Computational Statistics and Data Analysis 69 (2014) 243–254

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

# Approximate conditional least squares estimation of a nonlinear state-space model via an unscented Kalman filter[☆]

Kwang Woo Ahn[a,*], Kung-Sik Chan[b]

[a] Division of Biostatistics, Medical College of Wisconsin, Milwaukee, WI 53226, United States
[b] Department of Statistics and Actuarial Science, The University of Iowa, Iowa City, IA 52242, United States

ARTICLE INFO

*Article history:*
Received 28 January 2013
Received in revised form 3 July 2013
Accepted 29 July 2013
Available online 19 August 2013

*Keywords:*
Nonlinear time series
SIR model
State-space model
Unscented Kalman filter

ABSTRACT

The problem of estimating a nonlinear state-space model whose state process is driven by an ordinary differential equation (ODE) or a stochastic differential equation (SDE), with discrete-time data is studied. A new estimation method is proposed based on minimizing the conditional least squares (CLS) with the conditional mean function computed approximately via the unscented Kalman filter (UKF). Conditions are derived for the UKF–CLS estimator to preserve the limiting properties of the exact CLS estimator, namely, consistency and asymptotic normality, under the framework of infill asymptotics, i.e. sampling is increasingly dense over a fixed domain. The efficacy of the proposed method is demonstrated by simulation and a real application.



## 1. Introduction

The Kalman filter was proposed by Kalman (1960) who demonstrated the usefulness of the Kalman filter for drawing inference for a wide class of problems using the linear state-space model. However, many scientific studies require nonlinear state-space models with which the linear Kalman filter is inapplicable. For instance, the Susceptible–Infected–Recovered (SIR) model, which is the nonlinear ordinary differential equation, has been widely used in epidemiology, see Diekmann and Heesterbeek (2000).

Many statistical methods for estimating a nonlinear state-space model have been proposed in the literature. Durbin and Koopman (2000) proposed an iterative scheme for deriving an approximate linear-model whose likelihood approximately matches the mode of the log-likelihood function of the underlying nonlinear model. Using the linear approximate model, they proposed importance sampling schemes from classical and Bayesian perspectives. However, their method suffers from high computational cost and the need for calculating the Jacobian matrix, which may be infeasible. Other approaches, for example, particle filter, bootstrap filter, and sequential Monte Carlo, have been proposed in engineering and statistics, see Kitagawa (1996), Liu and Chen (1998), Pitt and Shephard (1999) and Doucet et al. (2001). Starting with random samples called particles from an initial distribution, these methods update the particles according to the conditional or the posterior distribution of the true states based on importance sampling, which is often computationally expensive as Durbin and Koopman (2001). On the other hand, the extended Kalman filter (EKF) approximates a nonlinear model by its first order Taylor expansion. However, the approximation error of the EKF becomes non-negligible with strongly nonlinear models.

☆ The Supplementary Materials available online include all the proofs of the theorems, verification of some properties of a model presented in the beginning of Section 5, and a simulation study using an SDE model.

* Corresponding author. Tel.: +1 414 955 7387; fax: +1 414 955 6513.
*E-mail addresses:* kwooahn@mcw.edu (K.W. Ahn), kung-sik-chan@uiowa.edu (K.-S. Chan).

0167-9473/$ – see front matter 
http://dx.doi.org/10.1016/j.csda.2013.07.038

Another drawback is that similar to the method of Durbin and Koopman (2000), the EKF requires calculating the Jacobian, which may be infeasible.

Besides these nonlinear Kalman filters, several methods have been proposed for specifically estimating a nonlinear state-space model driven by an ordinary differential equation (ODE). Ramsay et al. (2007) proposed the generalized profile estimation method. However, the method is computationally expensive as it requires profiling out the generally high-dimensional coefficient in the functional expansion. Liang and Wu (2008) proposed the two-step approach, but it is applicable only when the full state vector is observed with measurement errors over discrete time.

Instead of relying on linearization techniques employed by the EKF, the *unscented Kalman filter* (UKF) was proposed by Julier and Uhlmann (1997) for extending Kalman filter to nonlinear models using a deterministic sampling scheme consisting of the so-called "sigma points". In each step, the UKF generates a set of sigma points and updates the prediction formulas based on these sigma points. So the UKF bears resemblance to the particle filter. However, the sigma points are deterministic samples and the number of the sigma points is much less than the number of particles. Consequently, the UKF is computationally more efficient than the particle filter. Also, empirical results suggest that the UKF appears to be superior to EKF and works properly and satisfactorily from a Bayesian perspective, see Julier and Uhlmann (2004) and Wu and Smyth (2007). Yet the use of UKF in estimating a nonlinear state-space model from a frequentist perspective is largely unexplored.

The method of conditional least squares (CLS) provides a general approach for estimating a state-space model and it enjoys consistency and asymptotic normality, under some mild regularity conditions, see Klimko and Nelson (1978). However, its applicability is limited by the tractability of the conditional mean function of the future process given the data, which is generally intractable for nonlinear state-space models. The UKF provides a numerical scheme for computing the predictors with which we can calculate an approximation to the conditional least squares (CLS) objective function. We propose to estimate a model by minimizing the latter objective function; we shall refer to this method as the UKF–CLS method. The main purpose of our paper is to study conditions under which the UKF–CLS estimator preserves the consistency and the asymptotic distribution of the exact CLS estimator and its applications to ordinary differential equations (ODE) or stochastic differential equations (SDE).

In Section 2, we briefly review the limiting properties of the CLS method and study conditions under which the approximate CLS method preserves the asymptotic properties of the CLS method. We detail the UT and UKF in Section 3. In Section 4, we study that convergence rate of the UKF and the large-sample properties of the UKF–CLS. The efficacy of the proposed method is then illustrated by simulations in Section 5 and a real application in Section 6. We briefly conclude in Section 7. All the proofs of theorems in this paper can be found in the online Supplementary Materials.

## 2. Approximate CLS

Let the conditional expectation of $\mathbf{y}_t$ given $\mathbf{y}_1, \ldots, \mathbf{y}_{t-1}$ be denoted by $\mathrm{E}_{\boldsymbol{\theta}}(\mathbf{y}_t|\mathbf{y}_1, \ldots, \mathbf{y}_{t-1}) = \mathrm{E}_{\boldsymbol{\theta}}(\mathbf{y}_{t|t-1})$; the conditional mean is the best 1-step ahead predictor in the sense of minimizing the mean squared prediction error. Then, the CLS method estimates the $a \times 1$ unknown true parameter value $\boldsymbol{\theta}_0$ by the argument, denoted by $\hat{\boldsymbol{\theta}}$, that minimizes the conditional sum of squares

$$Q_n(\boldsymbol{\theta}) = \sum_{t=1}^{n} (\mathbf{y}_t - \mathrm{E}_{\boldsymbol{\theta}}(\mathbf{y}_{t|t-1}))^2.$$

Theorem 2.2 of Klimko and Nelson (1978) established some general large-sample properties for the CLS estimator, which we now briefly describe. Let $|\cdot|$ be the Euclidean norm. Assume that the conditional expectation $\mathrm{E}_{\boldsymbol{\theta}}(\mathbf{y}_{t|t-1})$ is twice continuously differentiable in $\boldsymbol{\theta}$. Taylor expansion implies that for $\delta > 0$ and $|\boldsymbol{\theta} - \boldsymbol{\theta}_0| < \delta$, there exists $\boldsymbol{\theta}^*$ with $|\boldsymbol{\theta}_0 - \boldsymbol{\theta}^*| < \delta$ such that

$$Q_n(\boldsymbol{\theta}) = Q_n(\boldsymbol{\theta}_0) + (\boldsymbol{\theta} - \boldsymbol{\theta}_0)^T \frac{\partial Q_n(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} + \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)^T \mathbf{V}_n (\boldsymbol{\theta} - \boldsymbol{\theta}_0) + \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)^T \mathbf{U}_n(\boldsymbol{\theta}^*)(\boldsymbol{\theta} - \boldsymbol{\theta}_0),$$

where

$$\mathbf{V}_n = \left(\frac{\partial^2 Q_n(\boldsymbol{\theta}_0)}{\partial \theta_i \partial \theta_j}\right), \qquad \mathbf{U}_n(\boldsymbol{\theta}^*) = \left(\frac{\partial^2 Q_n(\boldsymbol{\theta}^*)}{\partial \theta_i \partial \theta_j}\right) - \mathbf{V}_n.$$

Under some mild regularity conditions, Klimko and Nelson (1978) derived the consistency and the large-sample distribution of $\hat{\boldsymbol{\theta}}$:

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \rightsquigarrow \mathrm{N}(\mathbf{0}, \mathbf{V}^{-1}\mathbf{W}\mathbf{V}^{-1}).$$

The CLS method is, however, not always applicable for estimating a nonlinear state-space model because the 1-step ahead predictor is generally intractable. This problem may be overcome if the 1-step ahead predictor can be well approximated by some numerical scheme. We show below that the approximate CLS estimator enjoys the same large-sample properties of the (exact) CLS estimator, if the approximation error is of the order $o_p(n^{-1})$. Moreover, we shall demonstrate in latter sections that the UKF provides such an approximation scheme. Let $\hat{\mathrm{E}}_{\boldsymbol{\theta}}(\mathbf{y}_{t|t-1})$ be an approximation of $\mathrm{E}_{\boldsymbol{\theta}}(\mathbf{y}_{t|t-1})$. Then, we have the following theorem: