Journal of Multivariate Analysis 100 (2009) 1029–1043

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Extreme value theory for stochastic integrals of Legendre polynomials[☆]

Alexander Aue[a,*], Lajos Horváth[b], Marie Hušková[c]

[a] Department of Statistics, University of California, One Shields Avenue, Davis, CA 95616, USA
[b] Department of Mathematics, University of Utah, 155 South 1440 East, Salt Lake City, UT 84112, USA
[c] Department of Statistics, Charles University, Sokolovská 83, CZ-18600 Praha, Czech Republic

ARTICLE INFO

Article history:
Received 1 November 2007
Available online 15 October 2008

AMS 2000 subject classifications:
primary 60G70
secondary 62J02
62J12

Keywords:
Extreme value asymptotics
Gaussian processes
Gumbel distribution
Legendre polynomials
Polynomial regression

ABSTRACT

We study in this paper the extremal behavior of stochastic integrals of Legendre polynomial transforms with respect to Brownian motion. As the main results, we obtain the exact tail behavior of the supremum of these integrals taken over intervals $[0, h]$ with $h > 0$ fixed, and the limiting distribution of the supremum on intervals $[0, T]$ as $T \to \infty$. We show further how this limit distribution is connected to the asymptotic of the maximally selected quasi-likelihood procedure that is used to detect changes at an unknown time in polynomial regression models. In an application to global near-surface temperatures, we demonstrate that the limit results presented in this paper perform well for real data sets.

1. Introduction

This paper is concerned with the extremal behavior of a specific Gaussian process. While the study of extremes appears naturally in a number of application areas such as finance, insurance, and telecommunications (see Finkenstaedt and Rootzén [12] for a recent review), the results here are motivated by a problem arising in the context of change detection in polynomial regression models as originally discussed in Brown et al. [9] and MacNeill [25], among others. These contributions utilize residual based statistics for the change-point test, which were further analyzed in [15–17].

We are interested in the maximally selected quasi-likelihood procedure instead. This means in particular that the existing theory established in the aforementioned papers cannot be applied in our setting. On the contrary, our results are based on novel computations fusing the theoretical work on the extremal behavior of general stationary stochastic processes developed in Albin [1–3] with research on the detection of breaks in linear models in Jarušková [18,19], and Albin and Jarušková [4]. Noting that the maximally selected quasi-likelihood procedure can be rewritten in terms of a quadratic form maximum, the asymptotic analysis can be carried out by determining the large sample properties of the quadratic form supremum of a Gaussian process.

More precisely, this Gaussian process is defined as a multivariate stochastic integral with respect to standard Brownian motion involving Legendre polynomial transforms as coordinatewise integrands. In order to derive the extreme value asymptotics for the quadratic form supremum of these stochastic integrals, one has to establish sharp estimates for the

☆ Research partially supported by NSF grants DMS 0604670 and DMS 0652420, grant RGC-HKUST6428/06H, grant MSM 0021620839 and grant GACR 201/06/0186.
* Corresponding author.
E-mail addresses: alexaue@wald.ucdavis.edu (A. Aue), horvath@math.utah.edu (L. Horváth), huskova@karlin.mff.cuni.cz (M. Hušková).

0047-259X/$ – see front matter
doi:10.1016/j.jmva.2008.10.004

eigenvalues of their covariance matrices. The computations involved are complex and delicate, and have so far been carried out only for the bivariate case (thus corresponding to linear models) with the help of mathematical software in [4]. The main aim of this paper is thus to fill in this gap and to provide a rigorous mathematical analysis that will consequently lead to the limiting extreme value behavior in the general multivariate case (thus corresponding to polynomial regressions).

The paper is organized as follows. In Section 2, we define precisely the stochastic integrals of Legendre polynomial transforms and state the limit results for the extremal tail behavior and the global extremes. Section 3 establishes first the connection between the stochastic integrals and the maximally selected quasi-likelihood procedure. In the second part, we apply our main results to the climatological benchmark data set HadCRUT3 which contains averaged global near-surface temperatures from 1850 until 2008. Finally, all proofs are given in Section 4.

2. The large sample extremal behavior

For $j \geq 0$, let

$$p_j(x) = \sqrt{\frac{2j+1}{2}}\frac{1}{2^j j!}\frac{\mathrm{d}^j}{\mathrm{d}x^j}\left(x^2-1\right)^j, \quad x \in [-1, 1]$$

denote the jth order Legendre polynomial on $[-1, 1]$. The sequence of functions $(p_j(x): j \geq 0)$ forms an orthonormal basis in the Hilbert space $\mathcal{L}^2[-1, 1]$. Since $p_j(x)$ is a polynomial of order j, its leading coefficient, say, a_j (assigned to x^j) is nonzero and also strictly positive (see p. 36 of [13]). Therefore, the Legendre polynomial transforms onto the interval $[0, t]$ given by

$$\varphi_{j,t}(x) = \left(\frac{t}{2}\right)^j \frac{1}{a_j} p_j\left(\frac{2x}{t} - 1\right), \quad x \in [0, t],$$

are well defined. Let $(W(t): t \geq 0)$ be a standard Brownian motion. For $j \geq 0$ and $t > 0$, set

$$Z_j(t) = \int_0^t \varphi_{j,t}(x)\mathrm{d}W(x) \quad \text{and} \quad Q_j(t) = \frac{Z_j(t)}{\sqrt{\operatorname{Var} Z_j(t)}}. \tag{2.1}$$

As usual in the study of standardized variables, we shall work with the exponentially transformed processes $(X_j(t): t \geq 0)$ defined by

$$X_j(t) = Q_j(\mathrm{e}^t), \quad t \geq 0,\ j \geq 0.$$

It will be shown in Section 4 (see Lemma 4.1) that the processes $(X_j(t): t \geq 0)$ have a simpler covariance structure than their counterparts $(Q_j(t): t \geq 0)$. As a consequence of the exponential transformation, we obtain for example that the vector-valued process $(\boldsymbol{X}(t): t \geq 0)$ given by

$$\boldsymbol{X}(t) = (X_0(t), \ldots, X_p(t))', \quad t \geq 0,$$

is stationary for any fixed $p > 0$. Here, $\boldsymbol{x}'$ denotes the transpose of a vector $\boldsymbol{x}$. We are interested in the asymptotic behavior (as $T \to \infty$) of $M(T) = \sup_{0 \leq t \leq T} |\boldsymbol{X}(t)|$ and in the tail behavior of $M(h)$ for a fixed $h > 0$. To derive the limits, we shall apply the general method developed in Albin [1], which relies heavily on properties of the covariance matrix, say, $\mathbf{A}(t)$ of the random vector $(\boldsymbol{X}(0), \boldsymbol{X}(t))$, where $t > 0$. The structure of $\mathbf{A}(t)$ will be the subject of Lemmas 4.1 and 4.2. Albin and Jarušková [4] considered the special case $p = 1$ for which they computed the eigenvalues and the eigenfunctions of the covariance matrix $\mathbf{A}(t)$ with the computer software *Mathematica*. This technique, however, is not applicable in the case of an arbitrary $p > 1$. We shall therefore develop the necessary technical framework in this paper.

Our first result concerns the tail behavior of $M(h)$ for a given $h > 0$. Denote by $\Gamma(t) = \int_0^\infty \mathrm{e}^{-y}y^{t-1}\mathrm{d}y$ the Gamma function.

Theorem 2.1. *For any $h > 0$, it holds that*

$$\lim_{u\to\infty} \frac{\mathrm{e}^{u/2}}{u^{(p+1)/2}} P\left(M(h) > u\right) = C_p h \quad \text{with } C_p = \frac{p+1}{2^{(p+1)/2}\Gamma((p+1)/2)}.$$

While the high level exceedance statement of Theorem 2.1 is certainly of independent interest for statistical inference about the tail behavior of $M(h)$, it is utilized here mainly as an auxiliary tool to provide the following result for the global extremes of $M(T)$ as $T \to \infty$.

Theorem 2.2. *For any real x, it holds that*

$$\lim_{T\to\infty} P\left(M(T) - g(T) \leq 2x\right) = \exp\left(-\mathrm{e}^{-x}\right),$$

where $g(T) = 2\log T + (p+1)\log\log T + \log C_p$.

Both theorems are based on a series of lemmas which are stated and proved in Section 4. Similar results for the norm of Gaussian vector processes are available in the literature. Horváth [14] and Pitarbarg [26] discuss, for example, the case