Connecting the Dots: The Straight-Line Case Revisited

Burkhard Schaffrin

Summary

In the present contribution the author tries to compare newly developed procedures for straight-line adjustment in two dimensions with other methods that have proven rather popular in Geodetic Science when both coordinates of a point cloud are affected by random observation errors. An attempt is made for a rigorous and systematic treatment within the so-called EIV-Model (EIV: Errors-in-Variables), leading to the method of Total Least-Squares (TLS) adjustment. A weighted form of this procedure (WTLS: Weighted Total Least-Squares) will be presented, along with a discussion of a suitable choice for the weights.

Zusammenfassung

Im vorliegenden Beitrag bemüht sich der Autor um eine Gegenüberstellung neu entwickelter und in der Geodäsie gebräuchlicher Verfahren zur Bestimmung einer ausgleichenden Geraden durch eine zweidimensionale Punktwolke mit gemessenen – und daher mit Zufallsfehlern behafteteten – (x,y)-Koordinaten. Nach einer gewissen Systematisierung im EIV-Modell (EIV: Errors-in-Variables) wird besonders auf das Verfahren einer gewichteten TLS-Ausgleichung (TLS: Total Least-Squares) eingegangen und die Problematik einer geeigneten Gewichtswahl aufgezeigt.

Introduction

The problem of finding an adjusted straight-line through a two-dimensional point cloud is a very old one. Wolf (1968, p. 422), for instance, reports an approximate solution by L. Cauchy (1789–1857) on the straight-line case whereas, more recently, Schmid and Schmid (1971) even tackled the general problem of curve (resp. surface) fitting from a geodetic viewpoint. The textbooks that took up this question extensively include those by Wolf (1975, p. 224), Höpcke (1981, p. 176), Mikhail and Gracie (1981, pp. 237–241), Koch (1990, p. 105), and Wolf and Ghilani (1997, pp. 423–430), although not all of them really succeeded as will be shown later. While differences in the individual solutions are mostly due to the various assumptions on the random observation noise for the respective coordinates, it sometimes appears, however, that optimal estimates of the slope angle are compared with optimal estimates of the slope itself as, e. g., by Wolf (1968, pp. 417–421). It is of utmost importance to show that such a comparison – and thus the corresponding formulas by Koch (1990, p. 105) – can also work for other than small (!) slope angles where the "tan" function is nearly linear. Similarly the inverse optimal slope, by all means, ought to coincide with the optimal inverse slope.

Some clarification and enlightenment may well have come from the study of the *Total Least-Squares (TLS)* method that had originally been introduced by Golub and van Loan (1980); see also van Huffel and Vandewalle (1991), and Golub and van Loan (1996). TLS is somewhat of a misnomer as it represents a regular least-squares adjustment, but within an *Error-in-Variables (EIV) Model*, not unlike Least-Squares Collocation (LSC) which really represents the standard least-squares adjustment within a so-called Mixed Model; see Schaffrin (2001) and Schaffrin et al. (2006) for extensive explanations.

Here the TLS procedure will briefly be summarized, in order to find the "unique" solution of Lenzmann and Lenzmann (2001) and Reinking (2001) for the straight-line case in a more elegant fashion. Moreover, it will be shown that the traditional way to derive the famous formula by Schumann (1916) for the optimal slope angle – as quoted by Wolf (1968, p. 419) – can easily be modified to generate the Total Least-Squares Solution (TLSS) of the slope instead. Finally, an extension to a *weighted TLS adjustment technique* will be given, following Schaffrin and Wieser (2007), and applied to the straight-line adjustment again. It is emphasized, however, that the notion of a suitable weight choice is still unclear and needs to be investigated in the future.

1 Total Least-Squares Fit Within an EIV-Model

Throughout this contribution, only the *Errors-in-Variables (EIV) Model* will be used for the two-dimensional straight-line fit, thereby assuming that both coordinates of the point cloud are affected by random observation errors. It is important to note that the case of zero residuals after an adjustment, for either the x- or the y-coordinates, will here *not be interpreted* as result of a least-squares approach to a modified model, *but* as a modified estimation technique within the underlying EIV-Model. This view will, in particular, lead to precision measures that turn out different from those commonly in use; but such a discussion is beyond the scope of the present paper.

Let us now define the EIV-Model for a *straight-line fit* through a two-dimensional point cloud with *observed* (x,y)-coordinates. In the case of n points, we collect the x- and y-coordinates separately in two $n \times 1$ vectors, namely

$$x := \left[x_1, \cdots, x_n\right]^T \text{ and } y := \left[y_1, \cdots, y_n\right]^T. \tag{1.1}$$

These, after having removed the random observation errors by subtracting the (unknown) n × 1 vectors e_x and e_y, form the relationship

$$y - e_y = \tau \cdot \xi_1 + (x - e_x) \cdot \xi_2 \tag{1.2a}$$

where $\tau := [1, \cdots, 1]^T$ denotes the n × 1 "summation vector". In addition, we introduce first and second moments for the random error vectors:

$$e_x \sim (0, \sigma_o^2 Q_x) \quad \text{and} \quad e_y \sim (0, \sigma_o^2 Q_y) \tag{1.2b}$$

with, at this point, *no correlation* between them:

$$C\{e_x, e_y\} = E\{e_x \cdot e_y^T\} = 0 \ . \tag{1.2c}$$

Here σ_o^2 denotes a common (unknown) variance component that is unitless; Q_x and Q_y are the respective cofactor matrices (with proper units), both symmetric and positive-definite. Furthermore, E denotes "expectation", and C stands for "covariance"; later we shall use the symbol D for "dispersion".

Obviously, the parameter vector $\xi := [\xi_1, \xi_2]^T$ consists of two unknowns, usually denoted by "intercept" (ξ_1) and "slope" (ξ_2). We note, in particular, that the "slope angle" is not directly represented in formula (1.2a), and must be derived from the estimated ξ_2. Due to the nonlinear "tan" function, an optimally estimated slope may thus not necessarily generate an optimal estimate for the slope angle (and vice versa).

On the other hand, there exists a *dual form* of the relationship (1.2a), namely

$$x - e_x = \tau \cdot \eta_1 + (y - e_y) \cdot \eta_2 \tag{1.3a}$$

with the identities

$$\eta_1 := -(\xi_1 / \xi_2) \quad \text{and} \quad \eta_2 := 1/\xi_2 \tag{1.3b}$$

as *"reverse intercept"* and *"inverse slope"*. Of course, we cannot generally expect the optimal estimate of η_2 to be the inverse optimal estimate of ξ_2, and similarly for the relationship between η_1 and ξ_1.

A close look at model (1.2a-c) and model (1.3a-b) leads to the conclusion that they can both be classified as *nonlinear Gauss-Helmert Models* for which a Least-Squares Solution (LESS) may be formed by proper (!) iterative linearization. This, however, can be tricky as was pointed out already by Pope (1972), followed by another discussion by Lenzmann and Lenzmann (2004a,b; 2005), Koch (2004), and Kupferer (2004).

On the other hand, the solution of linear normal equations for iteratively linearized GH-Models only defines *one particular algorithm* to solve the nonlinear normal equations for the nonlinear GH-Model. This argument has actually been brought forward by Lenzmann and Lenzmann (2007) for the nonlinear Gauss-Markov Model, and is equally valid in the present context. In fact, we shall design a *quite efficient* algorithm of *Newton type* that solves the nonlinear normal equations for model (1.2a-c) and model (1.3a-b), respectively; for more details we refer to Schaffrin et al. (2006). There, an even more efficient algorithm of *Raleigh type* has also been presented which may, however, show convergence problems for inaccurate starting values.

In the following we adapt the derivations for the weighted TLS adjustment by Schaffrin and Wieser (2007) to the straight-line case by first allowing τ to be a random vector whose cofactor matrix is set to zero in the end. In this spirit, model (1.2a-c) can be rewritten as:

$$y = [\tau - e_\tau, x - e_x] \cdot \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + e_y \ , \tag{1.4a}$$

$$\begin{bmatrix} e_y \\ e_\tau \\ e_x \end{bmatrix} \sim \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \sigma_o^2 \begin{bmatrix} Q_y & 0 & 0 \\ 0 & \sigma_\tau^2 Q_x & 0 \\ 0 & 0 & Q_x \end{bmatrix} \right), \qquad \sigma_\tau^2 \to 0 \ . \tag{1.4b}$$

By using the notations

$$A := [\tau, x] \ , \ E_A := [e_\tau, e_x] \ \text{and} \ \xi := [\xi_1, \xi_2]^T \ , \tag{1.5a}$$

we obtain, for the vectorial form of E_A, the characterization

$$e_A := vecE_A = \begin{bmatrix} e_\tau \\ e_x \end{bmatrix} \sim \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma_o^2 Q_A \right) \tag{1.5b}$$

with

$$Q_A := Q_o \otimes Q_x \ , \quad Q_o := \begin{bmatrix} \sigma_\tau^2 & 0 \\ 0 & 1 \end{bmatrix}. \tag{1.5b}$$

Here, ⊗ denotes the "Kronecker-Zehfuss product" of matrices, defined by

$$G \otimes H := [g_{ij} \cdot H] \quad \text{if} \quad G = [g_{ij}] \ ; \tag{1.6}$$

for more details, we refer to Grafarend and Schaffrin (1993, p. 409).

By applying the *Euler-Lagrange approach* to model (1.4a-b) in combination with (1.5a-c), Schaffrin and Wieser (2007) succeeded in giving the corresponding *nonlinear normal equations* the form:

$$\hat{\xi} = \left[A^T (Q_y + (\hat{\xi}^T Q_o \hat{\xi}) \cdot Q_x)^{-1} A - \hat{\nu}' \cdot Q_o \right]^{-1} \cdot \left[A^T (Q_y + (\hat{\xi}^T Q_o \hat{\xi}) \cdot Q_x)^{-1} y \right] \tag{1.7a}$$

with

$$\hat{\nu}' := (y - A\hat{\xi})^T (Q_y + (\hat{\xi}^T Q_o \hat{\xi}) \cdot Q_x)^{-1} \cdot Q_x (Q_y + (\hat{\xi}^T Q_o \hat{\xi}) \cdot Q_x)^{-1} (y - A\hat{\xi}). \tag{1.7b}$$

After introducing the auxiliary vector $\hat{\lambda}$ (of Lagrange multipliers actually), namely

$$\hat{\lambda} := (Q_y + (\hat{\xi}^T Q_o \hat{\xi}) \cdot Q_x)^{-1}(y - A\hat{\xi}), \tag{1.8a}$$

the residual vector, resp. residual matrix result in:

$$\tilde{e}_y = Q_y \cdot \hat{\lambda} \quad \text{and} \quad \tilde{E}_A = -Q_x \hat{\lambda} \hat{\xi}^T Q_o \tag{1.8b}$$

so that the *weighted sum of squared residuals*, that has been minimized as our objective, becomes:

$$\hat{\lambda}^T Q_y \hat{\lambda} + \hat{\lambda}^T (\hat{\xi}^T Q_o \hat{\xi} \otimes Q_x)\hat{\lambda} = \hat{\lambda}^T (y - A\hat{\xi}) \; . \tag{1.8c}$$

By setting $Q_x = Q_y =: P^{-1}$, the system (1.7a-b) can be simplified into

$$\hat{\xi} = (N - \hat{v} \cdot Q_o)^{-1} \cdot c \quad \text{for} \quad [N, c] := A^T P [A, y] \tag{1.9a}$$

with $\sigma_\tau^2 = 0$ and

$$\hat{v} := (y - A\hat{\xi})^T P (y - A\hat{\xi}) / (1 + \hat{\xi}^T Q_o \hat{\xi}) = \hat{\lambda}^T (y - A\hat{\xi}) \; . \tag{1.9b}$$

After further translating these formulas into the original straight-line notation, we arrive at the system:

$$\begin{bmatrix} \hat{\xi}_1 \\ \hat{\xi}_2 \end{bmatrix} = \begin{bmatrix} \tau^T P \tau & \tau^T P x \\ x^T P \tau & x^T P x - \hat{v} \end{bmatrix}^{-1} \cdot \begin{bmatrix} \tau^T P y \\ x^T P y \end{bmatrix} \tag{1.10a}$$

where $\hat{v}$ may be rewritten, using (1.9a), into:

$$\begin{aligned} \hat{v} &= (y - A\hat{\xi})^T P (y - A\hat{\xi}) / (1 + \hat{\xi}^T Q_o \hat{\xi}) \\ &= \left[-\hat{\xi}^T (c - N\hat{\xi}) + y^T P (y - A\hat{\xi}) \right] \cdot (1 + \hat{\xi}^T Q_o \hat{\xi})^{-1} \\ &= \left[(\hat{\xi}^T Q_o \hat{\xi}) \cdot \hat{v} + (y^T P y - c^T \hat{\xi}) \right] \cdot (1 + \hat{\xi}^T Q_o \hat{\xi})^{-1} \end{aligned} \tag{1.10b}$$

or, even more simply, into:

$$\begin{aligned} \hat{v} &= (1 + \hat{\xi}^T Q_o \hat{\xi}) \cdot \hat{v} - (\hat{\xi}^T Q_o \hat{\xi}) \cdot \hat{v} \\ &= y^T P (y - A\hat{\xi}) = \hat{\lambda}^T (y - A\hat{\xi}) \; . \end{aligned} \tag{1.10c}$$

Apparently, $\hat{v}$ represents the *sum of squared residuals* here (in contrast to $\hat{v}'$ above) which has been minimized through the least-squares process. Therefore, $\hat{v}$ can as well be interpreted as *minimum (generalized) eigenvalue* to the system

$$\begin{bmatrix} N & c \\ c^T & y^T P y \end{bmatrix} \begin{bmatrix} \hat{\xi} \\ -1 \end{bmatrix} = \begin{bmatrix} Q_o & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\xi} \\ -1 \end{bmatrix} \cdot \hat{v}_{\min} \, , \tag{1.11}$$

which, for $P := I_n$ and $Q_o := I_m$, corresponds exactly to the original TLS solution by Golub and van Loan (1980). Here, however, the matrix Q_o is singular as soon as we set $\sigma_\tau^2 := 0$. Consequently, our algorithm of Newton type will rely on the equations (1.10a-c), here written up for the *homoscedastic* case (i.e., $P := I_n$).

After partial elimination of $\hat{\xi}_1$ from the system (1.10a), we first obtain, in the *i-th iteration*,

$$\begin{aligned} \hat{\xi}_2^{(i)} &= \left[x^T x - x^T \tau (\tau^T \tau)^{-1} \tau^T x - \hat{v}^{(i)} \right]^{-1} \\ &\quad \cdot \left[x^T y - x^T \tau (\tau^T \tau)^{-1} \tau^T y \right] \\ &= (x^T x - n \cdot \bar{x}^2 - \hat{v}^{(i)})^{-1} (x^T y - n \cdot \bar{x} \, \bar{y}) \end{aligned} \tag{1.12a}$$

with the center coordinates

$$\bar{x} := x^T \tau (\tau^T \tau)^{-1} \quad \text{and} \quad \bar{y} := y^T \tau (\tau^T \tau)^{-1} , \tag{1.12b}$$

and after convergence:

$$\hat{\xi}_1 = \bar{y} - \bar{x} \cdot \hat{\xi}_2 \, . \tag{1.13}$$

The coefficient $\hat{v}^{(i+1)}$ may be computed through (1.10a) via

$$\begin{aligned} \hat{v}^{(i+1)} &= y^T y - c^T \hat{\xi}^{(i)} = y^T y - y^T \tau \cdot \hat{\xi}_1^{(i)} - y^T x \cdot \hat{\xi}_2^{(i)} \\ &= (y^T y - n \cdot \bar{y}^2) - (y^T x - n \cdot \bar{x} \, \bar{y}) \cdot \hat{\xi}_2^{(i)} \end{aligned} \tag{1.14a}$$

or, in a more sophisticated manner, through the update from (1.10b) via

$$\begin{aligned} \hat{v}^{(i+1)} &= \hat{v}^{(i)} + (y^T y - c^T \hat{\xi}^{(i)} - \hat{v}^{(i)}) / [1 + (\hat{\xi}^{(i)})^T Q_o \hat{\xi}^{(i)}] \\ &= v^{(i)} + \left[(y^T y - n \cdot \bar{y}^2) - (y^T x - n \cdot \bar{x} \, \bar{y}) \cdot \hat{\xi}_2^{(i)} - \hat{v}^{(i)} \right] / [1 + (\hat{\xi}_2^{i})^2] \end{aligned} \tag{1.14b}$$

while using

$$\hat{v}^{(0)} := 0 \tag{1.14c}$$

as initial value.

It is interesting to see that (1.12a) and (1.14a) indeed lead to the (standard) *minimum eigenvalue problem*:

$$\begin{bmatrix} x^T x - n \cdot \bar{x}^2 & x^T y - n \cdot \bar{x} \, \bar{y} \\ x^T y - n \cdot \bar{x} \, \bar{y} & y^T y - n \cdot \bar{y}^2 \end{bmatrix} \begin{bmatrix} \hat{\xi}_2 \\ -1 \end{bmatrix} = \begin{bmatrix} \hat{\xi}_2 \\ -1 \end{bmatrix} \cdot \hat{v}_{\min} \tag{1.15}$$

as first described by Schaffrin et al. (2006, p. 148). Also note that, in case of convergence problems, formula (1.12a) may be replaced by

$$\hat{\xi}_2^{(i+1)} = (x^T x - n \cdot \bar{x}^2)^{-1} (x^T y - n \cdot \bar{x} \, \bar{y} + \hat{v}^{(i)} \cdot \hat{\xi}_2^{(i)}) \tag{1.16a}$$

with the initial solution

$$\hat{\xi}_2^{(i)} = (x^T x - n \cdot \bar{x}^2)^{-1} (x^T y - n \cdot \bar{x} \, \bar{y}) \, . \tag{1.16b}$$

Obviously, all the above schemes ought to result in the same TLS-Solutions (TLSS) $\hat{\xi}_1$ and $\hat{\xi}_2$ for intercept and slope.

Unfortunately, no statistical optimality properties are hitherto known for them (although *consistency* has meanwhile been established). The reason, of course, lies in the *strong nonlinearity* of the TLSS which also prevents us from discussing any precision measures for the estimates right now, with the exception of the variance component σ_o^2 that we propose to estimate through

$$\hat{\sigma}_o^2 = \hat{\nu}/(n-2)\,. \tag{1.17}$$

However, we cannot even claim unbiasedness for this estimate at this point.

As usually, the two residual vectors can be computed on the basis of the estimated parameters via (1.8a-b) which formulas, for the straight-line case, reduce to:

$$\begin{aligned}\hat{\lambda} &= (y - \tau\hat{\xi}_1 - x\hat{\xi}_2)/(1+\hat{\xi}_2^2)\\ &= \left[(y-\tau\bar{y}) - (x-\tau\bar{x})\hat{\xi}_2\right]/(1+\hat{\xi}_2^2) = \tilde{e}_y\end{aligned} \tag{1.18a}$$

and

$$\tilde{e}_x = -\hat{\lambda}\cdot\hat{\xi}_2 = -\tilde{e}_y\cdot\hat{\xi}_2\,, \tag{1.18b}$$

with the *control formula*

$$\hat{\lambda}^T\hat{\lambda}\cdot(1+\hat{\xi}_2^2) = \tilde{e}_y^T\tilde{e}_y + \tilde{e}_x^T\tilde{e}_x = \hat{\nu}\,. \tag{1.19}$$

Here we conclude our derivations on the basis of the EIV-Model (1.2a-c).

2 The Issue of Uniqueness

A long-standing issue surrounding the straight-line fit is the question whether the EIV-Model (1.2a-c) yields the same adjustment result as the EIV-Model (1.3a-b), *i.e.* after a formal exchange of the vectors x and y in their original relationship. We have already argued that such an expectation is unfounded in view of the *nonlinearity* of the TLS estimates. But the *geometric* point of view may suggest a different answer; thus, let us compare the TLS results for both models *analytically* in the *homoscedastic* case.

Given the EIV-Model (1.3a-b), the corresponding system of *nonlinear normal equations* would be formed by exchanging the respective quantities in formula (1.10a), thereby leading to

$$\begin{bmatrix}\hat{\eta}_1\\ \hat{\eta}_2\end{bmatrix} = \begin{bmatrix}\tau^T\tau & \tau^T y\\ y^T\tau & y^T y - \bar{\nu}\end{bmatrix}^{-1}\begin{bmatrix}\tau^T x\\ y^T x\end{bmatrix} \tag{2.1a}$$

with

$$\bar{\nu} = x^T(x - \tau\hat{\eta}_1 - y\hat{\eta}_2) \tag{2.1b}$$

in analogy to (1.10c). Partial elimination of $\hat{\eta}_1$ then generates the "reduced system" solution

$$\begin{aligned}\hat{\eta}_2 &= \left[y^T y - y^T\tau(\tau^T\tau)^{-1}\tau^T y - \bar{\nu}\right]^{-1}\\ &\quad\cdot\left[y^T x - y^T\tau(\tau^T\tau)^{-1}\tau^T x\right]\\ &= (y^T y - n\cdot\bar{y}^2 - \bar{\nu})^{-1}(y^T x - n\cdot\bar{x}\,\bar{y}),\end{aligned} \tag{2.2a}$$

followed by

$$\hat{\eta}_1 = \bar{x} - \bar{y}\cdot\hat{\eta}_2\,, \tag{2.2b}$$

while the coefficient $\bar{\nu}$ fulfills the identity

$$\begin{aligned}\bar{\nu} &= x^T x - x^T\tau\cdot\hat{\eta}_1 - x^T y\cdot\hat{\eta}_2\\ &= (x^T x - n\cdot\bar{x}^2) - (x^T y - n\cdot\bar{x}\,\bar{y})\cdot\hat{\eta}_2.\end{aligned} \tag{2.3}$$

Moreover, from the analogy to (1.15), we may conclude that $\bar{\nu}$ solves the *minimum eigenvalue problem*

$$\begin{bmatrix}y^T y - n\cdot\bar{y}^2 & y^T x - n\cdot\bar{x}\,\bar{y}\\ y^T x - n\cdot\bar{x}\,\bar{y} & x^T x - n\cdot\bar{x}^2\end{bmatrix}\begin{bmatrix}\hat{\eta}_2\\ -1\end{bmatrix} = \begin{bmatrix}\hat{\eta}_2\\ -1\end{bmatrix}\cdot\bar{\nu}_{\min} \tag{2.4}$$

where the matrix shows the same elements as the matrix in (1.15), with only the diagonal elements having changed their position. Consequently, both characteristic polynomial are *identical*:

$$\begin{aligned}(x^T x - n\cdot\bar{x}^2 - \hat{\nu})(y^T y - n\cdot\bar{y}^2 - \hat{\nu}) &= (x^T y - n\cdot\bar{x}\,\bar{y})^2\\ &= (x^T x - n\cdot\bar{x}^2 - \bar{\nu})(y^T y - n\cdot\bar{y}^2 - \bar{\nu})\end{aligned} \tag{2.5a}$$

and generate the *same minimum eigenvalues*, respectively:

$$\bar{\nu}_{\min} \equiv \hat{\nu}_{\min}\,. \tag{2.5b}$$

Hence, when combining (1.14a) with (2.3), we immediately arrive at the identity

$$\begin{aligned}&(x^T y - n\cdot\bar{x}\,\bar{y})^2\cdot\hat{\xi}_2\hat{\eta}_2\\ &= (x^T x - n\cdot\bar{x}^2 - \bar{\nu}_{\min})(y^T y - n\cdot\bar{y}^2 - \hat{\nu}_{\min})\,,\end{aligned} \tag{2.6a}$$

and with (2.5a-b) at

$$\hat{\xi}_2\hat{\eta}_2 \equiv 1\,, \quad\text{or}\quad \hat{\eta}_2 = 1/\hat{\xi}_2\,, \tag{2.6b}$$

which confirms that, in case of the TLS adjustment, the estimate of the inverse slope, $\hat{\eta}_2$, turns out to be inverse to the estimated slope, $\hat{\xi}_2$.

In addition, by combining this result with (1.13) and (2.2b), we obtain

$$\hat{\eta}_1 = \bar{x} - \bar{y}\cdot(1/\hat{\xi}_2) = -(\bar{y} - \bar{x}\cdot\hat{\xi}_2)\cdot(1/\hat{\xi}_2) = -(\hat{\xi}_1/\hat{\xi}_2) \tag{2.6c}$$

for the TLS estimates which is again in total agreement with (1.3b) for the parameters themselves.

Summarizing let us emphasize how instrumental the interpretation of the TLS procedure by means of a *minimum eigenvalue problem* was in order to prove the *identity of* *the adjustment* results in both EIV-Models (1.2a-c) and (1.3a-b) in a rather elegant way.

3 Interludium: A Numerical Example and a Discussion of Reinking's Algorithm

Let us use the same example here as in Schaffrin et al. (2006, pp. 148–153), again assuming equal weights for and no correlations among the observed coordinates of the following four points:

No. i	1	2	3	4
x_i	1.0	1.5	2.0	3.0
y_i	1.0	3.0	0.5	2.5

$$\Rightarrow \bar{x} = (\tau^T x)/n = 1.875\,, \quad \bar{y} = (\tau^T y)/n = 1.75\,. \tag{3.1}$$

Then the system (1.15) turns out as

$$\frac{1}{4}\begin{bmatrix} 8.75 & 3.5 \\ 3.5 & 17.0 \end{bmatrix}\begin{bmatrix} \hat{\xi}_2 \\ -1 \end{bmatrix} = \begin{bmatrix} \hat{\xi}_2 \\ -1 \end{bmatrix} \cdot \hat{\nu}_{\min} \tag{3.2}$$

with its *characteristic polynomial*

$$(8.75 - 4\hat{\nu})(17.0 - 4\hat{\nu}) - (3.5)^2 = 0 \tag{3.3a}$$

which has the roots (i.e., eigenvalues)

$$(\hat{\nu}_{\max} = 4.5712) \quad \text{and} \quad \hat{\nu}_{\min} = 1.8663\,, \tag{3.3b}$$

leading to the *estimated variance component*

$$\hat{\sigma}_o^2 = 1.8663/(4-2) = 0.933 \tag{3.4}$$

from (1.17), to the *estimated slope*

$$\hat{\xi}_2 = (8.75 - 4\hat{\nu}_{\min})^{-1}(3.5) = 2.7242 \tag{3.5a}$$

from (1.12a), and to the *estimated intercept*

$$\hat{\xi}_1 = 1.75 - (1.875)\cdot(2.7242) = -3.358 \tag{3.5b}$$

from (1.13). Finally, the *residual vectors* are computed from (1.18a-b) as follows:

$$\tilde{e}_y = \left[0.19400;\;\; 0.26977;\;\; -0.18890;\;\; -0.27487\right]^T \tag{3.6a}$$

and

$$\tilde{e}_x = -\tilde{e}_y \cdot \hat{\xi}_2 = \left[-0.52846;\;\; -0.73485;\;\; 0.51456;\;\; 0.74881\right]^T. \tag{3.6b}$$

The same results are obtained via the formulas (1.12a-b) through (1.16a-b) after eight cycles during which no matrices are to be inverted; for all the details, we refer the reader to Schaffrin et al. (2006, pp. 150–152).

In contrast, Reinking (2001) suggested to apply "Helmert's knack" (i.e., "Helmerts Kunstgriff", cf. Helmert (1907, p. 286)) to the EIV-Model (1.2a-c), thereby transforming the random error vector e_x into an additional $n \times 1$ parameter vector via

$$-\xi_o := e_x - \underset{\sim}{0}\,, \quad \text{resp.} \quad \underset{\sim}{0} = \xi_o + e_x \tag{3.7a}$$

where the "stochastic zero vector" $\underset{\sim}{0}$ strips the vector e_x of its randomness, due to its characteristics

$$\underset{\sim}{0} \sim (\xi_o, \sigma_o^2 Q_x)\,, \quad C\{\underset{\sim}{0}, e_y\} = 0\,. \tag{3.7b}$$

Consequently, the model equation (1.2a) becomes

$$y - e_y = \tau\cdot(\xi_1^o + \delta\xi_1) + \left[(x - \underset{\sim}{0}) + \xi_o\right]\cdot(\xi_2^o + \delta\xi_2)\,, \tag{3.8a}$$

or

$$\begin{aligned} \delta y &:= y - \tau\cdot\xi_1^o - (x - \underset{\sim}{0})\cdot\xi_2^o \\ &\approx (\xi_2^o \cdot I_n)\xi_o + \tau\cdot(\delta\xi_1) + (x - \underset{\sim}{0})\cdot(\delta\xi_2) + e_y \end{aligned} \tag{3.8b}$$

after neglecting the second order term, $\xi_o \cdot (\delta\xi_2)$. We note that the vector $\underset{\sim}{0}$ also removes the randomness from the vector x because of

$$D\{x - \underset{\sim}{0}\} = D\{(x - e_x) + (e_x - \underset{\sim}{0})\} = D\{E\{x\} - \xi_o\} = 0\,. \tag{3.8c}$$

The *combined observation equations* (3.7a) and (3.8b) have the size $2n \times (n+2)$ and read:

$$\begin{bmatrix} \delta y \\ \underset{\sim}{0} \end{bmatrix} = \begin{bmatrix} \xi_2^o \cdot I_n & \tau & x - \underset{\sim}{0} \\ I_n & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \xi_o \\ \delta\xi_1 \\ \delta\xi_2 \end{bmatrix} + \begin{bmatrix} e_y \\ e_x \end{bmatrix} \tag{3.9a}$$

with the random error characteristics:

$$\begin{bmatrix} e_y \\ e_x \end{bmatrix} \sim \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma_o^2 \begin{bmatrix} Q_y & 0 \\ 0 & Q_x \end{bmatrix}\right). \tag{3.9b}$$

This happens to be a classical *Gauss-Markov Model* where the weighted least-squares approach yields the $(n+2)\times(n+2)$ *normal equations system*:

$$\begin{bmatrix} Q_x^{-1} + (\xi_2^o)\cdot Q_y^{-1} & Q_y^{-1}\tau\cdot\xi_2^o & Q_y^{-1}(x - \underset{\sim}{0})\cdot\xi_2^o \\ \xi_2^o\cdot\tau Q_y^{-1} & \tau^T Q_y^{-1}\tau & \tau^T Q_y^{-1}(x - \underset{\sim}{0}) \\ \xi_2^o\cdot(x - \underset{\sim}{0})^T Q_y^{-1} & (x - \underset{\sim}{0})^T Q_y^{-1}\tau & (x - \underset{\sim}{0})^T Q_y^{-1}(x - \underset{\sim}{0}) \end{bmatrix} \cdot \begin{bmatrix} \hat{\xi}_o \\ \delta\hat{\xi}_1 \\ \delta\hat{\xi}_2 \end{bmatrix}$$

$$= \begin{bmatrix} Q_y^{-1}(\delta y)\cdot\xi_2^o + Q_x^{-1}\cdot\underset{\sim}{0} \\ \tau^T Q_y^{-1}(\delta y) \\ (x - \underset{\sim}{0})^T Q_y^{-1}(\delta y) \end{bmatrix} \tag{3.10}$$

which needs to be solved in each iteration step according to Reinking (2001).

For a large number of points, however, a much better strategy would consist in eliminating the n × 1 vector ξ_0 right from the observation equations (3.9a) and *only then* forming the normal equations. In this way, we are first led to the (equivalent) *Gauss-Helmert Model*

$$w := \delta y - \xi_2^o \cdot \underset{\sim}{0} = [\tau, x - \underset{\sim}{0}] \cdot \begin{bmatrix} \delta\xi_1 \\ \delta\xi_2 \end{bmatrix} + [I_n, \quad \xi_o^2 \cdot I_n] \begin{bmatrix} e_y \\ e_x \end{bmatrix} \quad (3.11)$$

and, afterwards, to the *normal equations*

$$\begin{bmatrix} \tau^T \\ (x - \underset{\sim}{0})^T \end{bmatrix} \left[Q_y + (\xi_2^o)^2 Q_x\right]^{-1} \left([\tau, (x - \underset{\sim}{0})] \begin{bmatrix} \delta\hat{\xi}_1 \\ \delta\hat{\xi}_2 \end{bmatrix} - w\right) = 0 \quad (3.12)$$

which could also be obtained by partially eliminating the vector ξ_0 from the system (3.10) in a much lengthier procedure.

We emphasize the structural similarity of (3.12) with the *nonlinear* normal equations (1.7a-b). For $Q_x = Q_y = P^{-1} := I_n$, the system (3.12) can be simplified further to provide the solution

$$\begin{bmatrix} \delta\hat{\xi}_1 \\ \delta\hat{\xi}_2 \end{bmatrix} = \begin{bmatrix} \tau^T\tau & \tau^T(x - \underset{\sim}{0}) \\ (x - \underset{\sim}{0})^T\tau & (x - \underset{\sim}{0})^T(x - \underset{\sim}{0}) \end{bmatrix}^{-1} \cdot \begin{bmatrix} \tau^T w \\ (x - \underset{\sim}{0})^T w \end{bmatrix} \quad (3.13)$$

which resembles (1.10a), but only furnishes *increments* to update the old linearization point. In the end, after convergence, this *“modified Reinking algorithm”* will indeed generate the correct Total Least-Squares Solution (TLSS), though with the highest costs of all the algorithms presented so far. It may, however, give us a relatively straight-forward approach to the still open question how to compute the variance-covariance matrix, resp. the *mean squared error matrix* of the (nonlinear) TLSS. This issue has to be treated in a separate paper though.

4 A Surprising Connection with Schumann's Estimate of the Slope Angle

Now let us discuss another approach that was actually recommended by Koch (2001) even though it would only generate an optimal estimate of the slope angle, say ϑ, and *not necessarily of the slope*, $\xi_0 = \tan\vartheta$ itself. Its beauty consists of the closed formula that had originally been presented by Schumann (1916); see also Wolf (1968, p. 419), or Koch (1990, p. 104). Here is a brief derivation within the framework of model (1.2a-c) under the assumption $Q_x = Q_y = P^{-1}$ for the weights.

First, from (1.2a), we obtain the *modified observation equations*

$$y \cdot \cos\vartheta - x \cdot \sin\vartheta - \tau \cdot (\xi_1 \cos\vartheta) = e_y \cdot \cos\vartheta - e_x \cdot \sin\vartheta \quad (4.1a)$$

where ϑ and $\xi_1 \cdot \cos\vartheta$ are the *new parameters*. Also, for the random errors we have

$$\begin{bmatrix} e_y \\ e_x \end{bmatrix} \sim \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \sigma_o^2 \begin{bmatrix} P^{-1} & 0 \\ 0 & P^{-1} \end{bmatrix} \right) \quad (4.1b)$$

so that

$$E\{e_y \cdot \cos\vartheta - e_x \cdot \sin\vartheta\} = 0 , \quad (4.1c)$$

and

$$D\{e_y \cdot \cos\vartheta - e_x \cdot \sin\vartheta\} = \sigma_o^2(\cos^2\vartheta + \sin^2\vartheta)P^{-1} = \sigma_o^2 P^{-1} . \quad (4.1d)$$

The weighted least-squares adjustment may thus be based on the *objective function*

$$\begin{aligned} \Phi(\vartheta, \xi_1\cos\vartheta) &:= (e_y^T P e_y) \cdot \cos^2\vartheta - 2(e_x^T P e_y) \cdot \sin\vartheta\cos\vartheta \\ &\quad + (e_x^T P e_x) \cdot \sin^2\vartheta \\ &= (y^T P y) \cdot \cos^2\vartheta - 2(x^T P y) \cdot \sin\vartheta\cos\vartheta \\ &\quad + (x^T P x) \cdot \sin^2\vartheta \\ &\quad - 2\left[(\tau^T P y) \cdot \cos\vartheta - (\tau^T P x) \cdot \sin\vartheta\right] \\ &\quad \cdot (\xi_1\cos\vartheta) + (\tau^T P \tau)(\xi_1\cos\vartheta)^2 = \min . \end{aligned} \quad (4.2)$$

The *necessary Euler-Lagrange conditions* are then easily derived as follows:

$$\begin{aligned} \frac{\partial\Phi}{\partial\vartheta} &= (y^T P y) \cdot (-2\cos\hat{\vartheta} \cdot \sin\hat{\vartheta}) \\ &\quad - 2(x^T P y) \cdot (\cos^2\hat{\vartheta} - \sin^2\hat{\vartheta}) \\ &\quad + (x^T P x) \cdot (2\sin\hat{\vartheta} \cdot \cos\hat{\vartheta}) \\ &\quad + 2\left[(\tau^T P y) \cdot \sin\hat{\vartheta} + (\tau^T P x) \cdot \cos\hat{\vartheta}\right] \cdot (\widehat{\xi_1\cos\vartheta}) \\ &= \left[(x^T P x) - (y^T P y)\right] \cdot \sin(2\hat{\vartheta}) - 2(x^T P y) \cdot \cos(2\hat{\vartheta}) \\ &\quad + 2 \cdot \tau^T P(y \cdot \sin\hat{\vartheta} + x \cdot \cos\hat{\vartheta}) \cdot (\widehat{\xi_1\cos\vartheta}) \doteq 0, \end{aligned} \quad (4.3a)$$

$$\begin{aligned} \frac{\partial\Phi}{\partial(\xi_1\cos\vartheta)} &= -2 \cdot \tau^T P(y \cdot \cos\hat{\vartheta} - x \cdot \sin\hat{\vartheta}) \\ &\quad + 2(\tau^T P \tau) \cdot (\widehat{\xi_1\cos\vartheta}) \doteq 0 . \end{aligned} \quad (4.3b)$$

Now (4.3b) furnishes, on the basis of $\hat{\vartheta}$, at first:

$$\widehat{\xi_1\cos\vartheta} = \left[(\tau^T P y) \cdot \cos\hat{\vartheta} - (\tau^T P x) \cdot \sin\hat{\vartheta}\right] / (\tau^T P \tau) \quad (4.4)$$

which, after inserting it into (4.3a), leads to:

$$2(x^T Py)\cdot\cos(2\hat{\vartheta}) - \left[(x^T Px)-(y^T Py)\right]\cdot\sin(2\hat{\vartheta})$$
$$= 2\left[(\tau^T Py)\cdot\sin\hat{\vartheta} + (\tau^T Px)\cdot\cos\hat{\vartheta}\right]$$
$$\cdot\left[(\tau^T Py)\cdot\cos\hat{\vartheta} - (\tau^T Px)\cdot\sin\hat{\vartheta}\right]/(\tau^T P\tau)$$
$$= 2(\tau^T Px)\cdot(\tau^T Py)\cdot(\cos^2\hat{\vartheta}-\sin^2\hat{\vartheta})/(\tau^T P\tau)$$
$$+\left[(\tau^T Py)^2-(\tau^T Px)^2\right]\cdot(2\sin\hat{\vartheta}\cos\hat{\vartheta})/(\tau^T P\tau)$$
$$= \left[2(\tau^T Px)\cdot(\tau^T Py)\cdot\cos(2\hat{\vartheta}) - \left[(\tau^T Px)^2-(\tau^T Py)^2\right]\sin(2\hat{\vartheta})\right]/(\tau^T P\tau), \tag{4.5}$$

and finally to:

$$\tan(2\hat{\vartheta}) = \frac{2\left[(x^T Py)(\tau^T P\tau)-(\tau^T Px)(\tau^T Py)\right]}{\left[(x^T Px)-(y^T Py)\right](\tau^T P\tau) - \left[(\tau^T Px)^2-(\tau^T Py)^2\right]} \tag{4.6}$$

which is a *weighted form* of the original formula by Schumann (1916); it provides an optimal estimate of the slope, ϑ.

Obviously, in the case of $P := I_n$, formula (4.6) reduces to:

$$\tan(2\hat{\vartheta}) = \frac{2\left[(x^T y)/n - \bar{x}\,\bar{y}\right]}{(x^T x - y^T y)/n - (\bar{x}^2-\bar{y}^2)} = \frac{2(x^T y - n\cdot\bar{x}\,\bar{y})}{(x^T x - y^T y) - n\cdot(\bar{x}^2-\bar{y}^2)} \tag{4.7a}$$

from which $\tan\hat{\vartheta}$ can be gained and inserted into (4.4) to give us the *practical solution*:

$$\hat{\hat{\xi}}_1 = (\widehat{\xi_1 \cos\vartheta})/(\cos\hat{\vartheta}) = \bar{y} - \bar{x}\cdot\tan\hat{\vartheta} \tag{4.7b}$$

which, however, is not automatically the optimal/TLS estimate of the intercept, ξ_1, within the EIV-Model (1.2a-c); only the composite parameter ($\xi_1\cos\vartheta$) has been optimally estimated through (4.4).

If we want to find the TLS estimate for the slope, $\xi_2 = \tan\vartheta$, itself and for the intercept, ξ_1, we could use the same technique as above while taking the derivatives of the objective function (4.2) with respect to $\xi_2 = \tan\vartheta$ and ξ_1 instead. We thus rewrite (4.2) as:

$$\Phi(\xi_2 = \tan\vartheta, \xi_1)$$
$$:= (e_y\cdot\cos\vartheta - e_x\sin\vartheta)^T P(e_y\cdot\cos\vartheta - e_x\sin\vartheta)$$
$$= (1+\tan^2\vartheta)^{-1}$$
$$\cdot\left[(y^T Py)-(x^T Py)\cdot 2\tan\vartheta + (x^T Px)\cdot\tan^2\vartheta\right]$$
$$-2\left[(\tau^T Py)-(\tau^T Px)\cdot\tan\vartheta\right](1+\tan^2\vartheta)^{-1}\cdot\xi_1$$
$$+(\tau^T P\tau)\cdot(1+\tan^2\vartheta)^{-1}\cdot\xi_1^2 = \min. \tag{4.8}$$

Now the *necessary Euler-Lagrange conditions* turn out as:

$$(1+\tan^2\vartheta)^2\cdot\frac{\partial\Phi}{\partial(\tan\vartheta)}$$
$$= (y^T Py)\cdot(-2\widehat{\tan\vartheta}) - (x^T Py)\cdot 2(1-\widehat{\tan^2\vartheta})$$
$$+(x^T Px)\cdot(2\widehat{\tan\vartheta})$$
$$+(\tau^T Px)\cdot\hat{\xi}_1\cdot 2(1-\widehat{\tan^2\vartheta})$$
$$+\left[(\tau^T P\tau)\cdot\hat{\xi}_1^2 - 2(\tau^T Py)\cdot\hat{\xi}_1\right]\cdot(-2\widehat{\tan\vartheta}) \doteq 0, \tag{4.9a}$$

and

$$(1+\tan^2\vartheta)\cdot\frac{\partial\Phi}{\partial\xi_1} = 2(\tau^T P\tau)\hat{\xi}_1 - 2\tau^T P(y - x\cdot\widehat{\tan\vartheta}) \doteq 0. \tag{4.9b}$$

From (4.9b), it follows immediately that

$$\hat{\xi}_1 = (\tau^T P\tau)^{-1}\cdot\tau^T P(y - x\cdot\widehat{\tan\vartheta}) \tag{4.10}$$

which would coincide with (4.4) if only the same estimates for ϑ had to be applied. Here, however, $\hat{\xi}_2 = \widehat{\tan\vartheta}$ comes from (4.9a) after implementing (4.10) via

$$-(y^T Py)\cdot\hat{\xi}_2 - (x^T Py)\cdot(1-\hat{\xi}_2^2) + (x^T Px)\cdot\hat{\xi}_2$$
$$-\hat{\xi}_1\cdot\hat{\xi}_2\cdot\left[(\tau^T P\tau)\cdot\hat{\xi}_1 - \tau^T P(y - x\cdot\hat{\xi}_2)\right]$$
$$+\hat{\xi}_i\cdot\left[(\tau^T Py)\cdot\hat{\xi}_2 + (\tau^T Px)\right] = 0, \tag{4.11a}$$

$$(x^T Py) + \left[(y^T Py)-(x^T Px)\right]\cdot\hat{\xi}_2 - (x^T Py)\cdot\hat{\xi}_2^2$$
$$= (\tau^T P\tau)^{-1}\cdot\tau^T P(y - x\cdot\hat{\xi}_2)\cdot\tau^T P(y\cdot\hat{\xi}_2 + x) - \hat{\xi}_1\cdot\hat{\xi}_2\cdot 0$$
$$= (\tau^T P\tau)^{-1}\left[(\tau^T Py)(\tau^T Px) + \left[(\tau^T Py)^2-(\tau^T Px)^2\right]\cdot\hat{\xi}_2\right.$$
$$\left. -(\tau^T Px)(\tau^T Py)\cdot\hat{\xi}_2^2\right], \tag{4.11b}$$

$$\left[(x^T Py)(\tau^T P\tau)-(\tau^T Px)(\tau^T Py)\right]\cdot\hat{\xi}_2^2$$
$$-(\left[(y^T Py)-(x^T Px)\right](\tau^T P\tau) - \left[(\tau^T Py)^2-(\tau^T Px)^2\right])\cdot\hat{\xi}_2$$
$$-\left[(x^T Py)(\tau^T P\tau)-(\tau^T Px)(\tau^T Py)\right] = 0 \tag{4.11c}$$

which represents a *quadratic equation* for the optimal slope, $\hat{\xi}_2$. Its solutions are apparently obtained either directly, or by employing the relationship (4.6) through

$$\hat{\xi}_2^2 + 2\hat{\xi}_2/(\tan 2\hat{\vartheta}) - 1 = 0 \tag{4.12}$$

which leads to

$$(\widehat{\tan\vartheta})_{1/2} = \frac{-1\pm\sqrt{1+(\tan 2\hat{\vartheta})^2}}{\tan(2\hat{\vartheta})}. \tag{4.13}$$

Interestingly enough, the optimal/TLS slope can be found in closed form from the optimal/TLS slope angle by applying the quite elementary formula (4.13), – another form of the uniqueness argument.

If we now return to the numerical example from the previous Section 3 where $P = I_n$ holds true, formula (4.7a) applies and provides the *optimal slope angle* through

$$\tan(2\hat{\vartheta}) = \frac{2\cdot(3.5-3.28125)}{-0.0625-0.453125} = -0.848485 \tag{4.14a}$$

as

$$2\hat{\vartheta} = 139.6859^\circ\ , \quad \text{resp.} \quad \hat{\vartheta} = 69.8430^\circ \tag{4.14b}$$

from which we obtain the possibly *"non-optimal slope"* as

$$\tan\hat{\vartheta} = 2.72422 \tag{4.14c}$$

and the possibly *"non-optimal intercept"* from (4.7b) as

$$\hat{\hat{\xi}}_1 = 1.75-(1.875)\cdot\tan\hat{\vartheta} = -3.3579\ . \tag{4.14d}$$

Alternatively, formula (4.13) gives us the two solutions

$$\begin{aligned}(\widehat{\tan\vartheta})_{1/2} &= (-1\pm\sqrt{1.72})/(-0.848485)\\ &\in\{2.72422;\quad(-0.36708)\}\end{aligned} \tag{4.15a}$$

of which the bigger one *coincides* indeed with the previous solution (3.5a) and, therefore, represents the *optimal/TLS slope*. Consequently, formula (4.10) provides the *optimal/TLS intercept* again, namely

$$\hat{\xi}_1 = \bar{y}-\bar{x}\cdot\hat{\xi}_2 = \bar{y}-\bar{x}\cdot(\widehat{\tan\vartheta}) = -3.358\ , \tag{4.15b}$$

exactly as in (3.5b). Thus, the residual vectors $\tilde{e}_x$ and $\tilde{e}_y$ will automatically result in (3.6a-b) as well.

5 The Case of Heteroscedastic Weights

Although the algorithm (1.7a-b) with (1.8a-c) works for arbitrary cofactor matrices Q_x and Q_y, let us first concentrate on the *heteroscedastic case* where we have

$$P = Q_x^{-1} = Q_y^{-1} = Diag(p_1,\cdots,p_n) \tag{5.1}$$

and the simplified algorithm (1.9a-b) could be applied. Alternatively, we may use (4.13) in connection with the weighted form (4.6) of *"Schumann's formula"*, or the *modified minimum eigenvalue problem* (1.11).

For the numerical example from Section 3 we may for instance define,

$$p_1 = p_4 := 4 \quad \text{and} \quad p_2 = p_3 := 1\ , \tag{5.2}$$

from which we obtain (for $\tau^T P\tau = 10$)

$$\bar{\bar{x}} := (\tau^T Px)/(\tau^T P\tau) = 1.95 \tag{5.3a}$$

and

$$\bar{\bar{y}} = (\tau^T Py)/(\tau^T P\tau) = 1.75\ , \tag{5.3b}$$

as well as

$$x^T Px = 46.25\ ,\quad x^T Py = 39.5\ ,\quad y^T Py = 38.25\ . \tag{5.4}$$

The analogous system to (1.15) thus turns out as:

$$\frac{1}{10}\begin{bmatrix}82.25 & 53.75\\ 53.75 & 76.25\end{bmatrix}\begin{bmatrix}\hat{\hat{\xi}}\\ -1\end{bmatrix} = \begin{bmatrix}\hat{\hat{\xi}}_2\\ -1\end{bmatrix}\cdot\hat{\hat{\nu}}_{\min} \tag{5.5}$$

with its *characteristic polynomial*

$$(82.25-10\hat{\hat{\nu}})(76.25-10\hat{\hat{\nu}})-(53.75)^2 = 0 \tag{5.6a}$$

whose roots are

$$(\hat{\hat{\nu}}_{\max} = 13.3084) \quad \text{and} \quad \hat{\hat{\nu}}_{\min} = 2.5416 \tag{5.6b}$$

Consequently, after inserting $\hat{\hat{\nu}}_{\min}$ into (1.10), we obtain the new *weighted TLS solution*

$$\begin{bmatrix}\hat{\hat{\xi}}\\ \hat{\hat{\xi}}\end{bmatrix} = \begin{bmatrix}10 & 19.5\\ 19.5 & 43.7084\end{bmatrix}^{-1}\begin{bmatrix}17.5\\ 39.5\end{bmatrix} = \begin{bmatrix}-0.0942\\ 0.94574\end{bmatrix} \tag{5.7}$$

for intercept and slope which appear considerably different from the homoscedastic results in (3.5a-b).

A check by formula (4.6) will first give us

$$\tan(2\hat{\hat{\vartheta}}) = \frac{2\cdot(395-341.25)}{(80-74)} = 17.9167\ , \tag{5.8}$$

and from (4.13)

$$(\widehat{\tan\vartheta})_{1/2} \in\{0.94574;\ (-1.05737)\}\ , \tag{5.9a}$$

before arriving at

$$\hat{\hat{\xi}}_1 = -0.09419 \tag{5.9b}$$

from (4.10). These results nicely coincide with (5.7) while the new *estimated variance component* becomes

$$\hat{\hat{\sigma}}_o^2 = 2.5416/(4-2) = 1.271\ . \tag{5.10}$$

We skip the computation of the residual vectors $\tilde{\tilde{e}}_x$ and $\tilde{\tilde{e}}_y$, in accordance with formulas (1.8a-c), at this point in order to draw the reader's attention to another application of the system (1.11). Since the setting of $Q_o := 0$ namely, indicates the case of *nonrandom x-coordinates* we easily obtain the *weighted Least-Squares Solution* (LESS) by solving

$$N\hat{\hat{\xi}}_{LESS} - c = 0 \quad (5.10a)$$

and setting

$$\hat{\hat{v}}_{LESS} = y^T P(y - A\hat{\hat{\xi}}_{LESS}) \,. \quad (5.10b)$$

This leads us numerically to

$$\begin{bmatrix} \hat{\hat{\xi}}_1 \\ \hat{\hat{\xi}}_2 \end{bmatrix}_{LESS} = \begin{bmatrix} 10 & 19.5 \\ 19.5 & 46.25 \end{bmatrix}^{-1} \begin{bmatrix} 17.5 \\ 39.5 \end{bmatrix} = \begin{bmatrix} 0.4757 \\ 0.6535 \end{bmatrix} \quad (5.11)$$

in the *heteroscedatic case*, with

$$\hat{\hat{v}}_{LESS} = 8.869 = 2(\hat{\hat{\sigma}}_o^2)_{LESS} \,. \quad (5.12)$$

In contrast, for the *homoscedastic case* ($P := I_4$), we know from Schaffrin et al. (2006) that

$$\begin{bmatrix} \hat{\xi}_1 \\ \hat{\xi}_2 \end{bmatrix}_{LESS} = \begin{bmatrix} 1.000 \\ 0.400 \end{bmatrix} \quad (5.13)$$

and

$$\hat{v}_{LESS} = 3.900 = 2(\hat{\sigma}_o^2)_{LESS} \,. \quad (5.14)$$

As a main result, we can see that the TLS adjustment provides always a *much better* fit than the standard Least-Squares adjustment (i.e., $Q_o := 0$). This is true for both, the *homoscedastic* and *heteroscedastic* case.

Moreover, although we generally lost somewhat of the quality of fit in the *heteroscedastic* case, it is interesting to note that the weighted TLS adjustment here turns out to be still *superior* to the (unweighted) Least-Squares adjustment, also known as Ordinary Least-Squares (OLS) adjustment. Whether this is a general feature, needs further investigations though.

6 Conclusions and Outlook

We have thoroughly studied the case of a two-dimensional straight-line fit within an Error-in-Variables (EIV) Model under a least-squares approach, here leading to the so-called *Total Least-Squares (TLS) adjustment* that enables us to accommodate random x- *and* y-coordinates, as well as *arbitrary cofactor matrices* for them. Incidentally, the case when only one coordinate is considered to be random, is also covered by specialization, and so is the homoscedastic (*i.e.*, "unweighted") case.

While offering an *iterative algorithm of Newton type* that improves "Reinking's algorithm" considerably in terms of numerical efficiency, we yet recommend the solution of a (generalized) *minimum eigenvalue problem* for this particular application, unless we resort to the *weighted analog* to "Schumann's formula" that provides both, the TLS estimate of the slope angle and the basis for a closed solution of the optimal/TLS estimate of the slope itself.

A rather simple *numerical example* finally allows to provide a first judgment on the behavior of the alternative algorithms although, in the future, further investigations appear necessary into the statistical properties of the *nonlinear* (!) TLS estimates. These include practical formulas for the *variances* and *covariances* of any TLS results and, perhaps, methods for hypothesis testing in this context. Moreover, the case of correlated x- and y-coordinate also deserves more attention.

Acknowledgement

This contribution was finalized during the author's visit to the Geodetic Institute of the University of Karlsruhe (TH) as a Mercator Professor (DFG), with Prof. Günter Schmitt as his host. This is most gratefully acknowledged.

I am also most indebted to my collobarators Yaron Felus (Ferris State University) and Andreas Wieser (Graz University of Technology) who helped to test some of the algorithms presented here.

References

Golub, G.H., and Ch. van Loan (1980): An analysis of the Total Least-Squares problem, SIAM J. Numer. Anal. 17 (1980), 883–893.

Golub, G.H., and Ch. van Loan (1996): Matrix Computations, 3rd edition, Johns Hopkins University Press: Baltimore/MD.

Grafarend, E., and B. Schaffrin (1993): Adjustment Computations in Linear Models (in German), Bibliograph. Inst.: Mannheim, Germany.

Helmert, F.R. (1907): Least-Squares Adjustment Computations (in German), 2nd edition, Teubner: Leipzig/Berlin.

Höpcke, W. (1981): Error Analysis and Adjustment Computations (in German), de Gruyter: Berlin/New York.

Huffel, S. van, and J. Vandewalle (1991): The Total Least-Squares Problem. Computational Aspects and Analysis, SIAM: Philadelphia/PA.

Koch, K.R. (1990): Bayesian Inference with Geodetic Applications, Springer Lect. Notes in Earth Science, vol.31, Berlin/Heidelberg/New York

Koch, K.R. (2001): Remarks on the publication "On the determination of unique transformation parameters" (in German), Z. für Vermessungswesen 126 (2001), 297.

Koch, K.R. (2004): Remarks on the publication "Rigorous adjustment of the nonlinear Gauss-Helmert Model" (in German), Allgem. Verm.-Nachr. 111 (2004), 276–277.

Kupferer, S. (2004): On the correct linearization of nonlinear Gauss-Helmert Models (in German), Allgem. Verm.-Nachr. 111 (2004), 394–396.

Lenzmann, E., and L. Lenzmann (2001): On the determination of unique transformation parameters (in German), Z. für Vermessungswesen 126 (2001), 138–142.

Lenzmann, L., and E. Lenzmann (2004a): Rigorous adjustment of the nonlinear Gauss-Helmert Model (in German) Allgem. Verm.-Nachr. 111 (2004), 68–73.

Lenzmann, L., and E. Lenzmann (2004b): Response to the remarks by K.R. Koch (in German), Allgem. Verm.-Nachr. 111(2004), 277–278.

Lenzmann, L., and E. Lenzmann (2005): Response to the contribution "On the correct linearization of nonlinear Gauss-Helmert Models" (in German), Allgem. Verm.-Nachr. 112 (2005), 114.

Lenzmann, L., and E. Lenzmann (2007): On the solution of nonlinear Gauss-Markov Models (in German), Z. für Vermessungswesen 132 (2007), 108–110.

Mikhail, E.M., and G. Gracie (1981): Analysis and Adjustment of Survey Measurements, van Nostrand and Reinhold: New York.

Pope, J.A. (1972): Some pitfalls to be avoided in the iterative adjustment of nonlinear problems; in: Proc. of the 38th ASPRS Meeting: Washington/DC, 449–477.

Reinking, J. (2001): Comment on "On the determination of unique transformation parameters" (in German), Z. für Vermessungswesen 126 (2001), 295–296.

Schaffrin, B. (2001): Equivalent systems for various forms of Kriging, including least-squares collocation, Z. für Vermessungswesen 126 (2001), 87–94.

Schaffrin, B., I.P. Lee, Y. Felus and Y.S. Choi (2006): Total Least-Squares for geodetic straight-line and plane adjustment, Boll. Geod. Sci. Aff. 65 (2006), 141–168.

Schaffrin, B., and A. Wieser (2007): On weighted Total Least-Squares adjustment for linear regression, J. of Geodesy, accepted for publication.

Schmid, D.E., and H.H. Schmid (1971): Curve and surface fitting with potential functions, Z. für Vermessungswesen 96 (1971), 488–497.

Schumann, R. (1916): Determination of a straight line by least-squares adjustment of its points' observed coordinates (in German), Session Reports of the K.u.K. Acad. of Sciences (Math. Nat. Class, Abt. IIa), Vol. 125, Vienna 1916, 1429 ff.

Wolf, H. (1968): Adjustment Computations by the Methods of Least-Squares (in German), Dümmler: Bonn, Germany.

Wolf, H. (1975): Adjustment Computations. Formulas for Practical Applications (in German), Dümmler: Bonn, Germany.

Wolf, P.R., and C.D. Ghilani (1997): Adjustment Computations. Statistics and Least-Squares in Surveying and GIS, Wiley: New York.

Author's address
Prof. Dr.-Ing. habil. Burkhard Schaffrin
School of Earth Sciences
The Ohio State University
125 South Oval Mall
Columbus/OH 43210, USA
schaffrin.1@osu.edu