## **Improved Total Least Square Algorithm**

Deng Yonghe\*,1,2,3

<sup>1</sup>College of Engineering and designing, Lishui College, Lishui, Zhejiang, 323000, P.R. China

**Abstract:** Aim to blemish of total least square algorithm based on error equation of virtual observation, this paper proposed a sort of improved algorithm which doesn't neglect condition equation of virtual observation, and considers both error equation and condition equation of virtual observation. So, the improved algorithm is better. Finally, this paper has fitted a straight line in three-dimensional space based on the improved algorithm. The result showed that the improved algorithm is viable and valid.

Keywords: total least square, error equation, condition equation, virtual observation, mean square error of unit weight

#### 1. INTRODUCTION

Total least square algorithm has obvious advantage [1], but there exists a sort of problem because it supposed every individual of designing matrix all has error. So, literatures [1, 2] made some reasonable improvement, put forward virtual observation method and listed error equations of virtual observation. Their works are useful, but there existed some questions, because they didn't consider condition equations among virtual observations which maybe existed. So, further improvement should be made.

This paper put forward condition equations of virtual observations, considered condition equations and error equations of virtual observations at the same time, and dealed with fitting a straight line in three-dimensional space based on the improved algorithm. The result showed that the improved algorithm is viable and valid.

# 2. THE PRINCIPLE DEALING WITH TOTAL LEAST SQUARE BASED ON VIRTUAL OBSERVATIONS [1, 2]

Literature [1, 2], which supported part of  $\hat{B}$  did not give any error, yielding the model of least squares as follows:

$$\hat{L} = \hat{B}\hat{X} - d \qquad P_1 \tag{1}$$

$$\hat{L}_{\scriptscriptstyle B} = \hat{X}_{\scriptscriptstyle B} \tag{2}$$

where,

Eq.1 is actual observation equation;

 $\hat{L}$  --- estimated value of column vector of observations;

 $\hat{B}$  --- estimated value of designing matrix;

 $\hat{X}$  ---estimated value of column vector of unknown parameters;

d --- column vector of constant;

Eq.2 is virtual observation equation;

 $\hat{L}_B$  ---estimated value of column vector from virtual observation.

 $\hat{X}_B$  ---estimated value of column vector of unknown parameters.

 $P_1$  and  $P_2$  are weight matrix.

Based on Eq.1 and Eq.2,we can obtain error equation such as Eq.3

$$V' = B'\hat{x}' - l \tag{3}$$

where.

$$V' = \left[ \begin{array}{c} V \\ V_B \end{array} \right];$$

$$B' = \left[ \begin{array}{cc} B^0 & A \\ 0 & E \end{array} \right];$$

$$\hat{x}' = \begin{bmatrix} \hat{x} \\ \hat{x}_B \end{bmatrix};$$

$$l = \begin{bmatrix} L + d - B^{0} X^{0} \\ L_{B} - X^{0}_{B} \end{bmatrix};$$

*V*--- residual of actual observation value;

 $V_{B}$ --- residual of virtual observation value;

<sup>&</sup>lt;sup>2</sup>School of Geodesy and Geomatics, Wuhan University, Wuhan, 430079, P.R. China

<sup>&</sup>lt;sup>3</sup>School of Continuous Education, Wuhan University, Wuhan, 430079, P.R. China

<sup>\*</sup>Address correspondence to this author at College of Engineering and Designing, Lishui College, Lishui, Zhejiang, 323000, P.R. China; Tel: +86 18969588403; E-mails: lsxydengyonghe@sina.com; a15925722009@163.com

 $B^0$  --- approximation of  $\hat{B}$ ;

$$\Delta \hat{B} = \hat{B} - B^0 ;$$

 $X^0$  --- approximation of X;

$$\hat{x} = \hat{X} - X^0;$$

A is easily solved out by  $\Delta \hat{B} X^0 = A \hat{\chi}_B [1]$ ;

0 of left bottom of  $\begin{bmatrix} B^0 & A \\ 0 & E \end{bmatrix}$  is zero matrix;

E is unit matrix.

Then,based on least square,we can obtain estimated value of unknown parameters as Eq.4 [1, 2],

$$\hat{x}' = \left(B'^{T} P B'\right)^{-1} B'^{T} P l \tag{4}$$

estimated value of variances of unit weight as Eq.5 [1, 2],

$$\hat{\sigma}_0^2 = \frac{V'^{T}PV'}{n-t} \tag{5}$$

and estimated value of cofactor matrix of unknown parameters as E.q.6[1-2]

$$Q_{\text{pp}} = \left(B^{\prime}^{T} P B^{\prime}\right)^{-1} \tag{6}$$

where

$$P = \left[ \begin{array}{cc} P_1 & 0 \\ 0 & P_2 \end{array} \right];$$

0 of 
$$\begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix}$$
 is zero matrix;

 $\hat{\sigma}_0^2$  --- estimated value of variances of unit weight;

 $Q_{\it XX}$  --- estimated value of cofactor matrix of unknown parameters.

*n*---quantity of actual observation values;

t---necessary quantity of actual observation values.

(Note:Literatures [1, 2] neglected many sign ' $\Lambda$ ' which means estimated value.)

# 3. THE PRINCIPLE OF IMPROVED ALGORITHM OF TOTAL LEAST SQUARE

We suppose there exists condition equation Eq.7 [3].

$$C\hat{\chi}_B - W_{\chi_B} = 0 \tag{7}$$

where:0 is matrix.

Eq.8 can be derived from Eq.7.

$$C'\hat{x}' - W_{x'} = 0 \tag{8}$$

where:

$$C' = (0 \quad C);$$

$$W_{x'} = \left[ \begin{array}{c} 0 \\ W_{X_B} \end{array} \right];$$

$$0 \text{ of } C' = \begin{pmatrix} 0 & C \end{pmatrix} \text{ is matrix;}$$

0 of 
$$W_{x'} = \begin{bmatrix} 0 \\ W_{X_B} \end{bmatrix}$$
 is matrix;

0 of 
$$C'\hat{x}' - W_{x'} = 0$$
 is matrix.

If we unite Eq.8and Eq.3,and compute them based on adjustment of indirect observations with conditions,we can obtain estimated value of unknown parameters as Eq.9[3],

$$x' = \left(N_{bb}^{-1} - N_{bb}^{-1}C'^{T}N_{cc}^{-1}C'N_{bb}^{-1}\right)W + N_{bb}^{-1}C'^{T}N_{cc}^{-1}W_{x'}$$
(9)

estimated value of variances of unit weight as Eq.10[3],

$$\hat{\sigma}_0^2 = \frac{{V'}^T P V'}{n - u + s} \tag{10}$$

and estimated value of cofactor matrix of unknown parameters as Eq.11[3].

$$Q_{\vec{x}\vec{x}} = N_{bb}^{-1} - N_{bb}^{-1} C^{\prime T} N_{cc}^{-1} C^{\prime} N_{bb}^{-1}$$
(11)

where.

$$N_{bb} = B'^T P B'$$
;

$$W = B^{\prime T} P l$$
;

$$N_{cc} = C' N_{bb}^{-1} C'^T$$
;

u---quantity of unknown parameters;

s--- quantity of condition equations of unknown parameters.

#### 4. EXAMPLE

This paper adopted the improved algorithm of total least square to data of Table 1 which came from Table 1 of literature [4].

Table 1. Measured Data of Spatial Straight Line

Assumed Point Name by me	x(m)	y(m)	z(m)
1	3.0036	2.9960	3.0041
2	4.0034	4.9980	6.0033
3	5.0050	6.9992	9.0050
4	5.9964	9.0036	11.9962
5	7.0032	10.9968	15.0024
6	8.0038	13.0001	18.0045
7	9.0010	15.0040	20.9996
8	9.9972	16.9993	23.9982
9	11.0011	18.9962	27.0001
10	11.9997	20.9980	29.9978

We still suppose equation of straight line in three-dimensional space such as Eq.12[4].

$$\frac{x - x_0}{A} = \frac{y - y_0}{B} = \frac{z - z_0}{C} \tag{12}$$

Then, Eq. 13 can come from Eq. 12[4].

$$\begin{cases} x = az + b \\ y = cz + d \end{cases}$$
 (13)

where,

$$a = \frac{A}{C}$$
;

$$b = \chi_0 - \frac{A}{C} Z_0;$$

$$c = \frac{B}{C}$$
;

$$d = y_0 - \frac{B}{C} z_0.$$

So, we can obtain error equation Eq.14 based on Tab.1.(Note:There is some neglect in Eq.4 and Eq.5 of literature[4].)

$$V = \begin{bmatrix} B_{11}^{0} + \hat{\chi}_{B11} & 1 & 0 & 0 \\ 0 & 0 & B_{23}^{0} + \hat{\chi}_{B23} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ B_{191}^{0} + \hat{\chi}_{B191} & 1 & 0 & 0 \\ 0 & 0 & B_{203}^{0} + \hat{\chi}_{B203} & 1 \end{bmatrix} \begin{bmatrix} a^{0} + \delta \hat{a} \\ b^{0} + \delta \hat{b} \\ c^{0} + \delta \hat{c} \\ d^{0} + \delta \hat{d} \end{bmatrix} - \begin{bmatrix} x_{1} \\ y_{1} \\ \vdots \\ x_{10} \\ y_{10} \end{bmatrix} (14)$$

where,

 $B_{ij}^{\ 0}$  ---approximation of  $B_{ij}$  in designing matrix;

$$\hat{\chi}_{B_{ii}}$$
 --- residual of  $B_{ij}^{\ 0}$ ;

$$\hat{B}_{ij} = \hat{\chi}_{R_{ii}} + B_{ij}^0;$$

 $a^0$ --- approximation of a;

 $b^0$ --- approximation of b;

 $c^0$ --- approximation of c;

 $d^0$ --- approximation of d;

 $\delta a$  --- residual of  $a^0$ ;

 $\delta b$  --- residual of  $b^0$ ;

 $\delta c$  --- residual of  $c^0$ ;

 $\delta d$  --- residual of  $d^0$ ;

$$\hat{a} = \delta a + a^0$$
;

$$\hat{b} = \delta b + b^0;$$

$$\hat{c} = \delta c + c^0$$
;

$$\hat{d} = \delta d + d^0$$

Obviously, Eq. 15 can come from Eq. 14.

$$V = \begin{bmatrix} B_{11}^{0} & 1 & 0 & 0 \\ 0 & 0 & B_{21}^{0} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ B_{191}^{0} & 1 & 0 & 0 \\ 0 & 0 & B_{201}^{0} & 1 \end{bmatrix} \begin{bmatrix} \delta \hat{a} \\ \delta \hat{b} \\ \delta \hat{c} \\ \delta \hat{d} \end{bmatrix} + \begin{bmatrix} a^{\circ} & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & c^{\circ} & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & a^{\circ} & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & c^{\circ} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & a^{\circ} & 0 \\ 0 & 0 & 0 & 0 & \cdots & a^{\circ} & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_{B1} \\ \hat{x}_{B23} \\ \hat{x}_{B31} \\ \vdots \\ \hat{x}_{B991} \\ \hat{x}_{B203} \end{bmatrix} = \begin{bmatrix} x_{1} B_{10}^{0} a^{0} b^{0} \\ y_{1} B_{20}^{0} c^{0} d^{0} \\ x_{2} B_{30}^{0} a^{0} b^{0} \\ y_{2} B_{40}^{0} c^{0} d^{0} \\ \vdots \\ x_{10} B_{199}^{0} a^{0} b^{0} \end{bmatrix}$$

then,we can obtain

When we consider residual of virtual observations we can obtain Eq.17.

$$V_{B} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda_{B11} - B_{11}^{0} \\ \lambda_{B23} - B_{23}^{0} \\ \lambda_{B_{31}} \\ \lambda_{B_{31}} \\ \lambda_{B_{31}} \\ \lambda_{B_{33}} \\ \vdots \\ \lambda_{B_{191}} - B_{191}^{0} \\ \lambda_{B_{303}} \end{bmatrix}$$
(17)

where,  $L_{B_{ii}}$  ---virtual observations of  $B_{ij}$ .

Finally, let

$$V' = \begin{bmatrix} V \\ VB \end{bmatrix} \tag{18}$$

Obviously, there exist condition equations among virtual observations.

$$\hat{x}_{B_{i1}} = \hat{x}_{B_{(i+1)3}}, \ i = 1, 3, 5, L, 19$$
(19)

Then, based on Eq.19, we can obtain Eq.20.

If we select

$$\begin{bmatrix} a^{0} & b^{0} & c^{0} & d^{0} & B_{11}^{0} & B_{23}^{0} & \cdots & B_{191}^{0} & B_{203}^{0} \end{bmatrix}^{T}$$

$$= \begin{bmatrix} \frac{x_{10} - x_{1}}{z_{10} - z_{1}} & x_{1} - \frac{x_{10} - x_{1}}{z_{10} - z_{1}} z_{1} & \frac{y_{10} - y_{1}}{z_{10} - z_{1}} & y_{1} - \frac{y_{10} - y_{1}}{z_{10} - z_{1}} z_{1} & z_{1} & z_{1} & z_{10} & z_{10} \end{bmatrix}^{T}$$

$$= \begin{bmatrix} 0.333267 & 2.002434 & 0.666896 & 0.992577 & 3.0041 & 3.0041 & \cdots & 29.9978 & 29.9978 \end{bmatrix}^{T}$$

and P = E (There, we suppose actual observation and virtual observation are all equal precision independent surveying values, So, we can suppose P = E.), and alternately compute for two times to obtain parameter estimation values such as Eq.22 and Eq.23,

$$\begin{bmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \\ \hat{d} \\ \hat{d} \\ \hat{b} \\ \hat{c} \\ \hat{d} \\ \hat{d} \\ \hat{B}11 \\ \hat{B}23 \\ \hat{B}31 \\ \hat{B}43 \\ \hat{B}51 \\ \hat{B}63 \\ \hat{B}71 \\ \hat{B}83 \\ \hat{B}91 \\ \hat{B}111 \\ \hat{B}123 \\ \hat{B}131 \\ \hat{B}143 \\ \hat{B}151 \\ \hat{B}163 \\ \hat{B}171 \\ \hat{B}183 \\ \hat{B}191 \\ \hat{B}203 \end{bmatrix}$$

$$(23) \\ (23) \\$$

and mean square errors of unit weight of two times are such as Eq.24 and Eq.25.

$$\hat{\sigma}_{0}(\text{the first time}) = \pm \sqrt{\frac{V'^{T}PV'}{n-u+s}} = \pm \sqrt{\frac{0.000149131}{(20+20)-(4+20)+10}} = \pm 0.00239495 \text{(m)}$$

$$\hat{\sigma}_{0}(\text{the sencond time}) = \pm \sqrt{\frac{V'^{T}PV'}{n-u+s}} = \\ \pm \sqrt{\frac{0.000149134}{(20+20)\cdot(4+20)+10}} = \\ \pm 0.00239498(\text{m})$$
 (25)

When time of alternate computing is two, we can obtain satisfying result, then, stop computing and select the first result as final result. So, we can obtain equation of straight line in three-dimensional space based on literature[4].

$$\begin{cases} x = 0.333248z + 2.002473 \\ y = 0.666823z + 0.995788 \end{cases}$$
 (26)

then, the standard equation of the straight line is

$$\frac{x - 2.002473}{0.333248} = \frac{y - 0.995788}{0.666823} = \frac{z - 0}{1}$$
 (27)

If we don't consider Eq.19 or Eq.20 like literatures[1-2,4], we still alternately compute for two times to obtain parameter estimation values such as Eq.28 and Eq.29 based on Eq.21 and P=E.

$$\begin{bmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \\ \hat{d} \\ \hat{b} \\ \hat{c} \\ \hat{d} \\ \hat{B}_{11} \\ \hat{B}_{23} \\ \hat{B}_{31} \\ \hat{B}_{43} \\ \hat{B}_{51} \\ \hat{B}_{63} \\ \hat{B}_{71} \\ \hat{B}_{103} \\ \hat{B}_{111} \\ \hat{B}_{123} \\ \hat{B}_{131} \\ \hat{B}_{133} \\ \hat{B}_{131} \\ \hat{B}_{143} \\ \hat{B}_{151} \\ \hat{B}_{163} \\ \hat{B}_{151} \\ \hat{B}_{163} \\ \hat{B}_{151} \\ \hat{B}_{163} \\ \hat{B}_{171} \\ \hat{B}_{183} \\ \hat{B}_{171} \\ \hat{B}_{183} \\ \hat{B}_{171} \\ \hat{B}_{163} \\ \hat{B}_{171} \\ \hat{B}_{183} \\ \hat{B}_{171} \\ \hat{B}_{183} \\ \hat{B}_{171} \\ \hat{B}_{183} \\ \end{pmatrix} \begin{bmatrix} 0.333248 \\ 2.002473(m) \\ 0.666823 \\ 0.995786(m) \\ 6.003402(m) \\ 6.002872(m) \\ 9.005488(m) \\ 9.004385(m) \\ 11.99507(m) \\ 12.00011(m) \\ 15.00276(m) \\ 15.00104(m) \\ 18.0038(m) \\ 20.99973(m) \\ 21.00199(m) \\ 23.99741(m) \\ 23.99864(m) \\ 27.00037(m) \\ 26.99831(m) \\ \end{bmatrix}$$

$$(29)$$

29.99795(m)

29.99733(m)

 $\hat{B}_{191}$ 

 $\hat{B}_{203}$ 

and mean square errors of unit weight of two times are such as Eq.30 and Eq.31.

$$\hat{\sigma}_{0}(\text{the first time})' = \pm \sqrt{\frac{V'^{T}PV'}{n-t}} = \pm \sqrt{\frac{0.000121321}{(20+20)\cdot(20+4)}} = \pm 0.00275364(m)$$
(30)

 $\hat{\sigma}_0$  (the sec ond time)' =

$$\pm\sqrt{\frac{V'^T P V'}{n-t}} = \pm\sqrt{\frac{0.000121327}{(20+20)-(20+4)}} = \pm0.00275372(m)$$
(31)

When time of alternate computing is two,we can still obtain satisfying result, then, stop computing and select the first result as final result. So, we can still obtain equation of straight line in three-dimensional space based on literature [4].

The difference of Eq.6 and Eq.11 is

 $N_{bb}^{-1}C^TN_{cc}^{-1}C'N_{bb}^{-1}$ , so, the difference of estimated value of cofactor matrix of unknown parameters is very obvious.

Though the equation of straight line based on method of literature [1, 2] and based on new method of this paper are almost same in certain accuracy, but there still exists some difference, and at the same time, estimated value of cofactor matrix of unknown parameters is different, mean square error of unit weight of the improved method is better, and theory of the improved method is more strict, so, the improved method of this paper is better.

### CONCLUSION

This paper put forward a sort of new algorithm of total least square and deduced homologous formula. The new algorithm can improve algorithm and mean square error of unit weight, because it considered condition equations among virtual observations in designing matrix, but the other literatures didn't.

Example of fitting a straight line in three-dimensional space showed that the improved algorithm of this paper is viable and valid.

### CONFLICT OF INTEREST

The author confirms that this article content has no conflict of interest.

#### **ACKNOWLEDGEMENTS**

Declared none.

#### REFERENCES

[1] E. H. Wei, Z. X, Yin, G. W. Li, and Z. Q. Li, "On 3D coordinate transformations with virtual observation method," *Geomatics and Information Science of Wuhan University*, vol. 39, no. 2, pp. 152-156, 2014.

- Y. B. Yao, and J. Kong, "A new combined LS method considering [2] random errors of design matrix," Geomatics and Information Science of Wuhan University, vol. 39, no. 9, pp. 1028-1032, 2014.
- Teaching and Research Group of Surveying Adjustment of Wuhan [3] Technical University of Surveying and Mapping, Basis of
- Surveying Adjustment 3rd ed. Wuhan Technical University, Beijing,
- pp.123-126, 1996. Y. B. Yao, S. H. Huang, J. Kong, and J. Q. He, "Total least squares [4] algorithm for fitting spatial straight lines," Geomatics and Information Science of Wuhan University, vol.39, no.5, pp.571-574, 2014.

Received: September 17, 2014 Revised: December 17, 2014 Accepted: December 23, 2014

© Deng Yonghe; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/ by-nc/3.0/) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.