

Inference of multivariate dependence structures in extreme value theory

Stima della dipendenza nei valori estremi multivariati

Giulia Marcon, Simone A. Padoan, Philippe Naveau, Pietro Muliere

Abstract Modelling dependence structures in high-dimensional problems of extreme events is of interest in several application areas. Current dependence models for multivariate extremes are based upon max-stable distributions and one approach is to investigate the Pickands dependence function through nonparametric estimators. In the bivariate setting, there exist several estimators while more problematic are the assumptions that must be satisfied in multivariate extremes. The aim is to briefly review an existing nonparametric inference method for estimating the Pickands function, which assume known marginal distributions, in the multivariate framework.

Abstract *L'analisi della dipendenza nella teoria dei valori estremi multivariati è di interesse in diversi settori applicativi. Alcuni dei più recenti modelli di dipendenza sono basati su distribuzioni max-stabile ed un particolare approccio prevede lo studio della funzione di dipendenza Pickands attraverso stimatori non parametrici. Nel caso bivariato esistono numerosi stimatori mentre sono più problematici i presupposti che devono essere verificati nel caso multivariato. L'obiettivo è quello di esaminare un metodo non parametrico per stimare la funzione Pickands, assumendo note le distribuzioni marginali, nel caso multivariato.*

Key words: Extreme values, Max-stable distribution, Pickands dependence function, Nonparametric estimation, Bernstein polynomials

Giulia Marcon
Università Bocconi, Milano, Italia e-mail: giulia.marcon@phd.unibocconi.it

Simone A. Padoan
Università Bocconi, Milano, Italia e-mail: simone.padoan@unibocconi.it

Naveau Philippe
Laboratoire des Sciences du Climat et l'Environnement CNRS, Gif-sur-Yvette, France e-mail: Philippe.Naveau@lsce.ipsl.fr

Pietro Muliere
Università Bocconi, Milano, Italia e-mail: pietro.muliere@unibocconi.it

1 Introduction

In the last years, interest in high-dimensional and multivariate problems of extreme events has increased. In some fields, as environment and economics, spatial analysis of dependence structures are more often needed. Extreme value theory provides the necessary tools to deal with an overall analysis of both marginal distributions and dependence structures. This work aims to be a contribution in the direction of proposing an overview of nonparametric methods for estimating the dependence function, assuming knowledge of the marginal distributions, in the multivariate framework.

1.1 Dependence function

Let $\mathbf{X}_i = (X_{i,1}, \dots, X_{i,d})$, $i = 1, \dots, n$, be a d -dimensional random variable, with continuous marginal distribution functions $F_1, \dots, F_d$, and assume it follows a multivariate max-stable distribution [16]. To make the problem more tractable, assume that the margins are unit Fréchet, that is $\mathbb{P}(X_j \leq x) = e^{-1/x_j}$, $x_j > 0$. Then, the joint distribution of $\mathbf{X}$ can be written as

$$G(\mathbf{x}) = \exp\{-V(\mathbf{x})\}, \quad V(\mathbf{x}) = \left(\frac{1}{x_1} + \dots + \frac{1}{x_d} \right) A(\mathbf{w}). \quad (1)$$

The exponent measure function $V(\mathbf{x})$ is an homogeneous function of order -1 which describes the dependence structure among data [4]. It can be rewritten in terms of the Pickands dependence function $A(\mathbf{w})$ which is a restriction in the unit simplex,

$$\mathcal{S}_{\bar{d}} := \left\{ (w_1, \dots, w_{\bar{d}}) \in [0, 1]^{\bar{d}} : \sum_{j=1}^{\bar{d}} w_j \leq 1, w_j \geq 0 \forall j \in \{1, \dots, \bar{d}\} \right\}, \quad (2)$$

where $w_j = \frac{x_j}{r}$, $r = x_1 + \dots + x_d$, are pseudo polar coordinates [4] and $\bar{d} = d - 1$. A Pickands function $A(\mathbf{w})$ needs to satisfy the following properties [8]:

1. $A(\mathbf{w})$ is continuous and convex;
2. $A(\mathbf{w})$ has lower limit at $\max(w_1, \dots, w_{\bar{d}}, 1 - \sum_{j \leq \bar{d}} w_j) \geq 1/d$ which corresponds to the case of complete dependence;
3. $A(\mathbf{w})$ has upper limit equal to one, for any $\mathbf{w} = (w_1, \dots, w_{\bar{d}}) \in \mathcal{S}_{\bar{d}}$ and it implies that each component of $\mathbf{X}$ is independent of all the others;
4. $A(\mathbf{w})$ at the extremal points of the simplex, $\mathbf{e}_j = (0, \dots, 0, 1, 0, \dots, 0)$ for $1 \leq j \leq \bar{d}$ and $\mathbf{0} = (0, \dots, 0)$, assumes the unit value.

In order to model the exponent measure function, many parametric models have been proposed in literature while they can be quite restrictive despite their flexibility. This reason leads many authors focusing on nonparametric estimates of the dependence function or of its restriction $A(\mathbf{w})$. Extending to the multivariate case, however, most of the estimators proposed provide results which not satisfied the

above properties 1 – 4. It implies that the resulting estimate of $V(\mathbf{x})$ in (1) does not match a proper distribution $G(\mathbf{x})$.

2 Nonparametric estimation

Several nonparametric estimators of the Pickands dependence function have been proposed. In the bivariate case the first examples are given in [15] and [5]. Further modifications have been proposed and discussed in [2] and [9], assuming the marginal to be known. The estimation problem of the Pickands dependence function in the case $d > 2$ was studied in [18] and [6], assuming knowledge of the marginal distributions. Their estimators are based on functionals of the transformed random variables $Y_{ij} = -\log F_j(X_{ij})$ ($i = 1, \dots, n$, $j = 1, \dots, d$), as in the estimators proposed by [15] and [2] for the bivariate case. In the multivariate case other extensions can be found in [1] and [7]. However, these estimators of the Pickands dependence functions are not in general valid Pickands functions themselves due to possible violation of necessary conditions. An alternative estimator is proposed in [12] as a solution of an optimization problem subject to constraints which guarantee the satisfaction of all required properties. The latter estimator is discussed in this review.

2.1 Nonparametric estimator using Bernstein polynomials

The methodology adapted in [12] is based on the projection of a pilot estimate $\hat{A}$, returned from some basic estimators, onto the set of functions $\mathcal{A}$ which satisfy properties 1 – 4. The representation of the Pickands function,

$$B_A(\mathbf{w}; k) = \sum_{\ell \in L_k} \beta_\ell b_\ell(\mathbf{w}; k) \quad (3)$$

is given through the Bernstein polynomials where $k \in \mathbb{N}$ determines the order of the polynomials, the vector of coefficients β defines the shape of the representation, with $\beta_\ell \in [0, 1]$. L_k is the set of all $\bar{d}$ -dimensional vectors with values in $\{0, \dots, k\}$, for which the components have sum less or equal to k . The polynomial basis $b_\ell(\mathbf{w}; k)$, with $\mathbf{w} \in \mathcal{S}_{\bar{d}}$, are the multivariate extension of $b_\ell(w; k) = \binom{k}{j} w^j (1-w)^{k-j}$, $j = 0, \dots, k$, $w \in [0, 1]$, as in [10] and [14]. This approach has been selected because Bernstein polynomials can uniformly approximate any functions [11] and since $A(\mathbf{w})$ is continuous on $\mathcal{S}_{\bar{d}}$, it follows that the sequence of Bernstein polynomials (3) converges uniformly to A on $\mathcal{S}_{\bar{d}}$ as k goes to infinity, in other words

$$\lim_{k \rightarrow \infty} B_A(\mathbf{w}; k) = A(\mathbf{w}). \quad (4)$$

Thus, the functional representation $B_A(\mathbf{w}; k)$ provides an approximation of A , for any given k ; larger is k and more accurate the approximation becomes. However, their flexibility is useful to provide estimates of functions subject to shape restrictions such as the Pickands, in fact Bernstein polynomials possess the optimal shape restriction property among all polynomials [3], so that an estimation procedure subject to shape restrictions can be easily accommodated. Another advantage of using the above methodology is that all the constraints are expressed taking into account only the coefficients β of the Bernstein representation (3), thus the estimation method appears to be computationally handy also in high dimensions. Finally, the resulting estimator $\tilde{A}$ is obtained solving a constrained optimization problem. The resulting estimator has nice asymptotic properties, under opportune conditions, as sample size increases.

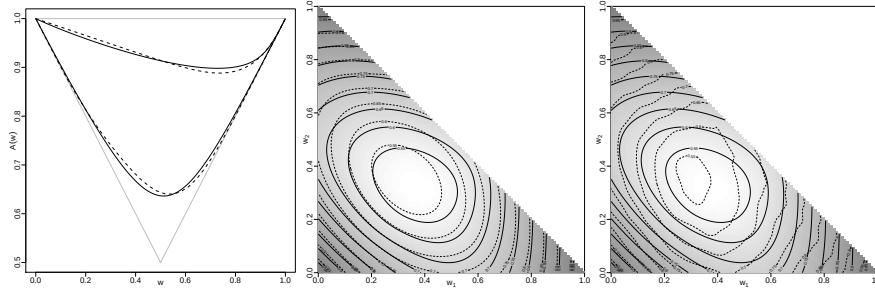


Fig. 1 Theoretical (solid line) and estimated (dashed line) Pickands dependence function of the extremal skew-t model: (a) bivariate case (b) trivariate case with $\tilde{A}(\mathbf{w})$ (c) trivariate case with $\hat{A}(\mathbf{w})$

Figure 1 shows the theoretical (solid line) and estimated (dashed line) Pickands functions corresponding to the extremal skew-t model (EST) [13] for the bivariate and trivariate case. $\tilde{A}$ is the estimator proposed in [12] with an order of the polynomial $k = 10$. In particular, the first plot refers to the EST model with $\alpha_1 = 3$, $\alpha_2 = -6$, $\nu = 3$, and two different correlation parameters, i.e. $\omega = -0.2$ and 0.9 which determine low and strong dependence respectively. The middle plot shows the estimated Pickands function for the EST model with $\alpha_1 = 7$, $\alpha_2 = -10$, $\alpha_3 = 1$, $\nu = 3$ and $\omega = 0.9$. The last plot represents, for the same model, the preliminary estimation $\hat{A}$, most commonly used in literature. Notice that $\hat{A}$ is not convex and therefore is not a proper Pickands function, whereas $\tilde{A}$ satisfies all the required constraints at each point of the domain.

References

1. Berghaus, B., Bücher, A., Dette, H.: Minimum distance estimators of the Pickands dependence function and related tests of multivariate extreme-value dependence. *Journal de la Société Française de Statistique* 154(1), 116-137 (2013)

2. Capéraà, P., Fougères, A.L., Genest, C.: A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika* 84(3), 567-577, (1997)
3. Carnicer, J., Peña J.: Shape preserving representations and optimality of the bernstein basis. *Adv. Comput. Math.* 1 (2), 173-196 (1993)
4. de Haan L., Ferreira A.: *Extreme Value Theory: An Introduction*. Springer, Boston (2006)
5. Deheuvels, P.: On the limiting behavior of the pickands estimator for bivariate extreme-value distributions. *Statist. Probab. Lett.* 12, 429-439 (1991)
6. Gudendorf, G., Segers, J.: Nonparametric estimation of an extreme-value copula in arbitrary dimensions. *J. Multivariate Anal.* 102, 3747 (2011)
7. Gudendorf, G., Segers, J.: Nonparametric estimation of multivariate extreme-value copulas. *Journal of Statistical Planning and Inference* 142 (12) 3073-3085 (2012)
8. Falk, M., Hüsler, J., Reiss, R.: *Laws of Small Numbers: Extremes and Rare Events* (2011)
9. Hall, P., Tajvidi, N.: Distribution and dependence-function estimation for bivariate extreme-value distributions. *Bernoulli* 6 (5), 835-844 (2000)
10. Li, F. J.: Interpolation and convergence of bernstein-bézier coefficients. *Acta Mathematica Sinica, English Series* 27 (9), 1769-1782 (2011)
11. Lorentz, G.: *Bernstein Polynomials*. Mathematical expositions. University of Toronto Press (1953)
12. Marcon, G., Padoan, S.A., Naveau, P., Muliere, P.: Nonparametric estimation of the Pickands dependence function using Bernstein polynomials. *Submitted* (2014)
13. Padoan, S.A.: Multivariate extreme models based on underlying skew-t and skew-normal distributions. *Journal of Multivariate Analysis* 102, 977-991 (2011)
14. Petrone, S.: On the role of mixtures in bayesian nonparametrics. *Atti della XLII Riunione Scientifica della Società Italiana di Statistica* 1, 257-268 (2004)
15. Pickands, J.: Multivariate extreme value distribution. In *Proceedings of the 43rd Session of the I.S.I. International Statistical Institute*, 859-878 (1981)
16. Resnick S.I.: *Extreme values, Regular Variation and Point Processes*. Berlin: Springer-Verlag (1987)
17. Wang, J., Ghosh, S.: Shape restricted nonparametric regression with bernstein polynomials. *Computational Statistics and Data Analysis* 56 (9), 2729-2741 (2012)
18. Zhang, D., Wells, M. T., Peng, L. Nonparametric estimation of the dependence function for a multivariate extreme value distribution. *J. Multivariate Anal.* 99, 577-588 (2008)