CrossMark

# Adjustment of non-typical errors-in-variables models

**V. Mahboub**[1] · **A. A. Ardalan**[2] · **S. Ebrahimzadeh**[3]

**Abstract**   In this contribution, non-typical errors-in-variables (EIV) model is introduced as a more general form of the so-called EIV model which is highly relevant for geodetic data processing. Total least-squares (TLS) algorithms within a typical EIV model cannot deal with the non-typical EIV models since the first order moment/mean of the random design matrix in a non-typical EIV model is not linear. In this paper, we propose a new algorithm besides a standard solution to deal with this model. To achieve this goal, first we review a classic algorithm in order to solve this problem using traditional non-linear least-squares method within a non-linear mixed model. Then by comparison, weighted TLS algorithm within the typical EIV model can be replaced by the proposed approach after some slight modifications which results in an excellent approximate solution. This foundation is important because there is no need for linearization in the TLS algorithms. The proposed way is not sensitive to the approximate initial values of the unknown parameters and it is applicable to curve fitting, surface reconstruction or other non-typical EIV models. Here we employ it to the curve fitting. Also two examples convincingly demonstrate that the standard TLS solution of the non-typical EIV model is not admissible when it is incorrectly considered as a typical EIV model; i.e., the non-linear relationships of the elements of the random design matrix are neglected.

**Keywords**   Typical EIV model · Non-typical EIV model · Weighted total least-squares algorithm · Curve fitting

✉ V. Mahboub
   vahid_mahboobk@ut.ac.ir; vahid_mahboobk@alumni.ut.ac.ir

[1]   Young Researchers Club, Lahijan Branch, Islamic Azad University, Lahijan, Iran

[2]   Geodesy Division, Department of Surveying and Geomatics Engineering, Faculty of Engineering, Tehran University, Tehran, Iran

[3]   Remote Sensing Division, Department of Surveying and Geomatics Engineering, Faculty of Engineering, Tehran University, Tehran, Iran

## 1 Introduction

The concept of typical/classic errors-in-variables (EIV) models has been already introduced by several authors in the last years. A typical EIV model is similar to a Gauss–Markov (GM) model but all the variables are subject to random errors. For further reading, see e.g., Van Huffel and Vandewalle (1991), Schaffrin and Wieser (2008), Felus (2004), Schaffrin et al. (2012a, b), Mahboub et al. (2012), Fang (2013, 2014), Snow and Schaffrin (2012), Snow (2012) and Mahboub (2014), etc. Meanwhile, some other researchers investigated this problem traditionally; see e.g., Neitzel (2010) and Shen et al. (2011). The term "total least-squares (TLS)" was coined in the field of numerical analysis by Golub and Van Loan (1980) as one of the standard solutions of this model. Also in several contributions particularly in geodetic literature, its applications have been investigated. Linear regression (Schaffrin and Wieser 2008; Fang 2011), geodetic resection (Schaffrin and Felus 2008), transformation (Mahboub 2012) and rapid satellite positioning (Mahboub and Sharifi 2013) are some examples. Nevertheless, there are some other problems such as curve fitting which all the variables are subject to errors but the model is not similar to GM model. In other words, the design matrix in these kinds of models is a non-linear function of random variables. Some examples are as follows (both y and x coordinates are subject to random errors and underlining indicates random variables):

$$\underline{y} = a\sin\underline{t} + b\cos\underline{t} + c = \begin{bmatrix} \sin\underline{t} & \cos\underline{t} & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = A(\underline{t}) \begin{bmatrix} a \\ b \\ c \end{bmatrix},$$

$$\underline{y} = a_0 + a_1\underline{t} + a_2\underline{t}^2 + a_3\underline{t}^3 = \begin{bmatrix} 1 & \underline{t} & \underline{t}^2 & \underline{t}^3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = A(\underline{t}) \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}.$$

In fact, in the non-typical EIV model, the first order moment/mean of the non-linear random design matrix $A(\underline{t})$ is not known. We only know the mean of the random variables ($\underline{t}$). In other words, in the non-typical EIV model, the functional dependence of $A$ on $\underline{t}$ is considered through $\underline{A} = A(\underline{t})$. Mathematically, one encounters the following functional models for the typical and non-typical EIV models:

$$\underline{y} = A(\underline{t})\xi \quad \text{with} \quad \begin{cases} E(A(\underline{t})) = A(E(\underline{t})) & \text{for typical EIV model (I),} \\ E(A(\underline{t})) \neq A(E(\underline{t})) & \text{for non-typical EIV model (II),} \end{cases}$$

is the vector of unknown parameters, where E(·) denotes expectation operator. In geodesy, the non-typical EIV models appear in some problems. Clearly curve fitting is one of them due to the above formulas and it will be discussed in Sect. 4 (numerical results and discussions). The other example is surface reconstruction. "One well established technique to construct a surface that best fits to an observed scattered point cloud is based on the Kriging methodology that uses semi-variograms. As this semi-variogram regularly turns out to have a steep slope near the origin—where it matters most—, a better idea seems to be seeking a best fit on the basis of the Total Least-Squares (TLS) principle" (Schaffrin and Uzun 2008). Although they correctly mentioned that considering the errors for both ordinate and abscissa provides an estimated semi-variogram that is "nearest" to the empirical values in the *geometric sense*, namely measured along perpendicular projections onto the graph of the semi-variogram, they should have considered the empirical semi-

variogram as a non-typical EIV model since the observed quantities in its linearized versions have non-linear relationships. The examples of this paper numerically indicate the significance of this problem. In other words, the standard TLS solution of the model (II) is not admissible when it is incorrectly considered as the model (I).

Although the method of least-squares is one of the oldest methods of estimation, it is still the automobile of modern statistical analysis (Stigler 1999); therefore, in this paper, first we review an algorithm in order to solve the non-typical EIV model which is a novel model using the traditional least-squares method with the linearization of a nonlinear mixed model and iterative improvement of the solution. We employ the traditional Lagrange approach to optimize the target function of this problem. Then by comparison, a modified weighted TLS (WTLS) algorithm is proposed in order to treat the non-typical EIV model (II) as a typical EIV model (I) overseeing the dependence of $A$ on $\underline{t}$ through $\underline{A} = A(\underline{t})$.

The TLS algorithms modify the elements of the design matrix in the typical and non-typical EIV models, see e.g., matrix $A(\underline{t})$ in models (I) and (II). This modification is desirable for the typical EIV models since the design matrix is a linear function of the random variables; however, it is not necessarily correct for the non-typical EIV models [model (II)] which have nonlinear design matrix. Therefore, the proposed way based on the modified WTLS algorithm is not necessarily equivalent to the former way but as it is shown by numerical examples it results in an excellent approximate solution. This foundation is important because there is no need for linearization in the TLS algorithms. Moreover, their rate of convergence is usually better than the traditional approach.

This paper is organized as follows. In the next section, the concepts of the non-typical EIV model are introduced. An algorithm in order to solve this problem is reviewed using the traditional non-linear least-squares method within a non-linear mixed model. We then propose the modified WTLS algorithm to deal with this problem. In a later section, two simulation studies give insight into the efficiency of the algorithms proposed. Finally we conclude the paper.

## 2 Non-typical EIV model and a review of its classic solution based on the traditional least-squares method with linearization of a non-linear mixed model: algorithm 1

The typical/classic EIV model has been introduced by several contributions. In this model the first order moment/mean of design matrix is linear, however, it is not true for non-typical EIV model in which the design matrix is a non-linear function of random variables.

The mathematical definitions of these two types of models clarify our discussion. The typical EIV model is given as follows:

$$\underline{y} = (A - \underline{E}_A)\xi + \underline{e}_y, \quad rank\,A = m < n, \tag{1}$$

$$\begin{bmatrix} \underline{e}_y \\ \underline{e}_A \end{bmatrix} := \begin{bmatrix} \underline{e}_y \\ \mathrm{vec}(\underline{E}_A) \end{bmatrix} \sim \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix} \middle| \sigma_0^2 \begin{bmatrix} Q_y & 0 \\ 0 & Q_A \end{bmatrix} \right), \quad \mathrm{P}_y = Q_y^{-1}, \quad \mathrm{P}_A = Q_A^{-1}. \tag{2}$$

Here $\underline{y}$ is the $n \times 1$ observation vector, $\underline{e}_y$ is the respective $n \times 1$ vector of observational noise, $A$ is the $n \times m$ coefficient matrix of input variables (observed), $\underline{E}_A$ is the corresponding $n \times m$ matrix of random noise, $\xi$ is the $m \times 1$ parameters vector

(unknown), $D(\underline{e}_y) = \sigma_0^2 Q_y$ and $D(\underline{e}_A) = \sigma_0^2 Q_A$ are the corresponding dispersion matrices of size $n \times n$ and $mn \times mn$ (partly known), $\sigma_0^2$ is the variance component (unknown).

We define the non-typical EIV model as follows:

$$\underline{y} = A(t - \underline{e}_t)\xi + \underline{e}_y, \quad rank\ A(t - \underline{e}_t) = m < n, \tag{3}$$

or

$$E(\underline{y}) = E(A(\underline{t}))\xi \quad \text{with} \quad E(A(\underline{t})) \neq A(E(\underline{t})), \tag{4}$$

$$\begin{bmatrix} \underline{e}_y \\ \underline{e}_t \end{bmatrix} \sim \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix} | \sigma_0^2 \begin{bmatrix} Q_y & 0 \\ 0 & Q_t \end{bmatrix} \right), \quad P_y = Q_y^{-1}, \quad P_t = Q_t^{-1}, \tag{5}$$

where $\underline{y}$ is the $n \times 1$ observation vector, $\underline{e}_y$ is the respective $n \times 1$ vector of observational noise, the coefficient matrix $A()$ is the $n \times m$ non-linear function of input random variables $\underline{t}$ of size $l \times 1$ (observed), $\underline{e}_t$ is the corresponding $l \times 1$ vector of random noise $\xi$ is the $m \times 1$ parameters vector (unknown), $D(\underline{e}_y) = \sigma_0^2 Q_y$ and $D(\underline{e}_t) = \sigma_0^2 Q_t$ are the corresponding dispersion matrices of size $n \times n$ and $l \times l$ (known), $\sigma_0^2$ is the variance component (unknown).

Equation (4) can be easily converted into a non-linear mixed model which we define as follows (see e.g., Leick 2004, Chap. 4):

$$f(\underline{y}, \underline{t}, \xi) = 0, \tag{6}$$

where $f()$ indicates an implicit non-linear relationship with two groups of random observed vectors $\underline{y}$ and $\underline{t}$ and one deterministic unknown vector $\xi$. Snow (2012) denoted all the observed quantities (here $\underline{y}$ and $\underline{t}$) by a vector $\mathbf{Y}$ and the unknown vector $\xi$ by the vector $\mathbf{\Xi}$ and introduced this model as the non linear *Gauss–Helmert* model.

In fact the old works by Deming (1931) treat very special cases which we might now classify as non-linear type of EIV models, but he does not treat what we call non-typical EIV model in all his generality. The suggested paper by Xu et al. (2012) offers a presentation of preceding work plus two new results in Sects. 2.2 and 2.3. The first (2.2) is a WTLS algorithm treating the classical EIV model where he assumes that all elements of $A$ are independent random variables in the sense that $Q_A$ is invertible ($\boldsymbol{\omega}^{-1}$ in their notation). This is hardly the case because we almost never observe all the elements of $A$ but only a much smaller number of variables $t$ on which $A(t)$ depends. As a consequence $Q_A$ is singular and one has to use Rao's unified theory for getting a proper weight matrix out of infinitely many. In this respect our approach is much more general and closer to the reality of actual applications.

Also we can say that the non-typical EIV model (3) or even its linearized version will be given by Eq. (7) differ from the partial EIV model proposed in Xu et al. (2012) because they merely partitioned the matrix A into a stochastic and a deterministic parts while in our case the observed vector $t$ in $A(t)$ [Eq. (4)] cannot be separated from design matrix before linearization since its elements are arbitrary non-linear functions. In other words, the deterministic matrix B in Eq. (25-a) (Xu et al. 2012) cannot be extracted from our non-typical EIV model. In fact, the approach of Xu et al. (2012) can solves only the typical EIV model. Their approach is useful for statistical analysis.

In order to adjust the non-typical EIV model defined by Eqs. (3)–(6) using traditional non-linear LS within a mixed model, first one must linearize it respect to the two random

unknowns which refer to observable quantities $\underline{y}$ and $\underline{t}$ and one deterministic unknown $\xi$ as follows:

$$y_0 - A(t_0)\xi_0 + I_n e_y - \left(\xi^T \otimes I_n\right)\frac{\partial A(t)}{\partial t}\bigg|_{\substack{t\,=\,t_0 \\ \xi\,=\,\xi_0}} e_t - A(t_0)\delta\xi = 0,$$

$$w + e_y - Be_t - A\delta\xi = 0, \tag{7}$$

with

$$w = y_0 - A(t_0)\xi_0, \quad B = \left(\xi^T \otimes I_n\right)\frac{\partial A(t)}{\partial t}\bigg|_{\substack{t\,=\,t_0 \\ \xi\,=\,\xi_0}}, \quad A = A(t_0).$$

The Lagrange target function can be set up as follows:

$$\Phi\left(e_y,\ e_t,\ \lambda,\ \delta\xi\right) := e_y^T P_y e_y + e_t^T P_t e_t + 2\lambda^T\left(w + e_y - Be_t - A\delta\xi\right), \tag{8}$$

where $\lambda$ is the $n \times 1$ (unknown) Lagrange multiplier vector.

Since the derivation of this problem is well-known, we only present the following algorithm for the non-typical EIV:

First step: input $y^{(0)} = y$(observed), $t^{(0)} = t$(observed) and

$$\hat{\xi}^{(0)} = \left(A\left(t^{(0)}\right)^T P_y A\left(t^{(0)}\right)\right)^{-1} A\left(t^{(0)}\right)^T P_y y^{(0)}.$$

Second step: for $i \in N$ compute:

$$w^{(i)} = y - A(t)\hat{\xi}^{(i-1)}, \quad B^{(i)} = \left(\xi^T \otimes I_n\right)\frac{\partial A(t)}{\partial t}\bigg|_{\substack{t\,=\,t^{(i-1)} \\ \xi\,=\,\hat{\xi}^{(i-1)}}}, \quad A^{(i)} = A\left(t^{(i-1)}\right), \tag{9}$$

$$R^{(i)} = \left[Q_y + B^{(i)} Q_t B^{(i)T}\right]^{-1}, \tag{10}$$

$$\delta\hat{\xi}^{(i)} = \left(A^{(i)T} R^{(i)} A^{(i)}\right)^{-1} A^{(i)T} R^{(i)} w^{(i)}, \tag{11}$$

$$\hat{\lambda}^{(i)} = R^{(i)}\left(w^{(i)} - A^{(i)}\delta\hat{\xi}^{(i)}\right), \tag{12}$$

$$\tilde{e}_y^{(i)} = -Q_y \hat{\lambda}^{(i)}, \tag{13}$$

$$\tilde{e}_t^{(i)} = Q_t B^{(i)T} \hat{\lambda}^{(i)}, \tag{14}$$

$$\hat{\xi}^{(i)} = \hat{\xi}^{(i-1)} + \delta\hat{\xi}^{(i)}, \tag{15}$$

$$y^{(i)} = y + \tilde{e}_y^{(i)}, \tag{16}$$

$$t^{(i)} = t + \tilde{e}_t^{(i)}. \tag{17}$$

Third step: repeat 2nd step until one sees convergence.

The variance component $\hat{\sigma}_0^2$ can be estimated based on the proposed algorithm by exploring Eqs. (13)–(15) in the following quadratic forms:

$$\hat{\sigma}_0^2 = \frac{\tilde{e}_y^T P_y \tilde{e}_y + \tilde{e}_t^T P_t \tilde{e}_t}{n - m} = \frac{\hat{\lambda}^T R^{-1} \hat{\lambda}}{n - m}. \qquad (18)$$

*Important note* since $Q_y$ and $Q_t$ are usually invertible and the matrix $B$ is full row rank, the normal equation of this algorithm has an excellent stability and consequently the algorithm is usually stable. The iterative method based on the GHM is also not sensitive to initial values, because the estimates in the first iteration can be a very good approximated solution, see Shen et al. (2011).

## 3 A total least-squares (TLS) algorithm to deal with non-typical EIV model: algorithm 2

Although there is only one least squares criterion, there are several techniques by which least squares may be applied. Regardless of which technique is applied, the results of an adjustment of a given set of measurements must be the same. The choice of a technique, therefore, is mostly a matter of convenience and/or computational economy (Mikhail 1976). In fact, both the TLS and traditional non-linear LS methods are based on the L2 norm estimator; however, some theoretical properties have the former be more interesting than the latter. For instance, the TLS algorithms do not need linearization and they are less sensitive to the approximate initial values of the unknown parameters than the traditional approaches. As a result, one is right to seek a TLS algorithm to deal with the non-typical EIV model although in this case, the first order moment/mean of the random design matrix is not linear and we cannot employ a TLS algorithm to solve it directly.

A comparison between the algorithm of previous section and the WTLS algorithm within the typical EIV model [given by Eqs. (1) and (2)] in Mahboub (2012) shows that this WTLS algorithm has potential to deal with the non-typical EIV model after some slight modifications. In other words, treat the non-typical EIV model as a typical EIV model. Its actual computation can be based on Monte-Carlo methods. However, a first order approximation can be obtained by applying the law of covariance propagation to the noisy design matrix $A(\underline{t})$.

Approximately the converted typical EIV model is: $y - e_y = \left( A(t) - \frac{\partial A(t)}{\partial t} e_t \right) \xi$, i.e., $E_A = \frac{\partial A(t)}{\partial t} e_t$ and the dispersion matrix $Q_A$ is easily derived by the law of error propagation as $Q_A = \frac{\partial A(t)}{\partial t} Q_t \left( \frac{\partial A(t)}{\partial t} \right)^T$. The converted typical EIV model can be solved with WTLS. Summarizing the following WTLS algorithm for the non-typical EIV is proposed:

First step: $[N, c] = A^T P_y [A, y]$, $\xi^{(0)} = N^{-1} c$.

Second step: $Q_A = \frac{\partial vec(A(t))}{\partial t}\Big|_{t=t^{(0)}} Q_t \left( \frac{\partial vec(A(t))}{\partial t}\Big|_{t=t^{(0)}} \right)^T$.

Third step: for $i \in N$ compute:

$$R_1^{(i)} = \left[ Q_y + \left( \hat{\xi}^{(i-1)T} \otimes I_n \right) Q_A \left( \hat{\xi}^{(i-1)} \otimes I_n \right) \right]^{-1},$$

$$\hat{\lambda}^{(i)} = R_1^{(i)}\left(y - A\hat{\xi}^{(i-1)}\right),$$

$$R_2^{(i)} = \left(I_m \otimes \hat{\lambda}^{(i)T}\right)Q_A\left(\hat{\xi}^{(i-1)} \otimes I_n\right)R_1^{(i)},$$

$$\hat{\xi}^{(i)} = \left(A^T R_1^{(i)} A + R_2^{(i)} A\right)^{-1}\left(A^T R_1^{(i)} + R_2^{(i)}\right)y.$$

Fourth step: repeat *third* step until one sees convergence

$$\left\|\hat{\xi}^{(i)} - \hat{\xi}^{(i-1)}\right\| < \delta.$$

## 4 Numerical results and discussions

The determination of the initial position and constant velocity from redundant position measurements is an example of curve fitting. Let $g(x)$ be an unknown function and we measure $m$ points of it which both coordinates $(x, g(x))$ are falsified by random noise (see Fig. 1)

$$y_i \approx g(x_i), \quad i = 1, \ldots, m. \tag{19}$$

Clearly, one is not able to reconstruct an arbitrary function $g(x)$ from a finite set of its samples and we require additional information about it. The additional information can be expressed by a linear combination of $n$ known base functions $b_r(x)$, $r = 1, \ldots, n$,

$$g(x) = \sum_{r=1}^{n} a_r b_r(x), \tag{20}$$

with the $n$ unknown coefficients $a_r$. Equations (19) and (20) give the following system of equations:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \approx \begin{bmatrix} b_1(x_1) & \cdots & b_n(x_1) \\ \vdots & & \vdots \\ b_1(x_m) & \cdots & b_n(x_m) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}. \tag{21}$$

This system is non-linear because of the randomness of the coordinates $x_i$ which appear in the base functions $b_r$. Therefore one is not allowed to solve this system using linear LS method within a GM model. Also as it has been discussed in the previous sections, the TLS

Fig. 1 The sampling of an unknown function g(x) for $x_i$, $i = 1, \ldots, m$
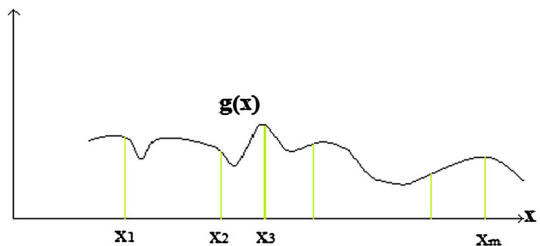
algorithms within a typical EIV model are not admissible theoretically. We should employ one of the two proposed ways in this research.

Due to the nature of the function $g(x)$, different base functions $b_r(x)$ can be used. Here we examine two sets of the useful base functions namely trigonometric and polynomial series.

## 4.1 Curve fitting using trigonometric base functions

Suppose that we measure the coordinates of 10 points of the function $g(x)$ which is given as follows:

$$\underline{y} = a \sin \underline{x} + b \cos \underline{x} + c = [\sin \underline{x} \quad \cos \underline{x} \quad 1] \begin{bmatrix} a \\ b \\ c \end{bmatrix} = A(\underline{x}) \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \qquad (22)$$

where $a$, $b$ and $c$ are the unknown coefficients and $(\underline{x}, \underline{y})$ denotes the noisy coordinates which have been observed with different precision. The coordinate of these samples with their weights are given in Table 1 and Fig. 2.

We adjust the system of Eq. (22) using four ways: 1, the linear LS method; 2, the standard TLS method on the assumption that the non-typical EIV is a typical EIV model; 3, algorithm 1 in Sect. 2 based on the traditional non-linear LS method within a mixed model; 4, algorithm 2 in Sect. 3 based on modified WTLS method within a typical EIV model. The estimated unknown parameters using five methods and the true value of the unknown parameters are given in Table 2.

From Table 2, one can clearly see that the linear LS method gives a bias estimation of non-typical EIV model (22) which in fact is a non-linear system of equations. Also one is not permitted to adjust this problem using the standard TLS method on the assumption that the non-typical EIV is a typical EIV model since as it has been proven theoretically, in such a case, the non-linear relationships of the elements of the random design matrix are neglected. The results of Linear Ls and TLS (third column) are incorrect. The correct solutions can be obtained by the two developed algorithms of this paper. Also algorithm 2 is more stable than algorithm 1 due to the number of iterations since the former converges after two iterations, while, starting by the same initial unknown parameters more than five iterations are required for the algorithm 2 to meet the same threshold; furthermore, it needs linearization while algorithm 2 does not require any linearization. Algorithm 2 is not

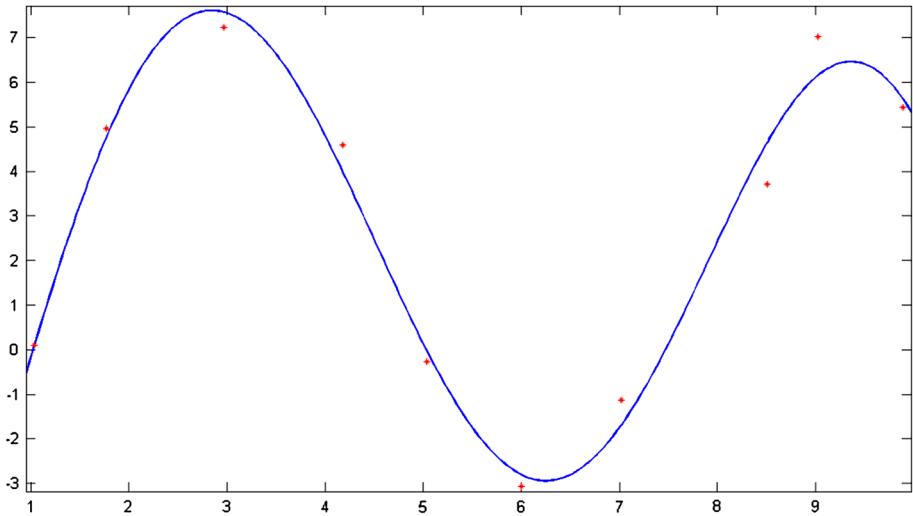| Table 1 Observed points and their weights | Point nos. | x | y | $w_x$ | $w_y$ |
|---|---|---|---|---|---|
| | 1 | 1.0388 | 0.10393 | $1^{-2}$ | $4^{-2}$ |
| | 2 | 1.774 | 4.9693 | $2^{-2}$ | $2^{-2}$ |
| | 3 | 2.9718 | 7.2256 | $1^{-2}$ | $3^{-2}$ |
| | 4 | 4.1803 | 4.5939 | $4^{-2}$ | $4^{-2}$ |
| | 5 | 5.0395 | −0.26773 | $2^{-2}$ | $4^{-2}$ |
| | 6 | 6.0017 | −3.0729 | $7^{-2}$ | $1^{-2}$ |
| | 7 | 7.0218 | −1.1306 | $1^{-2}$ | $1^{-2}$ |
| | 8 | 8.5085 | 3.7199 | $9^{-2}$ | $1^{-2}$ |
| | 9 | 9.0231 | 7.0219 | $3^{-2}$ | $5^{-2}$ |
| | 10 | 9.8862 | 5.4416 | $3^{-2}$ | $3^{-2}$ |

**Fig. 2** The trigonometric curve and the noisy coordinates of 10 *points* as its samples

**Table 2** The estimated unknown parameters using three methods and the true value of the unknown parameters

| Methods | Linear LS | TLS (on the assumption that the non-typical EIV is a typical EIV model) | Algorithm 1 | Algorithm 2 | Exact solution |
|---------|-----------|------------------------------------------------------------------------|-------------|-------------|----------------|
| *a* | 0.5884 | −0.3462 | 0.9769 | 0.9769 | 1 |
| *b* | −4.2861 | −0.39835 | −5.0649 | −5.0649 | −5 |
| *c* | 1.4642 | 1.62191 | 2.0316 | 2.0316 | 2 |

The results of Linear Ls and TLS (third column) are incorrect

sensitive to approximate initial values of the parameters which is the bottleneck problem that restricts the application of the nonlinear techniques and the solution of such methods is somewhat critical to handle due to the many pitfalls described by Pope (1972).

### 4.2 Curve fitting using polynomial base functions

Let $g(x)$ be an unknown function of which we measure the coordinates of 10 points:

$$\bar{y} = a_0 + a_1\underline{x} + a_2\underline{x}^2 + a_3\underline{x}^3 = \begin{bmatrix} 1 & \underline{x} & \underline{x}^2 & \underline{x}^3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = A(\underline{x}) \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}, \qquad (23)$$

where $a$, $b$ and $c$ are the unknown coefficients and $(\underline{x}, \underline{y})$ denotes the noisy coordinates which have been observed with different precision. The coordinate of these samples with their weights are given in Table 3 and Fig. 3.

Similarly we adjust the system of Eq. (23) using four ways: 1, the linear LS method; 2, the standard TLS method on the assumption that the non-typical EIV is a typical EIV

| **Table 3** Observed points and their weights | Point nos. | X | Y | $w_x$ | $w_y$ |
|---|---|---|---|---|---|
| | 1 | 1.004 | −1.007 | $1^{-2}$ | $4^{-2}$ |
| | 2 | 2.007 | −8.024 | $2^{-2}$ | $2^{-2}$ |
| | 3 | 3.003 | −13.01 | $1^{-2}$ | $3^{-2}$ |
| | 4 | 3.971 | −10.13 | $4^{-2}$ | $4^{-2}$ |
| | 5 | 5.007 | 6.957 | $2^{-2}$ | $4^{-2}$ |
| | 6 | 5.964 | 43.99 | $7^{-2}$ | $1^{-2}$ |
| | 7 | 6.991 | 107 | $1^{-2}$ | $1^{-2}$ |
| | 8 | 7.892 | 202 | $9^{-2}$ | $1^{-2}$ |
| | 9 | 9.031 | 335 | $3^{-2}$ | $5^{-2}$ |
| | 10 | 9.975 | 512.1 | $3^{-2}$ | $3^{-2}$ |
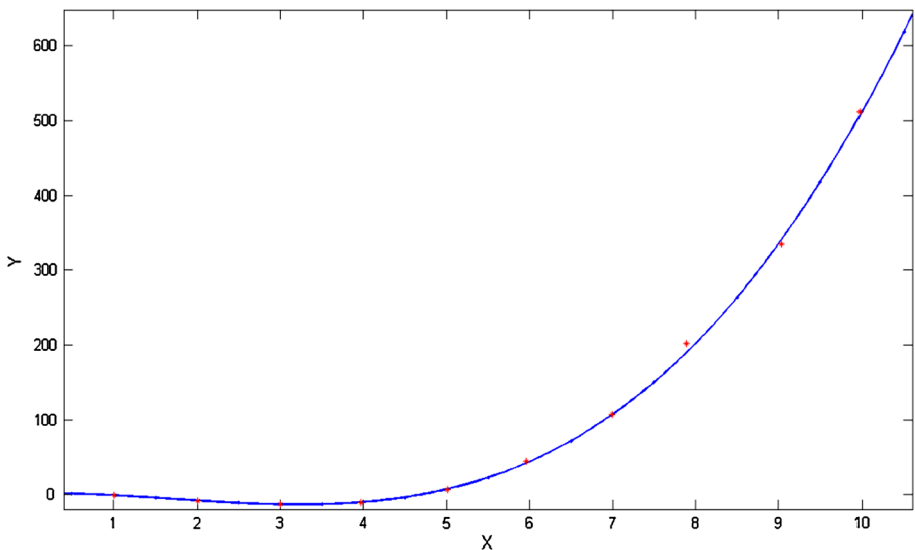


**Fig. 3** The polynomial curve and the noisy coordinates of 10 *points* as its samples

model; 3, algorithm 1 in Sect. 2 based on the traditional non-linear LS method within a mixed model; 4, algorithm 2 in Sect. 3 based on modified WTLS method within a typical EIV model. The estimated unknown parameters using four methods and the true value of the unknown parameters are given in Table 4.

Similar to trigonometric curve, the linear LS method gives a bias estimation in polynomial curve. By comparison, this bias is bigger than the bias of previous example. It can refer to different non-linear property of trigonometric and polynomial functions. The similar reasoning can be given for the standard TLS method on the assumption that the non-typical EIV is a typical EIV model. Also the reasonable solutions are obtained by the stable algorithms 1 and 2, although there is a negligible difference between the results.

**Table 4** The estimated unknown parameters using three methods and the true value of the unknown parameters

| Methods | Linear LS | TLS (on the assumption that the non-typical EIV is a typical EIV model) | Algorithm 1 | Algorithm 2 | Exact solution |
|---------|-----------|------------------------------------------------------------------------|-------------|-------------|----------------|
| $a_0$ | 17.6866 | 117.1073 | 1.9971 | 1.9971 | 2 |
| $a_1$ | −11.8162 | −8.45142 | 1.04638 | 1.04638 | 1 |
| $a_2$ | −2.4512 | 0.031834 | −5.0293 | −5.0293 | −5 |
| $a_3$ | 0.8662 | 0.00063 | 1.0046 | 1.0046 | 1 |

The results of Linear Ls and TLS (third column) are incorrect

## 5 Conclusions

In this paper, non-typical EIV model was introduced where the elements of the design matrix are non-linear functions of observed noisy quantities. As a result, the first order moment/mean of the design matrix is not directly known, consequently TLS algorithms within a typical/classic EIV model are not applicable to it.

The non-typical EIV model appears in some applications such as curve fitting and surface reconstruction. Two algorithms were presented to deal with it. Algorithm 1 is based on traditional non-linear LS method within a mixed model and algorithm 2 is a modified WTLS algorithm. Although the classic former way produces the optimal LS solution to this problem, traditional non-linear techniques usually have their own difficulties (see e.g., Pope 1972), while, algorithm 2 does not need linearization and its initial values can be easily computed by a linear LS estimation which is the bottleneck problem that restricts the application of the nonlinear techniques. Therefore a simple use of the linear LS method can produce these initial values.

As the numerical examples show, the linear LS method gives a bias estimation of the non-typical EIV model. The amount of this bias depends on non-linear property of the system of equations. This conclusion had been already obtained for the typical EIV model by several contributions.

Both examples convincingly demonstrate that the standard TLS solution of the non-typical EIV model is not admissible when it is incorrectly considered as a typical EIV model; i.e., the non-linear relationships of the elements of the random design matrix are neglected.

Although we do not claim that our results based on algorithm 2 are better than algorithm 1, algorithm 2 is a proper TLS approach which is more accurate than exiting TLS algorithms in dealing with the non-typical EIV model and also is simpler than the traditional non-linear LS method within a GH model (algorithm 1) since if we formulate the non-typical EIV model in terms of the non linear GH model, one has to solve a complicate model while in our method (algorithm 2) we can directly work with the design/coefficient matrix. We only need to compute the dispersion matrix $Q_A$. Furthermore, the examples demonstrated that algorithm 2 is converged in fewer numbers of iterations than algorithm 1.

Finally we emphasize that both the nonlinear LS and our algorithm (algorithms 1 and 2) are not unbiased estimators both for the typical and non-typical TLS problem even though the nonlinear function in the coefficient matrix is properly considered. One may design a bias corrected estimator for nontypical EIV model, see e.g., Box (1971) and Xu et al. (2012).

# References

Box MJ (1971) Bias in nonlinear estimation (with discussions). J R Stat Soc B 33:171–201

Deming WE (1931) The application of least squares. Philos Mag 11:146–158

Fang X (2011) Weighted total least squares solutions for applications in geodesy. PhD Dissertation number 294, Leibniz University, Hannover

Fang X (2013) Weighted total least squares: necessary and sufficient conditions, fixed and random parameters. J Geod. doi:10.1007/s00190-013-0643-2

Fang X (2014) A structured and constrained Total Least-Squares solution with cross-covariances. Stud Geophys Geod. doi:10.1007/s11200-012-0671-z

Felus A (2004) Application of total least squares for spatial point process analysis. J Surv Eng 130(3):126–133

Golub G, Van Loan C (1980) An analysis of the Total Least Squares problem. SIAM J Numer Anal 17:883–893

Leick A (2004) GPS satellite surveying. Wiley, New York

Mahboub V (2012) On weighted total least-squares for geodetic transformations. J Geod 86(5):359–367. doi:10.1007/s00190-011-0524-5

Mahboub V (2014) Variance component estimation in errors-in-variables models and a rigorous total least-squares approach. Stud Geophys et Geod 58(1):17–40

Mahboub V, Sharifi MA (2013) On weighted total least-squares with linear and quadratic constraints. J Geod. doi:10.1007/s00190-012-0598-8

Mahboub V, Amiri-Simkooei AR, Sharifi MA (2012) Iteratively reweighted total least-squares: a robust estimation in errors-in-variables models. Surv Rev. doi:10.1179/1752270612Y.0000000017

Mikhail E (1976) Observations and least squares. IEP-A dun-Donnelley Publisher, New York

Neitzel F (2010) Generalization of Total Least-Squares on example of unweighted and weighted similarity transformation. J Geod 84(2010):751–762

Pope AJ (1972) Some pitfalls to be avoided in the iterative adjustment of nonlinear problems. In: Proceedings of the 38th annual meeting of the American Society of Photogrammetry, Washington, DC, pp 449–477

Schaffrin B, Felus Y (2008) An algorithmic approach to the total least-squares problem with linear and quadratic constraints. Stud Geophys Geod 53(1):1–16

Schaffrin B, Uzun S (2008) Surface reconstruction via total least-squares adjustment of the semi-variogram. In International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences (Vol. 37, pp. 141–148). International Society for Photogrammetry and Remote Sensing.

Schaffrin B, Wieser A (2008) On weighted total least-squares adjustment for linear regression. J Geod 82(7):415–421

Schaffrin B, Neitzel F, Uzun S, Mahboub V (2012a) Modifying Cadzow's algorithm to generate the optimal TLS-solution for the structured EIV-model of a similarity transformation. J Geod Sci. doi:10.2478/v10156-011-0028-5

Schaffrin B, Snow K, Neitzel F (2012b) On the Errors-In-Variables Model with singular covariance matrices. In: 21st International workshop on matrices and statistics, Będlewo, Poland

Shen YZ, Li BF, Chen Y (2011) An iterative solution of weighted total least squares adjustment. J Geod 85:229–238

Snow K (2012) Topics in total least-squares adjustment within the errors-in-variables model: singular cofactor matrices and priori information. PhD Dissertation, report No., 502, Geodetic Science Program, School of Earth Sciences, The Ohio State University, Columbus

Snow K, Schaffrin B (2012) Weighted total least-squares collocation with geodetic applications. In: SIAM conference in applied linear algebra, Valencia, Spain, June 2012

Stigler F (1999) Statistics on the table. Harvard University Press, Cambridge

Van Huffel S, Vandewalle J (1991) The total least-squares problem. Computational aspects and analysis. SIAM, Philadelphia

Xu PL, Liu JN, Shi C (2012) Total least squares adjustment in partial errors-in-variables models: algorithm and statistical analysis. J Geod 86(8):661–675