Contribution ID: **63** Type: **not specified**

Cellwise Outliers: Challenges and Solutions

Tuesday, 17 September 2024 14:00 (1 hour)

It is well-known that real data often contain outliers. The term outlier typically refers to a case, corresponding to a row of the $n \times d$ data matrix. In recent times also cellwise outliers are being considered. These are suspicious cells (entries) that can occur anywhere in the data matrix. Even a relatively small proportion of outlying cells can contaminate over half the rows, which is a problem for existing rowwise robust methods. Challenges posed by cellwise outliers are discussed, and some methods developed so far to deal with them. On the one hand there has been work on the detection of outlying cells, after which one might replace them by missing values and apply techniques for incomplete data. On the other hand cellwise robust methods have been constructed, yielding estimates that are less affected by outlying cells. In lower dimensions the focus has been on robust estimation of location and covariance as well as linear regression, whereas in higher dimensions one needs cellwise robust principal components. Some real data examples are provided for illustration.

Type of presentation

Talk

Classification

Mainly methodology

Keywords

Anomaly detection; Multivariate methods; Robust statistics

Primary author: Prof. ROUSSEEUW, Peter (KU Leuven)

Presenter: Prof. ROUSSEEUW, Peter (KU Leuven)

Session Classification: George Box Award: Peter Rousseeuw

Track Classification: Box award