Multivariate Quantiles via Set Optimization

Andreas Hamel
Free University of Bozen-Bolzano,
Faculty of Economics and Management, Italy
E-mail: andreas.hamel@unibz.it

Abstract

What is a quantile, e.g., the median, of a random vector? This simple question does not have a textbook answer yet. In the statistical literature, depth regions and depth functions are used as multivariate substitutes for univariate quantiles and cumulative distribution functions but they are not straightforward generalizations of one-dimensional quantiles and cdf's.
The major difficulty is related to the order relation which is basic for defining quantiles: while in one dimension, only $\leq$ for real numbers makes sense, it is not clear which order should be used in higher dimensions. A new concept for multivariate quantiles is introduced based on vector orders generated by convex cones. A new statistical function, called lower cone distribution function, is defined as an alternative to the joint distribution function and then quantile sets as upper level sets of this function. In this way, quantile functions become set-valued, share almost all properties with univariate quantile functions and are indeed generalizations of the latter: this can be shown by set optimization techniques.
Applications are given to classification of multidimensional data points.

The talk is based on

- Hamel, Kostner, Cone distribution functions and quantiles for multivariate random variables, J. Multivariate Analysis 2018,
- Ararat, Hamel, Lower cone distribution functions and set-valued quantiles form Galois connections, Theory of Probability and Its Applications 2020,
- Hamel, Kostner, Computation of quantile sets for bi variate ordered data, Comp. Stat. & Data Analysis 2022.