

Nonparametric Tests for Detection of High Dimensional Outliers

Reza Modarres*

Department of Statistics, George Washington University, Washington, DC, USA

February 6, 2021

Abstract

Based on the ordered values of the total dissimilarity of each observation from all others, we present a nonparametric method for detection of high dimensional outliers. We provide algorithms to obtain the distribution of the test statistic based on the bootstrap and for the detection of outliers. To visually identify the suspect outliers a high dimensional outlier plot using the dissimilarity of the observations is presented. We compare the interpoint distance outlier test (IDOT) with five competing methods under different distributions. We show that Wilks outlier method is not robust and loses power as the number of the variables approaches the sample size. IDOT shows the best performance for outlier detection in terms of the average number of the outliers detected and the probability of the correct identification.

Keywords: Wilks Outlier Test; Dissimilarity Measure; Interpoint Distance; Bootstrap.

AMS Subject Classifications: 62H10, 62H15, 62E15.

*Corresponding author e-mail: reza@gwu.edu