FILOMAT (Niš) **13** (1999), 149–155

MAXIMUM LIKELIHOOD ESTIMATION FOR THE FAREX(1) MODEL

Biljana Č. Popović[1] and Miroslav M. Ristić

Abstract. The estimation of the unknown parameters of the model FAREX(1) are given in this paper. As the likelihood function of the model is not differentiable, the maximum of this function is determined by means of modified Hooke-Jeeves' method. Three different modifications of the method are compared and the results are given in tables and figures.

1. Introduction

The model FAREX(1) is one of the first order autoregressive time series with exponential marginal distribution and it has been defined by Mališić [1] as follows: The stationary sequence of random variables $\{X_n,\ n \in \{0, \pm 1, \pm 2, \dots\}\}$ which is defined by the equation

$$X_n = \begin{cases} \alpha X_{n-1} & \text{with probability } p, \\ \beta X_{n-1} + \varepsilon_n & \text{with probability } 1-p, \end{cases}$$

is so called FAREX(1) if and only if

$$\varepsilon_n = \begin{cases} 0 & \text{with probability } \frac{(\alpha-p)\beta}{(1-p)\alpha}, \\ E_n & \text{with probability } \frac{1-\beta}{1-p}, \\ \alpha E_n & \text{with probability } \frac{p(\beta-\alpha)}{(1-p)\alpha}, \end{cases}$$

where $0 < p < \alpha < \beta < 1$ and $\{E_n,\ n \in \{0, \pm 1, \pm 2, \dots\}\}$ is the sequence of i.i.d. random variables with $\varepsilon(\lambda)$ distribution.

Received February 5, 1999

2000 *Mathematics Subject Classification.* 62M10.

Key words and phrases. Autoregressive process, maximum likelihood estimation, FAREX(1) model.

[1]This research was supported by Grant of RFNS 04M03 through Math. Ins. SANU.

As the consequence of this definition, X_n has the same $\varepsilon(\lambda)$ marginal distribution for each n.

Further on we shall suppose that λ is known and, for the simplicity, is equal 1.

The likelihood function of the model FAREX(1) was presented by Popović [2]. Let $(X_1, X_2, \dots, X_n)$ be a sample of size n from this model. The likelihood function of the sample is given by

$$f(x_1,x_2,\dots,x_n) = e^{-x_1}u(x_1)\prod_{k=2}^{n}\Big\{p\delta(x_k-\alpha x_{k-1}) + (\alpha-p)\frac{\beta}{\alpha}\delta(x_k-\beta x_{k-1}) + \\ +u(x_k-\beta x_{k-1})\Big[(1-\beta)e^{-(x_k-\beta x_{k-1})} + p\frac{\beta-\alpha}{\alpha^2}e^{-\frac{(x_k-\beta x_{k-1})}{\alpha}}\Big]\Big\}, \qquad (1.1)$$

where $\delta(x)$ is Dirac's delta function and $u(x)$ is Heviside's unit function.

In this paper, we shall use the method of maximum likelihood estimation to estimate unknown parameters α, β and p which define autoregressive time series FAREX(1). The estimate of α was already presented in [2] and it was given in analytic manner but, the estimates of β and p will be defined by means of numerical methods.

Smith [4] has used numerical gradient method to solve the problem of maximum likelihood estimation. Some modifications of the likelihood function has produced a differentiable function. This will not be of any use when the parameters of FAREX(1) are to be discussed.

2. Maximum likelihood estimation

Let us suppose that we have a sample $(X_1, X_2, \dots, X_n)$. Popović [2] has shown that

$$\hat{\alpha} = \min_{2\le k\le n}\frac{X_k}{X_{k-1}}$$

is a good estimate of α. The goodness of fit of $\hat{\alpha}$ to α has been discussed in [2] also. So, we shall estimate p and β now.

First of all, we shall approximate the values of the function $\delta(x)$ by the values of the function $\delta_\epsilon(x)$ which follows

$$\delta_\epsilon(x) = \begin{cases} \frac{1}{2\epsilon} & , |x| \le \epsilon \\ 0 & , \text{otherwise.} \end{cases}$$

If we set this into (1.1), we shall have the function

$$f^*(x_1,x_2,\dots,x_n) = e^{-x_1}u(x_1)\prod_{k=2}^{n}\Big\{p\delta_\epsilon(x_k-\alpha x_{k-1}) + (\alpha-p)\frac{\beta}{\alpha}\delta_\epsilon(x_k-\beta x_{k-1}) + \\ +u(x_k-\beta x_{k-1})\Big[(1-\beta)e^{-(x_k-\beta x_{k-1})} + p\frac{\beta-\alpha}{\alpha^2}e^{-\frac{(x_k-\beta x_{k-1})}{\alpha}}\Big]\Big\}. \qquad (2.1)$$

which will be used as the approximation to the likelihood function of the model FAREX(1). Applying natural logarithm to both sides of (2.1), we shall have

$$\ln f^*(x_1, x_2, \ldots, x_n) = \ln\left(e^{-x_1} u(x_1)\right) + g(x_1, x_2, \ldots, x_n; \alpha, p, \beta),$$

where

$$g(x_1, x_2, \ldots, x_n; \alpha, p, \beta) = \sum_{k=2}^{n} \ln\left\{p\delta_\epsilon(x_k - \alpha x_{k-1}) + (\alpha - p)\frac{\beta}{\alpha}\delta_\epsilon(x_k - \beta x_{k-1}) + \right.$$
$$\left. +u(x_k - \beta x_{k-1})\left[(1-\beta)e^{-(x_k - \beta x_{k-1})} + p\frac{\beta - \alpha}{\alpha^2} e^{-\frac{(x_k - \beta x_{k-1})}{\alpha}}\right]\right\}. \tag{2.2}$$

So, we shall now search for the maximum of (2.2) instead for the maximum of (1.1).

As the original Hooke-Jeeves' method converges slowly, we shall use some modifications of this model which make the convergence to be faster. The first modification of Hooke-Jeeves' method (MHJ1) will be as follows:

1. Compute $\hat{\alpha}$.
2. Set the initial values for $\hat{p}$ and $\hat{\beta}$.
3. Set the step vector $\Delta = (\Delta_p, \Delta_\beta)$.
4. Set ϵ which will induce the approximation of δ-function.
5. Set θ which will be used as the upper bound of the error.
6. Compute max $= g(x_1, \ldots, x_n; \hat{\alpha}, \hat{p}, \hat{\beta})$.
7. While $\sqrt{\Delta_p^2 + \Delta_\beta^2} > \theta$ do:
   a. If there exist $i \in \{-1, 1\}$, $0 < \hat{p} + i\Delta_p < \hat{\alpha}$, such that
   $$g(x_1, \ldots, x_n; \hat{\alpha}, \hat{p} + i\Delta_p, \hat{\beta}) > g(x_1, \ldots, x_n; \hat{\alpha}, \hat{p}, \hat{\beta}),$$
   then set $\hat{p} = \hat{p} + i\Delta_p$.
   b. If there exist $j \in \{-1, 1\}$, $\hat{\alpha} < \hat{\beta} + j\Delta_\beta < 1$, such that
   $$g(x_1, \ldots, x_n; \hat{\alpha}, \hat{p}, \hat{\beta} + j\Delta_\beta) > g(x_1, \ldots, x_n; \hat{\alpha}, \hat{p}, \hat{\beta}),$$
   then set $\hat{\beta} = \hat{\beta} + j\Delta_\beta$.
   c. Else, set $\Delta_p = \Delta_p/2$ and $\Delta_\beta = \Delta_\beta/2$.
8. $\hat{p}$ and $\hat{\beta}$ will be the maximum likelihood estimators.

Now, for the same reason, to improve the rate of the convergence of MHJ1, the method MHJ1 will be modified to MHJ2:

1. Compute $\hat{\alpha}$.
2. Set the initial values for $\hat{p}$ and $\hat{\beta}$.
3. Set ϵ which will induce the approximation of δ-function.
4. Set θ which will be used as the upper bound of the error.
5. Compute max $= g(x_1, \ldots, x_n; \hat{\alpha}, \hat{p}, \hat{\beta})$.
6. Set change=true.

7. While change=true do:
   a. Set change=false.
   b. Set $\Delta = (\Delta_p, \Delta_\beta)$.
   c. While $\Delta_p > \theta$ do:
   If there exist $i \in \{-1, 1\}$, $0 < \hat{p} + i\Delta_p < \hat{\alpha}$, such that

$$g(x_1, \ldots, x_n; \hat{\alpha}, \hat{p} + i\Delta_p, \hat{\beta}) > g(x_1, \ldots, x_n; \hat{\alpha}, \hat{p}, \hat{\beta}),$$

   then set $\hat{p} = \hat{p} + i\Delta_p$ and change=true. Else, set $\Delta = \Delta_p/2$.
   d. While $\Delta_\beta > \theta$ done:
   If there exist $j \in \{-1, 1\}$, $\hat{\alpha} < \hat{\beta} + j\Delta_\beta < 1$, such that

$$g(x_1, \ldots, x_n; \hat{\alpha}, \hat{p}, \hat{\beta} + j\Delta_\beta) > g(x_1, \ldots, x_n; \hat{\alpha}, \hat{p}, \hat{\beta}),$$

   then set $\hat{\beta} = \hat{\beta} + j\Delta_\beta$ and change=true. Else, set $\Delta_\beta = \Delta_\beta/2$.
8. $\hat{p}$ and $\hat{\beta}$ will be the maximum likelihood estimators.

The modification MHJ2 is inadequate in some way. In fact, it can happen that $\hat{p}$ goes far away from the exact value for the certain value of $\hat{\beta}$, and after the value of $\hat{\beta}$ has been solved, $\hat{p}$ will be near to its exact value. It will be demonstrated in Section 3.

If we define the modification MHJ3 where we shall first search for maximum according to β and after that according to p, the similar problem can be recognized. But, in the both cases, the estimators for p and β will approach to $\hat{\alpha}$.

The modification MHJ3 follows the algorithm:

1. Compute $\hat{\alpha}$.
2. Set the initial values for $\hat{p}$ and $\hat{\beta}$.
3. Set ϵ which will induce the approximation of δ-function.
4. Set θ which will be used as the upper bound of the error.
5. Compute max $= g(x_1, \ldots, x_n; \hat{\alpha}, \hat{p}, \hat{\beta})$.
6. Set change=true.
7. While change=true do:
   a. Set change=false.
   b. Set $\Delta = (\Delta_p, \Delta_\beta)$.
   c. While $\Delta_\beta > \theta$ do:
   If there exist $j \in \{-1, 1\}$, $\hat{\alpha} < \hat{\beta} + j\Delta_\beta < 1$, such that

$$g(x_1, \ldots, x_n; \hat{\alpha}, \hat{p}, \hat{\beta} + j\Delta_\beta) > g(x_1, \ldots, x_n; \hat{\alpha}, \hat{p}, \hat{\beta}),$$

   then set $\hat{\beta} = \hat{\beta} + j\Delta_\beta$ and change=true. Else, set $\Delta_\beta = \Delta_\beta/2$.
   d. While $\Delta_p > \theta$ do:
   If there exist $i \in \{-1, 1\}$, $0 < \hat{p} + i\Delta_p < \hat{\alpha}$, such that

$$g(x_1, \ldots, x_n; \hat{\alpha}, \hat{p} + i\Delta_p, \hat{\beta}) > g(x_1, \ldots, x_n; \hat{\alpha}, \hat{p}, \hat{\beta}),$$

   then set $\hat{p} = \hat{p} + i\Delta_p$ and change=true. Else, set $\Delta = \Delta_p/2$.

8. $\hat{p}$ and $\hat{\beta}$ will be the maximum likelihood estimators.

All three modifications depend on the value of ϵ. The smaller it is, the slower the convergence is.

3. Example

We shall consider the time series FAREX(1) with true values $p = 0.8, \alpha = 0.9$ i $\beta = 0.98$. For all three methods $\hat{\alpha} = 0.9$, and that will be the estimator for α for all five samples which we shall consider of sample sizes 50, 100, 500, 1000 and 2000 respectively.

As it is $0 < p < \alpha$ and $\alpha < \beta < 1$, we can use $\hat{\alpha}/2$ and $(1 + \hat{\alpha})/2$, as the initial values $\hat{p}$ and $\hat{\beta}$ respectively. Further on we shall define vector $\Delta = (\hat{\alpha}/4, (1 - \hat{\alpha})/4)$ and the error bound $\theta = 0.000001$. As we have remarked the rate of convergence mostly depends on ϵ. We shall use $\epsilon = 0.005$ and compare the modifications. The results follow:

TABLE 1. The modification MHJ1.

Sample size	$\hat{p}$	$\hat{\beta}$
50	0.82567921	0.98912258
100	0.82919655	0.97999992
500	0.80836887	0.97999992
1000	0.80745735	0.97999992
2000	0.81301060	0.97999992

TABLE 2. The modification MHJ2.

Sample size	$\hat{p}$	$\hat{\beta}$
50	0.82587147	0.97999878
100	0.82800350	0.97423553
500	0.80836887	0.97999878
1000	0.80745735	0.97999878
2000	0.81301060	0.97999878

TABLE 3. The modification MHJ3.

Sample size	$\hat{p}$	$\hat{\beta}$
50	0.82587147	0.97999878
100	0.82800350	0.97423553
500	0.80836887	0.97999878
1000	0.80745735	0.97999878
2000	0.81301060	0.97999878

The model defined in this example is illustrated in Figure 1. for the sample siz 50. The rate of convergence for three modifications is presented in figures 2, 3 an 4.

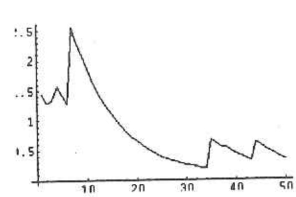

FIGURE 1. Sample simulation of size 50.

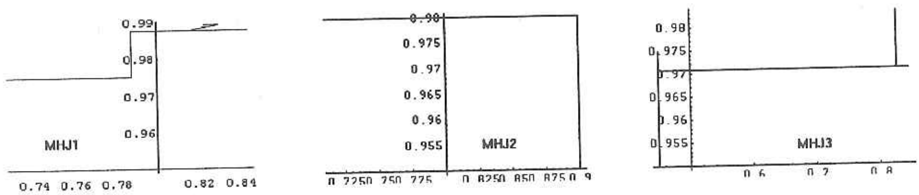

FIGURE 2. The rate of convergence of three modifications.

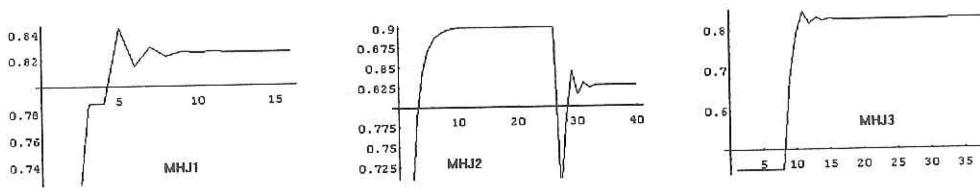

FIGURE 3. Convergence of the estimator $\hat{p}$.

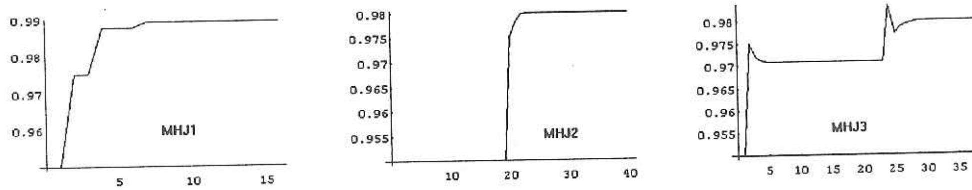

FIGURE 4. Convergence of the estimator $\hat{\beta}$.

If we set ϵ to be less, for example $\epsilon = 0.0005$, the convergence in all thr modifications will be slower. On the contrary, if we set ϵ to be greater than first proposed, the estimators of the parameters will oscillate.

If we consider FAREX(1) defined for the real values $\alpha = 0.04$, $\beta = 0.08$ a $p = 0.025$, then good results of the estimation procedure will be achieved for $\epsilon = 0$ The less values of ϵ will make the estimation procedure unsuccessful. This confir the importance of the choice of ϵ.

4. Discussion

For the estimation of the unknown parameters p and β we could use some other nongradient method, Powell's method for instance. But our efforts were concentrated to show what are the possibilities of solving the problem of maximum likelihood estimation of the parameters of the model FAREX(1) which is only one of the exponentially marginally distributed time series, but almost general one of order one. The Hooke-Jeeves' method was chosen because of its simplicity and reliability. The lack of this method is that the convergence depends too much on the choice of ϵ.

References

[1] Mališić J., *On exponential autoregressive time series models*, Math. and Prob. Theory, Vol. B, 1987, 147–153.

[2] Popović B., *On estimation of parameters of time series with exponential marginals*, VII Conference on Applied Mathematics, Osijek, 1989, 185–191.

[3] Raftery A., *Estimation efficace pour un processus autoregressif exponentiel a densite discontinue*, Pub. Inst. Stat. Univ. Paris, XXV, fasc. 1-2, 1980, 65–91.

[4] Smith R., *Maximum likelihood estimation for the NEAR(2) model*, J. R. Statist. Soc. A, **48**, no. 2, 1986, 251–257.

B. Č. Popović
Department of Mathematics, Faculty of Sciences and Mathematics, Univerity of Niš
Ćirila i Metodija 2, 18000 Niš, Yugoslavia
E-mail: bipop@archimed.filfak.ni.ac.yu

M. M. Ristić
Department of Mathematics, Faculty of Sciences and Mathematics, Univerity of Niš
Ćirila i Metodija 2, 18000 Niš, Yugoslavia
E-mail: mristic@archimed.filfak.ni.ac.yu