Multivariate sparse clustering for extremes

Nicolas Meyer[1]; Olivier Wintenberger[2]

[1] University of Copenhagen, Universitetsparken 5, 2100-DK Copenhagen, meyer@math.ku.dk
[2] LPSM, Sorbonne Université, 4 place Jussieu, 75005 Paris, France, olivier.wintenberger@upmc.fr

Abstract:
Studying the tail dependence of multivariate extremes is a major challenge in extreme value analysis. Under a regular variation assumption, the dependence structure of the positive extremes is characterized by a measure, the spectral measure, defined on the positive orthant of the unit sphere. This measure gathers information on the localization of large events and has often a sparse support since such events do not simultaneously occur in all directions. However, it is defined via weak convergence which does not provide a natural way to capture this sparsity structure. In this talk, we introduce the notion of sparse regular variation which allows to better learn the tail structure of a random vector X. We use this concept in a statistical framework and provide a procedure which captures clusters of extremal coordinates of X. This approach also includes the identification of a threshold above which the values taken by X are considered as extreme. It leads to an efficient algorithm called MUSCLE for MUltivariate Sparse Clustering for Extremes. We illustrate our method on numerical experiments and on real-world data sets.

Keywords:
Euclidean projection; model selection; multivariate extremes; regular variation