Outlier Detection in Large Sets of Multivariate Time Series

P. Galeano and D. Peña

1 Departamento de Estadística, Universidad Carlos III, Madrid, Spain; pedro.galeano@uc3m.es, daniel.pena@uc3m.es

Undergraduate Prep: Presenting author

Keywords. Dynamic Factor Models; Big Data; Additive Outlier; Principal Components, Projection Pursuit; Robust Estimation.

This article presents a procedure based on projections to find outliers in a large set of multivariate time series. It is assumed that the data have been generated by a Dynamic Factor Model and two types of outliers are considered. Common outliers, generated by the factors, which affect several or all of the time series, and specific or idiosyncratic outliers, which are generated by the specific components and affect a single time series. The outliers are identified by projecting the vector of time series into directions with some optimality properties and searching for univariate outliers in this directions. The procedure is fast to apply and does not require to specify a multivariate model for the data.