Geophysical Research Abstracts
Vol. 15, EGU2013-8709, 2013
EGU General Assembly 2013

Outlier Detection In Linear Regression Using Standart Parity Space Approach

Utkan Mustafa Durdag and Serif Hekimoglu

Yildiz Technical University, Civil Engineering Faculty, Department of Geomatic Engineering, Istanbul, Turkey (umdurdag@yildiz.edu.tr)

Despite all technological advancements, outliers may occur due to some mistakes in engineering measurements. Before estimation of unknown parameters, aforementioned outliers must be detected and removed from the measurements. There are two main outlier detection methods: the conventional tests based on least square approach (e.g. Baarda, Pope etc.) and the robust tests (e.g. Huber, Hampel etc.) are used to identify outliers in a set of measurement. Standart Parity Space Approach is one of the important model-based Fault Detection and Isolation (FDI) technique that usually uses in Control Engineering. In this study the standart parity space method is used for outlier detection in linear regression. Our main goal is to compare success of two approaches of standart parity space method and conventional tests in linear regression through the Monte Carlo simulation with each other. The least square estimation is the most common estimator as known and it minimizes the sum of squared residuals. In standart parity space approach to eliminate unknown vector, the measurement vector projected onto the left null space of the coefficient matrix. Thus, the orthogonal condition of parity vector is satisfied and only the effects of noise vector noticed. The residual vector is derived from two cases that one is absence of an outlier; the other is occurrence of an outlier. Its likelihood function is used for determining the detection decision function for global Test. Localization decision function is calculated for each column of parity matrix and the maximum one of these values is accepted as an outlier. There are some results obtained from two different intervals that one of them is between 3σ and 6σ (small outlier) the other one is between 6σ and 12σ (large outlier) for outlier generator when the number of unknown parameter is chosen 2 and 3. The measure success rates (MSR) of Baarda's method is better than the standart parity space method when the confidence intervals are chosen α=0.01 and α=0.001. The MSR of the standart parity space method is better than the Baarda method only when the confidence interval is chosen α=0.05. In the case of generating outliers in the range of 3σ-6σ, the percentage of the MSR of the standart parity space method and the Baarda method is 78.72, 66.38, respectively. Correspondingly in the case of 6σ-12σ interval, mean success rate of the standart parity space method is higher than the Baarda's. Also the Baarda method generates outliers more than the standart parity space method when the observations don't contain outlier.