

Single-linkage Method to Detect Multiple Outliers with Different Outlier Scenarios in Circular Regression Model

Siti Zanariah Satari^{1,2}, Nur Faraidah Muhammad Di¹ and Roslinazairimah Zakaria¹

¹Faculty of Industrial Sciences & Technology, Universiti Malaysia Pahang, 26300 Gambang, Pahang Darul Makmur, Malaysia.

²Corresponding author: zanariah@ump.edu.my

Abstract.

Single-linkage is one of the algorithms in agglomerative clustering technique that can be used to detect outliers. The single-linkage algorithm combines two clusters with the closest pair of observations. Then, the clusters are combined into larger clusters, until all the observations are formed in the same cluster. In this study, a single-linkage algorithm method that utilised a circular distance based on the City-block distance as the similarity distance is used. The performance of the method in detecting multiple outliers for a circular regression model is tested via simulation studies with three different outlier scenarios which are outliers in u -space only, v -space only and both uv -space. The performance is measured by calculating the “success” probability (p_{out}), masking error (p_{mask}) and swamping error (p_{swamp}) for both outlier scenarios. It is found that the single linkage method performed well in detecting outliers for both outlier scenarios and applicable for circular regression model.