Geodesy and Geodynamics 2012,3(4):39 – 46

http://www.jgg09.com
Doi:10.3724/SP.J.1246.2012.00039

Properties of the total least squares estimation

Wang Leyang[1,2]

[1] *Faculty of Geomatics, East China Institute of Technology, Nanchang* 330013, *China*

[2] *Jiangxi Province Key Lab for Digital Land, Fuzhou* 344000, *China*

Abstract: Through theoretical derivation, some properties of the total least squares estimation are found. The total least squares estimation is the linear transformation of the least squares estimation, and the total least squares estimation is unbiased. The condition number of the total least squares estimation is greater than the least squares estimation, so the total least squares estimation is easier to be affected by the data error than the least squares estimation. Then through the further derivation, the relationships of solutions, residuals and unit weight variance estimations between the total least squares and the least squares are given.

Key words: total least squares (TLS); least squares (LS); singular value decomposition (SVD); residuals; unit weight variance

1 Introduction

In the classical least squares (LS) approach the coefficient matrix is assumed to be free from error, and all errors are confined to be the observation vector. However, in surveying engineering application, this assumption is often unrealistic. The coefficient matrix is not a constant matrix and is not composed of constants. For example, the coefficient matrix is available by measurements or it is an idealized approximation of the true operator, then both the matrix and the observation vector are contaminated by some noise. So, in this condition, using the classical least squares approach to deal with the problem that both the matrix and the observation vector are contaminated by noise may be not reasonable. An appropriate approach to this problem is the total least squares (TLS) method. The name total least squares was given by Golub and Van Loan[1], which was developed in the field of numerical analysis; but under the names orthogonal regression or errors-in-variables (EIV), this fitting method has a long history in the statistical literature. Over the years, it had been rediscovered many times, often independently, but only in the last two decades, it started to be used in practical applications[2].

Now, there are many researches about the total least squares in algorithms and applications in the surveying engineering: the algorithms, such as the singular value decomposition (SVD) algorithm[1-3] and the algorithm based on the Lagrange approach[4-9], and so on[10,11]; the applications, such as space resection[12], spatial pattern analysis and fitting[3,9], coordinate transformations and datum conversion[6,8], geodetic inversion[13,14], and so on[10,11]. There are also many researches on the properties of the total least squares and the differences between the total least squares and the classical least squares. The connections between the TLS and the classical LS were discussed and the sensitivity of both solutions and the SVD was compared mainly from the viewpoint of a numerical analyst[2]. Then the statistical properties of the TLS were studied[2]. As soon as some prior infor-

Received:2012-07-25; Accepted:2012-09-27

Corresponding author: Wang Leyang, E-mail: wleyang@163.com

The research was supported by the National Natural Science Foundation of China (41204003), Scientific Research Foundation of ECIT (DHBK201113) and Scientific Research Foundation of Jiangxi Province Key Laboratory for Digital Land (DLLJ201207).

mation about the distribution of the errors in the data is available, statistical properties of the TLS can be derived, and they prove the expected quality of the TLS model estimate parameters. Given the frequency of practical situations in which the "independent" variables are recorded with error, the TLS should be proved to be a very useful tool to data analysts. The TLS gives better estimates of the parameters of the EIV model than the LS does. The relations between the weighted TLS and the weighted LS solutions are obtained[15], and the analysis is useful especially for rank deficient problems and generalizes the results of Golub and Van Loan, Van Huffel and Vandewalle. The existing bounds of differences between the LS residuals, the weighted squares residuals and the minimum norm correction matrices of the TLS and LS problems were improved[16]. Different from the previous studies of the properties of the TLS problem from the viewpoint of pure mathematics, the properties of the TLS solution are given and proved in our paper from the viewpoint of surveying adjustment.

2 The TLS solution of the adjustment problem

For a linear estimation problem, the function model is

$$AX \approx b \tag{1}$$

where $A \in R^{m \times n} (m > n)$ is a full column rank coefficient matrix; $X \in R^{n \times 1}$ is the parameter to be estimated; $b \in R^{m \times 1}$ is the observation vector.

The basic thought of TLS is consideration of the error E_A of coefficient matrix A and the observation error e of the observation vector at the same time. So the error equation is

$$[A + E_A \quad b + e] \begin{bmatrix} X \\ -1 \end{bmatrix} = 0$$

That is

$$\bar{v} = [I_m \ -(X^T \otimes I_m)] \begin{bmatrix} e \\ e_A \end{bmatrix} \tag{2}$$

where, $\bar{v} = AX - b$.

If the stochastic model of adjustment problem is

$$\begin{bmatrix} e \\ vec(E_A) \end{bmatrix} \sim \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma_0^2 \begin{bmatrix} I_m & 0 \\ 0 & I_n \otimes I_m \end{bmatrix} \tag{3}$$

where $vec(\cdot)$ denotes the operator that stacks one row of a matrix rearwards of the previous one, and then transposes it to obtain a column vector; $\otimes$ denotes "Kronecker-Zehfuss product".

Then the TLS adjustment criterion is

$$\Phi = vec(E_A)^T vec(E_A) + e^T e = \min \tag{4}$$

The SVD of $[A \quad b]$ *is*

$$[A \quad b] = U \Sigma V^T \tag{5}$$

where, $U = [U_1 \quad U_2]$, $U_1 = [u_1 \ \cdots \ u_n]$, $U_2 = [u_{n+1} \ \cdots \ u_m]$, $u_i \in R^{m \times 1}$, $U^T U = I_m$; $V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}\begin{matrix} n \\ 1 \end{matrix} = [v_1 \ \cdots \ v_{n+1}]$, $v_i \in R^{(n+1) \times 1}$, $V^T V = I_{n+1}$; $\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} = diag(\sigma_1, \cdots, \sigma_{n+1}) \in R^{m \times (n+1)}$, $\Sigma_1 = diag(\sigma_1, \cdots, \sigma_n) \in R^{n \times n}$, $\Sigma_2 = [\sigma_{n+1} \ 0 \ \cdots \ 0]^T \in R^{(m-n) \times 1}$; $\sigma_1 \geqslant \cdots \sigma \geqslant_{n+1} \geqslant 0$ are the singular values.

When $\sigma_n > \sigma_{n+1}$ and $v_{n+1,n+1} \neq 0$, the TLS solution is[2]

$$\hat{X}_{TLS} = (N - \sigma_{n+1}^2 I_n)^{-1} A^T b \tag{6}$$

where $N = A^T A$; I is a identity matrix.

If the condition $\sigma_n > \sigma_{n+1}$ and $v_{n+1,n+1} \neq 0$ is not satisfied, then the minimum norm TLS solution is computed[2].

The precision evaluation of the TLS is[4]

$$\hat{\sigma}_0^2(TLS) = \frac{\sigma_{n+1}^2}{m - n} \tag{7}$$

where $\hat{\sigma}_0^2(TLS)$ is the estimation of mean square error of unit weight.

$$D(\hat{X}) \approx \hat{\sigma}_0^2(TLS)(N - \sigma_{n+1}^2 I_n)^{-1} N (N - \sigma_{n+1}^2 I_n)^{-1} \tag{8}$$

where $D(\hat{X}$ is the mean square error of the TLS solution.

3 The properties of the TLS solution

3.1 The TLS solution is the linear transformation of the LS solution

When the error of the coefficient matrix is not considered, the LS solution is

$$\hat{X}_{LS} = N^{-1} A^{\mathrm{T}} b \tag{9}$$

where $N = A^{\mathrm{T}} A$.

From the equations (6) and (9), we can get

$$\hat{X}_{TLS} = (N - \sigma_{n+1}^2 I)^{-1} N N^{-1} A^{\mathrm{T}} b = Z \hat{X}_{LS} \tag{10}$$

where

$$Z = (N - \sigma_{n+1}^2 I)^{-1} N = (I - \sigma_{n+1}^2 N^{-1})^{-1} \tag{11}$$

So, from the equation (10), we can find that: the TLS solution is the linear transformation of the LS solution. The equation (10) is consistent with Van Huffel and Vandewalle[2].

3.2 The expectation of the TLS solution

Using matrix inversion formula[17]

$$(D + ACB)^{-1} = D^{-1} - D^{-1} A (C^{-1} + BD^{-1} A) BD^{-1} \tag{12}$$

From the equation (11), Z can be rewritten as

$$Z = (I - \sigma_{n+1}^2 I N^{-1} I)^{-1} = I + \sigma_{n+1}^2 (N - \sigma_{n+1}^2 I)^{-1} \tag{13}$$

From the equations (10) and (13)

$$\hat{X}_{TLS} = (I + \sigma_{n+1}^2 (N - \sigma_{n+1}^2 I)^{-1}) \hat{X}_{LS} \tag{14}$$

Expecting both sides of the formula (14), we obtain

$$E(\hat{X}_{TLS}) = (I + \sigma_{n+1}^2 (N - \sigma_{n+1}^2 I)^{-1}) E(\hat{X}_{LS}) \tag{15}$$

So

$$E(\hat{X}_{TLS}) - E(\hat{X}_{LS}) = \sigma_{n+1}^2 (N - \sigma_{n+1}^2 I)^{-1} E(\hat{X}_{LS}) \tag{16}$$

When the errors are confined to the observation vector and the coefficient matrix, the TLS solution is unbiased[2].

$$E(\hat{X}_{TLS}) = X \tag{17}$$

where $E(\cdot)$ denotes the expection of $\hat{X}_{TLS}$.

So

$$(I + \sigma_{n+1}^2 (N - \sigma_{n+1}^2 I)^{-1}) E(\hat{X}_{LS}) = X \tag{18}$$

From the equation (15), we can see that: the expection of the TLS solution is also the linear transformation of the expection of the LS solution. From the equation (16), we can know that: when σ_{n+1} becomes smaller gradually, the difference of the expection of the TLS solution and the expection of the LS solution is also becoming smaller gradually. From the equation (18), we can get that: only when $\sigma_{n+1}^2 = 0$ (that is $\sigma_{n+1} = 0$), the coefficient matrix is free from error, the LS solution is unbiased; and the TLS solution degenerates to the LS solution.

3.3 The eigenvectors of Z are the same to the eigenvectors of N, and they are unrelated with σ_{n+1}^2

If the characteristic roots of N are $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n > 0$, and the corresponding eigenvectors are A_1, A_2, $\cdots$, A_n, then

$$N A_i = \lambda_i A_i (i = 1, 2, \cdots, n) \tag{19}$$

Both sides of the equation (19) are divided by $\frac{1}{\lambda} N^{-1}$, then

$$N^{-1}A_i = \frac{1}{\lambda_i}A_i \tag{20}$$

From the equation (19)

$$(N - \sigma_{n+1}^2 I)A_i = (\lambda_i - \sigma_{n+1}^2)A_i \tag{21}$$

So

$$(N - \sigma_{n+1}^2 I)^{-1}A_i = \frac{1}{\lambda_i - \sigma_{n+1}^2}A_i \tag{22}$$

The SVD of the coefficient matrix A is

$$A = U'\Sigma' V'^{\mathrm{T}} \tag{23}$$

where $U'[U'_1 \quad U'_2]$, $U_1 = [u'_1 \quad \cdots \quad u'_n]$, $U_2 = [u'_{n+1} \quad \cdots \quad u'_m]$, $u' \in R^{m\times 1}$, $U'^{\mathrm{T}}U' = I_m$; $V = [v'_1 \quad \cdots \quad v'_n]$, $v'_i \in R^{n\times 1}$, $V'^{\mathrm{T}}V' = I_n$; $\Sigma' = diag(\sigma'_1, \cdots, \sigma'_n) \in R^{m\times n}$; $\sigma'_1 \geqslant \cdots \geqslant \sigma'_n > 0$.

From the interlacing theorem for singular values[18], we get[2]

$$\sigma_1 \geqslant \sigma'_1 \geqslant \cdots \geqslant \sigma_n \geqslant \sigma'_n \geqslant \sigma_{n+1} \tag{24}$$

and[2]

$$\sigma_n \geqslant \sigma_{n+1} \Leftrightarrow \sigma'_n \geqslant \sigma_{n+1} \text{ and } v_{n+1,n+1} \neq 0 \tag{25}$$

From the equations (24) and (25)

$$\lambda_i - \sigma_{n+1}^2 > 0 \tag{26}$$

From the equation (20)

$$(I - \sigma_{n+1}^2 N^{-1})A_i = (1 - \frac{\sigma_{n+1}^2}{\lambda_i})A_i \tag{27}$$

From the equations (11) and (27)

$$ZA_i = \frac{\lambda_i}{\lambda_i - \sigma_{n+1}^2}A_i \tag{28}$$

From the above derivation, we can see that: the characteristic roots of N are $\lambda_i (i = 1, 2, \cdots, n)$, the characteristic roots of $(N - \sigma_{n+1}^2 I)$ are $(\lambda_i - \sigma_{n+1}^2)$ $(i = 1, 2, \cdots, n)$, the characteristic roots of $(N - \sigma_{n+1}^2 I)^{-1}$ are $\frac{1}{\lambda_i - \sigma_{n+1}^2} (i = 1, 2, \cdots, n)$, and the characteristic roots of Z are $\frac{\lambda_i}{\lambda_i - \sigma_{n+1}^2} (i = 1, 2, \cdots, n)$; the eigenvectors of them are $A_i (i = 1, 2, \cdots, n)$, and are the same to the eigenvectors of N, but unrelated with σ_{n+1}^2.

3.4 The relations between the TLS solution $\hat{X}_{\mathrm{TLS}}$ and the LS solution $\hat{X}_{\mathrm{LS}}$

From the equations (9) and (14)

$$\hat{X}_{\mathrm{TLS}} - \hat{X}_{\mathrm{LS}} = \sigma_{n+1}^2 (N - \sigma_{n+1}^2 I)^{-1}\hat{X}_{\mathrm{LS}} \tag{29}$$

Norming both sides of the formula (29), we get

$$\| \hat{X}_{\mathrm{TLS}} - \hat{X}_{\mathrm{LS}} \| = \| \sigma_{n+1}^2 (N - \sigma_{n+1}^2 I)^{-1}\hat{X}_{\mathrm{LS}} \| \tag{30}$$

So

$$\| \hat{X}_{\mathrm{TLS}} - \hat{X}_{\mathrm{LS}} \| \leqslant \| \sigma_{n+1}^2 (N - \sigma_{n+1}^2 I)^{-1} \| \cdot \| \hat{X}_{\mathrm{LS}} \| \tag{31}$$

That is

$$\frac{\| \hat{X}_{\mathrm{TLS}} - \hat{X}_{\mathrm{LS}} \|}{\| \hat{X}_{\mathrm{LS}} \|} \leqslant \| \sigma_{n+1}^2 (N - \sigma_{n+1}^2 I)^{-1} \| \tag{32}$$

From the above derivation, we know that: for an estimation problem, the coefficient matrix A and observation b are fixed, so the maximum ration of the difference norm $\| \hat{X}_{\mathrm{TLS}} - \hat{X}_{\mathrm{LS}} \|$ between the TLS solution $\hat{X}_{\mathrm{TLS}}$ and the LS solution $\hat{X}_{\mathrm{LS}}$ and the norm $\| \hat{X}_{\mathrm{LX}} \|$ of the LS solution $\hat{X}_{\mathrm{LS}}$ is a fixed value $\| \sigma_{n+1}^2 (N - \sigma_{n+1}^2 I)^{-1} \|$.

3.5 The relations between the TLS residuals $\hat{V}_{\mathrm{TLS}}$ and the LS residuals $\hat{V}_{\mathrm{LS}}$

The LS criterion is

$$\begin{aligned}\varphi_{\mathrm{LS}} &= V_{\mathrm{LS}}^{\mathrm{T}} V_{\mathrm{LS}} = (AX - b)^{\mathrm{T}}(AX - b) \\ &= \| AX = b \|^2 = \min\end{aligned} \tag{33}$$

So, the LS residuals are

$$\hat{V}_{LS} = A\hat{X}_{LS} - b \tag{34}$$

From the equations (9) and (34)

$$V_{LS} = AN^{-1}A^{T}b - b = (AN^{-1}A^{T} - I)b \tag{35}$$

The TLS criterion (4) is also equal to

$$\varphi_{TLS} = V_{TLS}^{T}V_{TLS} = \frac{(AX-b)^{T}(AX-b)}{1+X^{T}X} = \frac{\| AX-b \|^{2}}{1+\| X \|^{2}} = \min \tag{36}$$

So, the TLS residuals are

$$\hat{V}_{TLS} = \frac{A\hat{X}_{TLS} - b}{\sqrt{1+\| \hat{X}_{TLS} \|^{2}}} \tag{37}$$

From the equations (14) and (37)

$$\hat{V} = \frac{A\hat{V}_{LS} - b + \sigma_{n+1}^{2}A(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS}}{\sqrt{1+\| \hat{X}_{TLS} \|^{2}}} \tag{38}$$

From the equations (37) and (38)

$$\hat{V}_{TLS} = \frac{\hat{V}_{LS} + \sigma_{n+1}^{2}A(N-\sigma^{2''}n+1I)^{-1}\hat{X}_{LS}}{\sqrt{1+\| \hat{X}_{TLS} \|}} \tag{39}$$

Because

$$\sqrt{1+\| \hat{X}_{TLS} \|^{2}} \geqslant 1 \tag{40}$$

Then

$$\frac{\hat{V}_{LS} + \sigma_{n+1}^{2}A(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS}}{\sqrt{1+\| \hat{X}_{TLS} \|^{2}}} \leqslant \hat{V}_{LS} + \sigma_{n+1}^{2}A(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS} \tag{41}$$

From the equations (39) and (41)

$$\hat{V}_{TLS} \leqslant \hat{V}_{LS} + \sigma_{n+1}^{2}A(N-\sigma_{n+1}^{2}I)\hat{X}_{LS} \tag{42}$$

From the equations (9) and (42)

$$\hat{V}_{TLS} - \hat{V}_{LS} \leqslant \sigma_{n+1}^{2}A(N-\sigma_{n+1}^{2}I)^{-1}N^{-1}A^{T}b \tag{43}$$

Norming both sides of the formula (39), we get

$$\| \hat{V}_{TLS} \| \leqslant \frac{1}{\left\| \sqrt{1+\| \hat{X}_{TLS} \|^{2}} \right\|}\left(\| \hat{V}_{LS} \| + \| \sigma_{n+1}^{2}A(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS} \| \right) \tag{44}$$

From the equations (40) and (44)

$$\| \hat{V}_{TLS} \| \leqslant \| \hat{V}_{LS} \| + \| \sigma_{n+1}^{2}A(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS} \| \tag{45}$$

That is

$$\| \hat{V}_{TLS} \| - \| \hat{V}_{LS} \| \leqslant \| \sigma_{n+1}^{2}A(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS} \| \tag{46}$$

Substituting equation (9) into equation (46),

$$\| \hat{V}_{TLS} \| - \| \hat{V}_{LS} \| \leqslant \| \sigma_{n+1}^{2}A(N-\sigma_{n+1}^{2}I)^{-1}N^{-1}A^{T}b \| \tag{47}$$

From the equations (43) and (47), we can see: the difference between the TLS residuals $\hat{V}_{TLS}$ and the LS residuals $\hat{V}_{LS}$ is smaller than a constant $\sigma_{n+1}^{2}A(N-\sigma_{n+1}^{2}I)^{-1}N^{-1}A^{T}b$; the difference between the TLS residuals norm $\| \hat{V}_{TLS} \|$ and the LS residuals norm $\| \hat{V}_{LS} \|$ is also smaller than a constant $\| \sigma_{n+1}^{2}A(N-\sigma_{n+1}^{2}I)^{-1}N^{-1}A^{T}b \|$.

3.6 The relations between the estimation of mean square error of unit weight of the TLS and the LS

The estimation of mean square error of unit weight of the TLS (7) is equal to

$$\hat{\sigma}_{0}^{2}(TLS) = \frac{V_{TLS}^{T}V_{TLS}}{m-n} = \frac{(A\hat{X}_{TLS}-b)^{T}(A\hat{X}_{TLS}-b)}{(m-n)(1+\hat{X}_{TLS}^{T}\hat{X}_{TLS})} \tag{48}$$

The estimation of mean square error of unit weight of the LS is

$$\hat{\sigma}(LS)=\frac{V_{LS}^{T}V_{LS}}{m-n}=\frac{(A\hat{X}_{LS}-b)^{T}(A\hat{X}_{LS}-b)}{m-n} \tag{49}$$

Because

$$1+\hat{X}_{TLS}^{T}\hat{X}_{TLS}\geqslant 1 \tag{50}$$

Then

$$\hat{\sigma}(TLS)\leqslant\frac{(A\hat{X}_{TLS}-b)^{T}(A\hat{X}_{TLS}-b)}{m-n} \tag{51}$$

From the equations (14) and (48)

$$(A\hat{X}_{TLS}-b)^{T}(A\hat{X}_{TLS}-b)=(A\hat{X}_{LS}-b)^{T}(A\hat{X}_{LS}-b)+(A\hat{X}_{LS}-b)^{T}(A\sigma_{n+1}^{2}(N-\sigma^{2}{}''n+1I)^{-1}\hat{X}_{LS})+(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS})^{T}\cdot(A\hat{X}_{LS}-\mathrm{b})+(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS})^{T}(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS} \tag{52}$$

From the equations (49), (51) and (52)

$$\hat{\sigma}_0^2(TLS)\leqslant\hat{\sigma}_0^2(LS)+\frac{1}{m-n}\{(A\hat{X}_{LS}-b(^{T}(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}I)^{-1}\hat{X})+(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS})^{T}(A\hat{X}_{LS}-b)+(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS})^{T}(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS})\} \tag{53}$$

$$(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS})^{T}(A\hat{X}_{LS}-b)=(A\hat{X}_{LS}-b)^{T}(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS}) \tag{54}$$

$$(A\hat{X}_{LS}-b)^{T}(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS})+(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS})^{T}(A\hat{X}_{LS}-b)+(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}\mathrm{I})^{-1}\hat{X}_{LS})^{T}(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS})=2(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS})^{T}(A\hat{X}_{LS}-b)+(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS})^{T}(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS}) \tag{55}$$

So

$$\hat{\sigma}_0^2(TLS)-\hat{\sigma}_0^2(LS)\leqslant\frac{1}{m-n}\{2(A\sigma_{n+1}^{2}(N-\hat{\sigma}_{n+1}^{2}I)^{-1}\hat{X}_{LS})^{T}(A\hat{X}_{LS}-b)+(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS})^{T}(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS})\} \tag{56}$$

For LS, there is

$$A^{T}(A\hat{X}_{LS}-b)=0 \tag{57}$$

So, the equation (56) can be written as

$$\hat{\sigma}_0^2(TLS)-\hat{\sigma}_0^2(LS)\leqslant\frac{1}{m-n}(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS})^{T}(A\sigma_{n+1}^{2}(N-\sigma_{n+1}^{2}I)^{-1}\hat{X}_{LS} \tag{58}$$

The equation (58) gives the upper limit of the difference between the estimation of mean square error of unit weight of the TLS and the LS. Equation (58) can also be obtained from the equations (47), (48) and (49).

The TLS can get better fitting data, and has a smaller fitting residual[2,19]

$$\| b-\hat{A}\hat{X}_{TLS}\|_2\leqslant\| b-A\hat{X}_{LS}\|_2 \tag{59}$$

Because

$$\hat{\sigma}_0^2(TLS)=\frac{\| b-\hat{A}\hat{X}_{TLS}\|_2^2}{m-n} \tag{60}$$

and

$$\hat{\sigma}_0^2(LS)=\frac{\| b-A\hat{X}_{LS}\|_2^2}{m-n} \tag{61}$$

Then

$$\hat{\sigma}_0^2(TLS)\leqslant\hat{\sigma}_0^2(LS) \tag{62}$$

From the equation (62), we can see: the estimation of mean square error of unit weight of the TLS $\hat{\sigma}_0^2(TLS)$ is smaller than the LS $\hat{\sigma}_0^2(LS)$, so the TLS can get a better fitting data.

3.7 The TLS solution is always more poorly conditioned than the LS solution

For the LS, from the equation (23), we can get

$$\lambda_{max}=(\sigma_1')^2,\ \lambda_{min}=(\sigma_n')^2 \tag{63}$$

where λ_{max} denotes the maximum eigenvalue; λ_{min} denotes the minimum eigenvalue.

From the equations (9) and (63), the condition number of the LS solution is

$$cond(\hat{X}_{LS}) = \frac{\lambda_{max}}{\lambda_{min}} \tag{64}$$

From the equations (6) and (63), the condition number of the TLS solution is

$$cond(\hat{X}_{TLS}) = \frac{\lambda_{max} - \lambda_{n+1}}{\lambda_{min} - \lambda_{n+1}} \tag{65}$$

where $\lambda_{n+1} = \sigma_{n+1}^2$.

From the equation (24)

$$\lambda_{max} - \lambda_{n+1} > 0,\ \lambda_{min} - \lambda_{n+1} \tag{66}$$

where $(N - \sigma_{n+1}^2 I_n)$ is not singular.

So, from the equations (64) and (65)

$$cond(\hat{X}_{TLS}) - cond(\hat{X}_{LS}) = \frac{\lambda_{n+1}(\lambda_{max} - \lambda_{n+1})}{\lambda_{min}(\lambda_{min} - \lambda_{n+1})} > 0 \tag{67}$$

From the equation (67), we can see: the condition number of the TLS solution $cond(\hat{X}_{TLS})$ is greater than the condition number of the LS solution $cond(\hat{X}_{LS})$; the TLS solution is always more poorly conditioned than the LS solution. The LS ridge estimation improves the ill-posed condition by adding a positive constant to the diagonal elements of N; while, the TLS approach subtracts a positive constant to the diagonal elements of N, so it is a irregular process. Compared to the LS, the TLS solution is easier to be affected by the data error. Branham[20] gives a review of the condition number of the TLS solution, here we give a detail derivation.

4 Conclusions

In the surveying engineering, the coefficient matrix may be affected by the sampling or modeling or measurement errors. So both the coefficient matrix and the observation vector are contaminated by some noise. The TLS method is particularly useful in modeling situations in which all the given variables in the system include errors and should be treated symmetrically. In these situations, the TLS approach yields more accurate estimations than the LS approach. Through theoretical derivation, many properties of the total least squares estimation are obtained. The total least squares estimation is the linear transformation of the least squares estimation, and the expection of the TLS solution is also the linear transformation of the expection of the LS solution. When the coefficient matrix is contaminated by some noise, the LS solution is biased, while the TLS solution is unbiased. The eigenvectors of Z are the same to the eigenvectors of N, and they are unrelated with σ_{n+1}^2. The relations between the TLS solution $\hat{X}_{TLS}$ and the LS solution $\hat{X}_{LS}$, and the relations between the TLS residuals $\hat{V}_{TLS}$ and the LS residuals $\hat{V}_{LS}$ are also researched in the paper. The TLS can get better fitting data, and has a smaller fitting residual, and the estimation of mean square error of unit weight of the TLS $\hat{\sigma}_0^2(TLS)$ is smaller than the LS $\hat{\sigma}_0^2(LS)$. The TLS solution is always more poorly conditioned than the LS solution. The LS ridge estimation improves the ill-posed condition by adding a positive constant to the diagonal elements of N; while, the TLS approach subtracts a positive constant to the diagonal elements of N, so it is a irregular process. Compared to the LS, the TLS solution is easier to be affected by the data error. This work will be useful to the researchers and practitioners in surveying engineering to understand and use the TLS approach, and to find efficient methods for implementing it.

References

[1] Golub G H and Van Loan C F. An analysis of the total least squares problem. SIAM J Numer Anal., 1980, 17(6): 883-893.

[2] Van Huffel S and Vandewalle J. The total least squares problem: Computational aspects and analysis. Society for Industrial and Applied Mathematics, Philadelphia. 1991.

[3] Felus Y A. Applications of total least-squares for spatial pattern analysis. J Surv Eng., 2004, 130(3): 126-133.

[4] Schaffrin B. A note on constrained total least-squares estimation. Linear Algebra Appl., 2006, 417(1): 245-258.

[5] Schaffrin B and Wieser A. On weighted total least-squares adjustment for linear regression. Journal of Geodesy, 2008, 82(7):

415 –421.

[6] Schaffrin B and Felus Y A. On the multivariate total least-squares approach to empirical coordinate transformations: Three algorithms. Journal of Geodesy, 2008, 82 (6): 373 –383.

[7] Schaffrin B and Felus Y A. An algorithmic approach to the total least-squares problem with linear and quadratic constraints. Studia Geophysica et Geodaetica, 2009, 53(1): 1 –16.

[8] Felus Y A and Burtch R C. On symmetrical three-dimensional datum conversion. GPS Solut. , 2009, 13: 65 –74.

[9] Schaffrin B and Snow K. Total Least-Squares regularization of Tykhonov type and an ancient racetrack in Corinth. Linear Algebra and its Applications, 2010, 432: 2061 –2076.

[10] Van Huffel S. Recent Advances in total least squares techniques and errors-in-variables modeling. Society for Industrial and Applied Mathematics, Philadelphia. 1997.

[11] Van Huffel S and Lemmerling P. Total least squares and errors-in-variables modeling: analysis, algorithms and applications. Kluwer Academic Publishers, Dordrecht. 2002.

[12] Chen Yi, Lu Jue and Zheng Bo. Application of total least squares to space resection. Geomatics and Information Science of Wuhan University, 2008, 33(12): 1271 –1274. (in Chinese)

[13] Wang Leyang, Xu Caijun and Lu Tieding. Inversion of strain parameter using distance changes based on total least squares. Geomatics and Information Science of Wuhan University, 2010, 35 (2): 181 –184. (in Chinese)

[14] Xu Caijun, Wang Leyang, Wen Yangmao and Wang Jianjun. Strain rates in the Sichuan-Yunnan region based upon the total least squares heterogeneous strain model from GPS data. Terr Atmos Ocean Sci. ,2011,22(2):133 –147.

[15] Wei Musheng and Chen Guoliang. Solution sets and property for weighted total least squares problem. Applied Mathematics-A Journal of Chinese Universities, 1994, 9(3): 304 –311. (in Chinese)

[16] Liu Yonghui and Wei Musheng. On the comparision of the total least squares and the least squares problems. Mathematica Numerica Sinica, 2003, 25(4): 479 –492. (in Chinese)

[17] Cui Xizhang, Yu Zongchou, Tao Benzao, et al. Generalized surveying adjustment (new edition). Wuhan Technical University of Surveying and Mapping, Wuhan. 2001. (in Chinese)

[18] Thompson R C. "Principal submatrices IX: interlacing inequalities for singular values of submatrices." Linear Algebra Appl. , 1972,5:1 –12.

[19] Zhang Hongyue, Huang Jindong and Fan Wenlei. Total least squares method and its application to parameter estimation. Acta Automatic Sinica, 1995, 21(1): 40 –48. (in Chinese)

[20] Branham R L J R. Multivariate orthogonal regression in astronomy. Celestial Mechanics & Dynamical Astronomy, 1995, 61: 239 –251.