Journal of Computational and Applied Mathematics 333 (2018) 200–214

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

Analysis, detection and correction of misspecified discrete time state space models

Salima El Kolei [a,*], Frédéric Patras [b]

[a] ENSAI, UBL, Campus de Ker Lann, rue Blaise Pascal, BP 37203, 35172, Bruz cedex, France
[b] Université Côte d'Azur, CNRS, UMR 7351, Parc Valrose, 06108 NICE Cedex 2, France

ARTICLE INFO

Article history:
Received 15 June 2017
Received in revised form 12 October 2017

Keywords:
Kalman filter
Extended Kalman filter
State space models
Misspecified models
Robust estimation

ABSTRACT

Misspecifications (i.e. errors on the parameters) of state space models lead to incorrect inference of the hidden states. This paper studies weakly nonlinear state space models with additive Gaussian noises and proposes a method for detecting and correcting misspecifications. The latter induce a biased estimator of the hidden state but also happen to induce correlation on innovations and other residues. This property is used to find a well-defined objective function for which an optimization routine is applied to recover the true parameters of the model. It is argued that this method can consistently estimate the bias on the parameter. We demonstrate the algorithm on various models of increasing complexity.

1. Introduction

This paper is concerned with the following family of discrete time state space models with additive Gaussian noises:

$$\begin{cases} x_t = b(\theta_0, x_{t-1}) + \beta_{\theta_0}\eta_t, \\ y_t = h(\theta_0, x_t) + \sigma_{\theta_0}\varepsilon_t. \end{cases} \tag{1}$$

The variables $\eta_t \sim \mathcal{N}(0, I_{n\times 1})$, $\varepsilon_t \sim \mathcal{N}(0, I_{m\times 1})$ are assumed to be independent standard normal variables, $t \in \mathbf{N}^*$, β_{θ_0} (resp. σ_{θ_0}) are $n\times n$ (resp. $m\times m$, with $\sigma_0\sigma_0^*$ positive definite) matrices, and θ_0 stands for the vector of parameters of the model. The functions b, h, β, α are assumed to be differentiable. The hidden states (or unobserved signal process) $\{x_t, t \in \mathbb{N}\}$ take value in $\mathcal{X} := \mathbf{R}^n$ and the observations $\{y_t, t \in \mathbf{N}^*\}$ in $\mathcal{Y} := \mathbf{R}^m$. We also denote the noise covariance matrices $R_{\theta_0} := \sigma_{\theta_0}\sigma_{\theta_0}^*$, $Q_{\theta_0} := \beta_{\theta_0}\beta_{\theta_0}^*$ where * stands for the transpose.

The aim of filtering is to make inference about the hidden state x_t conditionally to the observations $y_1, \dots, y_t$ denoted $y_{1:t}$ thereafter. In order to do so, there are various ways to estimate the parameters θ_0 that, in most situations of interest are unknown and have to be approximated. They may for example be estimated using standard techniques (MLEs...), or be incorporated to the set of random quantities to be estimated. To quote only one example in the recent literature, Particle Gibbs samplers have proven to be a good way to simulate the joint distribution of hidden processes and model parameters in hidden Markov chain models, see e.g. [1–3].

Here, we face a different problem: we consider the situation where θ_0 has been incorrectly estimated, for example using a given biased estimator $\hat{\theta}$ such that $\mathbb{E}[\hat{\theta}] = \theta = \theta_0 + \epsilon$ (the way the estimator has been devised is of no matter for our purposes). Our interest for these questions originated in the study of random volatility models such as Heston's, where some parameters are difficult to estimate. We wanted to understand how errors on the model parameters could impact the volatility estimates. The detection of errors method that is the purpose of the present article first arose from

* Corresponding author.
E-mail address: salima.el-kolei@ensai.fr (S.E. Kolei).

https://doi.org/10.1016/j.cam.2017.10.020

statistical phenomena detected in numerical simulations. We realized soon that the phenomena were universal, and related to theoretical properties of misspecified models. Application domains include for example engineering and control where the parameters θ may be known at inception but may change to a new value θ_0, for example due to a mechanical problem, so that θ becomes a wrong value for the true model parameters. Detecting the change from θ to θ_0 may then be useful not only to improve the inference process, but also to detect the underlying problem.

It is well-known that using such incorrect filter models deteriorates the filter performance and may even cause the filter to diverge. Various results have been obtained in the literature on the impact of ϵ on the estimator of the hidden state; error covariance matrices have been studied and compared with the covariance matrices of the conditional distribution of x_t and x_{t+1} knowing $y_{1:t}$. These results are described in [4], where the reader can also find a survey of the classical literature on the subject.

The aim of the present article is different: we want to take advantage of the theoretical properties of misspecified state space models, not only to understand the impact of ϵ on the estimation of the hidden states but also, ultimately, to use its statistical properties in order to get a correct set of parameters for the state space model.

The key result underlying our analysis is that misspecifications do certainly induce a biased estimator of the hidden state but also, and most importantly for our purposes, they happen to induce correlation on the innovations and other residues associated to observations. This property is used to find a well-defined objective function for which an optimization routine is applied to recover the true parameters of the model. It is argued that this method can consistently estimate the bias on the parameter. The method is easy to implement and runs fast. We demonstrate the algorithm on various models of increasing complexity.

Discrete time state space models are notoriously ubiquitous; their use is discussed in most textbooks on filtering from the early [5,4,6,7] to the recent literature we refer e.g. to [8] for a survey. Application domains of our results include, besides finance, control and engineering: ecology, economy, epidemiology, meteorology and neuroscience.

For example, our approach can be applied for neuronal dynamics where linear state space models are widely used due to their numerical efficiency. Parameter estimation in biophysical modeling of neurons and detection of changes in the behavior of the neuron cell membrane voltage constitute a problem of central importance (see [9,10] and [11]). For these models, we observe the time evolution of the neuron cell membrane voltage and can model the hidden variable, the variation of ion concentration, using a noisy discrete dynamical system. This variation of ion depends on various parameters θ_0 that control the voltage dependence of the steady states and relaxation times of activation and inactivation. For this model our method can be potentially useful to detect different patterns of electrical activity.

Another example of application where the approach can be useful is for multisensor-multitarget tracking. Typical applications of multisensor-multitarget tracking are in air traffic control [12], space surveillance and radar or visual tracking (see [13] and [14]). Most practical systems consist of a number of different sensors and the goal is to estimate the targets of interest from these different sensors. Many of these systems are modeled as discrete time linear or weakly nonlinear state space models as for example the target vehicle motion state radar estimation. The hidden variables, the position, the velocity and the acceleration of the target vehicle, can be estimated by the Kalman filter from the radar measurements (see [15]). For this kind of applications our approach leads to detect sudden vehicle movements, as for example an increasing velocity, or an abnormal deviation of the trajectory of the target due to a mechanical problem.

The paper is organized as follows. Section 2 presents the model assumptions and introduces various estimators and processes, including the "interpolation process" (3) that plays a central role in the article. Section 3 states the theoretical results. In Section 4, we describe the method and in the following one demonstrates the algorithms on three examples: the first application is largely pedagogical and studies an elementary autoregressive linear model for which our approach can be easily understood. We move then to a nonlinear (square root) model, and, to conclude, apply our approach to a complex and nonlinear model, that is the Heston model, widely used in finance for option pricing and portfolios hedging. The behavior of this last model when it comes to parameter estimation is notoriously difficult; our method behaves nevertheless quite satisfactorily. We compare finally our method and estimator (based on the interpolation process) with the estimator using the same strategy but based instead on innovations. Some concluding remarks are provided in Section 7. The technical proofs are gathered in Appendices A and B.

Notation: for any continuously differentiable function g, $[\partial g/\partial\theta]$ denotes the vector of the partial derivatives of g w.r.t θ.

2. The misspecified (extended) Kalman filter

In the linear case, the model (1) reads ($t \in \mathbf{N}^*$):

$$\begin{cases} x_t = u_t(\theta_0) + A_{\theta_0} x_{t-1} + \beta_{\theta_0}\eta_t \\ y_t = d_t(\theta_0) + C_{\theta_0} x_t + \sigma_{\theta_0}\varepsilon_t. \end{cases} \tag{2}$$

If the vector of parameters θ_0 is perfectly known, the optimal filtering $p_{\theta_0}(x_t|y_{1:t})$ is Gaussian and the Kalman filter gives exactly the two first conditional moments: $\hat{x}_t = \mathbb{E}[x_t|y_{1:t}]$ and $P_t = \mathbb{E}[(x_t - \hat{x}_t)(x_t - \hat{x}_t)^*|y_{1:t}]$. In particular, the Kalman filter estimator is the *BLUE* (Best Linear and Unbiased Estimator) among linear estimators.

In most real applications, the linearity assumption of the functions h and b is not satisfied. A linearization by a first order Taylor series expansion can be performed and the Extended Kalman filter (EKF) consists in applying the Kalman filter on this linearized model. Concretely, for the EKF, the matrix C_{θ_0} is the derivative of the function h with respect to (w.r.t.) x computed