The Rationality Inspection Global Method of Measuring Data

Zhao Liu[1, a], Maodong Pan[2], Shaolin Wang[1], Jiangtao Wei[1]

[1]China Satellite Maritime Tracking and Control Department, Jiangsu 214400, China;

[2]School of Mathematics Sciences, University of Science and Technology of China, Anhui 230026, China.

[a]zhaoliu@mail.ustc.edu.cn

Abstract. Least square method for polynomial extrapolation is one of data rationality inspection methods that are often used to identify outliers. However, the method is local and only inspects the rationality of current data points. If the outliers are accepted before, the method can cause the following normal points to be abandoned and the abnormal points to be received, which will lead to the accumulation of errors. Moreover, because of the method needs to accumulate points before inspection, the initial data points cannot be inspected. To this end, this paper presents a global data nationality inspection algorithm. The method is fast and accurate, and can effectively eliminate the accumulated error in global.

Introduction

The spacecraft orbit measurement and control is an important part of aerospace measurement and control system. In spacecraft trajectory measurement data processing section, the rationality inspection of the measurement data is crucial. Due to the various abnormal conditions of track radio measurement and data collection and transmission system (such as systematic error, random error, the deviation caused by the human factors, etc.), the obtained data often contain outliers. Outliers cannot be eliminated thoroughly, but the presence of outliers can cause serious impact on the measurement results [1]. For example, in the process of one measuring data points are shown in Figure 1.

As we can see from Figure 1, when measurement data contain the abnormal 10th point, the deviation of the fitting line L_2 is large. After excluding abnormal points, the result of the fitting line L_1 is far superior to L_2.
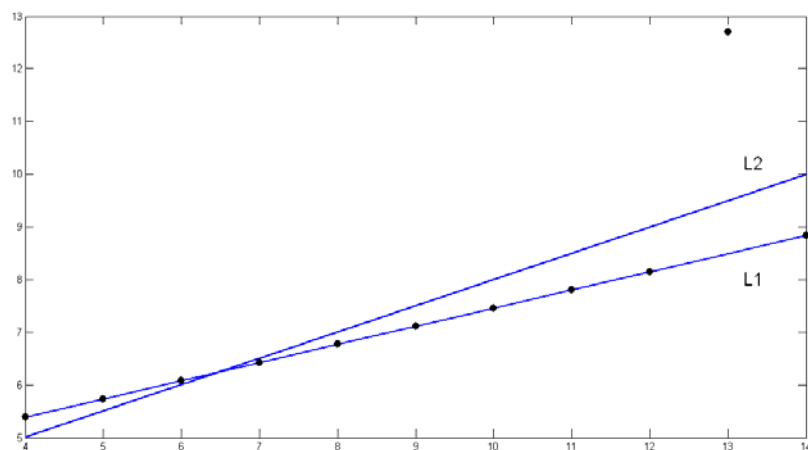

Fig. 1 Data fitting result

Therefore, identifying and eliminating outliers in measuring data is very crucial [3]. Sometimes, reasonable data are used to replace the outliers if necessary. Least square extrapolation method is widely used in aerospace measurement data rationality inspection.

Least Square Method for Data Rationality Inspection

In spacecraft measurement system, all measurement data should be inspected before used [2]. At present, the least square four points extrapolation are mainly used to forecast data for reasonable inspection [5]. The idea is: fitting curve by accumulated data, predicting the following data by least squares extrapolation [6]. If the difference between the actual measured value and the theoretical value is small, the actual measured data is treated reasonable [4]; Otherwise, the actual measured data is outlier and should be discarded or replaced by the theoretical data.

Suppose there are n pairs of experiment data $(t_i, y_i)(i=1,2,\cdots,n)$. We use these data to fit a curve which is $y = f(t) = a_0 + a_1 t + \cdots + a_m t^m$.

Define E as: $E = \sum\limits_{i=1}^{n} (y_i - \sum\limits_{j=0}^{m} a_j t^j)^2$. According to the least square theory, the best fitting curve can be obtained by minimizing E. For the sake of simplicity, we directly presents the result as following:

$$y_i' = 0.75 y_{i-4} - 1.25 y_{i-3} - 0.75 y_{i-2} + 2.25 y_{i-1} \qquad (1)$$

Set σ as the threshold value. If $|y_i - y_i'| < \sigma$, we treat the actual data y_i as the rational data; otherwise, y_i is treated as the outlier and should be replaced by y_i'.

Global Method For Data Rationality Inspection

However, the extrapolation method stated above has its weaknesses. First, four accumulated data points are needed before inspection, so the data points in the initial stage and every time after using predictive value replacing the outlier cannot be inspected. Second, the method always inspects the current data point. If there are outlier in the process of measurement was receiving, it may result in larger deviation and error accumulation.

Therefore, we propose a global method for data rationality inspection. The idea of the method is as following: in the inspection section, we inspect all measured data instead of current data. If the global error is larger than the threshold value, we treat the point, which has the greatest error, not only the current data, as the outlier. It can effectively avoid the error accumulation and simultaneously, all points could be inspected.

Suppose there are n pairs of experiment data $(t_i, y_i)(i=1,2,\cdots,n)$. We use a quadratic curve to fit these data points:

$$y = at^2 + bt + c \qquad (2)$$

Denote the new measured data point as (t_{n+1}, y_{n+1}), the global error of these data points is defined as following:

$$\delta = \sum_{i=1}^{n+1}(y_i - at_i^2 - bt_i - c)^2 = \sum_{i=0}^{n+1}\delta_i \qquad (3)$$

Where is the error of the corresponding data point. Set σ as the threshold value. We propose an global data rationality inspection algorithm as following:

1. Compute the fitting function(Equation 2) for the data point sequence $(t_i, y_i)(i=1,2,\cdots,n+1)$.

2. Compute the global error δ by Equation 3.

3. If $\delta \leq \sigma$, (t_{n+1}, y_{n+1}) is treated to be rational; otherwise, the point (t_j, y_j) is considered as the outlier where $j = \arg\max \delta_i$. Delete (t_j, y_j) from the data point sequence $(t_i, y_i)(i=1,2,\cdots,n+1)$ and go to step 2.

Results

We use an simulation experiment data to test the two methods introduced in prior section, see Figure 2. Because the least-square extrapolation method need to accumulate four initial data points before inspection, the outlier point P_1 cannot be detected. At the same time, the abnormal point P_2

cannot be detected correctly because of P_1 (extrapolation predicted value of P_2 depends on the value of P_1, see Equation 1).
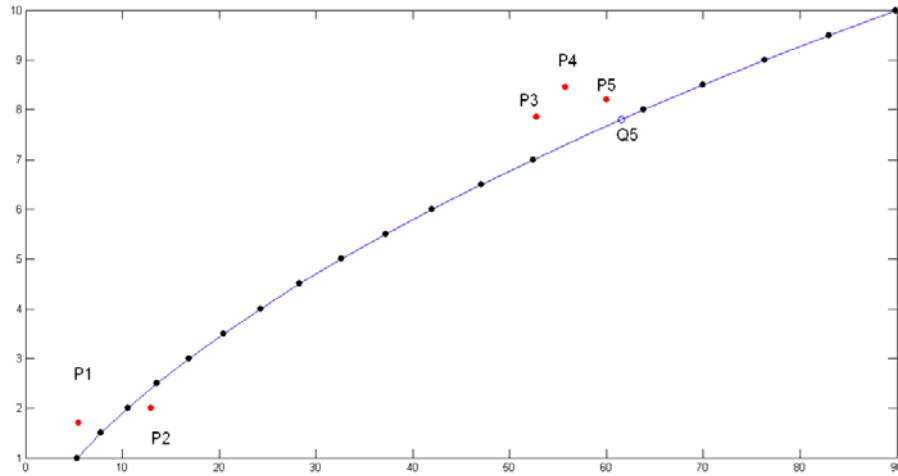

Fig. 2 Simulation results of a certain type equipment

In addition, the method only inspect whether the current point is reasonable, so the error could be accumulated. Abnormal P_3 was accepted within a certain threshold. Since then, the anomalous point P_4 was also accepted due to the impact of P_3. The errors of P_3 and P_4 are accumulated and lead to the incorrect predicted result P_5, which is far away from the normal data point Q_5. At last, Q_5 is treat as outlier and replaced but P_5.However, if we use the algorithm proposed in Section 3, the mistakes could be avoided. The outliers can be detected and it won't appear error accumulation. The algorithm is rapid, efficient and robust. Simultaneously, it can effectively avoid the error accumulation.
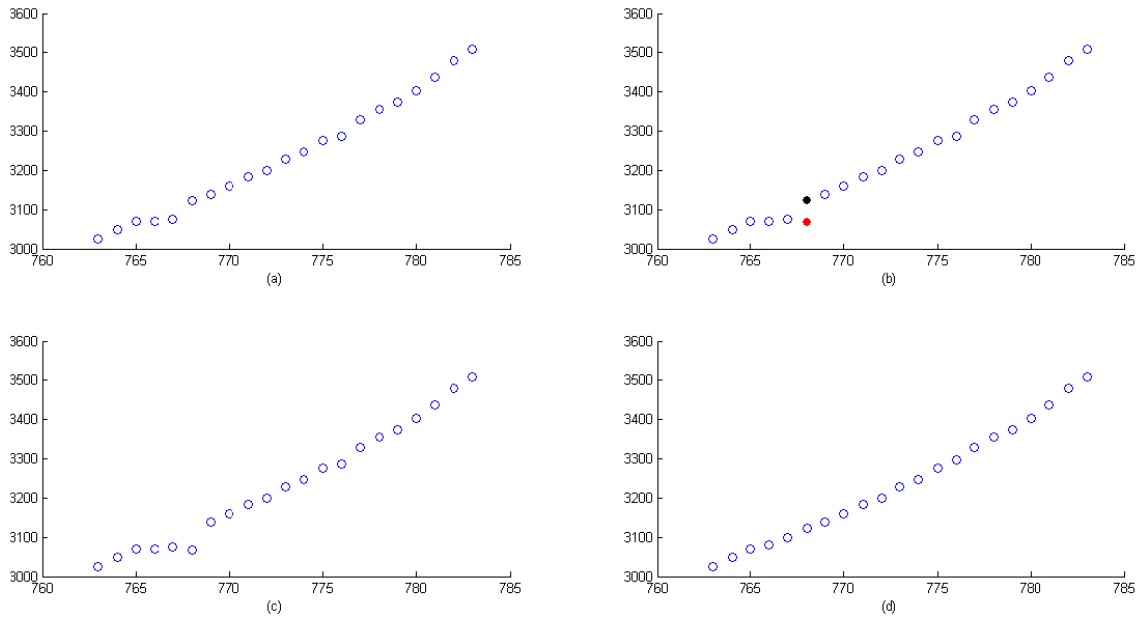

Fig. 3 Test results of a certain type equipment

The results shown in Figure 3 are actual experiment results of a certain type equipment. Figure (a) shows the actual measured data points. In figure (b), the least square method is used for data rationality inspection. The black point denotes the abnormal data detected by least square method and the red point is the result of extrapolation by Equation 1. It can be seen that due to the accumulation of measurement error, the subsequent normal measurement point (black point) is treated as outlier and replaced by the extrapolation point (red point) incorrectly. The least square method result is shown in figure (c). Figure (d) shows the inspection result of our algorithm. As can be seen from the result, our algorithm can effectively inspect the abnormal point and avoid the error accumulation.

Conclusion

In this paper, we discuss the data rationality inspection methods and propose a global data rationality inspection algorithm which can eliminate the accumulated error and improve the accuracy of measurement to overcome the weakness of the widely used least square method. The algorithm not only has the advantages of fast, accurate and robust, but also makes up for the deficiency of the least square method. Next, the rationality inspection for measurement data processing still can be improved, such as adaptive algorithm to eliminate the outliers and so on.

References

[1]   Zhi-Hua M U, Zhang H J. The Analysis System of Missile's Trajectory Accuracy Based on Least Squares Estimate[J]. Command Control & Simulation, 2007.

[2]   Li T, Zeng Y H, Wang Z M. PICKING OUT OUTLIER DATA AND ITS PARALLEL ALGORITHM[J]. Journal of Unmerical Methods & Computer Applications, 2000, (3):171-177.

[3]   Xing Ni, Yunhua Zhang, The Lgical Inspecting Method of Dynamic Measuring data. Operations Research and Management Science, 2011,(4),113-115.

[4]   Liu P C, Maddala G S. Rationality of survey data and tests for market efficiency in the foreign exchange markets[J]. Journal of International Money & Finance, 1992, 11(4):366–381.

[5]   Marquardt D W. An Algorithm for Least-Squares Estimation of Nonlinear Parameters[J]. Journal of the Society for Industrial & Applied Mathematics, 2006, 11(2):431-441.

[6]   Lawson C L, Hanson R J. Solving least squares problems.[J]. Prentice-Hall Series in Automatic Computation, 1974, 77(1):673-682.