ProbStat Forum, Volume 05, October 2012, Pages 107–111
ISSN 0974-3235

ProbStat Forum is an e-journal.
For details please visit www.probstat.org.in

Parameter estimation and dependence characterization of the MAR(1) process

Marta Ferreira

Center of Mathematics of Minho University, Campus de Gualtar 4710 - 057 Braga, Portugal

Abstract. Classical linear ARMA with normal distributed noises are not suitable for heavy tailed phenomena. MARMA processes obtained by replacing summation by the maximum operator are more appropriate. We consider unit Fréchet first order MARMA, denoted MAR(1), and present a characterization based on ordinal autocorrelation. An estimator of the model's parameter and respective consistency and asymptotic normality properties are also stated.

1. INTRODUCTION

It is now well recognized that heavy tailed phenomena occurs in nature and society and cannot be described by the normal distribution. A useful class of processes for modeling heavy tails and extremal structures of dependent events are the max-ARMA or MARMA processes introduced in Davis and Resnick [4], which are analogous to ARMA by replacing summation by the maximum operator:

$$X_n = \left(\bigvee_{i=1}^{p} \alpha_i X_{n-i} \right) \bigvee \left(\bigvee_{j=0}^{q} \beta_j Z_{n-j} \right)$$

where parameters $\alpha_i, \beta_j \geq 0$, $1 \leq i \leq p$, $0 \leq j \leq q$, are non-negative independent and identically distributed (i.i.d.) innovations. Besides a more suitable formulation than heavy tailed ARMA given easily definable finite dimensional distributions, they have a wide application to various natural phenomena (Helland and Nielsen [7], Daley and Haslett [5], Hooghiemstra and Scheffer [8], Todorovic and Gani [13], Coles [3]), reliability (Davis and Resnick ([4]) or financial series (Zhang and Smith [14]).

Several first order max-autoregressive formulations that include random coefficients and power transformations have also been considered in literature (Alpuim [1], Alpuim and Athayde [2], Ferreira and Canto e Castro [6]).

Stationary MARMA processes with Fréchet marginal d.f. are maximum-stable and thus convenient for calculations. Here we focus on first order max-autoregressive MARMA. Indeed, since our procedures also apply to continuous strictly increasing functions of the process and that the values can be normalized so that they come from an unit Fréchet model, we shall consider unit Fréchet margins, i.e., $F(x) = \exp(-1/x)$. For more details, see Lebedev ([11], [12]).

2010 *Mathematics Subject Classification.* Primary: 60G10; Secondary: 62M10
Keywords. Autoregressive processes, heavy tail, estimation of parameters, ordinal autocorrelation
Received: 11 July 2012; Revised: 05 September 2012; Accepted: 09 September 2012
Email address: msferreira@math.uminho.pt (Marta Ferreira)

Given the maximum stability property satisfied by independent unit Fréchet distributed r.v.'s, ξ_1, ξ_2 and ξ, i.e.,

$$c_1\xi_1 \bigvee c_2\xi_2 = (c_1 + c_2)\xi$$

we shall define the first order max-autoregressive process, denoted MAR(1), as

$$X_i = cX_{i-1} \bigvee (1-c)Z_i,\ 0 \le c < 1 \tag{1}$$

with margins and noises having unit Fréchet d.f. F.

An estimation procedure for model's parameter c have been stated in Lebedev [11] based on fluctuation probabilities,

$$f_1 := P(X_{n-1} < X_n) = \frac{1}{2-c}\,.$$

We derive estimators for c^m ($m \ge 1$) using a similar methodology and state consistency and asymptotic normality (Section 2). In particular, m can be chosen in order to obtain the smallest variance.

A drawback of very heavy tailed processes is that first and second order moments do not exist and the usual auto-correlation function (ACF) of classical linear models cannot be used. Alternatively, we consider ordinal autocorrelation based on *Spearman's rho* coefficient. Besides an alternative estimator for the parameter, it provides a characterization for MAR(1) processes (Section 3).

2. Estimation of the model's parameter

In the following we use the m-step ($m \ge 1$) transition probability function (tpf) of the process, given by

$$\begin{aligned} Q^m(x,]0,y]) = \; & P(X_{n+m} \le y | X_n = x) = \begin{cases} \prod_{j=0}^{m-1} F\big(\frac{y}{c^j(1-c)}\big) & , x \le y/c^m \\ 0 & , x > y/c^m, \end{cases} \\ = \; & \begin{cases} \exp\big(-\frac{1-c^m}{y}\big) & , x \le y/c^m \\ 0 & , x > y/c^m. \end{cases} \end{aligned} \tag{2}$$

Proposition 2.1. *Let $\{X_i\}_{i\ge1}$ be a MAR(1) process. Then, for each $m \ge 1$,*

$$c^m = 2 - 1/f_m.$$

Proof. Observe that the m-step fluctuation probabilities $f_m := P(X_{n-m} < X_n)$ are given by

$$\begin{aligned} f_m : &= P(X_n \le X_{n-m}) = \int_0^\infty P(X_n \le x | X_{n-m} = x) dF_X(x) \\ &= \int_0^\infty Q^m(x,]0,x] dF_X(x) = \int_0^\infty F(x)^{1-c^m} dF_X(x) \\ &= \frac{1}{2-c^m}. \end{aligned} \tag{3}$$

□

The fluctuation probabilities can be used to estimate the process parameter, c. Considering

$$\widehat{f}_m = \frac{1}{n-m}\sum_{j=m+1}^{n} \mathbf{1}_{\{X_j \le X_{j-m}\}},\ m \ge 1,$$

we will see that estimator

$$\widehat{c}^m = 2 - 1/\widehat{f}_m$$

is consistent and asymptotically normal.

Proposition 2.2. *Let $\{X_i\}_{i\ge1}$ be a MAR(1) process. Then, for each $m \ge 1$,*

$$n^{1/2}(\widehat{c}^m - c^m) \stackrel{D}{\to} N(0, \sigma_m^2/f_m^4) \tag{4}$$

where

$$\sigma_m^2 = f_m(1-f_m)(1-2f_m+\chi_m)/(1-\chi_m), \tag{5}$$

with f_m given in (3) *and*

$$\chi_m = \tfrac{c(2-c^m)}{2}\Big(\tfrac{1}{c^m-2c-2} - \tfrac{2}{c^m-2c-1} + \tfrac{1}{c^m+2c}\Big).$$

Proof. Just observe that $\widehat{f}_m$ corresponds to the mean of Bernoulli trials with Markov dependence and results on this can be found in Klotz [10]. More precisely, we have that, $n^{1/2}(\widehat{f}_m - f_m) \stackrel{D}{\to} N(0,\sigma_m^2)$ holds for σ_m^2 given in (5), where $\chi_m = P(X_{j-m} < X_j | X_{j-m-1} < X_{j-1})$ with $\max(0,(2f_m-1)/f_m) \le \chi_m \le 1$. The result in (4) is now straightforward by the Delta Method, i.e., $n^{1/2}(g(\widehat{f}_m) - g(f_m)) \stackrel{D}{\to} N(0,\sigma_m^2(g'(f_m))^2)$, with $g(x) = 2 - x^{-1}$. In the following we compute the variance, σ_m^2

$$P(X_{m+j} \le X_j, X_{m+j-1} \le X_{j-1}) = \int_0^\infty\int_0^\infty\int_0^x Q(w,]-\infty,y])Q^{m-1}(y,dw)Q(x,dy)K(dx).$$

Now considering (2), the following development holds:

$$P(X_{m+j} \le X_j, X_{m+j-1} \le X_{j-1}) = \int_0^\infty\int_0^\infty\int_0^{\min(x,y/c)} F(y)Q^{m-1}(y,dw)Q(x,dy)F(dx)$$

$$= \int_0^\infty\Big\{\int_0^{xc} F(y)Q^{m-1}(y,]-\infty,y/c]) + \int_{xc}^\infty F(y)Q^{m-1}(y,]-\infty,x])\Big\}Q(x,dy)F(dx)$$

$$= \int_0^\infty\Big\{\int_0^{xc} F(y)[F(y/c)]^{1-c^{m-1}}) + \int_{xc}^{x/c^{m-1}} F(y)[F(x)]^{1-c^{m-1}}\Big\}Q(x,dy)F(dx).$$

Considering f the density function of the unit Fréchet d.f. F, the transition density of $Q(x,]0,y])$ is given by $q(x,y) = f(y)\mathbf{1}_{\{xc<y\}} + F(xc)\mathbf{1}_{\{xc=y\}}$. Hence we have,

$$\begin{aligned} P(X_{m+j} \le X_j, X_{m+j-1} \le X_{j-1}) &= \int_0^\infty\Big\{F(xc)[F(x)]^{1-c^{m-1}} + \frac{[F(x/c^{m-1})]^2 - [F(xc)]^2}{2}[F(x)]^{1-c^{m-1}}\Big\}F(dx) \\ &= \frac{c}{2}\Big(\frac{1}{c^m-2c-2} - \frac{2}{c^m-2c-1} + \frac{1}{c^m+2c}\Big). \end{aligned}$$

Observe that, after some calculations, we obtain the variance of $\widehat{c}^m$ given by

$$\sigma_m^2/f_m^4 = -\frac{(c^m-2)^2(c^m-1)^2(c^{2m}-2c-2c^m(1+c)))}{2c+c^{3m}+c^m(2+c+6c^2+4c^3)-c^{2m}(3+2c(1+c))}.$$

□

In particular, m can be chosen in order to obtain the smallest variance, provided that $\widehat{f}_m \in [1/2,1)$. Note also that no definite results can be obtained for $\widehat{f}_m < 1/2$ since $f_m \in [1/2,1)$. Indeed, as observed in Lebedev [11], the probability of such events goes to zero, as $n \to \infty$, and this may also be an indication of an inconsistency in our choice of the model.

3. Ordinal autocorrelation

The main drawback of a max-autoregressive modeling is that the usual analysis methods based on the autocorrelation function (ACF) of classical linear models cannot be used here, since the first and second moments do not exist. Alternatively, we can use ordinal correlation. We consider the *Spearman's rho* coefficient which corresponds to the Pearson correlation coefficient applied to marginal transform $(F_1(X), F_2(Y))$ of random pairs (X, Y) with marginal d.f.'s F_1 and F_2, respectively, and thus also stated in Joe [9]

$$\rho_S = \rho_S(X, Y) = 12EF_1(X)F_2(Y) - 3.$$

For lag-m random pairs, we denote the lag-m *Spearman's rho* coefficient,

$$\rho_{S,m} = \rho_{S,m}(X_1, X_{1+m}).$$

Observe that, if $Y_i = F(X_i)$ and $U_i = F(Z_i)$, then MAR(1) process given in (1) can be rewritten as

$$Y_i = Y_{i-1}^{1/c} \bigvee U_i^{1/(1-c)}, 0 \leq c < 1 \tag{6}$$

where all the r.v.'s Y_i and U_i are uniformly distributed in the interval $(0, 1)$ (standard uniform).

Proposition 3.1. *The lag-m Spearman's rho coefficient of MAR(1) process is*

$$\rho_{S,m} = \frac{3c^m}{2 + c^m}.$$

Proof. We shall use a similar approach of Lebedev [11], who has only consider the case $m = 1$.

Note that, for any constants $a > 0$ and $0 \leq b \leq 1$, and W an r.v. with standard uniform d.f., we have

$$P\left(W^{1/a} \bigvee b \leq x\right) = x^a \mathbf{1}_{\{x>b\}} + b^a \mathbf{1}_{\{x=b\}}$$

and hence

$$E\left(W^{1/a} \bigvee b\right) = \int_b^1 ax^a dx + b^{a+1} = \frac{a + b^{a+1}}{a + 1}. \tag{7}$$

Now observe that process (6) correspond to

$$Y_{1+m} = Y_1^{1/c^m} \bigvee \left(\bigvee_{j=2}^{m+1} U_j^{1/(c^{m+1-j}(1-c))} \right)$$

and that

$$P\left(\bigvee_{j=2}^{m+1} U_j^{1/(c^{m+1-j}(1-c))} \leq x \right) = x^{\sum_{k=1}^{m-1} c^k (1-c)} = x^{1-c^m}$$

i.e., $\bigvee_{j=2}^{m+1} U_j^{1/(c^{m+1-j}(1-c))} \overset{d}{=} W^{1/(^{1-c^m})}$. Now we calculate

$$\begin{aligned} EY_1Y_{1+m} = \; & E_{Y_1}E(Y_1(Y_1^{1/c^m} \bigvee \textstyle\bigvee_{j=2}^{m+1} U_j^{1/(c^{m+1-j}(1-c))})|Y_1) \\ = \; & E_{Y_1}E(Y_1(Y_1^{1/c^m} \bigvee W^{1/(1-c^m)})|Y_1) \\ = \; & EY_1 \tfrac{1-c^m+Y_1^{2/c^m-1}}{2-c^m} \end{aligned}$$

where the last equality is due to (7). We have

$$EY_1 \frac{1 - c^m + Y_1^{2/c^m-1}}{2 - c^m} = \frac{1}{2 - c^m}\left((1 - c^m)EY_1 + EY_1^{2/c^m}\right) = \frac{1}{2 - c^m}\left(\frac{1 - c^m}{2} + \frac{c^m}{2 + c^m}\right) = \frac{1 + c^m}{2(2 + c^m)}.$$

□

Thus, we can also estimate c^m from the ordinal *Spearman's rho* correlation coefficient.

Corollary 3.2. *In the MAR*(1) *process we have, for* $\rho_{S,m} \in [0, 1)$,

$$c^m = \frac{2\rho_{S,m}}{3 - \rho_{S,m}}.$$

The following characterization relationship of MAR(1) is useful for model identification.

Corollary 3.3. *In the MAR*(1) *process we have, for* $f_m \in [1/2, 1)$,

$$\rho_{S,m} = \frac{3}{2}\Big(1 - \frac{7}{4f_m}\Big).$$

References

[1] M.T. Alpuim: An extremal markovian sequence, J. Appl. Probab. 26 (1989), 219–232.
[2] M.T. Alpuim, E. Athayde: On the stationary distribution of some extremal Markovian sequence, J. Appl. Probab. 27 (1990), 291–302.
[3] S. G. Coles, Regional modelling of extreme storms via max-stable processes, J. Roy. Statist. Soc. Ser. B 55 (4) (1993), 797–816.
[4] R. Davis, S. Resnick, Basic properties and prediction of max-ARMA processes, Adv. Appl. Probab. 21 (1989), 781–803.
[5] D. J. Daley, J. Haslett, A thermal energy storage process with controlled input, Adv. Appl. Probab. 14 (1982), 257–271.
[6] M. Ferreira, L. Canto e Castro, Modeling rare events through a pRARMAX process, J. Statist. Plann. Inference 140 11 (2010), 3552–3566.
[7] I. S. Helland, T. S. Nilsen, On a general random exchange model, J. Appl. Probab. 13 (1976), 781–790.
[8] G. Hooghiemstra, C.L. Scheffer, Some limit results for energy storage model, Stoch. Proc. Appl. 22 (1986) 121–127.
[9] H. Joe, Multivariate Models and Dependence Concepts, Chapman Hall , London, (1997).
[10] J. Klotz, Statistical inference in Bernoulli trials with dependence, Ann. Statist. 1 (1973), 373–379.
[11] A. V. Lebedev, Statistical analysis of first-order MARMA processes, Mat. Zametki 83 (2008), 4, 552–558.
[12] A. V. Lebedev, Nonlinear Prediction in Max-Autoregressive Processes, Mat. Zametki 85(4) (2009), 636–640.
[13] P. Todorovic, J. Gani, Modelling the effect of erosion on crop production, J. Appl. Prob 24 (1987), 787–797.
[14] Z. Zhang, R. L. Smith, Modelling Financial Time Series Data as Moving Maxima Processes, Technical Report Dept. Stat. (Univ. North Carolina, Chapel Hill, NC, 2001), http://www.stat.unc.edu/faculty/rs/ papers/RLS_Papers.html.