

# A simple method for processing data with least square method

Wang chunyan<sup>\*a</sup>, Qi liquan<sup>a</sup>, Chen yongxiang<sup>a</sup>, Pang guangning<sup>a</sup>

<sup>a</sup> College of Optoelectronic Engineering, Changchun University of Science and Technology;  
Changchun, Jilin, China, 130022

## ABSTRACT

The least square method is widely used in data processing and error estimation. The mathematical method has become an essential technique for parameter estimation, data processing, regression analysis and experimental data fitting, and has become a criterion tool for statistical inference. In measurement data analysis, the distribution of complex rules is usually based on the least square principle, i.e., the use of matrix to solve the final estimate and to improve its accuracy. In this paper, a new method is presented for the solution of the method which is based on algebraic computation and is relatively straightforward and easy to understand. The practicability of this method is described by a concrete example.

**Keywords:** Error theory, Least squares, Data processing

## 1. INTRODUCTION

In order to obtain the most reliable results, the number of measurements is often more than the numbers of unknown parameters, which means that the number of residual equations is more than the number of unknown numbers. The method of general solving algebraic equations cannot be used to solve these problems. The least squares method convert the residual equation into an algebraic equation with a definite solution and the unknown parameters can be solved. This algebraic equation with definite solution is called the normal equations of least squares estimation. In this paper, we build the basis of linear algebra, and propose another method for solving the least squares estimator.

## 2. THE INTRODUCTION OF THE LEAST SQUARES PRINCIPLE

The least square method is a method that estimates the regression equation using the sample data according to the least squares criterion.

Suppose  $L_i$  is the  $i$ -th sample observed value and  $\hat{X}$  is the corresponding  $i$ -th sample best estimate. The residual between  $L_i$  and  $\hat{X}$  is denoted by  $e_i$ .

The criteria that apply the quadratic sum of the residual of all the observed values  $\sum e_i^2$  is minimized[4],  $\min \sum e_i^2 = \min \sum (L_i - X_i)^2$ , to determine the estimates of the unknown parameter  $x_1, x_2, x_3, \dots, x_k$ . And the the least squares criterion is

$$\begin{aligned} \frac{\partial \min \sum e_i^2}{\partial \hat{X}} &= \frac{\partial}{\partial \hat{X}} (L^T L - 2 \hat{X} A^T L + X^T A^T A \hat{X}) \\ &= -2 A^T L + A^T A \hat{X} \end{aligned} \quad (1)$$

and

$$\hat{X} = (A^T A)^{-1} A^T L \quad (2)$$

The standard deviation[5] of the direct measurement result is

\*245044961@qq.com; phone 13504325709

$$s = \sqrt{\frac{\sum_i e_i^2}{n-k}} \quad (3)$$

The corresponding standard deviation estimate can be obtained according to the equation

$$\sigma_{x_i} = \sigma \sqrt{d_{jj}} \quad (4)$$

where  $d_{jj} (j=1,2,3,4,\dots)$  is the value of diagonal of  $(A^T A)^{-1}$  obtained above.

### 3. Two methods of calculating $x_i$ and $\sigma_{x_i}$

#### 3.1 First method

Assume observation equation is:

$$A = \begin{Bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{Bmatrix} \quad (5)$$

where  $a_{ii} (i=1,2,3)$ —constant,  $l_j (j=1,2,3,4)$ —measurement result, which is also a constant.

List the coefficient matrix  $A$  and the measurement result matrix  $L$  [3]

$$A = \begin{Bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{Bmatrix} \quad (6)$$

$$L = \begin{Bmatrix} l_1 \\ l_2 \\ l_3 \\ l_4 \end{Bmatrix} \quad (7)$$

Eq.(1) shows that  $A^T A$  and  $A^T L$  should be known, then determine the inverse matrix  $C$  of  $A^T A$  :

$$C^{-1} = (A^T A)^{-1} \quad (8)$$

$$C^{-1} = (A^T A)^{-1} = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1t} \\ d_{21} & d_{22} & \dots & d_{2t} \\ \dots & \dots & \dots & \dots \\ d_{t1} & d_{t2} & \dots & d_{tt} \end{bmatrix} \quad (9)$$

The least square estimator of  $\hat{X}$  are obtained by multiplying  $C$  by  $AL$  .

#### 3.2 Second method

First, make a table according to the observation equation, where  $a_{ii} (i=1,2,3,4)$  and  $l_j (j=1,2,3,4)$  are all four numbers of the corresponding columns of the constants  $Ai (i=1,2,3)$  and  $Y$  , which are given by the four formulas above are assumed (Note: They are not a vector).

Next, make another form firstly.  $Ai \times Aj$  in first row represents the four numbers, which multiply the corresponding the four numbers that are represented by  $Ai$  and  $Aj$  , respectively. For example,  $A1 \times A1$  means the number of  $a_{11} \times a_{11}$  ,

$a_{21} \times a_{21}$ ,  $a_{31} \times a_{31}$ ,  $a_{41} \times a_{41}$ . And  $\sum i (i = 1, 2, 3, \dots)$  in last row represent the sum of the first four digits of the corresponding column.

Table 1 Calculation table of normal equation coefficient

$A1 \times A1$	$A1 \times A2$	$A1 \times A3$	$A2 \times A1$	$A2 \times A2$	$A2 \times A3$	$A3 \times A1$	$A3 \times A2$	$A3 \times A3$	$A1 \times Y$	$A2 \times Y$	$A3 \times Y$
$a_{11} \times a_{11}$	$a_{11} \times a_{12}$										
$a_{21} \times a_{21}$											
$a_{31} \times a_{31}$											
$a_{41} \times a_{41}$											
$\sum 1$	$\sum 2$	$\sum 3$	$\sum 4$	$\sum 5$	$\sum 6$	$\sum 7$	$\sum 8$	$\sum 9$	$\sum 10$	$\sum 11$	$\sum 12$

Then the last line of the table three per group divided into four groups and the three equations are written

$$\begin{cases} \sum 1X_1 + \sum 2X_2 + \sum 3X_3 = \sum 10 \\ \sum 4X_1 + \sum 5X_2 + \sum 6X_3 = \sum 11 \\ \sum 7X_1 + \sum 8X_2 + \sum 9X_3 = \sum 12 \end{cases} \quad (10)$$

Third, the least squares estimator could be calculated by solving the equations obtained in the second step.

**Prove:** The matrix  $X$  is grouped

$$X = \begin{pmatrix} a1 & a2 & a3 \end{pmatrix} \quad (11)$$

$$\text{So } X^T = \begin{pmatrix} b1 \\ b2 \\ b3 \end{pmatrix} \quad (\text{a is a column vector of } 4 \times 1, \text{ b is a row vector of } 1 \times 4), \text{ and so}$$

$$X^T X = \begin{bmatrix} a1b1 & a1b2 & a1b3 \\ a2b1 & a2b2 & a2b3 \\ a3b1 & a3b2 & a3b3 \end{bmatrix} \quad (12)$$

$$X^T Y = \begin{pmatrix} b1Y \\ b2Y \\ b3Y \end{pmatrix} \quad (13)$$

According to the basic characteristics of transposed and matrix

$$a1b1 = a_{11}a_{11} + a_{21}a_{21} + a_{31}a_{31} + a_{41}a_{41} \quad (14)$$

If  $\sum 1 = a1b1$ , the sum of the above table corresponds to the elements in the  $X^T X$  and  $X^T Y$  matrices. So the finally corresponding equation set in the third step is:

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (15)$$

Forth, the standard deviation  $S$  of the measured value is also obtained first, and the value of  $d_{jj}$  is calculated by the equation set obtained by the above form

$$\begin{cases} \sum 1C_1 + \sum 2C_2 + \sum 3C_3 = 1 \\ \sum 4C_1 + \sum 5C_2 + \sum 6C_3 = 0 \\ \sum 7C_1 + \sum 8C_2 + \sum 9C_3 = 0 \end{cases} \quad (16)$$

And the value of  $C_1$  obtained by the above formula is the value of  $d_{11}$ . Similarly, replace the value of the right side of the equation by 0 1 0 when calculating the value of  $d_{22}$ , and replace the value of the right side of the equation by 0 0 1

when calculating the value of  $d_{33}$ .  $d_{jj} = \begin{bmatrix} C_1 \\ C_2 \\ C_3 \end{bmatrix}$

Finally, the corresponding standard deviation is obtained by  $\sigma_{x_i} = \sigma \sqrt{d_{jj}}$ .

**Prove:** In the calculation of the matrix of the inverse matrix, the defining equation is  $CC^{-1} = E$ , from the above we already know

$$C = X^T X = \begin{bmatrix} a1b1 & a1b2 & a1b3 \\ a2b1 & a2b2 & a2b3 \\ a3b1 & a3b2 & a3b3 \end{bmatrix} \quad (17)$$

(This is a symmetric matrix, and its inverse matrix is symmetric matrix as well.) Assume its inverse matrix is

$C^{-1} = \begin{bmatrix} t1 & t4 & t5 \\ t4 & t2 & t6 \\ t5 & t6 & t3 \end{bmatrix}$ , then  $E$  is a unit matrix of  $3 \times 3$ . By matrix division, when  $C$  is multiplied by the first column of

$C^{-1}$ , the resulting value is the first column of  $E$ ,

$$\begin{bmatrix} a1b1 & a1b2 & a1b3 \\ a2b1 & a2b2 & a2b3 \\ a3b1 & a3b2 & a3b3 \end{bmatrix} \times \begin{bmatrix} t1 \\ t4 \\ t5 \end{bmatrix} = \begin{bmatrix} t1 \\ t4 \\ t5 \end{bmatrix} \quad (18)$$

The values of  $t1$ 、 $t2$ 、 $t3$  are the diagonal of the inverse matrix, that is, the value of  $d_{jj}$  which we need.

#### 4. Conclusion

As can be seen from the above two proofs, the last value obtained by the second method is exactly the same as the matrix algorithm. The first method of the least squares calculation given above is entirely based on the solution of the matrix, but the second method is more use of our more familiar algebra method. In the calculation of large amount of data, usually based on the guidance of the first method using MATLAB<sup>[1]·[2]</sup> and other computer software to solve the final value, when you encounter small data, or in the answer needs, the second method if you master, it will greatly reduce the difficulty of the calculation, and not easy to miscalculate, very practical. Finally, I would like to offer this document to teachers and students puzzled by the least squares of the matrix algorithm.

#### REFERENCES

- [1] CHEN Lanfeng, YANG Jingyu. [MATLAB simulation of curve fitting based on least-squares] Journal of Shenyang Normal University ( Natural Science Edition), (2014)
- [2] XU Yitang, [Curve Fitting Based on Least Squares Method and Its Application in MATLAB] Jiangsu University of Science and Technology, School of Electronic and Information (2012)
- [3] WANG Peng, "Strategy of Stimulating Studying Interest for Professional Course Teacher in Food Science and Engineering," Journal of Anqing Teachers College (Natural Science Edition), 21(1) (2015)
- [4] YUE Kui, "Programming for Evaluation of Roundness Error by Least Square Mean Circle Method," Journal of Hefei University of Technology, 4, (2006)
- [5] LIU Zhi-ping, SHI Lin-ying, "The Principle of Least Square Algorithm and its Achievement by MATLAB," 33-34, 7(19), (2008)