility of this new algorithm. Secondly, we show that HOSPoF achieves a higher order estimation accuracy as compared to UKF for smooth scalar nonlinearities. We believe that this new filter provides a new and powerful alternative heuristic to existing filtering algorithms and is useful especially in econometrics and in engineering applications.

1. Introduction

Consider the following state space form for a nonlinear time series:

$$\mathcal{X}(k+1) = \mathbf{f}(\mathcal{X}(k)) + \mathbf{Q}(\mathcal{X}(k))\mathbf{w}(k+1), \tag{1}$$

$$\mathcal{Y}(k) = \mathbf{h}(\mathcal{X}(k)) + \mathbf{v}(k), \tag{2}$$

where $\mathcal{X}(k)$ and $\mathcal{Y}(k)$ are the respective state vector and measurement vector at time $t(k)$; $\mathbf{f}$, $\mathbf{h}$ are given vector-valued deterministic functions; $\mathbf{Q}$ is a matrix valued deterministic function; and $\mathbf{v}(k), \mathbf{w}(k)$ are vector-valued random variables. The time increment $t(k) - t(k-1)$ is assumed constant for all k. The latent state estimation problem is the problem of constructing an estimate of the random vector $\mathcal{X}(k), k \geqslant 1$, based on the noisy time series data $\mathcal{Y}(1), \mathcal{Y}(2), \ldots, \mathcal{Y}(k)$. In the special case when $\mathbf{f}, \mathbf{h}$ are affine in $\mathcal{X}(k), \mathbf{Q}$ is an identity matrix and $\mathbf{v}(k), \mathbf{w}(k)$ are Gaussian, the optimal recursive solution to the state estimation problem is given by linear Kalman filter, as first outlined in [2]. The optimal recursive solution to the state estimation problem in nonlinear systems is usually not available in closed form. The first current approach that addresses the nonlinear filtering problems is extended Kalman filter (EKF), where Eq. (1) or its continuous time analogue is locally linearized resulting in a linear state space system. A Kalman filter is then employed to obtain the conditional state density of $\mathcal{X}(k)$. Standard textbooks such as [3] carry an extensive discussion of its theoretical underpinnings and implementation; also see [4,5]. The

* Corresponding author.
E-mail address: ksenia.ponomareva@brunel.ac.uk (K. Ponomareva).

0096-3003/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.amc.2013.06.084

second method is unscented Kalman filter (UKF), where a set of particles – or *sigma points* – and weights are used to evaluate the terms in closed-form expressions for updating the state estimate. Several applications of UKF in communication, tracking and navigation are discussed in [6,7], among others. The *ensemble filter* (EF) used in climatology is closely related to UKF; see [8] and references therein. An algorithm which combines some of the desirable properties of both UKF and EF has been proposed in [9]. In [10], approximate methods are developed to deal with the multiplicative uncertainty in the observation equation under sigma point filtering framework. The third common method is sequential Monte Carlo filter or *particle filter* (PF). For this technique, the required conditional density function of $\mathcal{X}(k)$ given measurement $\mathcal{Y}(k)$ at time $t(k)$ is represented by a set of random samples (or *particles*) and associated probability weights; see [11,12] and references therein for more details on PF. Particle filters need a specification of approximate posterior density, called the *proposal density*. This may itself be derived from EKF, UKF or the known state transition density. We will call the versions of particle filters as PF-EKF, PF-UKF and PF-T respectively.

The rest of this paper is organized as follows. Section 2 briefly reviews the traditional unscented Kalman filter. Section 3 introduces the algorithm for the unscented filter with higher order moment matching, which was first proposed by the authors in [1]. The traditional particle filter is discussed in Section 4.1. Sections 4.2 and 6 represent the main contribution of this paper. Section 4.2 introduces the new proposal distribution that uses and propagates the information about higher order moments. This section represents a major modification on the algorithm proposed in [1] since it allows random draws of sigma points. We are aware that matching of the average of higher order moments may not add value if the average is over a very large state dimension. However, the class of applications where the dimension is five or less is still very large; in fact, it is unusual to find time series models with more than four latent states in econometrics and finance. The proposed algorithm can outperform traditional filtering algorithms in latent state estimation of nonlinear time series models where the departure from conditional Gaussianity of prior distribution is quite significant and the state dimension is low enough to make matching of the average kurtosis and the average skewness useful. This is illustrated by an example in Section 5, where the utility of our method is compared with PF-T, PF-EKF and PF-UKF in a multivariate case on a real financial data set. The theoretical accuracy of the conditional mean and the conditional variance estimation using the new method for the univariate case is discussed in Section 6. Section 7 summarizes the results of the paper.

2. Unscented Kalman filter

Consider the system of Eqs. 1,2 with nonlinear functions $\mathbf{f}$ and $\mathbf{h}$. The unscented filtering algorithm can be briefly described as follows. Suppose that at time $t(k)$, the mean $\hat{\mathcal{X}}(k|k)$ and the covariance $\mathbf{P}_{xx}(k|k)$ are available for the system in Eq. (1). Then $2n+1$ symmetric *sigma points* are chosen in the following way:

$$\mathcal{X}^{(0)}(k|k) = \hat{\mathcal{X}}(k|k), \quad \mathcal{X}^{(i)}(k|k) = \hat{\mathcal{X}}(k|k) \pm (\sqrt{(n+\kappa)\mathbf{P}_{xx}})_i, \tag{3}$$

where $i = 1, 2, \dots, n$, κ is a scaling parameter and $(\sqrt{\mathbf{P}_{xx}})_i$ is the ith column of the matrix square root of $\mathbf{P}_{xx}$. The probability weights vector $\mathbf{W}$, where the ith component $\mathbf{W}_i$ is associated with the ith sigma point $\mathcal{X}^{(i)}(k|k)$, is defined as:

$$\mathbf{W}_0 = \frac{\kappa}{n+\kappa}, \quad \mathbf{W}_i = \frac{1}{2(n+\kappa)}, \quad i = 1, 2, \dots, 2n. \tag{4}$$

The following result can then be verified by a straightforward algebraic manipulation (see, e.g. [11]):

Proposition 1. *Sigma points and corresponding probability weights defined in (3) and (4) match the mean $\hat{\mathcal{X}}(k|k)$ and the covariance $\mathbf{P}_{xx}(k|k)$ exactly.*

We compute the predicted mean of $\mathcal{X}(k+1|k)$ using

$$\mathcal{X}^{(i)}(k+1|k) = \mathbf{f}(\mathcal{X}^{(i)}(k|k)), \quad \hat{\mathcal{X}}(k+1|k) = \sum_{i=0}^{2N} \mathbf{W}_i \mathcal{X}^{(i)}(k+1|k), \tag{5}$$

where $\mathbf{W}_i$ are defined in (4). Covariance matrices $\mathbf{P}_{xy}(k+1|k)$ and $\mathbf{P}_{yy}(k+1|k)$ are calculated as

$$\mathbf{P}_{xy}(k+1|k) = \sum_{i=0}^{2N} \mathbf{W}_i (\mathcal{X}^{(i)}(k+1|k) - \hat{\mathcal{X}}(k+1|k)) \mathbf{v}^{(i)}(k)^T, \quad \mathbf{P}_{yy}(k+1|k) = \sum_{i=0}^{2N} \mathbf{W}_i \mathbf{v}^{(i)}(k) \mathbf{v}^{(i)}(k)^T,$$

where $\mathbf{v}^{(i)}(k) = \mathcal{Y}^{(i)}(k+1) - \hat{\mathcal{Y}}(k+1), \mathcal{Y}^{(i)}(k+1) = \mathbf{h}(\mathcal{X}^{(i)}(k+1|k))$ and $\hat{\mathcal{Y}}(k+1) = \sum_{i=0}^{2N} \mathbf{W}_i \mathcal{Y}^{(i)}(k+1)$. $\mathbf{P}_{xx}(k+1|k)$ is computed similarly. Once the true measurement $\mathcal{Y}_{k+1}$ becomes available, we can update the mean estimate in (5) as $\hat{\mathcal{X}}(k+1|k+1) = \hat{\mathcal{X}}(k+1|k) + \mathbf{K}(k+1)(\mathcal{Y}_{k+1} - \hat{\mathcal{Y}}(k+1))$, where $\mathbf{K}(k+1) = \mathbf{P}_{xy}(k+1|k)\mathbf{P}_{yy}^{-1}(k+1|k)$. More details on this algorithm can be found in [11]. UKF has been successfully used as an alternative to EKF; see [7–14] and references therein. Besides being used as a stand-alone filtering algorithm, it has also been used to produce a proposal distribution for PF, see [13].