Available online at www.sciencedirect.com
ScienceDirect
Journal of Statistical Planning and Inference 137 (2007) 3113–3126
journal of statistical planning and inference
www.elsevier.com/locate/jspi

A two-state regime switching autoregressive model with an application to river flow analysis

Krisztina Vasas*, Péter Elek, László Márkus

Department of Probability Theory and Statistics Eötvös Loránd University, Pázmány Péter sétány 1/C, H-1117 Budapest, Hungary

Received 11 October 2005; accepted 9 May 2006
Available online 16 March 2007

Abstract

We propose a regime switching autoregressive model and apply it to analyze daily water discharge series of River Tisza in Hungary. The dynamics is governed by two regimes, along which both the autoregressive coefficients and the innovation distributions are altering, moreover, the hidden regime indicator process is allowed to be non-Markovian. After examining stationarity and basic properties of the model, we turn to its estimation by Markov Chain Monte Carlo (MCMC) methods and propose two algorithms. The values of the latent process serve as auxiliary parameters in the first one, while the change points of the regimes do the same in the second one in a reversible jump MCMC setting. After comparing the mixing performance of the two methods, the model is fitted to the water discharge data. Simulations show that it reproduces the important features of the water discharge series such as the highly skewed marginal distribution and the asymmetric shape of the hydrograph.

MSC: 60G07; 62F15; 62M05; 62M09

Keywords: Change point detection; Reversible jump MCMC; Regime switching autoregressive model; River flow analysis

1. Introduction

In the analysis of time series data one can often encounter examples when a phenomenon features different behaviors in distinct but randomly changing time periods. A traditional model with fixed dynamical structure fitted to such data often leads to a complicated and artificial model structure. A straightforward solution is to allow the dynamics to change across regimes, governed by a latent external force. (We do not consider here models where the regimes are determined by observable states of the process itself, e.g. self-exciting threshold autoregressive (SETAR) processes.) These regime switching models have the advantage that their latent regime structure can often be interpreted as a proxy for unobserved phenomena. Within this family the regime switching autoregressive models deserve particular attention as they represent a straightforward generalization to linear models.

Among the numerous fields where regime switching models are applied, hydrology and in particular river flow analysis is one of the best known. Forecasting water discharge series and simulating realistic river flow scenarios are

* Corresponding author. Tel.: +36 1 381 2151; fax: +36 1 381 2152.
E-mail addresses: v_krisz@cs.elte.hu (K. Vasas), elekpeti@cs.elte.hu (P. Elek), markus@cs.elte.hu (L. Márkus).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2006.05.019

very important both for flood control and for irrigation purposes. (A general introduction to stochastic hydrology is given in Lawrence and Kottegoda, 1977; Hipel and McLeod, 1994; Kottegoda, 1980.) As documented in e.g. Elek and Márkus (2004), traditional linear models—even after incorporating long range dependence, seasonal effects and non-Gaussian innovations—do not provide good hydrological forecasts and simulations. Alternatively, one can develop conditionally heteroscedastic models (Elek and Márkus, 2004, 2007), shot-noise processes (e.g. Konecny, 1992), neural networks (Hsu et al., 1995) or regime switching models (e.g. Aksoy, 2003; Szilágyi et al., 2006; Lu and Berliner, 1999) to improve model performance. Among these, regime switching models have the advantage that they can be easily interpreted in physical terms with the latent regimes corresponding to wet and dry, or, alternatively, rising and falling periods. In fact, since precipitation data are rarely available for the whole river catchment, the use of latent regimes as proxies of rainfall is often the only way to incorporate the physical properties into the modeling process.

In this paper we introduce a two-state regime switching autoregressive model and apply it to daily water discharge series of River Tisza in Hungary. Compared to the more commonly used models, ours has two distinguishing features. First, the innovation distribution changes across the regimes along with the autoregressive coefficients. The process has independent, gamma distributed increments in the so-called ascending regime—representing the positive shocks to the system—while it behaves as a Gaussian autoregression in the other (descending) regime. This choice gives the possibility to model the short rising and longer falling periods separately, which is a distinguishing feature of river flow series.

As a second characteristics of our model, the latent regime process is allowed to be non-Markovian. In the Markov-switching case the regime durations are independent geometrically distributed random variables whereas in our more general setting they are independent negative binomially distributed ones.

Estimation of the model is complicated by the presence of the latent structure but can be carried out using Markov Chain Monte Carlo (MCMC) methods. There are two possible approaches. In the first one the values of the latent state process are regarded as auxiliary parameters and the full conditionals for them are obtained in terms of the neighboring state values. This approach is easy to implement in the Markov-switching case and can be applied to the more general setting as well. Alternatively, the change points of the regimes can be chosen as auxiliary parameters, hence the problem can be regarded as change point detection. The real interest, however, lies in the underlying structural parameters and not in the actual positions of the change points. In this case, reversible jump MCMC (introduced in Green, 1995) is needed to alter the number of change points. (See Punskaya et al., 2002 for the use of this method in a similar problem.) Our approach is related to Lavielle and Labarbier (2001) as they too give both a fixed-dimension and a variable-dimension MCMC-algorithm for a change point detection problem.

The paper is organized as follows. Section 2 introduces the model and discusses its basic properties such as stationarity, moments and autocorrelation structure. Section 3 describes the estimation for the Markov-switching case and Section 4 contains details about the estimation in the general case. Section 5 applies the model to river flow data and compares the convergence properties of the MCMC algorithms. Finally, Section 6 concludes and summarizes the results. Proofs of the model properties are deferred to Appendix A.

2. The model and its basic properties

Assume that the process Y_t is governed by the hidden regime process I_t in the following way:

$$Y_t = Y_{t-1} + \varepsilon_{1,t} \quad \text{if } I_t = 0, \tag{1}$$

$$Y_t = a(Y_{t-1} - c) + c + \varepsilon_{2,t} \quad \text{if } I_t = 1, \tag{2}$$

where $\varepsilon_{1,t}$ is an i.i.d. sequence distributed as $\Gamma(\alpha, \lambda)$ (i.e. as a gamma distribution with shape parameter α and scale parameter λ) and $\varepsilon_{2,t}$ is an i.i.d. Gaussian sequence with zero mean and σ^2 variance. α, λ and σ are positive real numbers and we assume that $0 < a < 1$. The duration of the $I_t = 0$ regime is distributed as negative binomial with parameters (b, p_0) and the duration of the $I_t = 1$ regime is geometrically distributed with parameter p_1, where $b > 0$ and $0 < p_i < 1$ $(i = 0, 1)$. The negative binomial distribution is used in the following parametrization (here Γ denotes the usual Γ-function):

$$P(N = k) = \frac{\Gamma(b + k - 1)}{\Gamma(b)\Gamma(k)} p^b (1 - p)^{k-1}, \quad k = 1, 2, \ldots .$$