Robust Variable Selection in Linear regression for Compositional Data

Alexander Braumann, Fabian Schroeder, Peter Filzmoser

June 26, 2012

Abstract

This paper considers robust variable selection for a multiple linear regression model, where the regressors are compositional and the response is a real variable. As compositional data cannot be used directly in this setting, we use an approach based on the isometric log-ratio (ilr) transformation. For model selection purposes we adapted robust forward and backward stepwise selection algorithms. Simulation results confirm that the method can successfully select good models when outliers are present. The approach was furthermore applied to a real data set.
Keywords: robust variable selection, compositional data, isometric logration transformation