

# The Location and Evaluation of Gross Error Based on the Partly Least Square Ridge Estimation

Cheng Zhang

School of Arts and Communication, China University of Geosciences, Wuhan 430074, China

Lihong Jin

Department of Basic, City College Wuhan University of Science and Technology, Wuhan, 430083, China

Yuting Lv

Patent Examination Cooperation Center of the Patent Office, SIPO, Sichuan, Chengdu 610213, China

**Abstract**—By using the principle of partly least square to locate and value the gross errors in the observational data, the matrix equation of the adjustment model may have ill-posed problems, which results in the locations and evaluations of the gross error unreliable. For the observational data which contain gross errors and have ill-posed problems, firstly, this paper chooses the ridge estimation to deal with the ill-posed problems, then proposes a method based on the partly least square ridge estimation to locate and value the gross errors, and the specific search steps for gross errors are given, the locations and evaluations of the gross error can be gotten by the iterative algorithm. At last, by using the method of partly least square and the partly least square ridge estimation to locate and value the gross errors in one simulating computation, and analyzing the two methods' results, the proposed method's validity is verified. This method discusses the gross error approach from another perspective, it extends the existing error theory.

**Keywords**—gross error; ill-posed problem; partly least square; ridge estimation; location and evaluation.

## I. INTRODUCTION

In the observational data process, the presence of gross error has bad effect on the adjustment result<sup>[1]</sup>, in order to eliminate or weaken the effect, extensive researches have been carried out by domestic and foreign scholars, then a lot of solutions to the problem of gross error have been gotten, such as the data detecting method<sup>[2]</sup>, the partly least square method<sup>[3,4,5]</sup>, the mean shift method<sup>[6]</sup>, LEGE<sup>[7,8,9]</sup>, QUAD<sup>[10]</sup>, robust estimation<sup>[11,12]</sup> and so on. Among these methods, the principle of partly least square method is straightforward, which has been used for handling large amount of observations, the references [3] studied the way to locate and value gross error when observational data are independent or related, the references [4] aimed at realizing fast searching gross error from the relevant observational data. Actually, inadequate observational data or approximate correlation between the parameters in the adjustment model makes the matrix of observation equation appear ill-posed problem. If the observational data contain gross error and ill-posed problem, the partly least square method fails to locate and value gross error.

In order to get the positions and values of gross error when observations have ill-posed problem, this paper introduces the ridge estimate<sup>[13,14]</sup> into the adjustment model, and establishes a theory based on partly least square ridge estimate. Firstly, the observational data are grouped, and the ill-posed problem is solved with the ridge estimate, then the partly least square ridge estimators and the expression of gross error are gotten; Secondly, the gross error search steps are given in detail, which can help achieve locating gross error; Lastly, the effectiveness of this method is demonstrated with a simulating example. This method provides a new way to handle the observation which contain simultaneously gross error and ill-posed problem, and expands the application scope of partly least square.

## II. VALUE OF GROSS ERROR

Based on the system clustering, the observational data are grouped as two parts, the first group contains  $r$  observations without gross error, the second group contain  $q$  observations with gross error, then the error equation of two groups of observations based on the Gauss-Markov adjustment model can be written as:

$$V_r = B_r \hat{X} - L_r \quad (1)$$

$$V_q = B_q \hat{X} - L_q \quad (2)$$

Where  $n$  is a number of total observations, satisfies  $r + q = n$ ,  $L_r$  is a  $r \times 1$  dimension observation vector,  $B_r$  is a known  $r \times t$  nonsingular matrix,  $t$  is a number of necessary observations,  $V_r$  is a  $r \times 1$  residual vector of the first group of observations,  $\hat{X}$  is a  $t \times 1$  dimension unknown parameter vector, the marks in the Eq.(2) have the similar definition, the observations are assumed to be independent, the diagonal matrix  $P_{rr}$ ,  $P_{qq}$

represent the right matrix of two groups of observations respectively.

If the observational data have a ill-posed problem, the coefficient matrix of the first group observations appears multicollinearity, the matrix  $N_r = B_r^T P_{rr} B_r$  closes to singular, then a small error of observations may bring sharp change to the parameter solution, which makes the adjustment estimators unreliable, and results in a failure to value gross error further.

In order to get the locations and values of gross error again, this paper brings the ridge to adjustment model, then the adjustment criterion for the first group of observations is

$$V_r^T P_{rr} V_r + \kappa X^T M X = \min \quad (3)$$

Where  $\kappa$  is a ridge parameter,  $M$  is a symmetric positive definite matrix, for the ridge estimate, it is a unit matrix, under such a condition, the estimators based on the Eq.(1) are named the partly least square ridge estimate, they can be deduced as follows:

Firstly, the minimum value of Eq.(3) on parameter  $X$  can be gotten by the follow equation:

$$\begin{aligned} & \frac{\partial (V_r^T P_{rr} V_r + \kappa X^T M X)}{\partial X} \\ &= \frac{\partial V_r^T P_{rr} V_r}{\partial V_r} \cdot \frac{\partial V_r}{\partial X} + 2\kappa X^T M \\ &= 2V_r^T P_{rr} B_r + 2\kappa X^T M = 0 \end{aligned} \quad (4)$$

Secondly, by bringing the Eq.(1) into the Eq.(4),we can get the parameter estimator:

$$\hat{X}^i = (B_r^T P_{rr} B_r + \kappa M)^{-1} B_r^T P_{rr} L_r \quad (5)$$

According to the Eq.(1) and Eq.(5),the residuals of the first group of observations are gotten:

$$V_r^i = B_r (B_r^T P_{rr} B_r + \kappa M)^{-1} B_r^T P_{rr} L_r - L_r \quad (6)$$

Then the corresponding unit weight mean square error estimate is:

$$m_r^i = \sqrt{\frac{V_r^{iT} P_{rr} V_r^i}{r - t}} \quad (7)$$

Lastly, by bringing the Eq.(5) into the Eq.(2),we can get the corrections of the second observations is:

$$V_q^i = B_q (B_r^T P_{rr} B_r + \kappa M)^{-1} B_r^T P_{rr} L_r - L_q \quad (8)$$

Through choosing appropriate ridge parameter, the parameter estimator  $\hat{X}^i$  can avoid the effects of ill-posed problem and gross error, the way to select the ridge parameter are ridge trace method, GCV method, L curve method and so on<sup>[14]</sup>.The residuals  $V_q^i$  gotten with Eq.(8) is the difference between the second group of observations  $L_q$  and the estimator  $B_q \hat{X}^i$ ,which is approximate to the true value of gross error, then the gross error estimator is  $-V_q^i$ .

### III. POSITION OF GROSS ERROR

From the Eq.(8),we know the value of gross error depends on the parameter estimator  $\hat{X}^i$  and right grouping, the parameter estimator avoids suffering from the ill-posed problem with ridge estimate, so how to properly make observations into two groups is the primary problem.

Actually, we can not learn the number and positions of gross error in advance, this paper locates gross errors through the progressive search until all the positions of gross error are gotten.

For the unit weight mean square error (MSE) of observations without gross error is smaller than that of observations with gross error, so we can decide whether to continue searching gross error through the level of reduction about the smallest unit weight mean square error during the search process. According to the reference [3] and the partly least square estimate, the criteria of the level of reduction about unit weight mean square error is:

$$bizhi = \frac{m_r^i(k-1)}{m_r^i(k)} > 2 \quad (9)$$

Where  $k$  is the number of gross errors, if the total observations contain  $k$  or  $k-1$  gross errors,  $m_r^i(k)$  or  $m_r^i(k-1)$  is the smallest unit weight mean square error, which is gotten by adjusting the first group of observations based on the partly least square ridge estimate. The specific gross error search steps are:

1) For the observations with gross error and ill-posed problem, it is assumed that the observations are without gross error, and taken as the first group of observations to adjust, through GCV method the ridge parameter is determined, and according to the Eq.(5) to Eq.(7), the partly least square ridge estimators are gotten, where  $k=0$ ,  $m_r^i(0)$  is the corresponding unit weight mean square error estimator.

2) It is assumed that the observations contain one gross error, that is  $k=1$ , the 1th, 2th,  $\dots$  nth observation are taken to the second group in turn, at the same time, the corresponding remaining  $n-1$  observations are taken to the first group, there are  $n$  kinds of grouping. Similarly, through adjusting the first group observations, the corresponding partly least square ridge estimators are gotten, the smallest unit weight mean square error estimator  $m_r^i(1)$  is chosen from the  $n$  estimators, which is taken into the Eq.(9) to calculate the *bizhi*. With the calculation results, we decide whether to continue to research the gross errors:

If  $bizhi > 2$ , the unit weight mean square error turn small obviously, it means the observations contain only one gross error, that is the observation which has the smallest unit weight mean square error,  $q_1$  represents the corresponding gross error position, then the search ends;

If  $bizhi < 2$ , though the unit weight mean square error turn small, but the level of reduction is not significant, it shows the observations contain two gross errors at least, so it is necessary to continue to search the gross errors, then the search process goes to the 3th step.

3) It is assumed that the observations contain two gross errors, that is  $k=2$ . The position of one gross error is  $q_1$ , this gross error should be deleted from the total observations before searching the remaining  $n-1$  observations. Similarly, the 1th, 2th,  $\dots$   $i$  th ( $i \neq q_1$ ),  $\dots$  nth observations are taken to the second group in turn, and the corresponding remaining  $(n-1)-1$  observations are taken to the first group to adjust, the smallest unit weight mean square error estimator  $m_r^i(2)$  is chosen from the  $n-1$  groups of partly least square ridge estimators, then the  $m_r^i(1)$  and  $m_r^i(2)$  are taken into the Eq.(9) to calculate *bizhi*. If  $bizhi > 2$ , it concludes that the  $n-1$  observations have one gross error, that is the observation corresponding to the  $m_r^i(2)$ ,  $q_2$  represents the position of this gross error. Then there are two gross errors in the observations in total, and the positions are  $q_1$  and  $q_2$  respectively; if  $bizhi < 2$ , it means that the  $n$  observations contain three gross errors, it needs to search continually. Similarly, the second gross error should be

deleted from the  $n-1$  observations, and the remaining observations are searched by repeating the above steps, until the unit weight mean square error turn small obviously.

At last, according to the positions of gross error, the observations can be classified to two groups, by adjusting the first group of observations again, the partly least square ridge estimators are gotten, then we can get the values of gross error.

#### IV. SIMULATING COMPUTATION

One simulating computation is given to certify this paper's method. The observations can be described by an equation, that is  $L = BX + \Delta$ , where  $\Delta \sim N(0,1)$ , the number of observations satisfies  $n=10$ , the true value of parameter  $X$  satisfies  $X = [11111]^T$ , the number of necessary observations is  $t=5$ . The corresponding matrix is:

$$B = \begin{bmatrix} 2 & -5 & 1 & 1 & -9.5 \\ -2 & 4 & 1 & -1.05 & 8.5 \\ -2 & 1 & 1 & -1 & 2.4 \\ -1 & 2.5 & 4 & -0.5 & 7 \\ -1 & 3.2 & 4 & -0.5 & 8.4 \\ 1 & 1 & -3 & 0.4 & 0.49 \\ 3 & 7 & -3 & 1.5 & 12.7 \\ 5 & -1 & -2 & 2.5 & -3 \\ 4 & 2 & -2 & 2.01 & 3 \\ 4 & 3 & -2 & 2 & 5 \end{bmatrix}$$

Now, the 2th, 6th observations are added with gross errors with same value 10.

According to the condition number method[14] which is used to diagnosis the ill-posed problem, the condition number of the normal equation coefficient matrix  $N = B^T P B$  is  $1.2892 \times 10^5$ , it shows the observations have a bad ill-posed problem. In order to verify the method, the simulating computation will be computed with the partly least square method and the partly least square ridge estimate method.

##### A. The Partly Least Square Method

During the gross error search steps, when we assume that the observations contain not gross error, after the adjustment, the unit weight mean square error estimator is 1.6866; when we assume the observations contain one gross error, we can get the smallest unit weight mean square error estimator is 0.8323, the position is the 2th observation, the ratio of two unit weight mean square

error estimator is greater than 2, the search ends, we judge the second observation has gross error only, it turns out that the position is different from the true positions of gross error. The parameter estimator based on the location of the gross error is :

$$\hat{X} = [60.7360 \ 11.2029 \ 3.8774 \ -118.4024 \ -3.9899]^T$$

It is far from the true value of gross error.

#### B. The Partly Least Square Ridge Estimate Method

Firstly, it is assumed that the total observations contain not gross error, after adjusting the observations, the corresponding partly least square ridge estimators are gotten, among them , the unit weight mean square error is  $m_r^i(0) = 4.7751$ ;

Secondly, it is assumed that the observations have one gross error, according to the step 2) , we can get 10 unit weight mean square error estimators  $m_r^i$ , they are showed in figure 1, it is obvious that the 6th observation has the smallest unit weight mean square error estimator  $m_r^i(1) = 3.4746$  , and the corresponding unit weight mean square error estimator ratio is,

$$bizhi = \frac{m_r^i(0)}{m_r^i(1)} = 1.3743 < 2$$

It shows that the observations have 2 gross errors at least, and the position of the first gross error is 6, it needs to search continually.

Thirdly, it is assumed that the observations have two gross errors, before searching, the 6th gross error is deleted from the observations, and the remaining 9 observations are adjusted to get the corresponding unit weight mean square error estimators  $m_r^i$ , they are showed in the figure 2, the 2th observation has the smallest unit weight mean square error estimator 0.4760, and the ratio is, that means the total observations have two gross errors, and the position of the second gross error is 2, the search ends.

$$bizhi = \frac{m_r^i(1)}{m_r^i(2)} = 7.2995 > 2$$

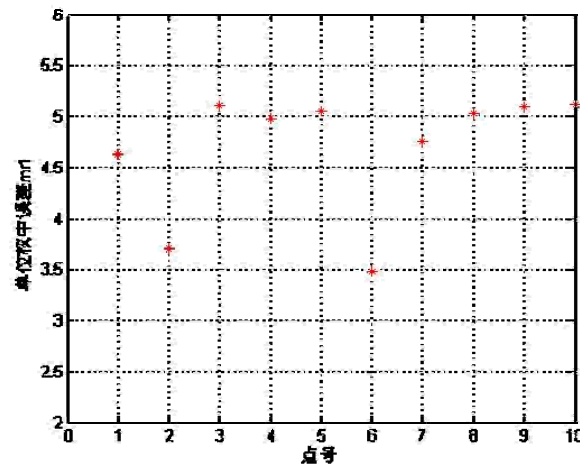


Figure 1. The unit weight MSE when the observations assumed with one gross error

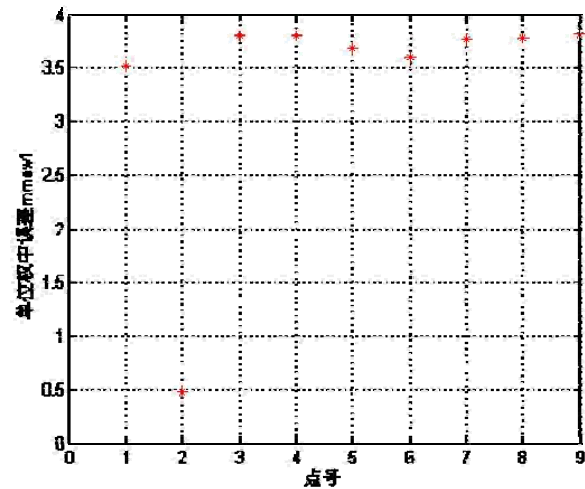


Figure 2. The unit weight MSE when the observations assumed with two gross error

Above all, the positions of gross error are the 2th, 6th observations, which satisfy the true positions. The two observations are taken to the second group, the remaining observations are taken to the first group to adjust. With the GCV method, the ridge parameter  $\kappa = 0.14$ , on this basis, the parameter estimator is:

$$\hat{X}^i = [1.1920 \ 0.5457 \ 1.1979 \ 0.6409 \ 1.2477]^T$$

Which closes to its true value, then with the Eq.(8) , we can get the gross error estimators are 9.4347, 10.7585 respectively, and the final unit weight mean square error is  $\hat{\sigma}_0 = 0.4337$ , the precision is improved obviously.

The computation results show that if the observations have ill-posed problem and gross error, the traditional partly least square method fails to locate and value gross

error, however, the partly least square ridge estimate method can get the positions and values of gross error.

## V. CONCLUSION

For realizing locating and valuing gross error when the observations have ill-posed problem, this paper proposed a method based on the partly least square ridge estimate, by combing the advantages of ridge estimate and partly least square, it can get the positions and values of gross error while overcoming the ill-posed problem. Some conclusions obtained in the deduced formula and the examples are given:

1) If observations contain gross error and ill-posed problem at the same time, the partly least square method fails to locate and value the gross error for the influence of ill-posed problem. This paper introduces the ridge estimate to deal with ill-posed problem, then the partly least square ridge estimators are deduced based on the principle of the partly least square.

2) The expression of gross error estimator is derived on the basis of the partly least square ridge estimate, and the position of gross error can be gotten by searching successively. The simulating computation certifies this paper's method is effective to locate and value gross error with high precision.

This method enriches the theory of the partly least square, extends its application scope, and provides a new way to deal with the observations with ill-posed problem and gross error.

## REFERENCES

- [1] Pan Xiong, Sun Haiyan. Semiparametric p-norm Maximum Likelihood Regression [J]. *Acta Geodaetica et Cartographica Sinica*, 2005, 34(1):30~34.
- [2] Baarda W. A Test Procedure for Use in Geodetic Networks [J]. *Netherlands Geod. Comm.*, 1968, 2(5):577.
- [3] Wang Aisheng, Ou Jikun. Method of Partly Least Squares and Application in Gross Error Location and Estimation [J]. *Science of Surveying and Mapping*, 2005, 30(2):70~72.
- [4] Lin Guoqing, Fan Dongming, Xie Yong. Method of Partly Least Squares in Fast Gross Error Location and Estimation [J]. *Surveying and Mapping*, 2010, 33(4):162~164.
- [5] Wang Haitao, Ou Jikun, Yuan Yunbin, Guo Jianfeng. On Equivalence of Three Estimators for Outliers in Linear Model [J]. *Geomatics and Information Science of Wuhan University*, 2013, 38(2): 162~166.
- [6] Song Lijie. On Theory Analysis for Expectation on Drift Model [J]. *Journal of Institute of Surveying and Mapping*, 2009, 17(3):160~162.
- [7] Yu Zongchou, Li Mingfeng. Simultaneous Location and Evaluation of Multidimensional Gross Errors [J]. *Journal of Wuhan Technical University of Surveying and Mapping*, 1996, 21(4):323~329.
- [8] Yu Zongchou, Li Mingfeng. A Further Discussion on the Nature of the LEGE Method and Its Improved Searching Method [J]. *Journal of Wuhan Technical University of Surveying and Mapping*, 1998, 23(3):238~243.
- [9] Song Lijie, Yang Yuanxi. Comparison between Data Snooping and LEGE [J]. *Acta Geodaetica et Cartographica Sinica*, 1999, 28(4):295~300.
- [10] Ou Jikun. Quasi-Accurate Detection of gross errors (QUAD) [J]. *Acta Geodaetica et Cartographica Sinica*, 1999, 28(1):15~20.
- [11] Yu Xuexiang, Lv Weicai. The Application of Robust Estimation to the Gross Error Detection and Adjustment Calculation [J]. *Engineering of Surveying and Mapping*, 1996, 5(3):40~44.
- [12] He Wei, Liming, Kan Qiyuan. Robust weighted non-interval GM(1,1) model in the large building subsidence prediction [J]. *Engineering of Surveying and Mapping*, 2014, 23(4):62~68.
- [13] Gui Qingming, Li Guozhong. Combined ridge with Shrunken Estimator and its Applications in Geodetic Adjustment [J]. *Journal of Geodesy and Geodynamics*, 2002, 22(1):16~21.
- [14] Wang Zhenjie. Research on the Regularization Solutions of Ill-posed Problems in Geodesy [D]. *Institute of Geodesy and Geophysics Chinese Academy of Sciences*, 2003.