Robustness and efficiency of multivariate coefficients of variation

S. Aerts[1,*], G. Haesbroeck[2] and C. Ruwet[2,3]

[1] *HEC-ULg, University of Liege, Belgium; stephanie.aerts@ulg.ac.be,*
[2] *Department of Mathematics, University of Liege, Belgium; g.haesbroeck@ulg.ac.be, c.ruwet@ulg.ac.be*
[3] *Service de mathématiques, Haute Ecole de la Province de Liège, Belgium; christel.ruwet@hepl.be*
[*] *Corresponding author*

Abstract. The coefficient of variation is a well-known measure used in many fields to compare the variability of univariate variables having really different means or expressed in different scales. However, when the dimension of the problem is greater than one, comparing the variability only marginally may lead to controversial results. Several multivariate extensions of the univariate coefficient of variation have been introduced in the literature in order to summarize global relative variability in one single index (see Albert & Zhang, 2010, for a review). These multivariate coefficients are defined in terms of the covariance matrix (via its trace, determinant,...) and of the mean vector of the underlying distribution.
In practice, these coefficients can be estimated by plugging any pair of location and covariance estimators in their definitions. However, as soon as the classical mean and covariance matrix are under consideration, the influence functions are unbounded, while the use of any robust estimators yields bounded influence functions.
While useful in their own right, the influence functions of the multivariate coefficients of variation will be further exploited in this talk to derive a general expression for the corresponding asymptotic variances under elliptical symmetry. Simulations compare the finite-sample efficiency of the classical estimator and the robust approach based on the Minimum Covariance Determinant estimator.
Then, focusing on two of the considered multivariate coefficients, a diagnostic tool based on their influence functions, as suggested in Pison & Van Aelst (2004) is derived. It allows to detect those observations having the tendency to increase or decrease the relative dispersion and it will be compared, on a real-life dataset, with the usual distance-plot.

Keywords. *Coefficient of Variation ; Influence Function ; Minimum Covariance Determinant Estimator*

References

Albert, A. & Zhang, L. (2010) A novel definition of the multivariate coefficient of variation. *Biometrical Journal*, **52**, 667–675.

Pison, G. and Van Aelst, S. (2004). Diagnostic plots for robust multivariate methods. *J. Comput. Graph. Statist.*, **13**, 310–329.