Automatica 47 (2011) 2245–2250

Contents lists available at SciVerse ScienceDirect

Automatica

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

Brief paper

Cubature Kalman smoothers[☆]

Ienkaran Arasaratnam [1], Simon Haykin

Cognitive Systems Laboratory, McMaster University, Hamilton, ON L8S 4K1, Canada

ARTICLE INFO

Article history:
Received 25 May 2010
Received in revised form
22 January 2011
Accepted 9 April 2011
Available online 31 August 2011

Keywords:
Cubature Kalman filter
Fixed-interval smoothing
Rauch–Tung–Striebel Smoothing
Square-root filtering

ABSTRACT

The cubature Kalman filter (CKF) is a relatively new addition to derivative-free approximate Bayesian filters built under the Gaussian assumption. This paper extends the CKF theory to address nonlinear smoothing problems; the resulting state estimator is named the fixed-interval cubature Kalman smoother (FI-CKS). Moreover, the FI-CKS is reformulated to propagate the square-root error covariances. Although algebraically equivalent to the FI-CKS, the square-root variant ensures reliable implementation when committed to embedded systems with fixed precision or when the inference problem itself is ill-conditioned. Finally, to validate the formulation, the square-root FI-CKS is applied to track a ballistic target on reentry.

1. Introduction

In Arasaratnam and Haykin (2009), Arasaratnam and Haykin described a new nonlinear filter named the *Cubature Kalman Filter* (CKF), for hidden state estimation based on nonlinear *discrete-time* state-space models. Like the celebrated Kalman filter for linear Gaussian models, an important virtue of the CKF is its mathematical rigor. This rigor is rooted in the *third-degree spherical-radial cubature rule* for numerically computing Gaussian-weighted integrals. A unique characteristic of the CKF is the fact that the spherical-radial cubature rule leads to an even number of equally-weighted cubature points ($2n$ points, where n is the dimensionality of the state vector). These cubature points are distributed uniformly on a sphere centered at the origin.

In a related context, the *unscented Kalman filter* (UKF) due to Julier et al. has an odd number of sigma points ($(2n + 1)$ points). These sigma points are distributed on an ellipsoid with a non-zero center point (Julier, Ulhmann, & Durrant-Whyte, 2000). Whereas the cubature points of the CKF follow rigorously from the spherical-radial cubature rule, the sigma points of the UKF are the result of the so-called *unscented transformation* applied to inputs. Theoretically, there is a fundamental difference between the CKF and the UKF. The CKF follows directly from the cubature rule whose important property is that it does not entail any free parameter. In contrast, the UKF purposely introduces a nonzero scaling parameter, commonly denoted by κ. Due to κ, a nonzero center point is often associated with a weight higher than that of the remaining set of sigma points (see Section VII of Arasaratnam and Haykin (2009) and Section III of Arasaratnam, Haykin, and Hurd (2010) for more details). Although the inclusion of the free parameter κ gives freedom to the UKF when it is non-zero, it destroys many desired numerical and theoretical properties of the UKF.

The parameter $\kappa = (3 - n)$ of the 'plain' UKF is zero only when the state dimensionality is three (by 'plain' we mean the UKF without using a scaled unscented transformation). For this special case, what is truly interesting is that the sigma point set boils down to the cubature point set and the algorithmic steps of the plain UKF become identical to that of the CKF. As such, the CKF may be considered as a special case of the UKF in an algebraic sense. However, it is ironic that the observation for setting κ equal to zero has been largely overlooked in the literature on nonlinear filtering for the past many years. Note that the authors of Wu, Hu, Wu, and Hu (2006) attempted to rederive the UKF from the integration perspective using monomial rules. However, a direct use of monomial rules leads to a free parameter similarly to the original derivation proposed by Julier et al. (2000). It is unfortunate again that the authors have completely ignored the possibility of setting the free parameter u_1 in (13) of Wu et al. (2006) to be $u_1 = \sqrt{d}$. This choice leads to the CKF equations, which in turn solves the inherent stability issue of the UKF.

It is well-known that the state estimate of a smoother algorithm is more accurate than that of the corresponding filter counterpart (Meditch, 1969). The motivation of this paper is to derive a

[☆] The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Giancarlo Ferrari-Trecate under the direction of Editor Ian R. Petersen.

E-mail addresses: haran@ieee.org (I. Arasaratnam), haykin@mcmaster.ca (S. Haykin).

[1] Tel.: +1 519 567 5738; fax: +1 519 944 8750.

0005-1098/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2011.08.005

CKF-based smoothing algorithm. The novel contributions of this paper are as follows. (i) Application of cubature integration to existing integration-based smoothing theory. The resulting algorithm is named the *fixed-interval cubature Kalman smoother* (FI-CKS). (ii) For improved numerical stability in systems with limited precision, we go on to develop a square-root version of the FI-CKS. The square-root FI-CKS propagates the square-roots of the error covariances. (iii) Application of the square-root FI-CKS to target tracking. The paper is organized as follows. Section 2 reviews the optimal yet conceptual Bayesian inference solution. Section 3 reviews the CKF briefly. Section 4 derives a suboptimal fixed-interval smoother, which we have named the Fixed-Interval Cubature Kalman Smoother (FI-CKS). We go on to modify the FI-CKS in a way that it propagates the square-roots of covariances for improved numerical stability in Section 5. Section 6 validates the square-root FI-CKS formulation by applying it to a target tracking problem. Section 7 concludes the paper with final remarks.

2. Optimal Bayesian smoother

Consider the following discrete-time nonlinear state-space model as shown by

$$\text{Process equation: } \mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{v}_{k-1} \quad (1)$$

$$\text{Measurement equation: } \mathbf{z}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{w}_k, \quad (2)$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is a hidden or latent variable, called the 'state' of the system at time k; $\mathbf{z}_k \in \mathbb{R}^m$ is the measurement at time k; $\mathbf{f}(.)$ and $\mathbf{h}(.)$ are some known nonlinear functions, and $\mathbf{v}_{k-1}$ and $\mathbf{w}_k$ are noise samples from two independent zero-mean Gaussian processes with covariances $\mathbf{Q}_{k-1}$ and $\mathbf{R}_k$, respectively. They account for process model uncertainty and the inaccuracy of a measuring device. Based on the state-space model, this section reviews how fixed-interval smoothing is performed.

The optimal solution of fixed-interval smoothing can be obtained in two different ways—two-filter smoothing (Fraser & Potter, 1969) and forward–backward smoothing (Rauch, Tung, & Striebel, 1965). For computational reasons, we will focus only on forward–backward smoothing. Given measurements up to time $N(> k), D_N = \{\mathbf{z}_1, \mathbf{z}_2, \ldots \mathbf{z}_N\}$, in forward–backward smoothing, using Bayes' rule the smoothed density $p(\mathbf{x}_k|D_N)$ is factored as follows:

$$\begin{aligned} p(\mathbf{x}_k|D_N) &= \int_{\mathbb{R}^n} p(\mathbf{x}_k, \mathbf{x}_{k+1}|D_N)\mathrm{d}\mathbf{x}_{k+1} \\ &= \int_{\mathbb{R}^n} p(\mathbf{x}_{k+1}|D_N)p(\mathbf{x}_k|\mathbf{x}_{k+1}, D_N)\mathrm{d}\mathbf{x}_{k+1}. \end{aligned} \quad (3)$$

The Markovian nature of the state-space model implies that given knowledge of D_k and $\mathbf{x}_{k+1}$, the state $\mathbf{x}_k$ is independent of future measurements $\{\mathbf{z}_{k+1}, \ldots \mathbf{z}_N\}$. That is, we may write

$$p(\mathbf{x}_k|\mathbf{x}_{k+1}, D_N) = p(\mathbf{x}_k|\mathbf{x}_{k+1}, D_k). \quad (4)$$

Substituting (4) into (3), we get the smoothed density

$$\begin{aligned} p(\mathbf{x}_k|D_N) &= \int_{\mathbb{R}^n} p(\mathbf{x}_{k+1}|D_N)p(\mathbf{x}_k|\mathbf{x}_{k+1}, D_k)\mathrm{d}\mathbf{x}_{k+1} \\ &= p(\mathbf{x}_k|D_k)\int_{\mathbb{R}^n} \frac{p(\mathbf{x}_{k+1}|D_N)p(\mathbf{x}_{k+1}|\mathbf{x}_k)}{p(\mathbf{x}_{k+1}|D_k)}\mathrm{d}\mathbf{x}_{k+1}. \end{aligned} \quad (5)$$

It is understood from (5) that the smoother has to perform two different passes. In the *forward filtering pass*, it computes the posterior density $p(\mathbf{x}_k|D_k)$ and the predictive density $p(\mathbf{x}_{k+1}|D_k)$ until the final time step; in the *backward smoothing pass*, it recursively computes the smoothed density backward in time starting from $k = N$.

For linear Gaussian systems in both the discrete and continuous time domains, the solution to forward–backward smoothing can be exactly found and is given by the *Rauch–Tung–Striebel* (RTS) smoother (Rauch et al., 1965). For nonlinear systems, however, the optimal smoothing solution is intractable for two reasons. (i) For a multi-dimensional system, we must compute the multi-dimensional integral (5). (ii) Even after this integral is computed, it may be difficult to propagate the smoothed density through subsequent time steps because the new smoothed density is not guaranteed to remain closed with a finite summary statistic. For these reasons, we resort to approximations to obtain a suboptimal smoother.

In the past, researchers have resorted to numerical methods to obtain approximate smoothing solutions. One of the well known approximate smoothers is the extended Kalman smoother, the basic idea of which is to apply the Kalman (or RTS) smoother theory by linearizing the nonlinear process and measurement functions using the first-order Taylor series expansion evaluated at the current estimate of the state (Bar Shalom, Li, & Kirubarajan, 2001). The RTS smoother theory can be well adopted to nonlinear Gaussian filters. In Särkkä and Hartikainen (2010) and Šimandl and Duník (2009), derivative-free RTS smoothers based on the unscented transformation, central differences (or the second order Stirling's interpolation) and Gauss–Hermite quadrature are presented in a unified framework. The CKF, a relatively new filter, yields reasonably accurate and numerically stable state estimates at a minimal cost (Arasaratnam & Haykin, 2009). In the subsequent sections, we derive the CKF-based square-root fixed-interval smoother.

3. Cubature Kalman filtering

In this section, before proceeding to the development of the Fixed-Interval Cubature Kalman Smoother (FI-CKS), we briefly review the CKF first (Arasaratnam & Haykin, 2009). The CKF is derived under the assumption that the predictive density of the joint state-measurement random variable is Gaussian (Arasaratnam & Haykin, 2009). This assumption naturally leads to a Gaussian predictive and filtering density of the state. Under this assumption, the Bayesian filter reduces to the problem of how to compute integrals whose integrands are all of the form *nonlinear function* × *Gaussian*. The CKF uses a third-degree cubature rule to numerically compute the above Gaussian-weighted integrals. For example, the cubature rule approximates an n-dimensional Gaussian weighted integral as follows:

$$\int_{\mathbb{R}^n} \mathbf{f}(\mathbf{x})\mathcal{N}(\mathbf{x}; \mu, \Sigma)\mathrm{d}\mathbf{x} \approx \frac{1}{2n}\sum_{i=1}^{2n} \mathbf{f}(\mu + \sqrt{\Sigma}\xi_i)$$

where a square-root factor of the covariance Σ satisfies the relationship $\Sigma = \sqrt{\Sigma}\sqrt{\Sigma}^T$ and the set of $2n$ cubature points are given by

$$\xi_i = \begin{cases} \sqrt{n}\mathbf{e}_i, & i = 1, 2 \ldots n \\ -\sqrt{n}\mathbf{e}_{i-n}, & i = n+1, n+2 \ldots 2n. \end{cases}$$

with $\mathbf{e}_i \in \mathbb{R}^n$ denoting the i-th elementary column vector. That is, the i-th entry of $\mathbf{e}_i$ is unity and all other entries are zero. The third-degree cubature rule is exact for Gaussian-weighted integrals whose integrands are written in the form of a linear combination of monomials up to the third degree or any odd-degree (Arasaratnam & Haykin, 2009). Assuming at time k that the posterior density $p(\mathbf{x}_k|D_k) = \mathcal{N}(\hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k})$ is known, we summarize the steps involved in the time-update and the measurement-update of the CKF as shown in Table 1 that was derived in Arasaratnam and Haykin (2009).