STATISTICAL ANALYSIS OF STOCHASTIC VOLATILITY MODELS

Ph.D. Thesis Summary

Eötvös Loránd University, Faculty of Sciences

ZSANETT ORLOVITS

SUPERVISOR: László Gerencsér, D.Sc.
Scientific Adviser

DOCTORAL SCHOOL OF MATHEMATICS
Director: Miklós Laczkovich, member of HAS

DOCTORAL PROGRAMME OF APPLIED MATHEMATICS
Director: György Michaletzky, D.Sc.

Computer and Automation Institute of the Hungarian Academy of Sciences,
Stochastic Systems Research Group
and
Budapest University of Technology and Economics
Mathematical Institute

Budapest, July 2011

1 Introduction

This thesis deals with some aspects of the statistical analysis of certain class of stochastic volatility models. Stochastic volatility models most frequently appears in modelling financial time series. Several models have been suggested to capture special features of financial time series such as volatility clustering, which means that long periods of low volatility are followed by short periods of high volatility. This immediately implies that a minimum requirement for a mathematical model of these time-series is to ensure that the conditional variance of the observation process is time-varying. Nowadays, the most widely accepted non-linear stochastic volatility model for modelling financial time series is the generalized autoregressive conditional heteroscedastic (GARCH) model introduced by Bollerslev [4].

The literature on the estimation, or identification, of the parameters of GARCH processes is almost exclusively devoted to off-line quasi maximum likelihood method. However, given that financial time series are often sampled at *high frequency*, a more convenient, and less expensive approach would be to use an *on-line* or recursive method. While the off-line estimation have been analysed under a variety of conditions in the literature, see e.g. Berkes et al. [3] and the references therein, the recursive estimation of GARCH processes has attracted little attention until recently, see [1, 6, 10].

The main contribution of this thesis is to propose a new approach to the recursive estimation of GARCH processes which is a natural adaptation of the off-line quasi-maximum likelihood method. The theoretical analysis of this method requires a deep mathematical technology developed in Benveniste et al. [2]. We present the proof of almost sure convergence and L_q-convergence up to certain q-s under reasonable and verifiable technical conditions, using the results of Gerencsér and Mátyás [9]. The successful adaptation of the theory of Benveniste et al. [2] has led to another powerful result: using the results of the above theory and the techniques of Gerencsér [8] we prove a strong approximation theorem for the error term of the off-line maximum likelihood estimator.

2 Basic properties of GARCH models

In this thesis we analyse the most widely used model in financial application, namely the GARCH model. The precise definition of a GARCH(r, s) process is given as follows: (y_n), $n \in \mathbb{Z}$ is called a GARCH(r, s) process if it satisfies

$$y_n = \sigma_n \varepsilon_n, \quad n \in \mathbb{Z}, \tag{1}$$

$$(\sigma_n^2 - \gamma^*) = \sum_{i=1}^{r} \alpha_i^* (y_{n-i}^2 - \gamma^*) + \sum_{j=1}^{s} \beta_j^* (\sigma_{n-j}^2 - \gamma^*), \quad n \in \mathbb{Z}, \tag{2}$$

where σ_n^2 is the conditional variance of y_n given its own past up to time $(n-1)$, (ε_n) is an independent and identically distributed (i.i.d.) sequence of random variables with zero mean and unit variance, $\gamma^* = \mathrm{E}y_{n-i}^2 = \mathrm{E}\sigma_{n-j}^2 > 0$ and $\alpha_i^*, \beta_j^* \geq 0$, $i = 1,\ldots,r$, $j = 1,\ldots,s$ denote the true, unknown parameters of the model. Defining the polynomials

$$C^*(z^{-1}) = \sum_{i=1}^{r} \alpha_i^* z^{-i}, \quad D^*(z^{-1}) = 1 - \sum_{j=1}^{s} \beta_j^* z^{-j},$$

equation (2) can be written in a compact form as

$$D^*(z^{-1})(\sigma^2 - \gamma^*) = C^*(z^{-1})(y^2 - \gamma^*),$$

where z^{-1} is the backward shift operator. In the following we will assume that the polynomials C^* and D^* are stable and relative prime. Defining the state vector

$$X_n^* = (y_n^2, \ldots, y_{n-r+1}^2, \sigma_n^2, \ldots, \sigma_{n-s+1}^2)^T,$$

it is easy to see that X_n^* satisfies a linear state equation

$$X_{n+1}^* = A_{n+1}^* X_n^* + u_{n+1}^*, \quad n \in \mathbb{Z}, \tag{3}$$

with

$$A_n^* = \begin{pmatrix} \eta^* \varepsilon_n^2 & \xi^* \varepsilon_n^2 \\ \bar{S} & 0 \\ \eta^* & \xi^* \\ 0 & \bar{S} \end{pmatrix} \quad \text{and} \quad u_n^* = \begin{pmatrix} \alpha_0^* \varepsilon_n^2 \\ 0 \\ \alpha_0^* \\ 0 \end{pmatrix},$$

where $\eta^* = (\alpha_1^*, \ldots, \alpha_r^*)$, $\xi^* = (\beta_1^*, \ldots, \beta_s^*)$ and $\bar{S}$ is the shift matrix having 1-s in the sub-diagonal, and 0-s elsewhere. It is easy to see that X_n^* is a Markov process.

A necessary and sufficient condition for the existence of a second order stationary solution of (1) and (2) is that

$$\sum_{i=1}^{r} \alpha_i^* + \sum_{j=1}^{s} \beta_j^* < 1, \tag{4}$$

see Bollerslev [4]. Conditions for the existence of a unique, strictly stationary and causal solution of (1) and (2), without restrictions on the moments, can be formulated in terms of a state-space representation, (3). Thus, consider the so-called generalized linear stochastic model (3) without specifying the matrices $A_n^* = A_n$ and the vectors $u_n^* = u_n$, instead, assuming only the following condition:

Condition 2.1 *Assume that* (A_n, u_n) *is an i.i.d. sequence of random matrices, and the* σ*-fields*

$$\mathcal{F}^A_{n+} = \sigma\{A_i : i > n\}, \quad \mathcal{F}^u_{n-} = \sigma\{u_i : i \leq n\}$$

are independent for any n*.*

Theorem 2.1 (Bougerol and Picard [5]) *Consider the generalized linear stochastic model (3) satisfying Condition 2.1. Assume that the model (3) is irreducible and that both* $\mathrm{E}\log^+ \|A_0\|$ *and* $\mathrm{E}\log^+ |u_0|$ *are finite. Then (3) has a strictly stationary solution if and only if the top-Lyapunov exponent*

$$\lambda(\mathcal{A}) := \lim_{n\to\infty} \frac{1}{n} \log \mathrm{E}\|A_n \ldots A_1\|$$

is strictly negative.

It is often not only necessary for the model in consideration to have a stationary and ergodic solution, but also it has finite moments of appropriate order. The existence of finite even order moments of X_n^* was studied by Feigin and Tweedie [7], using Markov-chain techniques, under the following condition, which is more restrictive than Condition 2.1:

Condition 2.2 *Assume that* (A_n, u_n) *is an i.i.d. sequence of random matrices and* (A_n) *and* (u_n) *are also independent of each other.*

In this thesis we provide a simple proof of the key step in deriving the result of Feigin and Tweedie [7] under Condition 2.1, which is weaker then Condition 2.2.

Theorem 2.2 *Let* (A_n) *be an i.i.d. sequence of random matrices such that* $||A_1|| \in L_q$*. Assume that for some even integer* $q \geq 2$

$$\rho\ [\ \mathrm{E}(A_1^{\otimes q})\ ]\ < 1,$$

where ρ *is the spectral radius. Then*

$$\lambda_q = \lambda_q(\mathcal{A}) := \lim_{n\to\infty} \frac{1}{n} \log \mathrm{E}\|A_n \ldots A_1\|^q < 0.$$

It follows that for any $\varepsilon > 0$ *we have*

$$\mathrm{E}\|A_n \ldots A_1\|^q \leq C e^{(\lambda_q + \varepsilon)n}$$

with some $C = C(\varepsilon) > 0$*.*

The above theorem is a direct corollary of the following proposition.

Proposition 2.1 *Let* (A_n) *be an i.i.d. sequence of random matrices as above. Then*

$$\mathrm{E}\|A_n \ldots A_1\|^q = (\mathrm{vec}I^T) \cdot\ [\ \mathrm{E}(A_1^{\otimes q})\ ]^n \cdot (\mathrm{vec}I),$$

where I is a unit matrix of appropriate dimension.

3 Estimation of the parameters of GARCH processes

For the identification of the parameters of GARCH processes we proceed similarly to ARMA processes. Write $\theta = (\alpha_0, \alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_s)^T$ and let $K \subset \mathbb{R}^{r+s+1}$ denote the set of θ-s such that the stability condition (4) holds and the polynomials C and D, defined by these θ-s, are stable. Let $K_0 \subset \text{int}\, K$ be a compact domain such that $\theta^* \in \text{int} K_0$. For a fixed tentative value of the system parameters, say θ, we *invert* the system

$$\bar{\sigma}_n^2(\theta) - \gamma = \sum_{i=1}^{r} \alpha_i (y_{n-i}^2 - \gamma) + \sum_{j=1}^{s} \beta_j (\bar{\sigma}_{n-j}^2(\theta) - \gamma)$$

to get the *frozen parameter* process $\bar{\sigma}_n^2(\theta)$, using the initial values $y_n = 0$ and $\bar{\sigma}_n^2(\theta) - \gamma = 0$ for all $n \leq 0$. Then we compute the estimated driving noise $\bar{\varepsilon}_n(\theta)$ by the inverse equation

$$\bar{\varepsilon}_n(\theta) = \frac{y_n}{\bar{\sigma}_n(\theta)}$$

with $n \geq 0$. Assuming now that ε_n is standard normal, the (conditional) log-likelihood function, modulo a constant, can be expressed via $\bar{\sigma}_n^2(\theta)$ as

$$L_N = \sum_{n=1}^{N} l_n = \sum_{n=1}^{N} -\frac{1}{2} \left(\log \bar{\sigma}_n^2(\theta) + \frac{y_n^2}{\bar{\sigma}_n^2(\theta)} \right)$$

where N denotes the sample size. The right hand side depends on θ^* via (y_n^2). To stress dependence of L_N on both θ and θ^*, we shall write $L_N = L_N(\theta, \theta^*)$. Then the general abstract estimation problem, assuming for a moment *stationary initialization*, is to find the solution of the non-linear algebraic equation

$$\mathrm{E}_{\theta^*} \frac{\partial}{\partial \theta} l_0(\theta, \theta^*) = 0. \tag{5}$$

The conditional quasi-maximum likelihood estimation $\hat{\theta}_N$ of θ^* is defined as the solution of the equation

$$\frac{\partial}{\partial \theta} L_N(\theta, \theta^*) = L_{\theta N}(\theta, \theta^*) = 0.$$

Define the asymptotic cost function, (a negative log-likelihood for the gaussian case) as

$$W(\theta, \theta^*) = \lim_{n \to \infty} \mathrm{E} \left[\frac{1}{2} \left(\log \bar{\sigma}_n^2(\theta) + \frac{y_n^2}{\bar{\sigma}_n^2(\theta)} \right) \right].$$

Then the asymptotic estimation problem is

$$\frac{\partial}{\partial\theta}W(\theta,\theta^*) = \lim_{n\to\infty} \mathrm{E}\frac{\bar{\sigma}_{\theta,n}(\theta)}{\bar{\sigma}_n^3(\theta)}\left(1 - \frac{y_n^2}{\bar{\sigma}_n^2(\theta)}\right) = 0. \tag{6}$$

The on-line algorithm: For the solution of the general estimation problem (5) the following stochastic approximation procedure is proposed: starting with some initial condition $\theta_0 \in K_0$ we define recursively

$$\theta_n = \theta_{n-1} - \frac{1}{n}\frac{\sigma_{\theta,n}}{\sigma_n}\left(1 - \frac{y_n^2}{\sigma_n^2}\right), \tag{7}$$

where σ_n and $\sigma_{\theta,n}$ denote the on-line estimates of $\bar{\sigma}_n(\theta_{n-1})$ and $\bar{\sigma}_{\theta,n}(\theta_{n-1})$, respectively. Thus, at time n, the volatility process σ is generated via the feedback equation

$$[D_{n-1}(z^{-1})(\sigma^2 - \gamma_{n-1})]_n = [C_{n-1}(z^{-1})(y^2 - \gamma_{n-1})]_n$$

with $D_{n-1} = D(z^{-1}, \theta_{n-1})$, and similarly for C_{n-1}. The algorithm is then completed by adding a resetting mechanism following Gerencsér and Mátyás [9].

The analysis of algorithm (7) is based on the stochastic approximation theory developed by Benveniste et al. [2], (will be shortly called as BMP), in which a methodology is based on the theory of Markov processes, and its modification by a resetting mechanism, given in Gerencsér and Mátyás [9]. To formulate the problem of estimating the GARCH parameters into a form that fits the framework of the BMP theory, we need to modify the state vector X_n^* first by the vector $\bar{Z}_n(\theta) = (\bar{\sigma}_n^2(\theta), \ldots, \bar{\sigma}_{n-s+1}^2)^T$. While the algorithm contains the derivative of the log-likelihood function with respect to θ, the resulting state vector $\bar{X}_n^e(\theta) = (X_n^*, \bar{Z}_n(\theta))^T$ will be further extended by its derivative with respect to θ. This leads us to a parameter dependent Markov process $\bar{\psi}_n(\theta)$ which follows a linear dynamics

$$\bar{\psi}_{n+1}(\theta) = P_{n+1}(\theta)\bar{\psi}_n(\theta) + w_n(\theta), \tag{8}$$

with a *block-triangular* state transition matrix

$$P_n(\theta) = \begin{pmatrix} A_n^e(\theta) & 0 \\ x & \tilde{A}_n^e(\theta) \end{pmatrix},$$

where $\tilde{A}_n^e(\theta) = \mathrm{diag}(A_n^e(\theta), \ldots, A_n^e(\theta))$ and $A_n^e(\theta)$ denotes the state-matrix of $\bar{X}_{n,\theta}^e(\theta)$. Note that $A_n^e(\theta)$ is stable due to the assumed stability of $D(z^{-1})$.

This implies that the asymptotic estimation problem (6) can be formulated in terms of $\bar{\psi}_n(\theta)$, thus the BMP-theory is applicable.

4 The BMP scheme

This section summarizes the basics of the general theory of recursive estimation as presented in Benveniste et al. [2], Chapter 2, Part II.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $(\bar{X}_n(\theta))$, with $\theta \in D \subset \mathbb{R}^d$, with D being a measurable open set in $\mathbb{R}^d$, be an $\mathbb{R}^k$-valued Markov-chain over $(\Omega, \mathcal{F}, P)$ with transition kernel $\Pi_\theta(x, A)$, having a unique invariant measure μ_θ. The initial state $\bar{X}_0(\theta)$ is assumed to have distribution μ_θ. Let H be a mapping from $D \times \mathbb{R}^k$ to $\mathbb{R}^d$. Then the basic estimation problem of the BMP-theory is to solve the equation

$$E_{\mu_\theta} H(\theta, \bar{X}_n(\theta)) = 0,$$

using observed values of $H(\theta, \bar{X}_n(\theta))$, or their computable approximations. We assume that a unique solution $\theta^* \in D$ exists. For the solution of the above problem the following stochastic approximation procedure is proposed: starting with some initial condition $\theta_0 = \xi$ define recursively

$$\theta_n = \theta_{n-1} + \frac{1}{n} H(\theta_{n-1}, X_n), \tag{9}$$

with $X_0 = x_0 \in \mathbb{R}^k$ being a possibly random initial state, and X_n is a non-homogeneous Markov chain defined by

$$P(X_{n+1} \in A | \mathcal{F}_n) = \Pi_{\theta_n}(X_n, A).$$

Here $\mathcal{F}_n$ is the σ-field of events generated by the random variables $X_0, \ldots, X_n$, and A is any Borel subset of $\mathbb{R}^k$.

Convergence of the estimator sequence (θ_n) with probability *strictly less than 1* has been proved in Benveniste et al. [2]. A BMP-scheme modified by an appropriate resetting mechanism developed by Gerencsér and Mátyás [9] provides the convergence of the estimator sequence with *probability* 1 to θ^*, under reasonable technical conditions.

The technically most demanding condition of the BMP theory requires the L_q-stability of the underlying Markov chain for an appropriately large q. The foundations of the verification of this condition are examined in the next section.

5 Stability properties of block-triangular stationary random matrices

Since the asymptotic estimation problem (6) is formulated in terms of a linear stochastic system with block-triangular state matrix, our goal is to extend the stability properties of general linear stochastic systems to those of the extended system (8). Motivated by this problem, the first result of this section is connected to the top-Lyapunov exponent of the extended system (8).

Theorem 5.1 *Let*

$$A_n = \begin{pmatrix} A_n^1 & 0 \\ B_n & A_n^2 \end{pmatrix}$$

be a stationary, ergodic sequence of $(d_1+d_2)\times(d_1+d_2)$ *matrices, satisfying* $\mathrm{E}\log^+\|A_n\| < +\infty$, *with* A_n^1 *and* A_n^2 *being square matrices. Then*

$$\lambda(\mathcal{A}) = \max(\lambda(\mathcal{A}^1), \lambda(\mathcal{A}^2)).$$

This result was the first attempt to verify the BMP-conditions, but it is an interesting and useful result on its own. In our setup, the key step of the verification of the BMP conditions is the L_q-stability of products of block-triangular stationary random matrices. Theorem 2.2 shows that a sufficient condition for L_q-stability of a product of an i.i.d. sequence of random matrices A_n, with even q, is that $\rho[\mathrm{E}(A_n^{\otimes q})] < 1$. By combining this result and the arguments of Theorem 5.1 we prove that the validity of the above condition for the diagonal blocks implies its validity for the full matrix.

Theorem 5.2 *Let*

$$A = \begin{pmatrix} A_1 & 0 \\ B & A_2 \end{pmatrix}$$

be a random $(d_1+d_2)\times(d_1+d_2)$ *matrix in* $L_2(\Omega, \mathcal{F}, P)$, *with* A_1 *and* A_2 *being square matrices. Then*

$$\rho[\ \mathrm{E}(A\otimes A)\ ] \ = \max\{\rho[\ \mathrm{E}(A_1\otimes A_1)\ ]\ ; \rho[\ \mathrm{E}(A_2\otimes A_2)\ ]\ \}.$$

Similarly, let q *be a positive integer, even or odd and let us assume that* $A \in L_q(\Omega, \mathcal{F}, P)$. *Then*

$$\rho[\ \mathrm{E}(A^{\otimes q})\ ] \ = \max\{\rho\ [\ \mathrm{E}(A_1^{\otimes q})\ ]\ ; \rho\ [\ \mathrm{E}(A_2^{\otimes q}\ )\ ]\}.$$

6 Convergence of the recursive algorithm

The main contribution of this part of the thesis is a rigorous convergence analysis of the recursive estimation method for the parameters of GARCH processes. To prove the convergence of (7), suitably modified with resetting, we need to verify the general conditions of the BMP-theory. By exploiting the special block-triangular structure of the extended system (8) and combining the results of Theorem 2.2 and 5.2, we prove the following theorem:

Theorem 6.1 *Let* $D^*(z^{-1})$ *be stable, and let* ε, *the limiting rate of* θ_n, *be sufficiently small. Let the truncation domain* D_0 *be such that for all* $\theta \in D_0$ *the corresponding polynomial* $D(z^{-1})$ *is stable, and let Condition 5.4.1 be satisfied. Assume further that with some even* $q \geq 6$

$$\rho[\mathrm{E}(A_0^*)^{\otimes q}] < 1 \tag{10}$$

is satisfied. Then the estimator sequence θ_n given by (7), and modified by a resetting mechanism converges to θ^* with probability 1*, and also in L_q, with rate*

$$\mathrm{E}^{1/q}|\theta_n - \theta^*|^q = O(n^{-(\alpha\wedge\frac{1}{2})}),$$

where $-\alpha < 0$ is the Lyapunov-exponent of the associated ODE (ordinary differential equation).

For Condition 5.4.1 and the definition of the limiting rate ε see Section 5.4.1 of the thesis or Gerencsér and Mátyás [9].

Finally, we also test the performance of our algorithm via simulated data. The first empirical studies indicate that the accuracy of the final estimator depends strongly on the initial values of the estimator. Secondly, we consider GARCH$(1,1)$ processes with stability margins $1-\alpha_1^*-\beta_1^*$ as small as 2% and 1%, driven by Gaussian white noise with zero mean and unit variance. In both cases the proposed algorithm performs remarkably well in spite of the fact that condition (10) is unlikely to be satisfied even with $q = 6$, since $\mathrm{E}{\varepsilon_0}^6 = 15$. Finally, we test the viability of our algorithm for real data by fitting a GARCH$(1,1)$ model. The results show remarkable accuracy of even such as simple model, and a good agreement in the performance of the off-line and the on-line estimator.

7 Strong approximation of GARCH processes

In this section we present a strong approximation result for GARCH processes with error terms, the moments of which are bounded from above by a tight upper bound, under conditions that is required by the techniques of the BMP-theory. Our result is an extension of a similar result for ARMA processes, given in Gerencsér [8].

Theorem 7.1 *Let $D^*(z^{-1})$ be stable, and assume further that for some positive even q*

$$\rho[\mathrm{E}(A_0^*)^{\otimes q}] < 1$$

is satisfied. Then

$$\hat{\theta}_N - \theta^* = -(R^*)^{-1}\frac{1}{N}\sum_{n=1}^{N} l_{\theta,n}(\theta^*,\theta^*) + r_N,$$

where R^* is the Fisher-information matrix and for the error term r_N we have for all $1 \le k \le \frac{q}{2d}$, $d > \dim\theta$,

$$\mathrm{E}^{1/k}|r_N|^k = O(N^{-1}).$$

The main contribution of this result is that the moments of the error term up to certain order is shown to be $O(N^{-1})$, and the dominant term is a martingale. This ensures that all basic limit theorems, that are known for the dominant term, which is a martingale, are also valid for the difference $\hat{\theta}_N - \theta^*$. The key step of the proof is a uniform strong law of large numbers for the log-likelihood function and its derivatives with respect to θ.

8 Conclusions

In this thesis, a new approach to the recursive estimation of GARCH processes has been proposed which is a natural adaptation of the off-line quasi-maximum likelihood method. Surprisingly, the theoretical analysis of this method requires a deep mathematical technology developed in Benveniste et al. [2]. Convergence in almost sure sense and in L_q has been established under reasonable and verifiable technical conditions, with the exception of (10). The best alternative to our method is the two pass recursive estimation method due to Aknouche and Guerbyenne [1]. However, the technical conditions under which their results are valid are not fully specified.

Our method exploits the relatively simple structure of the dynamics of the model, allowing its inversion. This observation implies further applicability of our method, especially for the several generalizations of the GARCH model, since many of them can be formulated in the form of the linear stochastic system (3). A further application would be the class of general bilinear models.

A second effort of this research was to develop a characterization theorem for the error term of the off-line maximum likelihood estimation following the arguments of Gerencsér [8]. Motivated by this result a further development of our research would be the detection of the changes of the parameters in the GARCH model in the off-line case using exponential forgetting.

Publications of the author

The thesis is based on the following journal articles and on one yet unpublished manuscript:

- L. Gerencsér, Gy. Michaletzky and Zs. Orlovits, Stability of block-triangular stationary random matrices, *Systems and Control Letters* 57 (2008) 620-625.
- L. Gerencsér and Zs. Orlovits, L_q-stability of products of block-triangular stationary random matrices, *Acta Scientiarum Mathematicarum (Szeged)* 74 (2008) 927-944.
- L. Gerencsér and Zs. Orlovits, Real time estimation of stochastic volatility processes, *Annals of Operations Research*, Special Volume: Stochastic Modeling and Optimization. Dedicated to the 80th Birthday of Professor András Prékopa, accepted for publication.
- L. Gerencsér and Zs. Orlovits, Strong approximation of GARCH models, manuscript.

References

[1] A. Aknouche and H. Guerbyenne, Recursive estimation of GARCH models, Communication in Statistics-Simulation and Computation 35 (2006) 925-938.

[2] A. Benveniste, M. Métivier and P. Priouret, Adaptive Algorithms and Stochastic Approximations, Springer Verlag, Berlin, 1990.

[3] I. Berkes, L. Horváth, P.S. Kokoszka, GARCH processes: structure and estimation, Bernoulli, 9 (2003) 201-217.

[4] T. Bollerslev, Generalized autoregressive conditional heteroscedasticity, Journal of Econometrics, 31 (1986) 307-327.

[5] P. Bougerol, N. Picard. Stationarity of GARCH processes and of some nonnegative time series, Journal of Econometrics, 52 (1992) 115-127.

[6] R. Dahlhaus and S. Subba Rao, A recursive online algorithm for the estimation of time-varying ARCH parameters, Bernoulli 13(2) (2007) 389-422.

[7] P.D. Feigin and R.L. Tweedie, Random coefficient autoregressive processes: a Markov chain analysis of stationarity and finiteness of moments, Journal of Time Series Analysis, 6(1) (1985) 1-14.

[8] L. Gerencsér, On the Martingale approximation of the estimation error of ARMA parameters. Systems & Control Letters 15 (1990) 417-423.

[9] L. Gerencsér and Z. Mátyás, Almost Sure and L_q Convergence of the Re-Initialized BMP scheme, In: Proc. of the 46th Conference on Decision and Control CDC'07, New Orleans, LA, USA, Dec. 12-14, 2007, WeB11.4, pp. 969-974.

[10] J.L. Kierkegaard, J.N. Nielsen, L. Jensen and H. Madsen, Estimating GARCH models using recursive methods, http://citeseer.comp.nus.edu.sg/261917.html, (2000).