

Multivariate ARIMA Compositional Time Series Analysis

Aguilar, L.¹ y Barceló-Vidal, C.²

¹ Dept. de Matemáticas, Escuela Politécnica, Univ. de Extremadura, 10071 Cáceres, Spain

² Dept. Informàtica i Matemàtica Aplicada, Campus de Montilvi, Univ. de Girona, E-17071 Girona, Spain

A compositional time series is obtained when a compositional data vector is observed at different points in time. Inherently, then, a compositional time series is a multivariate time series with important constraints on the variables observed at any instance in time. Although this type of data frequently occurs in situations of real practical interest, a trawl through the statistical literature reveals that research in the field is very much in its infancy and that many theoretical and empirical issues still remain to be addressed. Any appropriate statistical methodology for the analysis of compositional time series must take into account the constraints which are not allowed for by the usual statistical techniques available for analysing multivariate time series. One general approach to analyzing compositional time series consists in the application of an initial transform to break the positive and unit sum constraints, followed by the analysis of the transformed time series using multivariate ARIMA models. In this paper we discuss the use of the additive log-ratio, centred log-ratio and isometric log-ratio transforms. We also present results from an empirical study designed to explore how the selection of the initial transform affects subsequent multivariate ARIMA modelling as well as the quality of the forecasts.