Research results in 2015

Stéphane Girard

Inria Grenoble Rhône-Alpes & LJK (team MISTIS).

655, avenue de l'Europe, Montbonnot. 38334 Saint-Ismier Cedex, France

Stephane.Girard@inria.fr

Abstract

This note reports my research results in 2015. Three main research topics are addressed: High dimensional statistical learning, Copulas and Extreme-value analysis.

1 Extreme-value analysis

The decay of the survival function is driven by a real parameter called the extreme-value index. When this parameter is positive, the survival function is said to be heavy-tailed. The estimation of extreme risk measures is addressed in this context in [1, 2]. Applications are found in hydrology [3, 4, 5] and more generally in the modelling of environmental risks [6]. We also refer to [7] for an application to extreme-value methods to intrinsic dimensionality estimation.

A popular way to study the tail of a distribution function is to consider its high or extreme quantiles. While this is a standard procedure for univariate distributions, it is harder for multivariate ones, primarily because there is no universally accepted definition of what a multivariate quantile should be. In this paper, we focus on extreme geometric quantiles. Their asymptotics are established, both in direction and magnitude, under suitable integrability conditions, when the norm of the associated index vector tends to one [8, 9].

2 Copulas

Copula provides a relevant tool to build multivariate probability laws, from fixed marginal distributions and required degree of dependence. From Sklar's Theorem, the dependence properties of a continuous multivariate distribution can be entirely summarized, independently of its margins, by a copula. I proposed a new families of multivariate copulas adapted to high-dimensional problems [10, 11]. The estimation may be performed using a moments method [12]. I also addressed the nonparametric estimation of the conditional tail copula [13].

3 High dimensional statistical learning

Besides, I developped dimension reduction methods for high dimensional regression problems [14] with some applications in astrophysics.

Finally, I have proposed a parametrization of the Gaussian mixture model for classification purposes. It is assumed that the high-dimensional data live in subspaces with intrinsic dimensions smaller than the dimension of the original space and that the data of different classes live in different subspaces with different intrinsic dimensions. New high-dimensional data classifiers are introduced on the basis of this model. The use of kernel methods permits to extend the classifiers to the non-Gaussian framework [15]. See [16] for an application to the classification of hyperspectral images. An application of multinomial mixture models to the classification of verbal autopsy data has also been developped [17, 18].

References

[1] S. Girard and S. Louhichi. On the strong consistency of the kernel estimator of extreme conditional quantiles. In E. Ould-Said et al., editor, *Functional Statistics and Applications*, pages 59–77. Springer, 2015.

[2] E. Deme, S. Girard, and A. Guillou. Reduced-bias estimators of the conditional tail expectation for heavy-tailed distributions. In M. Hallin et al., editor, *Mathematical Statistics and Limit Theorems*, pages 105–123. Springer, 2015.

[3] B. Barroca, P. Bernardara, S. Girard, and G. Mazo. Considering hazard estimation uncertain in urban resilience strategies. *Natural Hazards and Earth System Sciences*, 15:25–34, 2015.

[4] J. El Methni, L. Gardes, and S. Girard. Estimation de mesures de risque pour des pluies extrêmes dans la région Cévennes Vivarais. *La Houille Blanche*, 4:46–51, 2015.

[5] J. El-Methni, L. Gardes, and S. Girard. Estimation non-paramétrique de mesures de risque pour des lois conditionnelles à queues lourdes avec application à des extrêmes pluviométriques. In *Congrès de la SMAI*, Les Karellis, juin 2015.

[6] S. Girard, A. Rivera, M. Stehlìk, and S. Torres. Extreme-value modelling of some glacial processes in Chilean Andes. In *International Conference on Risk Analysis*, Barcelona, Spain, mai 2015.

[7] L. Amsaleg, O. Chelly, T. Furron, S. Girard, M. Houle, and K. Kawarabayashi. Estimating local intrinsic dimensionality. In *21st ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 29–38, Sydney, Australia, aout 2015.

[8] S. Girard and G. Stupfler. Extreme geometric quantiles in a multivariate regular variation framework. *Extremes*, 18(4):629–663, 2015.

[9] S. Girard and G. Stupfler. On the asymptotic behaviour of extreme geometric quantiles. In *9th International Conference on Extreme Value Analysis*, Ann Arbor, USA, juin 2015.

[10] F. Durante, S. Girard, and G. Mazo. Copulas based on Marshall-Olkin machinery. In U. Cherubini et al., editor, *Marshall-Olkin Distributions. Advances in Theory and Applications*, volume 141 of *Springer Proceedings in Mathematics and Statistics*, pages 15–31. Springer, 2015.

[11] G. Mazo, S. Girard, and F. Forbes. A class of multivariate copulas based on products of bivariate copulas. *Journal of Multivariate Analysis*, 140:363–376, 2015.

[12] G. Mazo, S. Girard, and F. Forbes. Weighted least square inference based on dependence coefficients for multivariate copulas. *ESAIM: Probability and Statistics*, 19:746–765, 2015.

[13] L. Gardes and S. Girard. Nonparametric estimation of the conditional tail copula. *Journal of Multivariate Analysis*, 137:1–16, 2015.

[14] A. Chiancone, J. Chanussot, and S. Girard. Collaborative Sliced Inverse Regression. In *20th Young Statistical Meeting*, Vorau, Austria, octobre 2015.

[15] C. Bouveyron, M. Fauvel, and S. Girard. Kernel discriminant analysis and clustering with parsimonious Gaussian process models. *Statistics and Computing*, 25:1143–1162, 2015.

[16] M. Fauvel, C. Bouveyron, and S. Girard. Parsimonious Gaussian process models for the classification of hyperspectral remote sensing images. *IEEE Geoscience and Remote Sensing Letters*, 12(12):2423–2427, 2015.

[17] S. Sylla, S. Girard, A. Diongue, A. Diallo, and C. Sokhna. A classification method for binary predictors combining similarity measures and mixture models. *Dependence Modeling*, 3:240–255, 2015.

[18] S. Sylla, S. Girard, A. Diongue, A. Diallo, and C. Sokhna. Classification de données binaires via l'introduction de mesures de similarités dans les modèles de mélange. In *47èmes Journées de Statistique organisées par la Société Française de Statistique*, Lille, juin 2015.