Asymptotic and pre-asymptotic tail behavior of a power max-autoregressive model

Marta Ferreira

Sciences School, Department of Mathematics, Campus de Gualtar, University of Minho. 4710-057 Braga, Portugal.

Keywords: Markov chain. Extreme value theory. Stationary sequence. Extremal index. Tail dependence index. Asymptotic independence.

AMS: 60G70

Abstract

Nowadays it cannot be ignored the demand of various areas, like hydrology, geophysics or finances, on modeling extreme data or exceedance data above certain high levels. Most of the time, temporal dependence is present and it is usual to associate markovian sequences to data therein. Here it is presented the process $\text{ARMAX}_p$ whose parameter relates directly with the Ledford and Tawn coefficient of tail dependence, $\eta$, (Ledford and Tawn [5]). This index characterizes the penultimate tail dependence of a process and can be related with a threshold-dependent extremal index, which assumes an important role when extending discussions of extreme values from i.i.d. sequences to stationary ones (Bortot and Tawn [3]).

1. Introduction

Extreme Value Theory (EVT) became widely used in many applied sciences when faced with modeling high values. Recently, models for extreme values have been constructed on the assumption of temporal dependence like stationary Markov chains whose extremal properties can have a nice treatment. Consider a stationary first order Markov chain, $\{X_i\}_{i \geq 1}$, with continuous state-space. In what concerns the tail statistical modeling of $\{X_i\}_{i \geq 1}$, it is important to distinguish between asymptotic dependence or asymptotic independence, since an asymptotically independent Markov chain usually presents an extremal feature increasingly resembling that of an i.i.d. sequence at high thresholds. More precisely, consider the limit

$$\lim_{x \to x^*} P(X_2 > x | X_1 > x) = b,$$

where $b > 0$ corresponds to asymptotic dependence and $b = 0$ asymptotic independence. Procedures like proposed by Smith et al. [7] are not suitable since, above a fixed high threshold, they always consider a dependent structure given by a bivariate extreme value distribution. On the other hand, when dealing with bivariate extremes, Ledford and Tawn [5] suggested a model with a tail dependence coefficient, $\eta$, that is related to the above limit. More precisely, considering a random pair, $(W, V)$, with marginal d.f.’s $F_1$ and $F_2$, Ledford
and Tawn model assumes that \( P(1 - F_1(W) < t, 1 - F_2(V) < t) \) is a regularly varying function at 0 with index \( 1/\eta \), that is,
\[
\frac{P(1 - F_1(W) < tx, 1 - F_2(V) < ty)}{P(1 - F_1(W) < t, 1 - F_2(V) < t)} \to h(x, y), \quad t \downarrow 0,
\]
where \( h \) is an homogeneous function of order \( 1/\eta \), i.e., \( h(sx, sy) = s^{1/\eta}h(x, y) \) (Bingham et al. [2]). Furthermore, it is assumed that the convergence is uniform on \( \{(x, y) | \max(x, y) = 1\} \). Thus being, as \( t \downarrow 0 \), \( P(1 - F_2(V) < t(1 - F_1(W) < t) \sim t^{1/\eta - 1}L(1/t) \), where \( L \) is a slowly varying function. So, when \( F = F_1 = F_2 \), denoting \( X_1 = W \) and \( X_2 = V \), then \( P(X_2 > x | X_1 > x) \sim (t(x))^{1/\eta - 1}L(t(x)^{-1}) \), as \( x \rightarrow x^* \), with \( t(x) = 1 - F(x) \). Hence, for the pair \( (X_1, X_2) \), if \( \eta = 1 \) and \( L(1/t) \rightarrow 0 \) or if \( 0 < \eta < 1 \), there is an asymptotic independence since the limit above is 0, whilst the case \( \eta = 1 \) and \( L(1/t) \not\rightarrow 0 \) leads to a positive limit meaning asymptotic dependence. Moreover, \( \eta \in (1/2, 1) \) reveals a positive association, \( \eta \in (0, 1/2) \) a negative association and \( \eta = 1/2 \) an (almost) independence. The extremal index, \( \theta \), is another measure of dependence, more precisely, it measures the degree of local dependence in the extreme values of a stationary process. This parameter plays a very important role when estimating extremal properties of weakly mixing stationary series since, for certain sequences of levels, \( \{u_n\}_{n \geq 1} \),
\[
\lim_{n \to \infty} P(\max(X_1, ..., X_n) \leq u_n) = \lim_{n \to \infty} \left( P(X_1 \leq u_n) \right)^{n\theta}.
\]

Though i.i.d. processes have \( \theta = 1 \), the converse is false. Several processes that would be regarded as strongly dependent by other measures, have \( \theta = 1 \) (e.g. autoregressive Gaussian processes). An unit extremal index means that asymptotically extreme events occur singly. However, it can be possible to observe clustering of exceedances for levels of practical interest. These features motivated a threshold-based extremal index, \( \theta(u) \), in Bortot and Tawn [3]. If \( \theta \) is replaced by \( \theta(u) \) in (2), it will occur an improvement on all the estimations that can be based on this result, like high quantiles or return periods. Another nice feature is that \( \theta(u) \) can be related with the coefficient of tail dependence, \( \eta \) (Bortot and Tawn [3]). In this paper, an ARMAX\(_p\) process will be considered for the Markov chain \( \{X_i\}_{i \geq 1} \). This process has a parameter \( c \in (0, 1) \) that relates directly with the coefficient of tail dependence of Ledford and Tawn and has \( \theta = 1 \). The existence of stationary distribution will be proved first and the threshold-based extremal index will be derived.

2. The ARMAX\(_p\) process

Consider a sequence of i.i.d. r.v.’s, \( \{Z_i\}_{i \geq 1} \), having support contained in \([1, \infty]\) and common d.f. \( F_2 \). A sequence \( \{X_i\}_{i \geq 1} \) is said to be an ARMAX\(_p\) process if,
\[
X_i = \max(X_{i-1}, Z_i), \quad 0 < c < 1,
\]
with \( X_i \) independent from \( Z_i \), for all \( i \geq 1 \). Let \( \{X_i\}_{i \geq 1} \) be an ARMAX\(_p\) process as defined above. Suppose that \( K_n \) is the d.f. of \( X_n \). Then, \( K_n(x) = P(X_n \leq x) = P(X_n^c \leq x) = K_{n-1}(x^{1/c})F_Z(x) \). Similarly, \( K_{n-1}(x) = P(X_{n-2}^c \leq x, Z_{n-1} \leq x) = K_{n-2}(x^{1/c})F_Z(x) \) and solving iteratively, it is obtained, \( K_n(x) = P(X_0 \leq x^{1/c^n}) \prod_{j=0}^{n-1} F_Z(x^{1/c^j}) \). Since for any \( x > 1 \), \( P(X_0 \leq x^{1/c^n}) \to 1 \), as \( n \to \infty \), then \( K(x) = \lim_{n \to \infty} K_n(x) = \prod_{j=0}^{\infty} F_Z(x^{1/c^j}) \). Thus being, \( \{X_i\}_{i \geq 1} \) admits only one stationary distribution given by \( K(x) = \prod_{j=0}^{\infty} F_Z(x^{1/c^j}) \), which is non degenerate because \( F_Z(\cdot) \) is non degenerate. Also observe that \( K \) must verify relation

\[
K(x) = K(x^{1/c})F_Z(x).
\]

Now, it will be proved that, if \( F_Z \) is in the Fréchet max-domain of attraction, then the same happens with \( K \). By hypothesis, \( 1 - F_Z(\cdot) \) is regularly varying at \( \infty \) and hence,

\[
1 - F_Z(x) = x^{-1/\gamma}L_Z(x), \text{ where } L_Z \text{ is a slowly varying function},
\]

or equivalently, \(- \log F_Z(x) = x^{-1/\gamma}L(x)\) for a convenient slowly varying function \( L \). Based on the definition of \( K \), \(- \log K(x) = \sum_{j=0}^{\infty} x^{-1/(\gamma c^j)}L(x^{1/c^j})\), as \( x \to \infty \). In order to state an approach to the latest sum, it is used the criterion of the integral and get an approach to \( \int_0^\infty x^{-1/(\gamma c^j)}L(x^{1/c^j})dy \). Taking \( u = x^{1/c^j} \) and applying Karamata’s Theorem,

\[
\int_x^\infty u^{-1/\gamma}L(u)\left(-\frac{1}{u \log u \log c}\right)du = \int_x^\infty u^{-1/\gamma-1}L^*(u)du \sim x^{-1/\gamma}L^*(x),
\]

where \( L^* \) and \( L^{**}(x) = -1/\gamma L^*(x) \) are slowly varying functions. Therefore,

\[- \log K(x) \text{ is regularly varying or, equivalently,}
1 - K(x) = x^{-1/\gamma}L_K(x) \text{where } L_K \text{ is a slowly varying function}.
\]

Hence, the assertion follows. From now on, \( K \) is the d.f. of \( \{X_i\}_{i \geq 1} \), verifying the above conditions.

In what respects the dependence structure, first it will be proved that the process is regenerative and aperiodic and so it is strong mixing. From Asmussen [1], \( \{X_i\}_{i \geq 1} \) is regenerative if for some \( m > 0 \) and \( \epsilon \in (0,1) \),

\[
Q^m(x,B)\geq\lambda(B), \quad x \in R, \quad \text{for all real borelian } B \text{ and some distribution } \lambda,
\]

where \( Q^m(x,B) \) is the \( m \)-step transition probability function from state \( x \) to \( B \) and \( R \) is a recurrent set (\( H(R) > 0 \)). The transition probability function of the process ARMAX\(_p\) for \( y \geq x^c \) is \( Q(x,|x|,y):= P(X_n \leq y | X_{n-1} = x) = F_Z(y) \).

Let \( R = [r, +\infty[ \) be a proper subset of the support of \( Z_i \). Then \( K(R) = 1 - \prod_{j=0}^{\infty} F_Z(t^{1/c^j}) = 1 - F_Z(r) \prod_{j=1}^{\infty} F_Z(t^{1/c^j}) > 0 \) since \( 0 < t^{1/c^j} < 1 \) (\( r \) is in the support of \( Z_i \)) and hence, \( R \) is recurrent. Let \( x \in R, B \in B(\mathbb{R}) \), and let \( S \) be a compact subset of \( R \) such that, \( \forall y \in S, y > x^c \). We have that,

\[
Q(x,B) = \int_B dQ(x,z) = \int_{B \cap S} dQ(x,z) = \int_{B \cap S} dF_Z(z) = P(Z \in B \cap S) = \epsilon \lambda(B),
\]
where the distribution $\lambda(\cdot) = P(Z \in \cdot \cap S)/P(Z \in S)$ and $\epsilon = P(Z \in S)$. So, $m = 1$.

The condition below is sufficient to prove aperiodicity (Asmussen [1]):

$$Q^{n+1}(x, B) \geq \epsilon_1 \lambda(B) \quad \text{and} \quad Q^n(x, B) \geq \epsilon_2 \lambda(B), \quad \forall x \in R, \quad \epsilon_1, \epsilon_2 \in (0, 1).$$

Note that,

$$Q^2(x, B) \geq \int_B Q(z, dQ(x, z) \geq \lambda(B) \int_B dQ(x, z) = \epsilon \lambda(B) \int_B F_z(z) = \epsilon \lambda(B) F_z(B \cap S).$$

Therefore, just take $m = 1$, $\epsilon_1 = \epsilon P(Z \in B \cap S)$ and $\epsilon_2 = \epsilon$.

Next, it is shown that condition $D''(u_n)$ (Leadbetter and Nandagopalan [4]) also holds for some increasing real sequences $\{u_n\}_{n \geq 1}$. Note that, for $j \geq 2$,

$$P(X_1 > u_n, X_j \leq u_n < X_{j+1}) = \int_{u_n}^{\infty} P(X_{j+1} > u_n | X_j = z) Q^{j-1}(y, dz) K(dy)$$

$$= \int_{u_n}^{\infty} \sum_{j = 3}^{n/k_n - 1} P(X_1 > u_n, X_j \leq u_n < X_{j+1}) \leq \frac{n}{k_n} (1 - F_Z(u_n))(1 - K(u_n)).$$

By (3), $H(x) \leq F_z(x)$, then it is enough to consider $1 - H(u_n) = O(1/n)$ in order to obtain a null limit in the last expression.

The next calculation is the function $h(x, y)$ in (1) and the value of $\eta$. Taking $a_{x, t} = K^{-1}(1 - tx)$ and $a_{y, t} = K^{-1}(1 - ty)$, then, as $t \downarrow 0$,

$$h(x, y) \sim P(X_1 > a, x, X_1 > a, y) = \frac{t x - K(a, x) - K(a, y)}{t - K(a, x) - K(a, y)} K(a, y) K(a, y) / K(a, x)$$

$$= \frac{t x - K(a, x) - K(a, y)}{t - K(a, x) - K(a, y)} K(a, y) K(a, y) / K(a, x).$$

Since $\gamma > 0$, then $F^{-1}(1 - \gamma) \sim x^{-1/\gamma} F^{-1}(1 - t)$, as $t \downarrow 0$, $x > 0$. Thus being and using (5), after some calculations, $h(x, y) = xy$ if $c^m < 1/2$ and $h(x, y) = y^{1+c}$ if $c^m > 1/2$, which implies $\eta = 1/2$ and $\eta = c^m$, respectively. This result also tells us, that the bigger the value of parameter $c$, the bigger the lag $m$ must be chosen, in order to get asymptotically independent observations.

Based on O’Brien [6] extremal index characterization, Bortot and Tawn [3] stated a threshold-dependent form for $\theta$. Denote $r_{[u]}$ the length of the block used to define a cluster of high level exceedances above $u$ ($|[x]|$ represents the integer part of $x$), the threshold-dependent form of $\theta$ is like follows:

$$\theta(u, r_{[u]}) = P(Y_i \leq u, 2 \leq i \leq r_{[u]} | Y_1 > u)$$

$$= 1 - P(Y_2 > u | Y_1 > u) - \sum_{j=3}^{r_{[u]}} P(\max_{t=2} Y_i \leq u, Y_j > u | Y_1 > u), \quad (6)$$

where $u \geq 0$ and the sum is considered null if $r_{[u]} = 2$. We have $\theta = \lim_{u \to \infty} \theta(u, r_{[u]})$. Therefore, replacing in (2) the value of $\theta$ by a penultimate value, $\theta(u)$, leads to an improvement of the estimations referred above. In this case, as conditions $D'(u_n)$ and $D''(u_n)$ hold, $\theta$ can be calculated by the following limit (Leadbetter and Nandagopalan [4]):

$$\theta = \lim_{n \to \infty} P(X_2 \leq u_n | X_1 > u_n) = \lim_{n \to \infty} \int_{u_n}^{1/c} F_z(u_n) K(dy) = \lim_{n \to \infty} \left[ \frac{1 - K(u_n)}{1 - K(u_n)} \right]^{1 - \frac{1}{c}}.$$
Finally, an approach for $\theta(u, r_{|u|})$ in (6) is computed for the process ARMAX$_p$. The general term of the summation in (6) becomes,
\[
P(\max_{i=2}^{j-1} X_i \leq u, X_j > u|X_1 > u)\]
\[
= [P(\max_{i=2}^{j-1} X_i \leq u, X_j > u) - P(\max_{i=2}^{j-1} X_i \leq u, X_j \leq u, X_1 > u)]P(X_1 > u)^{-1} \]
\[
= \{[K(u^{1/\gamma}) - K(u)][F_x(u)]^{-2} - [K(u^{1/\gamma}) - K(u)][F_x(u)]^{-1}\}[1 - K(u)]^{-1}. \]
and
\[
P(X_x > u|X_1 > u) = 1 - \int_{\infty}^{\infty} P(w < u, Z \leq u)K(dw) = 1 - \frac{F_x(u)[K(u^{1/\gamma}) - K(u)]}{1 - K(u)}.
\]
The replacement of the two latter expressions in (6), leads to,
\[
\theta(u, r_{|u|}) = [K(u^{1/\gamma}) - K(u)][1 - K(u)]^{-1} \left(\frac{F_x(u)}{1 - K(u)}\right)^{r_{|u|}}. \]
and by (4), (5) and applying the first order Taylor approximation, the threshold-dependent extremal index for an ARMAX$_p$ process becomes, as $u \to \infty$,
\[
\theta(u, r_{|u|}) \approx \left(1 - u^{-\frac{1}{\gamma}} L_K(u^{1/\gamma})\right)[1 - u^{-1/\gamma} L_x(u)]^{-r_{|u|} - 1} \]
\[
\sim 1 - u^{-\frac{1}{\gamma}} \frac{L_K(u^{1/\gamma})}{L_K(u)} - (r_{|u|} - 1)u^{-1/\gamma} L_x(u) - u^{-\frac{1}{\gamma}} L_K(u^{1/\gamma}) L_x(u) \frac{L_K(u)}{L_K(u)}.
\]
Since $\lim_{u \to \infty} \theta(u, r_{|u|}) = 1$, then it must be considered $r_{|u|} = o(u^{1/\gamma})$.

3. Bibliography


