Journal of Statistical Planning and Inference 140 (2010) 1889–1902

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

Parameter estimation of state space models for univariate observations

Marco Costa [a,*], Teresa Alpuim [b]

[a] Higher School of Technology and Management of Águeda, University of Aveiro, Portugal
[b] Department of Statistics and Operations Research, Faculty of Sciences of University of Lisbon, Portugal

ARTICLE INFO

Article history:
Received 14 May 2008
Received in revised form
10 December 2009
Accepted 19 January 2010
Available online 28 January 2010

MSC:
60G25
60G35

Keywords:
Kalman filter
State space model
Parameters estimation
Area rainfall estimates

ABSTRACT

This paper contributes to the problem of estimation of state space model parameters by proposing estimators for the mean, the autoregressive parameters and the noise variances which, contrarily to maximum likelihood, may be calculated without assuming any specific distribution for the errors. The estimators suggested widen the scope of the application of the generalized method of moments to some heteroscedastic models, as in the case of state-space models with varying coefficients, and give sufficient conditions for their consistency. The paper includes a simulation study comparing the proposed estimators with maximum likelihood estimators. Finally, these methods are applied to the calibration of the meteorological radar and estimation of area rainfall.

1. Introduction

The Kalman filter technique, Kalman (1960), has been used in many different areas to describe the evolution of dynamic systems. The main goal of the algorithm is to find estimates of unobservable variables based on related observable variables through a set of equations called a state space model (SSM). Such models are defined by the equations

$$\mathbf{Y}_t=\mathbf{H}_t\boldsymbol{\beta}_t+\mathbf{e}_t \tag{1}$$

$$\boldsymbol{\beta}_t=\boldsymbol{\Phi}\boldsymbol{\beta}_{t-1}+\boldsymbol{\varepsilon}_t \tag{2}$$

Eq. (1) is called the measurement equation and relates the $n\times 1$ vector of observable variables, $\mathbf{Y}_t$, with the $m\times 1$ vector of unobservable variables, $\boldsymbol{\beta}_t$, called states. The $n\times m$ matrix $\mathbf{H}_t$ is a matrix of known coefficients and $\mathbf{e}_t$ is a white noise $n\times 1$ vector, called the measurement error, with covariance matrix $E(\mathbf{e}_t\mathbf{e}_t')=\boldsymbol{\Sigma}_{\mathbf{e}}$. Further, the vector of states β_t varies in time according to Eq. (2), the transition or state equation. In this, $\boldsymbol{\Phi}$ is an $m\times m$ matrix of autoregressive coefficients and $\boldsymbol{\varepsilon}_t$ is a white noise $m\times 1$ vector with covariance matrix $E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t')=\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$. The disturbances $\mathbf{e}_t$ and $\boldsymbol{\varepsilon}_t$ are assumed to be uncorrelated, that is, $E(\mathbf{e}_t\boldsymbol{\varepsilon}_s')=\mathbf{0}$ for all t and s. One class of models with particular interest arises when the state vector is a stationary process with mean $E(\boldsymbol{\beta}_t)=\boldsymbol{\mu}$.

* Corresponding author. Tel.: +351 234 611 500; fax: +351 234 611 540.
E-mail address: marco@ua.pt (M. Costa).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2010.01.036

Briefly, the Kalman filter is an iterative algorithm that produces an estimator of the state vector $\boldsymbol{\beta}_t$ at each time t, which is given by the orthogonal projection of the state vector onto the observed variables up to that time. Thus, let $\hat{\boldsymbol{\beta}}_{t|t-1}$ represent the estimator of β_t based on the information up to time $t-1$, that is, based on $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_{t-1}$, and let $\mathbf{P}_{t|t-1}$ be its mean squared error (MSE) matrix. As the orthogonal projection is a linear estimator, the predictor for the next variable, $\mathbf{Y}_t$, is given by

$$\hat{\mathbf{Y}}_{t|t-1} = \mathbf{H}_t \hat{\boldsymbol{\beta}}_{t|t-1}.$$

When, at time t, $\mathbf{Y}_t$ is available, the prediction error or innovation, $\boldsymbol{\eta}_t = \mathbf{Y}_t - \hat{\mathbf{Y}}_{t|t-1}$, is used to update the estimate of $\boldsymbol{\beta}_t$ trough the equation

$$\hat{\boldsymbol{\beta}}_{t|t} = \hat{\boldsymbol{\beta}}_{t|t-1} + \mathbf{K}_t \eta_t,$$

where $\mathbf{K}_t$ is called the Kalman gain matrix and is given by

$$\mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{H}_t' (\mathbf{H}_t \mathbf{P}_{t|t-1}\mathbf{H}_t' + \boldsymbol{\Sigma}_{\mathbf{e}})^{-1}.$$

Further, the MSE of the updated estimator $\hat{\boldsymbol{\beta}}_{t|t}$ represented by $\mathbf{P}_{t|t}$ verifies the relationship

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{K}_t \mathbf{H}_t \mathbf{P}_{t|t-1}.$$

In its turn, at time t, the forecast for the state vector $\boldsymbol{\beta}_{t+1}$ is given by the equation

$$\hat{\boldsymbol{\beta}}_{t+1|t} = \boldsymbol{\Phi}\hat{\boldsymbol{\beta}}_{t|t}$$

and its MSE matrix is

$$\mathbf{P}_{t+1|t} = \boldsymbol{\Phi}\mathbf{P}_{t|t}\boldsymbol{\Phi}' + \boldsymbol{\Sigma}_{\mathbf{e}}.$$

This recursive process needs initial values for the state vector, $\hat{\boldsymbol{\beta}}_{1|0}$, and for its MSE, $\mathbf{P}_{1|0}$ that will be seen with more detail in Section 2. As it is well known, the orthogonal projection corresponds to the best linear unbiased predictor. When the disturbances $\mathbf{e}_t$ and $\boldsymbol{\varepsilon}_t$ are normally distributed the state vector and the observed variables are also normal. Therefore, in this case, the orthogonal projection is also the conditional mean value and the Kalman filter is optimal.

A problem that frequently arises in practical applications of this algorithm is that the vector of parameters $\boldsymbol{\Theta} = \{\boldsymbol{\Phi}, \boldsymbol{\Sigma}_{\mathbf{e}}, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}\}$ is not known and needs to be estimated. Under the assumption of normality, the log-likelihood of a sample $(\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_n)$ can be written through conditional distributions, that is

$$\log L(\boldsymbol{\Theta}; \mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_n) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{n}\log(|\boldsymbol{\Omega}_t|) - \frac{1}{2}\sum_{t=1}^{n}\eta_t' \boldsymbol{\Omega}_t^{-1}\eta_t$$

where $\boldsymbol{\Omega}_t = \mathbf{H}_t \mathbf{P}_{t|t-1}\mathbf{H}_t' + \boldsymbol{\Sigma}_{\mathbf{e}}$. Thus, it is possible to obtain the maximum likelihood estimates maximizing the loglikelihood in order to the unknown parameters using numerical algorithms, namely, the EM algorithm (Dempster et al., 1977) or the Newton–Raphson algorithm (Harvey, 1996).

However, in some cases, the assumption of normality is not most appropriate as it may be seen, for example, in Alpuim and Barbosa (1999) or in Alpuim and Ribeiro (2003). Moreover, the use of numerical methods to maximize this likelihood function may be a difficult and complex task, very often without satisfactory results. These may occur either because the numerical iterative techniques do not converge or because of the existence of multiple critical points. The SSM with errors distributed accordingly the exponential family has been studied since the pioneering work of West et al. (1985), and later on, in the book of West and Harrison (1989). Kitagawa (1987) and Fahrmeir (1992) developed the SSM with non-Gaussian errors. Nevertheless, these authors do not study in detail the parameters estimation method.

The parameters estimation problem in the SSM with non-Gaussian errors was treated in more detail in later works which focus on the method of maximum likelihood estimation associated with Bayesian techniques (Carlin et al., 1992; Shephard and Pitt, 1997). Although Alpuim (1999) considers estimators for the errors variances based on the method of moments, the estimation of other parameters, as the autoregressive parameters, has not an easy solution.

In this paper, we suggest estimators based on the generalized method of moments (GMM) adjusted to accommodate some models with heteroscedasticity and give sufficient conditions for its consistency. We design a simulation study to compare the performance of the proposed methods of estimation with others, namely, the maximum likelihood and the estimators suggested by Alpuim (1999).

Hansen (1982) made the first important contribution to the GMM proving the strong consistency and asymptotic normality of these estimators under the assumption of stationarity and ergodicity of the observable variables. This result allowed the application of GMM in many different contexts and Hansen and West (2002) consider the use of GMM in econometric time series as a major development in this area. To reinforce this statement, we cite the problem of the estimation of a first-order condition or decision rule in a dynamic optimization problem which leading example was an Euler equation for consumption (Dynan, 2000) or the problem of forecasting ability over a multiperiod horizon of a financial variable (Mishkin, 1990). Another example, in actuarial sciences, is given by Alpuim and Ribeiro (2003) where the authors suggest a space state model to describe run-off triangles and use the Kalman filter to predict claim reserves. They use the GMM approach to estimate the parameters of the state space model.