On Weighted Total Least-Squares Adjustment with Multiple Constraints and Singular Dispersion Matrices

Shahram Jazaeri, Burkhard Schaffrin and Kyle Snow

Summary

The Total Least-Squares (TLS) adjustment with multiple constraints (including a quadratic constraint, e. g.) has seen increased attention in Geodetic Science over the last five years. We only refer to the contributions by Schaffrin and Felus (2009) and by Fang (2013) who both provided different algorithms for such cases as long as the dispersion matrices involved are proportional to identity matrices (homoscedastic case), or happen to be positive-definite (weighted, resp. structured case). Here, a new algorithm is presented that tolerates positive-semidefinite dispersion matrices as well, provided that a certain rank condition is fulfilled to guarantee the uniqueness of the constrained Weighted TLS solution.

Zusammenfassung

Die Ausgleichung nach der Total Least-Squares (TLS) Methode unter gleichzeitiger Berücksichtigung von mehreren Restriktionen (darunter z. B. einer quadratischen) ist in jüngster Vergangenheit verschiedentlich untersucht worden, besonders zum Gebrauch in der Geodäsie. Dazu nennen wir bloß die Beiträge von Schaffrin und Felus (2009) sowie von Fang (2013), die jeweils eigene Algorithmen präsentiert haben, zunächst für den homoskedastischen Fall, wo die Kovarianz-Matrizen proportional zur Einheitsmatrix sind, und dann für den regulär gewichteten Fall, wo diese Matrizen positiv-definit sind. Hier wird ein neuer Algorithmus vorgestellt, der auch positiv-semidefinite Kovarianz-Matrizen zulässt, vorausgesetzt, dass eine bestimmte Rang-Bedingung erfüllt ist, die die Eindeutigkeit der restringierten und passend gewichteten TLS-Lösung garantiert; dabei existieren die Gewichtsmatrizen im traditionellen Sinne natürlich nicht mehr.

Keywords: Errors-In-Variables Model, Weighted Total Least-Squares, Multiple Constraints, Singular Dispersion Matrices, Uniqueness Condition

1 Introduction

In Geodetic Science, the Total Least-Squares (TLS) approach to adjust so-called Errors-In-Variables (EIV) Models *without linearization* has been well established over the last decade. Here, we only refer to the contribution by Felus and Schaffrin (2002, 2005), Schaffrin et al. (2006), and Schaffrin (2007), all of which benefited generally from the seminal paper by Golub and Van Loan (1980), and the subsequent summary by Van Huffel and Vandewalle (1991). Linear constraints, both fixed and stochastic, were later added by Schaffrin and Felus (2005) and Schaffrin (2006) before the case of multiple constraints was treated by Schaffrin and Felus (2009), allowing both linear(ized) constraints and one quadratic constraint.

In all contributions so far, the *homoscedastic case* was assumed, were the dispersion matrices for the observation vector and the data matrix are proportional to identity matrices, with zero cross-covariances. Even the so-called Generalized TLS approach, stressed by Van Huffel and Vandewalle (1991), does not take more general weight matrices into account, but is more concerned with balanced estimation results.

Pioneering the use of *diagonal weight matrices* were Markovsky et al. (2006) who, in their terminology, solved the "elementwise weighted TLS problem," surmising that this might be the most general form of weighting that still allows a "closed solution." This statement, however, soon became obsolete when Schaffrin and Wieser (2008) presented closed formulas and a weighted TLS regression algorithm for fairly general positive-definite weight matrices, assuming only a certain Kronecker product structure for the dispersion matrix of the vectorized data matrix. More recently, Fang (2011), Mahboub (2012), and Schaffrin et al. (2012) were indeed able to drop the Kronecker product requirement altogether, while Fang (2011) and Snow (2012) even designed algorithms for a weighted TLS adjustment with cross-correlated observation vector and data matrix.

So far, however, the dispersion matrices and consequently the weight matrices had to be *positive-definite* so that one can be chosen as the inverse of the other. On the other hand, Neitzel and Schaffrin (2013) had already found a rank condition for the Gauss-Helmert Model (GH-Model) that guarantees a unique LEast-Squares Solution (LESS) even for a *singular dispersion matrix*, i. e., without the need of an inverse weight matrix. Knowing that the EIV-Model is a special case of a nonlinear GH-Model, as shown by Schaffrin and Snow (2010) for instance, this rank criterion was adopted by Snow (2012) and Schaffrin et al. (2014) in order to find several algorithms that would tolerate even *positive-semidefinite dispersion matrices* in a (properly weighted) TLS adjustment. Since the TLS approach by Schaffrin and Felus (2009) to the *EIV-Model with multiple constraints* was recently extended to the case of *positive-definite* dispersion, resp. weight matrices by Fang (2014), in this paper an attempt will be made to further extend the existing solutions for the case of *singular*, i. e., positive-semidefinite dispersion matrices in the presence of *multiple constraints*, provided that the above-mentioned rank condition (in generalized form) is met.

DOI 10.12902/zfv-0017-2014

After a comprehensive review of the various TLS algorithms to treat the unconstrained EIV-Model with singular dispersion matrices in chapter 2, the TLS approach by Fang (2014) to the case of multiple constraints will be reviewed in chapter 3, and eventually extended to the situation with singular dispersion matrices in chapter 4. Then the new algorithm will be tested in chapter 5 with the examples from Schaffrin and Felus (2009), Fang (2013), Schaffrin et al. (2014), as well as two examples that could not have been solved by any of the existing TLS algorithms. Finally, some conclusions will be drawn.

2 A comprehensive review of the TLS approach to the unconstrained EIV-Model with singular dispersion matrices

Let the EIV-Model be defined, as in Snow (2012), by

$$\boldsymbol{y} = (\underset{n\times m}{\boldsymbol{A}} - \boldsymbol{E}_A)\boldsymbol{\xi} + \boldsymbol{e}_y, \quad \operatorname{rk}\boldsymbol{A} = m, \tag{1a}$$

$$\begin{bmatrix}\boldsymbol{e}_y\\ \boldsymbol{e}_A\end{bmatrix} := \begin{bmatrix}\boldsymbol{e}_y\\ \operatorname{vec}\boldsymbol{E}_A\end{bmatrix} \sim (\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{0}\end{bmatrix}, \sigma_0^2\boldsymbol{Q} := \sigma_0^2\begin{bmatrix}\boldsymbol{Q}_y & \boldsymbol{Q}_{yA}\\ \boldsymbol{Q}_{yA}^T & \boldsymbol{Q}_A\end{bmatrix}), \tag{1b}$$

where

- $\boldsymbol{y}$ denotes the $n\times 1$ observation vector,
- $\boldsymbol{\xi}$ the $m\times 1$ (unknown) parameter vector,
- $\boldsymbol{A}$ the $n\times m$ data matrix with $n > m = \operatorname{rk}\boldsymbol{A}$,
- $\boldsymbol{e}_y$ the $n\times 1$ (unknown) random error vector associated with $\boldsymbol{y}$,
- $\boldsymbol{E}_A$ the $n\times m$ (unknown) random error matrix associated with $\boldsymbol{A}$,
- $\boldsymbol{e}_A$ the $nm\times 1$ vectorized form of $\boldsymbol{E}_A$,
- σ_0^2 the (unknown) variance component, and
- $\boldsymbol{Q}$ the $n(m+1)\times n(m+1)$ symmetric non-negative-definite cofactor matrix with $\operatorname{rk}\boldsymbol{Q} \le n(m+1)$.

In case that the cofactor matrix $\boldsymbol{Q}$ is *non-singular*, a unique weight matrix can be defined as its inverse:

$$\boldsymbol{P} := \boldsymbol{Q}^{-1} = \begin{bmatrix}\boldsymbol{P}_{11} & \boldsymbol{P}_{12}\\ \boldsymbol{P}_{12}^T & \boldsymbol{P}_{22}\end{bmatrix} \text{ if } \operatorname{rk}\boldsymbol{Q} = n(m+1), \tag{2}$$

which leads to the well-defined *Weighted Total Least-Squares (WTLS)* objective function

$$\begin{bmatrix}\boldsymbol{e}_y^T, & \boldsymbol{e}_A^T\end{bmatrix}\begin{bmatrix}\boldsymbol{P}_{11} & \boldsymbol{P}_{12}\\ \boldsymbol{P}_{12}^T & \boldsymbol{P}_{22}\end{bmatrix}\begin{bmatrix}\boldsymbol{e}_y\\ \boldsymbol{e}_A\end{bmatrix} = \min. \quad \text{s.t. } \boldsymbol{y} - (\boldsymbol{A}-\boldsymbol{E}_A)\boldsymbol{\xi} - \boldsymbol{e}_y = \boldsymbol{0}. \tag{3}$$

Based on the equivalent *Lagrange target function*

$$\Phi(\boldsymbol{e}_y, \boldsymbol{e}_A, \boldsymbol{\xi}, \boldsymbol{\lambda}) = \boldsymbol{e}_y^T\boldsymbol{P}_{11}\boldsymbol{e}_y + 2\boldsymbol{e}_y^T\boldsymbol{P}_{12}\boldsymbol{e}_A + \boldsymbol{e}_A^T\boldsymbol{P}_{22}\boldsymbol{e}_A + \\ +2\boldsymbol{\lambda}^T(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{\xi} - \boldsymbol{e}_y + (\boldsymbol{\xi}^T\otimes \boldsymbol{I}_n)\boldsymbol{e}_A) = \text{stationary}, \tag{4}$$

where $\boldsymbol{\lambda}$ denotes the $n\times 1$ vector of Lagrange multipliers, and $\otimes$ denotes the Kronecker-Zehfuss product of matrices, defined by $\boldsymbol{G}\otimes\boldsymbol{H} := [g_{ij}\boldsymbol{H}]$ if $\boldsymbol{G} = [g_{ij}]$. The Euler-Lagrange necessary conditions (or *nonlinear normal equations*) are obtained as

$$\frac{1}{2}\frac{\partial\Phi}{\partial\boldsymbol{e}_y} = \boldsymbol{P}_{11}\tilde{\boldsymbol{e}}_y + \boldsymbol{P}_{12}\tilde{\boldsymbol{e}}_A - \hat{\boldsymbol{\lambda}} \doteq \boldsymbol{0}, \tag{5a}$$

$$\frac{1}{2}\frac{\partial\Phi}{\partial\boldsymbol{e}_A} = \boldsymbol{P}_{12}^T\tilde{\boldsymbol{e}}_y + \boldsymbol{P}_{22}\tilde{\boldsymbol{e}}_A + (\hat{\boldsymbol{\xi}}\otimes\boldsymbol{I}_n)\hat{\boldsymbol{\lambda}} \doteq \boldsymbol{0}, \tag{5b}$$

$$\frac{1}{2}\frac{\partial\Phi}{\partial\boldsymbol{\xi}} = -\boldsymbol{A}^T\hat{\boldsymbol{\lambda}} + \tilde{\boldsymbol{E}}_A^T\hat{\boldsymbol{\lambda}} \doteq \boldsymbol{0}, \tag{5c}$$

$$\frac{1}{2}\frac{\partial\Phi}{\partial\boldsymbol{\lambda}} = \boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{\xi}} - \tilde{\boldsymbol{e}}_y + (\hat{\boldsymbol{\xi}}^T\otimes\boldsymbol{I}_n)\tilde{\boldsymbol{e}}_A \doteq \boldsymbol{0}. \tag{5d}$$

The corresponding sufficient condition is fulfilled since

$$\frac{1}{2}\frac{\partial^2\Phi}{\partial\begin{bmatrix}\boldsymbol{e}_y\\ \boldsymbol{e}_A\end{bmatrix}\partial\begin{bmatrix}\boldsymbol{e}_y^T \mid \boldsymbol{e}_A^T\end{bmatrix}} = \begin{bmatrix}\boldsymbol{P}_{11} & \boldsymbol{P}_{12}\\ \boldsymbol{P}_{12}^T & \boldsymbol{P}_{22}\end{bmatrix} = \boldsymbol{Q}^{-1} \tag{6}$$

is positive-definite, under (2), which guarantees the minimum in (3).

Now, the *residual vectors* $\tilde{\boldsymbol{e}}_y$ and $\tilde{\boldsymbol{e}}_A$ are directly derived from (5a) and (5b) in terms of $\hat{\boldsymbol{\lambda}}$ via

$$\begin{bmatrix}\tilde{\boldsymbol{e}}_y\\ \tilde{\boldsymbol{e}}_A\end{bmatrix} = \begin{bmatrix}\boldsymbol{P}_{11} & \boldsymbol{P}_{12}\\ \boldsymbol{P}_{12}^T & \boldsymbol{P}_{22}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{I}_n\\ -(\hat{\boldsymbol{\xi}}\otimes\boldsymbol{I}_n)\end{bmatrix}\hat{\boldsymbol{\lambda}} = \\ = \begin{bmatrix}[\boldsymbol{Q}_y - \boldsymbol{Q}_{yA}(\hat{\boldsymbol{\xi}}\otimes\boldsymbol{I}_n)]\hat{\boldsymbol{\lambda}}\\ [\boldsymbol{Q}_{yA}^T - \boldsymbol{Q}_A(\hat{\boldsymbol{\xi}}\otimes\boldsymbol{I}_n)]\hat{\boldsymbol{\lambda}}\end{bmatrix}. \tag{7}$$

Using (7) in (5d) leads to

$$\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{\xi}} = [\boldsymbol{Q}_y - \boldsymbol{Q}_{yA}(\hat{\boldsymbol{\xi}}\otimes\boldsymbol{I}_n)]\hat{\boldsymbol{\lambda}} - \\ - (\hat{\boldsymbol{\xi}}\otimes\boldsymbol{I}_n)^T[\boldsymbol{Q}_{yA}^T - \boldsymbol{Q}_A(\hat{\boldsymbol{\xi}}\otimes\boldsymbol{I}_n)]\hat{\boldsymbol{\lambda}} = \boldsymbol{Q}_1\cdot\hat{\boldsymbol{\lambda}} \tag{8a}$$

with

$$\boldsymbol{Q}_1 := [\boldsymbol{Q}_y - \boldsymbol{Q}_{yA}(\hat{\boldsymbol{\xi}}\otimes\boldsymbol{I}_n) - (\hat{\boldsymbol{\xi}}\otimes\boldsymbol{I}_n)^T\boldsymbol{Q}_{yA}^T + \\ + (\hat{\boldsymbol{\xi}}\otimes\boldsymbol{I}_n)^T\boldsymbol{Q}_A(\hat{\boldsymbol{\xi}}\otimes\boldsymbol{I}_n)] = \boldsymbol{Q}_1(\hat{\boldsymbol{\xi}}), \tag{8b}$$

or

$$\boldsymbol{Q}_1 = [\boldsymbol{I}_n \mid -(\hat{\boldsymbol{\xi}}\otimes\boldsymbol{I}_n)^T]\begin{bmatrix}\boldsymbol{Q}_y & \boldsymbol{Q}_{yA}\\ \boldsymbol{Q}_{yA}^T & \boldsymbol{Q}_A\end{bmatrix}\begin{bmatrix}\boldsymbol{I}_n\\ -(\hat{\boldsymbol{\xi}}\otimes\boldsymbol{I}_n)\end{bmatrix} =: \hat{\boldsymbol{B}}\boldsymbol{Q}\hat{\boldsymbol{B}}^T \tag{8c}$$

for

$$\hat{\boldsymbol{B}} := [\boldsymbol{I}_n \mid -(\hat{\boldsymbol{\xi}}\otimes\boldsymbol{I}_n)^T] = \boldsymbol{B}(\hat{\boldsymbol{\xi}}). \tag{8d}$$

It is worth noting that $\boldsymbol{Q}_1$ will be *nonsingular* for a positive-definite matrix $\boldsymbol{Q}$, which may even be true for *some* positive-semidefinite matrices $\boldsymbol{Q}$. In this case, (8a) can be solved for

$$\boxed{\hat{\boldsymbol{\lambda}} = \boldsymbol{Q}_1^{-1}(\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{\xi}})}, \tag{9}$$

which provides the estimated Lagrange multipliers – and thus both residual vectors – as a function of the estimated parameters. Further exploiting (5c), in connection with (7) and (9), now yields

$$A^T Q_1^{-1}(y - A\hat{\xi}) = A^T\hat{\lambda} = \tilde{E}_A^T\hat{\lambda} = (\hat{\lambda}^T \otimes I_m)\,\mathrm{vec}(\tilde{E}_A^T) =$$
$$= (\hat{\lambda}^T \otimes I_m)K_{nm}\tilde{e}_A = (I_m \otimes \hat{\lambda})^T\tilde{e}_A = \tag{10a}$$
$$= (I_m \otimes \hat{\lambda})^T[Q_{yA}^T - Q_A(\hat{\xi} \otimes I_n)]\,Q_1^{-1}(y - A\hat{\xi}) =$$
$$= -R_1 \cdot (y - A\hat{\xi}) \tag{10b}$$

with

$$R_1 := (I_m \otimes \hat{\lambda})^T[-Q_{yA}^T Q_1^{-1} + Q_A(\hat{\xi} \otimes Q_1^{-1})] =$$
$$= R_1(\hat{\xi}, \hat{\lambda}),\ (R_1 Q_1 + \tilde{E}_A^T)\hat{\lambda} = 0. \tag{10c}$$

Here, K_{nm} denotes an $nm \times nm$ commutation (or vec-permutation) matrix; for more details, see Magnus and Neudecker (2007).

Finally, (10b) can be converted into the so-called *"generalized normal equations"* of Schaffrin et al. (2012), namely

$$\boxed{(A^T Q_1^{-1} + R_1)A \cdot \hat{\xi} = (A^T Q_1^{-1} + R_1)y}\,, \tag{11}$$

which provide the estimated parameters as a function of the estimated Lagrange multipliers (9) that enter the definition for R_1 in (10c). This leads to the iterative *Algorithm 1 of Mahboub type*, as presented by Snow (2012, chapter 2.1).

Employing the second identity in (10c), the system (11) may be rewritten as

$$(A - \tilde{E}_A)^T Q_1^{-1} A \cdot \hat{\xi} = (A - \tilde{E}_A)^T Q_1^{-1} y, \tag{12}$$

which corresponds nicely to a result by Fang (2011). By symmetric extension, (12) can be converted into the more familiar system

$$\boxed{(A - \tilde{E}_A)^T Q_1^{-1}(A - \tilde{E}_A) \cdot \hat{\xi} = (A - \tilde{E}_A)^T Q_1^{-1}(y - \tilde{E}_A\hat{\xi})}\,, \tag{13}$$

which translates into the iterative *Algorithm 2 of Fang type*, as described by Snow (2012, chapter 2.1.1).

Note that (9), together with (5c), can *equivalently* be written as

$$\begin{bmatrix} Q_1 & A - \tilde{E}_A \\ (A - \tilde{E}_A)^T & 0 \end{bmatrix}\begin{bmatrix} \hat{\lambda} \\ \hat{\xi} \end{bmatrix} = \begin{bmatrix} y - \tilde{E}_A\hat{\xi} \\ 0 \end{bmatrix} \tag{14a}$$

with

$$\mathrm{vec}\,\tilde{E}_A = \tilde{e}_A = [Q_{yA}^T - Q_A(\hat{\xi} \otimes I_n)]\hat{\lambda} \tag{14b}$$

from (7), only requiring the invertibility of Q_1 (not of Q itself).

It is now straight-forward to obtain the Total Sum of (weighted) Squared Residuals (TSSR) from (7) as

$$\Omega(\mathrm{TSSR}) = \begin{bmatrix} \tilde{e}_y^T \mid \tilde{e}_A^T \end{bmatrix}\begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}\begin{bmatrix} \tilde{e}_y \\ \tilde{e}_A \end{bmatrix} =$$
$$= \hat{\lambda}^T(\hat{B}Q\hat{B}^T)\hat{\lambda} = \hat{\lambda}^T Q_1 \hat{\lambda} = \hat{\lambda}^T(y - A\hat{\xi}), \tag{15a}$$

leading to the *variance component estimate*

$$\boxed{\hat{\sigma}_0^2 = (n - m)^{-1} \cdot \hat{\lambda}^T(y - A\hat{\xi})}\,, \tag{15b}$$

which is *unique* whenever Q_1 is nonsingular, or whenever $\mathrm{rk}(\hat{B}Q) = \mathrm{rk}\,\hat{B} = n$ in case of a *positive-semidefinite* matrix Q.

However, as Schaffrin et al. (2013) have recently proven, a unique weighted TLS solution can even be obtained for a positive-semidefinite dispersion matrix Q as long as the more general *rank condition*

$$\boxed{\mathrm{rk}\,[BQ \mid A] = n} \tag{16}$$

of Neitzel and Schaffrin (2013) is fulfilled, still assuming that $\mathrm{rk}\,A = m < n$ as defined in (1a). Under the criterion (16), the system (14a) and (14b) can be rewritten as

$$\begin{bmatrix} Q_3 & A - \tilde{E}_A \\ (A - \tilde{E}_A)^T & 0 \end{bmatrix}\begin{bmatrix} \hat{\lambda} \\ \hat{\xi} \end{bmatrix} = \begin{bmatrix} y - \tilde{E}_A\hat{\xi} \\ 0 \end{bmatrix} \tag{17a}$$

with

$$Q_3 := Q_1 + (A - \tilde{E}_A)S(A - \tilde{E}_A)^T = Q_3(\hat{\xi}, \hat{\lambda}) \tag{17b}$$

for any chosen *positive-definite* $m \times m$ matrix S, and with

$$\mathrm{vec}\,\tilde{E}_A = \tilde{e}_A = [Q_{yA}^T - Q_A(\hat{\xi} \otimes I_n)]\hat{\lambda}. \tag{17c}$$

Note that Q_3 will be *nonsingular* under (16) as long as $\mathrm{rk}[\hat{B}Q \mid A - \tilde{E}_A] = \mathrm{rk}[\hat{B}Q \mid A] = n$, even though Q_1 may be singular, leading to the iterative *Algorithm 3 of Fang type* as described in Snow (2012, chapter 3.2.1). An alternative algorithm of Mahboub type, using the *nonsingular* matrix

$$Q_2 := Q_1 + ASA^T = Q_2(\hat{\xi}) \tag{18}$$

instead of Q_3 in case of a singular matrix Q_1, may be developed along the lines of Snow (2012, chapter 3.1) as well, eventually leading to the system

$$\boxed{(A^T Q_2^{-1} A + R_2) \cdot \hat{\xi} = A^T Q_2^{-1} y} \tag{19a}$$

with

$$R_2 := (A^T Q_2^{-1} AS - I_m) \cdot [(I_m \otimes \hat{\lambda})^T Q_A (I_m \otimes \hat{\lambda})] =$$
$$= R_2(\hat{\xi}, \hat{\lambda}) \tag{19b}$$

and

$$\boxed{\hat{\lambda} = \{Q_2 + AS \cdot [(I_m \otimes \hat{\lambda})^T Q_A(\hat{\xi} \otimes I_n)]\}^{-1}(y - A\hat{\xi})} \tag{19c}$$

in the case of *zero cross-correlations*, i. e., $Q_{yA} = 0$. An alternative development leads to

$$\boxed{\begin{aligned} \{[A^T - (I_m \otimes \hat{\lambda})^T Q_{yA}^T](Q'_2)^{-1}A + R'_2\} \cdot \hat{\xi} = \\ = [A^T - (I_m \otimes \hat{\lambda})^T Q_{yA}^T](Q'_2)^{-1}y \end{aligned}} \tag{20a}$$

with

$$Q'_2 := Q_1 + [A - Q_{yA}(I_m \otimes \hat{\lambda})] S \cdot \\ \cdot [A - Q_{yA}(I_m \otimes \hat{\lambda})]^T = Q'_2(\hat{\xi}, \hat{\lambda}), \tag{20b}$$

$$R'_2 := \{[A^T - (I_m \otimes \hat{\lambda})^T Q_{yA}^T](Q'_2)^{-1} \cdot \\ \cdot [A - Q_{yA}(I_m \otimes \hat{\lambda})] S - I_m\} \cdot (I_m \otimes \hat{\lambda})^T Q_A (I_m \otimes \hat{\lambda}) = \\ = R'_2(\hat{\xi}, \hat{\lambda}), \tag{20c}$$

and

$$\boxed{\hat{\lambda} = \{Q'_2 + [A - Q_{yA}(I_m \otimes \hat{\lambda})] \cdot S \cdot \\ \cdot [(I_m \otimes \hat{\lambda})^T Q_A (\hat{\xi} \otimes I_n)]\}^{-1} (y - A\hat{\xi})} \tag{20d}$$

when *cross-correlations* are taken into account.

3 The weighted TLS adjustment in the presence of multiple constraints – a brief review when $\mathrm{rk}(BQ) = \mathrm{rk}\, B$

In the following, linear constraints as well as one quadratic constraint are added to the model (1a) and (1b) resulting in the *EIV-Model with multiple constraints:*

$$y = (\underset{n \times m}{A} - E_A)\xi + e_y, \quad \mathrm{rk}\, A = m, \tag{21a}$$

$$\begin{bmatrix} e_y \\ e_A \end{bmatrix} := \begin{bmatrix} e_y \\ \mathrm{vec}\, E_A \end{bmatrix} \sim (\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma_0^2 Q := \sigma_0^2 \begin{bmatrix} Q_y & Q_{yA} \\ Q_{yA}^T & Q_A \end{bmatrix}), \tag{21b}$$

$$\kappa_0 = \underset{l \times m}{K} \xi, \quad \xi^T M \xi = \alpha_0^2, \tag{21c}$$

where

- κ_0 is a given $l \times 1$ vector,
- K is a fixed $l \times m$ matrix with $\mathrm{rk}\, K = l < m$,
- M is a fixed $m \times m$ symmetric, positive-(semi)definite matrix, and
- α_0^2 is a given positive constant.

In contrast to (1b), however, where the dispersion matrix Q could be singular as long as the rank condition (16) would hold, the singularity of Q is now somewhat more restricted to the case

$$\mathrm{rk}(BQ) = \mathrm{rk}\, B = n \quad \text{for} \quad B := [I_n | -(\xi \otimes I_n)^T], \tag{22}$$

similar to (8c) and (8d) where Q_1 is supposedly nonsingular. This case was recently discussed by Fang (2013), but the derivations here will follow more the original approach by Schaffrin and Felus (2009) in the homoscedastic situation. The new algorithm for the weighted TLS adjustment with multiple constraints under the more general rank condition (16) will then be developed in the next chapter.

Let us temporarily assume that Q is positive-definite and hence nonsingular; then, the *Weighted Total Least-Squares (WTLS)* objective function is well defined by

$$\begin{bmatrix} e_y^T \mid e_A^T \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{bmatrix} \begin{bmatrix} e_y \\ e_A \end{bmatrix} = \min. \\ \text{s.t. } y - A\xi - e_y + E_A\xi = 0, \ \kappa_0 - K\xi = 0, \\ \alpha_0^2 - \xi^T M \xi = 0, \tag{23}$$

and can equivalently be translated into the Lagrange target function

$$\Phi(e_y, e_A, \xi, \lambda, \mu_1, \mu_2) := e_y^T P_{11} e_y + 2e_y^T P_{12} e_A + \\ + e_A^T P_{22} e_A + 2\lambda^T (y - A\xi - e_y + (\xi^T \otimes I_n) e_A) - \\ - 2\mu_1^T (\kappa_0 - K\xi) - \mu_2(\alpha_0^2 - \xi^T M \xi) = \text{stationary}. \tag{24}$$

The corresponding Euler-Lagrange necessary conditions will read

$$\frac{1}{2}\frac{\partial \Phi}{\partial e_y} = P_{11}\tilde{e}_y + P_{12}\tilde{e}_A - \hat{\lambda} \doteq 0, \tag{25a}$$

$$\frac{1}{2}\frac{\partial \Phi}{\partial e_A} = P_{12}^T\tilde{e}_y + P_{22}\tilde{e}_A + (\hat{\xi} \otimes I_n)\hat{\lambda} \doteq 0, \tag{25b}$$

$$\frac{1}{2}\frac{\partial \Phi}{\partial \xi} = -A^T\hat{\lambda} + \tilde{E}_A^T\hat{\lambda} + K^T\hat{\mu}_1 + M\hat{\xi} \cdot \hat{\mu}_2 \doteq 0, \tag{25c}$$

$$\frac{1}{2}\frac{\partial \Phi}{\partial \lambda} = y - A\hat{\xi} - \tilde{e}_y + (\hat{\xi}^T \otimes I_n)\tilde{e}_A \doteq 0, \tag{25d}$$

$$\frac{1}{2}\frac{\partial \Phi}{\partial \mu_1} = -\kappa_0 + K\hat{\xi} \doteq 0, \tag{25e}$$

$$\frac{\partial \Phi}{\partial \mu_2} = -\alpha_0^2 + \hat{\xi}^T M \hat{\xi} \doteq 0, \tag{25f}$$

and provide the system of *nonlinear normal equations* for the estimated parameters $\hat{\xi}$ and the residual vectors $\tilde{e}_y$ and $\tilde{e}_A$, in particular since the sufficient condition for a minimum in (23) is fulfilled as

$$\frac{1}{2}\frac{\partial^2 \Phi}{\partial \begin{bmatrix} e_y \\ e_A \end{bmatrix} \partial \begin{bmatrix} e_y^T \mid e_A^T \end{bmatrix}} = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{bmatrix} = Q^{-1} \tag{26}$$

is positive-definite.

First, from (25a) and (25b), the two residual vectors are directly derived in terms of $\hat{\lambda}$ through inversion as in (7), namely as

$$\begin{bmatrix} \tilde{e}_y \\ \tilde{e}_A \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{bmatrix}^{-1} \begin{bmatrix} I_n \\ -(\hat{\xi} \otimes I_n) \end{bmatrix} \cdot \hat{\lambda} = \\ = (Q\hat{B}^T)\hat{\lambda} = \begin{bmatrix} [Q_y - Q_{yA}(\hat{\xi} \otimes I_n)]\hat{\lambda} \\ [Q_{yA}^T - Q_A(\hat{\xi} \otimes I_n)]\hat{\lambda} \end{bmatrix}. \tag{27}$$

Inserting (27) into (25d) now leads to

$$y - A\hat{\xi} = (\hat{B}Q\hat{B}^T)\hat{\lambda} =: Q_1 \cdot \hat{\lambda} \tag{28}$$

just as in (8a) to (8d), with a non-singular matrix Q_1 owing to the assumption (22). Note that this assumption will be given up in the next chapter; but, here, the inversion of (28) into

$$\boxed{\hat{\lambda} = Q_1^{-1}(y - A\hat{\xi})} \tag{29}$$

is still allowed, providing the estimated Lagrange multipliers as a function of the estimated parameters.

Using (29) in (25c) then yields the identity

$$(A - \tilde{E}_A)^T Q_1^{-1}(y - A\hat{\xi}) = K^T\hat{\mu}_1 + M\hat{\xi}\cdot\hat{\mu}_2, \tag{30a}$$

which immediately implies the identity

$$\begin{aligned}&(A - \tilde{E}_A)^T Q_1^{-1}(A - \tilde{E}_A)\cdot\hat{\xi} + K^T\hat{\mu}_1 = \\ &= (A - \tilde{E}_A)^T Q_1^{-1}(y - \tilde{E}_A\hat{\xi}) - M\hat{\xi}\cdot\hat{\mu}_2\end{aligned} \tag{30b}$$

that needs to be solved along with (25e) and (25f). This, equivalently, means that the estimated parameters from the extended system

$$\begin{aligned}&\begin{bmatrix}(A - \tilde{E}_A)^T Q_1^{-1}(A - \tilde{E}_A) & K^T \\ K & 0\end{bmatrix}\begin{bmatrix}\hat{\xi} \\ \hat{\mu}_1\end{bmatrix} = \\ &= \begin{bmatrix}(A - \tilde{E}_A)^T Q_1^{-1}(y - \tilde{E}_A\hat{\xi}) - M\hat{\xi}\cdot\hat{\mu}_2 \\ \kappa_0\end{bmatrix}\end{aligned} \tag{31a}$$

must simultaneously satisfy the "*secular equation*" from (25f), namely

$$\hat{\xi}^T M\hat{\xi} = \alpha_0^2, \text{ where } \hat{\xi} = \hat{\xi}(\hat{\lambda}, \hat{\mu}_1, \hat{\mu}_2). \tag{31b}$$

By introducing the abbreviations

$$\begin{aligned}&[\hat{N}_1 \mid \hat{c}_1] := (A - \tilde{E}_A)^T Q_1^{-1}\left[(A - \tilde{E}_A) \mid y - \tilde{E}_A\hat{\xi}\right] = \\ &= \left[\hat{N}_1(\hat{\xi}, \hat{\lambda}) \mid \hat{c}_1(\hat{\xi}, \hat{\lambda})\right],\end{aligned} \tag{32}$$

the estimated parameter vector from (31a) may be formally described as

$$\boxed{\begin{aligned}\hat{\xi} = {} & \left[\hat{N}_1^{-1}\hat{c}_1 + \hat{N}_1^{-1}K^T(K\hat{N}_1^{-1}K^T)^{-1}(\kappa_0 - K\hat{N}_1^{-1}\hat{c}_1)\right] \\ & - \hat{N}_1^{-1}\left[I_m - K^T(K\hat{N}_1^{-1}K^T)^{-1}K\hat{N}_1^{-1}\right]M\hat{\xi}\cdot\hat{\mu}_2 = \quad (33a) \\ = {} & : w_1 - z_1\cdot\hat{\mu}_2 \quad (33b)\end{aligned}}$$

so that the "secular equation" (31b) turns into the quadratic equation

$$\boxed{(z_1^T M z_1)\cdot\hat{\mu}_2^2 - 2(w_1^T M z_1)\cdot\hat{\mu}_2 + (w_1^T M w_1 - \alpha_0^2) = 0}. \tag{33c}$$

Once (33a) to (33c) is solved, the Total Sum of (weighted) Squared Residuals (TSSR) can be expressed as in (15a) by

$$\boxed{\begin{aligned}\Omega(\text{TSSR}) &= \hat{\lambda}^T(\hat{B}Q\hat{B}^T)\hat{\lambda} = \hat{\lambda}^T(y - A\hat{\xi}) = \\ &= \hat{\sigma}_0^2(n - m + l + 1),\end{aligned}} \tag{34}$$

thus providing a *variance component estimate* for this case.

Obviously, the system (33a) to (33c) is *fully consistent* with the one provided by Fang (2013, chapter 3). More importantly, in the *homoscedastic case*, the system (33a) to (33c) can as well be transformed into the implicit eigenvalue problem provided by Schaffrin and Felus (2009, eq. 24) when the residual matrix $\tilde{E}_A$ and the (hidden) vector $\hat{\lambda}$ are both expressed in terms of the factor $\hat{\nu} := \Omega$ from (34). A proof will be published elsewhere.

4 The new algorithm for a weighted TLS adjustment with multiple constraints when Q is singular with $\operatorname{rk}(BQ) < \operatorname{rk} B$

In this chapter, the weighted TLS adjustment with multiple constraints will be extended to the case where the dispersion matrix Q is *singular and reduces the rank of* B, *i.e.*, $\operatorname{rk}(BQ) < \operatorname{rk} B = n$, but still maintains the rank condition (16), namely $\operatorname{rk}[BQ \mid A] = n = \operatorname{rk} B$, following similar lines as in chapter 2 for the unconstrained weighted TLS adjustment. For this purpose, the system (31a) and (31b) in conjunction with (29) is rewritten in the form

$$\begin{bmatrix}Q_1 & A - \tilde{E}_A & 0 \\ (A - \tilde{E}_A)^T & 0 & -K^T \\ 0 & -K & 0\end{bmatrix}\begin{bmatrix}\hat{\lambda} \\ \hat{\xi} \\ \hat{\mu}_1\end{bmatrix} = \begin{bmatrix}y - \tilde{E}_A\hat{\xi} \\ M\hat{\xi}\cdot\hat{\mu}_2 \\ -\kappa_0\end{bmatrix}, \tag{35a}$$

$$\hat{\xi}^T M\hat{\xi} = \alpha_0^2 \quad \text{where} \quad \hat{\xi} = \hat{\xi}(\hat{\lambda}, \hat{\mu}_1, \hat{\mu}_2), \tag{35b}$$

which nicely reflects the equation (25c) again. Obviously, by eliminating $\hat{\lambda}$ from (35a), the system (31a) is reproduced, where $\tilde{E}_A$ comes from (27) and $\hat{\mu}_2$ from (33c). In the present case, however, it is not so easy to eliminate $\hat{\lambda}$ directly since now the $n \times n$ matrix $Q_1 = \hat{B}Q\hat{B}^T$ has only $\operatorname{rk} Q_1 = \operatorname{rk}(\hat{B}Q) < n$ and, thus, is *singular*.

On the other hand, due to the rank condition (16), the matrix $Q_3 = Q_3(\hat{\lambda}, \hat{\xi})$ as defined in (17b) should be *invertible* as long as

$$\operatorname{rk} Q_3 = \operatorname{rk}\left[\hat{B}Q \mid (A - \tilde{E}_A)\right] = \operatorname{rk}\left[\hat{B}Q \mid A\right] = n \tag{36}$$

can be maintained for the residual matrix $\tilde{E}_A$. It is, therefore, suggested to employ the middle part of (35a) in the forms

$$KS(A - \tilde{E}_A)^T\cdot\hat{\lambda} = KS(K^T\hat{\mu}_1 + M\hat{\xi}\cdot\hat{\mu}_2), \tag{37a}$$

respectively

$$(A - \tilde{E}_A)S(A - \tilde{E}_A)^T\cdot\hat{\lambda} = (A - \tilde{E}_A)S(K^T\hat{\mu}_1 + M\hat{\xi}\cdot\hat{\mu}_2), \tag{37b}$$

where the *positive-definite* $m \times m$ matrix S may be chosen in a suitable manner. After adding (37b) to the upper part of (35a), and (37a) to the lower part, the new system reads:

$$\begin{bmatrix} Q_3 & A-\tilde{E}_A & -(A-\tilde{E}_A)SK^T \\ (A-\tilde{E}_A)^T & 0 & -K^T \\ -KS(A-\tilde{E}_A)^T & -K & KSK^T \end{bmatrix} \begin{bmatrix} \hat{\lambda} \\ \hat{\xi} \\ \hat{\mu}_1 \end{bmatrix} = \\ = \begin{bmatrix} (y-\tilde{E}_A\hat{\xi})+(A-\tilde{E}_A)SM\hat{\xi}\cdot\hat{\mu}_2 \\ M\hat{\xi}\cdot\hat{\mu}_2 \\ -(\kappa_0+KSM\hat{\xi}\cdot\hat{\mu}_2) \end{bmatrix}, \tag{38}$$

which needs to be solved in conjunction with (35b). Due to the *nonsingularity* of Q_3, now $\hat{\lambda}$ may be eliminated via

$$\boxed{\hat{\lambda}=Q_3^{-1}[(y-\tilde{E}_A\hat{\xi})-(A-\tilde{E}_A)\hat{\xi}+ \\ +(A-\tilde{E}_A)S(K^T\hat{\mu}_1+M\hat{\xi}\cdot\hat{\mu}_2)],} \tag{39}$$

which compares nicely to equation (29). Feeding (39) back into the system (38) now leads to

$$\left[\begin{array}{c|c} (A-\tilde{E}_A)^TQ_3^{-1}(A-\tilde{E}_A) & [I_m-(A-\tilde{E}_A)^TQ_3^{-1}(A-\tilde{E}_A)S]K^T \\ K[I_m-S(A-\tilde{E}_A)^TQ_3^{-1}(A-\tilde{E}_A)] & -K[I_m-S(A-\tilde{E}_A)^TQ_3^{-1}(A-\tilde{E}_A)]SK^T \end{array}\right]\cdot\begin{bmatrix} \hat{\xi} \\ \hat{\mu}_1 \end{bmatrix} = \\ = \begin{bmatrix} (A-\tilde{E}_A)^TQ_3^{-1}(y-\tilde{E}_A\hat{\xi})- \\ -[I_m-(A-\tilde{E}_A)^TQ_3^{-1}(A-\tilde{E}_A)S]M\hat{\xi}\cdot\hat{\mu}_2 \\ \kappa_0-KS(A-\tilde{E}_A)^TQ_3^{-1}(y-\tilde{E}_A\hat{\xi})+ \\ +K[I_m-S(A-\tilde{E}_A)^TQ_3^{-1}(A-\tilde{E}_A)]SM\hat{\xi}\cdot\hat{\mu}_2 \end{bmatrix} \tag{40}$$

or, perhaps preferably, to

$$\begin{bmatrix} \hat{N}_3 & K^T \\ K & 0 \end{bmatrix}\begin{bmatrix} \hat{\xi} \\ \hat{\mu}_1 \end{bmatrix} := \left[\begin{array}{c|c} (A-\tilde{E}_A)^TQ_3^{-1}(A-\tilde{E}_A) & K^T \\ K & 0 \end{array}\right]\begin{bmatrix} \hat{\xi} \\ \hat{\mu}_1 \end{bmatrix} = \\ = \begin{bmatrix} (A-\tilde{E}_A)^TQ_3^{-1}[y-\tilde{E}_A\hat{\xi}+ \\ +(A-\tilde{E}_A)S(K^T\hat{\mu}_1+M\hat{\xi}\cdot\hat{\mu}_2)]-M\hat{\xi}\cdot\hat{\mu}_2 \\ \kappa_0 \end{bmatrix} =: \tag{41a}$$

$$=: \begin{bmatrix} \hat{c}_3-[I_m-(A-\tilde{E}_A)^TQ_3^{-1}(A-\tilde{E}_A)S]M\hat{\xi}\cdot\hat{\mu}_2 \\ \kappa_0 \end{bmatrix} =: \\ =: \begin{bmatrix} \hat{c}_3-\hat{d}_3\cdot\hat{\mu}_2 \\ \kappa_0 \end{bmatrix}, \tag{41b}$$

where $\hat{d}_3 = \hat{d}_3(\hat{\xi},\hat{\lambda})$ and $\hat{N}_3 = \hat{N}_3(\hat{\xi},\hat{\lambda})$, but $\hat{c}_3 = \hat{c}_3(\hat{\xi},\hat{\lambda},\hat{\mu}_1)$ also depends on $\hat{\mu}_1$. Equation (41a) allows the formal solution

$$\hat{\xi}=[\hat{N}_3^{-1}\hat{c}_3+\hat{N}_3^{-1}K^T(K\hat{N}_3^{-1}K^T)^{-1}(\kappa_0-K\hat{N}_3^{-1}\hat{c}_3)]- \\ -\hat{N}_3^{-1}[I_m-K^T(K\hat{N}_3^{-1}K^T)^{-1}K\hat{N}_3^{-1}]\cdot \\ \cdot[I_m-(A-\tilde{E}_A)^TQ_3^{-1}(A-\tilde{E}_A)S]M\hat{\xi}\cdot\hat{\mu}_2 = \tag{42a}$$

$$=: w_3 - z_3\cdot\hat{\mu}_2, \tag{42b}$$

with $w_3 = w_3(\hat{\xi},\hat{\lambda},\hat{\mu}_1)$ and $z_3 = z_3(\hat{\xi},\hat{\lambda})$, which, when reinserted into (35b), provides the *"secular equation"*

$$\boxed{(z_3^TMz_3)\cdot\hat{\mu}_2^2-2(w_3^TMz_3)\cdot\hat{\mu}_2+(w_3^TMw_3-\alpha_0^2)=0}\,. \tag{42c}$$

Of the two solutions, the one that *minimizes the TSSR* will be selected, which can be computed as in (34) when taking $\hat{\lambda}$ from (39). Alternatively, the system (40) can be handled in the same way as the system (31a) before.

Based on (42a) to (42c), the *novel Algorithm 1* is proposed in equations (43a) to (43p).

5 Numerical examples

In this chapter, we apply the new Algorithm 1 to a few numerical examples. These include a geodetic resection problem from Schaffrin and Felus (2009), a 2-dimensional rigid-transformation problem presented by Fang (2014), and a 2-dimensional similarity-transformation problem from Schaffrin et al. (2014). In all cases, a "model check" is made after the adjustment, which confirms that

$$\|y-A\hat{\xi}+\tilde{E}_A\hat{\xi}-\tilde{e}_y\|\approx 0,\ \|K\hat{\xi}-\kappa_0\|\approx 0, \\ \text{and } |\hat{\xi}^TM\hat{\xi}-\alpha_0^2|\approx 0, \tag{44}$$

where the approximation signs reflect the finiteness of machine precision and the fact that the algorithm's convergence threshold δ is nonzero.

5.1 Two examples with multiple constraints, but nonsingular dispersion matrices

5.1.1 An example from Schaffrin and Felus (2009), homoscedastic case

Schaffrin and Felus (2009) solved an overdetermined "simplified geodetic resection" problem, which essentially amounts to estimating coordinates of a physical location based on geodetic measurements made from it. In their case, 3-dimensional coordinates were estimated based on four (simulated) measurements; see their chapter 5 for a more detailed description.

The matrix $[y \mid A]$, comprised of the observation vector y and the data matrix A, together with its associated (nonsingular) error dispersion matrix $D\{\text{vec}[e_y \mid E_A]\}$, are given by

$$[y\mid A] := \left[\begin{array}{c|ccc} 6 & -0.5 & 1 & 0 \\ 3 & 0 & 1 & 0 \\ 4 & 0 & 0 & 1 \\ 10 & 1 & 0 & 1 \end{array}\right] \quad \text{and} \tag{45a}$$

$$D\{\begin{bmatrix} e_y \\ e_A \end{bmatrix}\} = \sigma_0^2 I_{n(m+1)},$$

Algorithm 1
For the weighted TLS adjustment with multiple constraints when $\boldsymbol{Q}$ is singular with $\operatorname{rk}(\boldsymbol{BQ}) < \operatorname{rk}\boldsymbol{B}$

Step 1: Compute an initial solution

$$\begin{bmatrix}\hat{\boldsymbol{\xi}}^{(0)}\\ \hat{\boldsymbol{\mu}}_1^{(0)}\end{bmatrix} := \begin{bmatrix}(\boldsymbol{A}^T\boldsymbol{Q}_y^+\boldsymbol{A}) & \boldsymbol{K}^T\\ \boldsymbol{K} & \boldsymbol{0}\end{bmatrix}^+ \begin{bmatrix}\boldsymbol{A}^T\boldsymbol{Q}_y^+\boldsymbol{y}\\ \boldsymbol{\kappa}_0\end{bmatrix}, \quad \tilde{\boldsymbol{E}}_A^{(0)} := \boldsymbol{0}, \quad \hat{\mu}_2^{(0)} := 0, \tag{43a}$$

where $\boldsymbol{Q}_y^+$ represents the “pseudo-inverse” of $\boldsymbol{Q}_y$;

Step 2: Repeat for $i \in \mathbb{N}$:

$$\hat{\boldsymbol{B}}^{(i)} := \left[\boldsymbol{I}_n \,\middle|\, -(\hat{\boldsymbol{\xi}}^{(i-1)} \otimes \boldsymbol{I}_n)^T\right], \quad \boldsymbol{Q}_1^{(i)} := \hat{\boldsymbol{B}}^{(i)}\boldsymbol{Q}(\hat{\boldsymbol{B}}^{(i)})^T, \tag{43b}$$

$$\boldsymbol{Q}_3^{(i)} := \boldsymbol{Q}_1^{(i)} + (\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)})\, \boldsymbol{S}(\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)})^T, \tag{43c}$$

$$\hat{\boldsymbol{N}}_3^{(i)} := (\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)})^T(\boldsymbol{Q}_3^{(i)})^{-1}(\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)}), \tag{43d}$$

$$\hat{\boldsymbol{c}}_3^{(i)} := (\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)})^T(\boldsymbol{Q}_3^{(i)})^{-1}\cdot[\boldsymbol{y} - \tilde{\boldsymbol{E}}_A^{(i-1)}\hat{\boldsymbol{\xi}}^{(i-1)} + (\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)})\boldsymbol{S}\boldsymbol{K}^T\hat{\boldsymbol{\mu}}_1^{(i-1)}], \tag{43e}$$

$$\hat{\boldsymbol{d}}_3^{(i)} := [\boldsymbol{I}_m - (\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)})^T(\boldsymbol{Q}_3^{(i)})^{-1}(\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)})\boldsymbol{S}]\boldsymbol{M}\cdot\hat{\boldsymbol{\xi}}^{(i-1)}, \tag{43f}$$

$$\boldsymbol{w}_3^{(i)} := (\hat{\boldsymbol{N}}_3^{(i)})^{-1}\hat{\boldsymbol{c}}_3^{(i)} + (\hat{\boldsymbol{N}}_3^{(i)})^{-1}\boldsymbol{K}^T[\boldsymbol{K}(\hat{\boldsymbol{N}}_3^{(i)})^{-1}\boldsymbol{K}^T]^{-1}[\boldsymbol{\kappa}_0 - \boldsymbol{K}(\hat{\boldsymbol{N}}_3^{(i)})^{-1}\hat{\boldsymbol{c}}_3^{(i)}], \tag{43g}$$

$$\boldsymbol{z}_3^{(i)} := (\hat{\boldsymbol{N}}_3^{(i)})^{-1}[\boldsymbol{I}_m - \boldsymbol{K}^T[\boldsymbol{K}(\hat{\boldsymbol{N}}_3^{(i)})^{-1}\boldsymbol{K}^T]^{-1}\boldsymbol{K}(\hat{\boldsymbol{N}}_3^{(i)})^{-1}]\cdot\hat{\boldsymbol{d}}_3^{(i)}, \tag{43h}$$

$$a_1^{(i)} := (\boldsymbol{z}_3^{(i)})^T\boldsymbol{M}\boldsymbol{z}_3^{(i)},\ a_2^{(i)} := (\boldsymbol{w}_3^{(i)})^T\boldsymbol{M}\boldsymbol{z}_3^{(i)},\ a_3^{(i)} := (\boldsymbol{w}_3^{(i)})^T\boldsymbol{M}\boldsymbol{w}_3^{(i)}, \tag{43i}$$

$$(\hat{\mu}_2^{(i)})_{1/2} = \frac{a_2^{(i)}}{a_1^{(i)}} \pm \sqrt{\left(a_2^{(i)}/a_1^{(i)}\right)^2 + (\alpha_0^2 - a_3^{(i)})/a_1^{(i)}}, \tag{43j}$$

$$(\hat{\boldsymbol{\xi}}^{(i)})_{1/2} := \boldsymbol{w}_3^{(i)} - \boldsymbol{z}_3^{(i)}\cdot(\hat{\mu}_2^{(i)})_{1/2}, \tag{43k}$$

$$(\hat{\boldsymbol{\mu}}_1^{(i)})_{1/2} := -[\boldsymbol{K}(\hat{\boldsymbol{N}}_3^{(i)})^{-1}\boldsymbol{K}^T]^{-1}[\boldsymbol{\kappa}_0 - \boldsymbol{K}(\hat{\boldsymbol{N}}_3^{(i)})^{-1}(\hat{\boldsymbol{c}}_3^{(i)} - \hat{\boldsymbol{d}}_3^{(i)}(\hat{\mu}_2^{(i)})_{1/2})], \tag{43l}$$

$$\begin{aligned}(\hat{\boldsymbol{\lambda}}^{(i)})_{1/2} = (\boldsymbol{Q}_3^{(i)})^{-1}[&(\boldsymbol{y} - \tilde{\boldsymbol{E}}_A^{(i-1)}\hat{\boldsymbol{\xi}}^{(i-1)}) - (\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)})(\hat{\boldsymbol{\xi}}^{(i)})_{1/2} + \\ &+ (\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)})\boldsymbol{S}(\boldsymbol{K}^T(\hat{\boldsymbol{\mu}}_1^{(i)})_{1/2} + \boldsymbol{M}(\hat{\boldsymbol{\xi}}^{(i)})_{1/2}\cdot(\hat{\mu}_2^{(i)})_{1/2})],\end{aligned} \tag{43m}$$

$$\Omega^{(i)}(\text{TSSR}) := \min\{(\hat{\boldsymbol{\lambda}}_{1/2}^{(i)})^T(\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{\xi}}^{(i)})\} = (\hat{\sigma}_0^2)^{(i)}\cdot(n - m + l + 1) \Rightarrow \{\hat{\boldsymbol{\xi}}^{(i)}, \hat{\boldsymbol{\mu}}_1^{(i)}, \hat{\mu}_2^{(i)}, \hat{\boldsymbol{\lambda}}^{(i)}\}, \tag{43n}$$

$$\tilde{\boldsymbol{e}}_A^{(i)} = [\boldsymbol{Q}_{yA}^T - \boldsymbol{Q}_A(\hat{\boldsymbol{\xi}}^{(i)} \otimes \boldsymbol{I}_n)]\cdot\hat{\boldsymbol{\lambda}}^{(i)},\ \tilde{\boldsymbol{E}}_A^{(i)} = \operatorname{Invec}\tilde{\boldsymbol{e}}_A^{(i)}, \tag{43o}$$

equivalent to $\tilde{\boldsymbol{e}}_A^{(i)} = \operatorname{vec}\tilde{\boldsymbol{E}}_A^{(i)}$;

Until

$$\|\hat{\boldsymbol{\xi}}^{(i)} - \hat{\boldsymbol{\xi}}^{(i-1)}\| < \delta \text{ and } \|\hat{\boldsymbol{\lambda}}^{(i)} - \hat{\boldsymbol{\lambda}}^{(i-1)}\| < \delta \text{ for a chosen threshold } \delta. \tag{43p}$$

respectively, with $m = 3$ and $n = 4$. The quadratic constraints are provided by

$$\frac{\xi_1^2}{12^2} + \frac{\xi_2^2}{8^2} + \frac{\xi_3^2}{12^2} = 1 \Rightarrow$$
$$\Rightarrow \boldsymbol{M} = \begin{bmatrix}(1/12)^2 & 0 & 0\\ 0 & (1/8)^2 & 0\\ 0 & 0 & (1/12)^2\end{bmatrix} \text{ and } \alpha_0^2 = 1. \tag{45b}$$

An additional linear constraint is given by

$$-2\xi_1 + 3\xi_3 = 16 \Rightarrow \boldsymbol{K} = \begin{bmatrix}-2 & 0 & 3\end{bmatrix} \text{ and } \kappa_0 = 16, \tag{45c}$$

leading to a system redundancy of $4 - 3 + 1 + 1 = 3$. A convergence threshold of $\delta = 1 \times 10^{-14}$ was used for Algorithm 1, and the identity matrix was chosen for $\boldsymbol{S}$, i. e., $\boldsymbol{S} = \boldsymbol{I}_3$.

Tab. 1: Solution for the geodetic resection problem

Quantity	Estimate/Calculation
$\hat{\xi}_1$	2.597297
$\hat{\xi}_2$	6.230453
$\hat{\xi}_3$	7.064865
Ω (TSSR)	0.218544
$\hat{\sigma}_0^2$	$(0.269904)^2$
redundancy	3
$\left\Vert y - A\hat{\xi} + \tilde{E}_A\hat{\xi} - \tilde{e}_y \right\Vert$	1.3×10^{-15}
$K\hat{\xi}$	16.000000
$\hat{\xi}^T M \hat{\xi}$	1.000000
δ (convergence threshold)	1×10^{-14}
S	I_3
Number of iterations	16

The results of Algorithm 1 are shown in Tab 1. The solution agrees with that of Schaffrin and Felus (2009) to within 1×10^{-5} for all estimated parameters and the TSSR. We do not expect an exact match since their convergence threshold was not reported. It is noted that the number of iterations drops from 16 to 12 when the convergence threshold for Algorithm 1 is increased to $\delta = 1 \times 10^{-10}$. The augmented matrix of residuals turns out to be

$$[\tilde{e}_y \mid \tilde{E}_A] = \left[\begin{array}{r|rrr} 0.0111 & -0.0288 & -0.0690 & -0.0782 \\ -0.0335 & 0.0870 & 0.2086 & 0.2366 \\ -0.0318 & 0.0825 & 0.1979 & 0.2244 \\ 0.0035 & -0.0091 & -0.0218 & -0.0247 \end{array}\right].$$

5.1.2 An example from Fang (2014), weighted case for nonsingular Q_1

Next we consider a 2-dimensional rigid-transformation problem solved by Fang (2014), who used the horizontal coordinates of a 3-dimensional data set presented by Felus and Burtch (2009). Given a set of coordinates (x_i, y_i), $i = 1, \ldots, n/2$, determined for $n/2$ locations in a source coordinate system and a corresponding set of coordinates (X_i, Y_i) determined in a target coordinate system, together with their respective error dispersion matrices, the problem is to estimate three transformation parameters between the source and target systems, namely a rotation angle α and two translation parameters t_1 and t_2.

Rather than to estimate the rotation angle α directly, it is common to estimate the trigonometric terms $\cos\alpha$ and $\sin\alpha$ so that the unknown parameter vector is defined in terms of the transformation parameters as $\xi := [\cos\alpha, \sin\alpha, t_1, t_2]^T$; thus $m = 4$. Following Fang's ordering of the data variables, this approach permits a linear model of the form

$$y := \begin{bmatrix} X_1 \\ \vdots \\ X_{n/2} \\ \hline Y_1 \\ \vdots \\ Y_{n/2} \end{bmatrix} \approx A\xi := \begin{bmatrix} x_1 & y_1 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ x_{n/2} & y_{n/2} & 1 & 0 \\ \hline y_1 & -x_1 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ y_{n/2} & -x_{n/2} & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\alpha \\ \sin\alpha \\ t_1 \\ t_2 \end{bmatrix}, \tag{46a}$$

where the "equals approximately" sign is used due to random errors in the data. This problem does not include a linear constraint on the parameters; however, in order to enforce a "rigid" transformation, the quadratic constraint

$$\cos^2\alpha + \sin^2\alpha = 1 \Rightarrow \xi^T \operatorname{Diag}([1,1,0,0])\xi = 1 =: \alpha_0^2 \tag{46b}$$

is applied. Of course, such a problem can still be solved using Algorithm 1 provided that all terms in K are removed from the algorithm.

In this problem, $n = 8$, and since there are no linear constraints involved, the system redundancy is $n - m + 1 = 8 - 4 + 1 = 5$. Fang's data are listed in Tab. 2.

Let the random errors of the coordinates from the source and target systems be represented by the $n \times 1$ vectors e_{xy} and e_{XY}, respectively. For the problem at hand, the error distribution is described by $D\{\operatorname{vec}[e_{xy}, e_{XY}]\} = \sigma_0^2 \cdot I_{2n}$. Due to the structure of the data matrix A, the cofactor matrix Q (introduced in (1b) and (21b) is generated by $Q = Z \cdot I_{2n} \cdot Z^T$, where the $n(m+1) \times 2n$ matrix Z is defined as

$$Z := \left[\begin{array}{c|cc} I_n & 0 & 0 \\ \hline 0 & I_{n/2} & 0 \\ 0 & 0 & I_{n/2} \\ 0 & 0 & I_{n/2} \\ 0 & -I_{n/2} & 0 \\ \hline 0_{2n\times n} & 0 & 0 \end{array}\right] \Rightarrow$$

$$\Rightarrow Q = \left[\begin{array}{c|cccc|c} I_n & \multicolumn{5}{c}{0_{n\times 4n}} \\ \hline & I_{n/2} & 0 & 0 & -I_{n/2} & \\ 0_{4n\times n} & 0 & I_{n/2} & I_{n/2} & 0 & 0_{2n} \\ & 0 & I_{n/2} & I_{n/2} & 0 & \\ & -I_{n/2} & 0 & 0 & I_{n/2} & \\ \hline & \multicolumn{4}{c|}{0_{2n}} & 0_{2n} \end{array}\right], \tag{46c}$$

Tab. 2: Data for the 2-dimensional rigid-transformation problem

Point No.	x_i	y_i	X_i	Y_i
1	30	40	290	150
2	100	40	420	80
3	100	130	540	200
4	30	130	390	300

with $\mathbf{0}_{2n}$ representing a $2n \times 2n$ matrix of zeros. Note that, though the cofactor matrix $\boldsymbol{Q}$ is singular, the matrix $\boldsymbol{Q}_1$ (defined in (8b) resp. (28)) turns out to be nonsingular in this case.

A convergence threshold of $\delta = 1 \times 10^{-12}$ and the matrix $\boldsymbol{S} = 1 \times 10^{-4} \cdot \boldsymbol{I}_4$ were used in Algorithm 1. The results of the algorithm are listed in Tab. 3. These results match exactly those presented in the independent work by Fang (2014). The very large value for the TSSR reveals that the data do not fit well to the rigid-transformation model, which may be a consequence of applying a different model than that used by Felus and Burtch (2009) for the 3-dimensional superset, as was already suggested by Fang. Nevertheless, our purpose for presenting this problem has been satisfied, which is to show the flexibility of our new algorithm.

Finally, the predicted random errors (residuals)

$$[\tilde{e}_y \mid \tilde{\boldsymbol{E}}_A] = \left[\begin{array}{r|rrrr} -32.6402 & 21.6305 & 25.7997 & 0 & 0 \\ 3.9843 & -16.5566 & 16.1227 & 0 & 0 \\ 37.6402 & -30.0749 & -22.6464 & 0 & 0 \\ -8.9843 & 25.0009 & -19.2760 & 0 & 0 \\ -8.2535 & 25.7997 & -21.6305 & 0 & 0 \\ -22.7637 & 16.1227 & 16.5566 & 0 & 0 \\ 0.7535 & -22.6464 & 30.0749 & 0 & 0 \\ 30.2637 & -19.2760 & -25.0009 & 0 & 0 \end{array}\right].$$

confirm that the structure of the data matrix $\boldsymbol{A}$ has been replicated exactly in the residual matrix $\tilde{\boldsymbol{E}}_A$.

Tab. 3: Solution for the 2-dimensional rigid-transformation problem

Quantity	Estimate/Calculation
$\hat{\xi}_1 := \widehat{\cos\alpha}$	0.810728
$\hat{\xi}_2 := \widehat{\sin\alpha}$	0.585423
$\hat{\xi}_3 := \hat{t}_1$	307.541719
$\hat{\xi}_4 := \hat{t}_2$	151.640630
Ω (TSSR)	8163.065565
$\hat{\sigma}_0^2$	$(40.405607)^2$
redundancy	5
$\left\Vert y - A\hat{\xi} + \tilde{E}_A\hat{\xi} - \tilde{e}_y \right\Vert$	2.4×10^{-13}
$\hat{\xi}^T M \hat{\xi}$	1.00000000000000
δ (convergence threshold)	1×10^{-12}
$\boldsymbol{S}$	$1 \times 10^{-4} \cdot \boldsymbol{I}_3$
Number of iterations	3

5.2 Some new examples, with singular dispersion matrices, not solvable by previously existing algorithms

In this section, we again turn our attention to a 2-dimensional transformation problem but now of type *similarity transformation*, which differs from the rigid transformation in that a transformation scale-factor is also estimated. In these experiments, the random errors associated with the data variables are not iid. Instead, the dispersion matrices for the random errors in the variables of both the source and target coordinate systems are full and turn out to be singular such that the matrix $\boldsymbol{Q}_1$ is also singular. The singular nature of $\boldsymbol{Q}_1$ makes this problem unsolvable by previously existing algorithms, which motivated our design of Algorithm 1.

The measurement data are taken from Snow (2012) and are listed in Tab. 4. The 10×10 cofactor matrices $\boldsymbol{Q}_{xy}$ and $\boldsymbol{Q}_{XY}$ are also taken from Snow (2012, Appendix A.2), but are not listed here for the sake of space. The most important characteristic of these matrices is their rank; both have a rank of seven (nullity of size three), giving rise to rk $\boldsymbol{Q}_1 = 8$ for the 10×10 matrix $\boldsymbol{Q}_1$.

Tab. 4: Coordinate estimates in source and target systems

Point	x_i [m]	y_i [m]	X_i [m]	Y_i [m]
1	453.8001	137.6099	400.0040	100.0072
2	521.2865	350.7972	500.0019	299.9994
3	406.8728	433.9247	399.9925	399.9933
4	110.5545	386.9880	100.0059	400.0022
5	157.4861	90.6802	99.9956	99.9978

5.2.1 An unconstrained similarity transformation

A variation of Algorithm 1 can be used for models without constraints. The algorithm is modified by omitting all terms associated with the linear and quadratic constraints, thereby resulting in the same algorithm as that presented by Snow (2012, chapter 3.2.1).

The model details are left for the next section, where linear and quadratic constraints are incorporated. Here we simply present the solution for the unconstrained case in Tab. 5, followed by a listing of the residuals (cf. Snow 2012, chapter 6.3).

The residuals are given by

$$[\tilde{e}_y \mid \tilde{\boldsymbol{E}}_A] = \\ = \left[\begin{array}{r|rrrr} 1.0204 & 0 & 0 & -4.4026 & 5.3231 \\ 0.8998 & 0 & 0 & -5.3231 & -4.4026 \\ 0.3453 & 0 & 0 & -1.8617 & -0.5454 \\ -0.1634 & 0 & 0 & 0.5454 & -1.8617 \\ -1.5805 & 0 & 0 & 7.1386 & -6.2318 \\ -0.9923 & 0 & 0 & 6.2318 & 7.1386 \\ 1.0399 & 0 & 0 & -4.2616 & 6.8490 \\ 1.2009 & 0 & 0 & -6.8490 & -4.2616 \\ -0.8250 & 0 & 0 & 3.3873 & -5.3948 \\ -0.9450 & 0 & 0 & 5.3948 & 3.3873 \end{array}\right] \times 10^{-3}\text{m},$$

where the structure of the data matrix $\boldsymbol{A}$ is replicated in the residual matrix $\tilde{\boldsymbol{E}}_A$.

Tab. 5: Solution for the 2-dimensional similarity transformation without constraints. Linear quantities are shown in units of meters.

Quantity	Estimate/Calculation
$\hat{\xi}_1 := \hat{t}_1$	−69.726354
$\hat{\xi}_2 := \hat{t}_2$	35.078215
$\hat{\xi}_3 := \widehat{\omega\cos\alpha}$	0.98765502
$\hat{\xi}_4 := \widehat{\omega\sin\alpha}$	−0.15642921
$\hat{\omega}$	0.99996626
Ω (TSSR)	6.164035
$\hat{\sigma}_0^2$	$(1.01357734)^2$
redundancy	6
$\left\Vert \boldsymbol{y} - \boldsymbol{A}\hat{\xi} + \tilde{\boldsymbol{E}}_A\hat{\xi} - \tilde{\boldsymbol{e}}_y \right\Vert$	$< 2 \times 10^{-11}$
δ (convergence threshold)	1×10^{-10}
$\boldsymbol{S}$	$1 \times 10^{-2} \cdot \boldsymbol{I}_4$
Number of iterations	4

5.2.2 A constrained similarity transformation

Following Snow's (2012) arrangement of the variables and parameters, the EIV-Model with linear and quadratic constraints is provided by

$$\underset{n\times 1}{\boldsymbol{y}} := \begin{bmatrix} X_1 \\ Y_1 \\ \cdots \\ X_{n/2} \\ Y_{n/2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & x_1 & -y_1 \\ 0 & 1 & y_1 & x_1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & x_{n/2} & -y_{n/2} \\ 0 & 1 & y_{n/2} & x_{n/2} \end{bmatrix} \boldsymbol{\xi} - $$

$$- \begin{bmatrix} 0 & 0 & e_{x_1} & -e_{y_1} \\ 0 & 0 & e_{y_1} & e_{x_1} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & e_{x_{n/2}} & -e_{y_{n/2}} \\ 0 & 0 & e_{y_{n/2}} & e_{x_{n/2}} \end{bmatrix} \boldsymbol{\xi} + \begin{bmatrix} e_{X_1} \\ e_{Y_1} \\ \cdots \\ e_{X_{n/2}} \\ e_{Y_{n/2}} \end{bmatrix}$$

$$=: (\underset{n\times m}{\boldsymbol{A} - \boldsymbol{E}_A}) \cdot \boldsymbol{\xi} + \boldsymbol{e}_y \quad \text{with} \quad \operatorname{rk} \boldsymbol{A} = m = 4, \tag{47a}$$

$$\begin{bmatrix} \boldsymbol{e}_y \\ \boldsymbol{e}_A \end{bmatrix} := \begin{bmatrix} \boldsymbol{e}_y \\ \operatorname{vec} \boldsymbol{E}_A \end{bmatrix} \sim \left(\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}, \sigma_0^2 \begin{bmatrix} \boldsymbol{Q}_y & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q}_A \end{bmatrix} =: \sigma_0^2 \boldsymbol{Q} \right), \tag{47b}$$

$$\boldsymbol{K}\boldsymbol{\xi} = \boldsymbol{\kappa}_0 \quad \text{and} \quad \boldsymbol{\xi}^T \boldsymbol{M} \boldsymbol{\xi} = \alpha_0^2. \tag{47c}$$

The parameter vector $\boldsymbol{\xi}$ can be expressed in terms of the 2-dimensional transformation parameters by $\boldsymbol{\xi} := [t_1, t_2, \omega\cos\alpha, \omega\sin\alpha]^T$, where the transformation scale-factor ω appears in the last two elements. Letting the random errors of the coordinates from the source and target systems be denoted by $\boldsymbol{e}_{xy}$ and $\boldsymbol{e}_{XY}$, respectively, their given dispersion matrices can be expressed as $D\{\boldsymbol{e}_{xy}\} = \sigma_0^2 \boldsymbol{Q}_{xy}$ and $D\{\boldsymbol{e}_{XY}\} = \sigma_0^2 \boldsymbol{Q}_{XY}$, respectively. Note that there are no cross-correlations between the two dispersion matrices, as the measurements made in the respective source and target systems come from independent sources.

Then, the dispersion matrix in (47b) can be expressed in terms of the given dispersion matrices as

$$\sigma_0^2 \boldsymbol{Q} := \sigma_0^2 \begin{bmatrix} \boldsymbol{Q}_y & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q}_A \end{bmatrix} :=$$

$$:= \sigma_0^2 \left[\begin{array}{c|cccc} \boldsymbol{Q}_{XY} & & & \boldsymbol{0} & \\ \hline & \boldsymbol{0} & \cdots & & \\ \boldsymbol{0} & \vdots & \boldsymbol{0} & \cdots & \\ & & \vdots & \boldsymbol{Q}_{xy} & \boldsymbol{Q}_{xy}\boldsymbol{T}^T \\ & & & \boldsymbol{T}\boldsymbol{Q}_{xy} & \boldsymbol{T}\boldsymbol{Q}_{xy}\boldsymbol{T}^T \end{array} \right], \tag{48a}$$

where

$$\underset{n\times n}{\boldsymbol{T}} := \operatorname{Diag}(\boldsymbol{T}', \ldots, \boldsymbol{T}'), \text{ with } \boldsymbol{T}' := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \tag{48b}$$

See Snow (2012) or Schaffrin et al. (2014) for further details.

Now, to conduct our experiment, let us suppose that the ratio between the transformation shift parameters t_1 and t_2 is known and that the transformation scale-factor is known, too. The given ratio can be imposed via a linear constraint in the model, while the given scale factor can be imposed via a quadratic constraint. It would be interesting to use, for the linear constraint, the ratio of $\hat{t}_1$ and $\hat{t}_2$ estimated within the unconstrained model from chapter 5.2.1 and, likewise, take the estimate of $\hat{\omega}$ within that unconstrained model for the quadratic constraint. These values are taken from Tab. 5 and are used in the constraint equations (47c) as follows:

$$t_1/t_2 = -1.9877395 \Rightarrow \boldsymbol{K} := [1,\ 1.9877395,\ 0,\ 0]$$
and $\quad \boldsymbol{\kappa}_0 := 0;$
$$(\omega\cos\alpha)^2 + (\omega\sin\alpha)^2 = \omega^2 := (0.99996626)^2 \Rightarrow$$
$$\Rightarrow \boldsymbol{M} := \operatorname{Diag}([0,0,1,1]) \quad \text{and} \quad \alpha_0^2 := (0.99996626)^2.$$

The solution based on Algorithm 1 is shown in Tab. 6. The results match those from the unconstrained case presented in chapter 5.2.1, thereby confirming that the constraints in the model have been fulfilled. The convergence threshold was set to $\delta = 1 \times 10^{-9}$, and the matrix $\boldsymbol{S}$ was defined as $\boldsymbol{S} := 1 \times 10^3 \cdot \boldsymbol{I}_4$. The solution converged in five iterations.

Finally, the residuals are given by

$$[\tilde{e}_y \mid \tilde{E}_A] = $$
$$= \left[\begin{array}{r|rrrr} 1.0204 & 0 & 0 & -4.4030 & 5.3227 \\ 0.8997 & 0 & 0 & -5.3227 & -4.4030 \\ 0.3454 & 0 & 0 & -1.8623 & -0.5452 \\ -0.1634 & 0 & 0 & 0.5452 & -1.8623 \\ -1.5805 & 0 & 0 & 7.1384 & -6.2314 \\ -0.9923 & 0 & 0 & 6.2314 & 7.1384 \\ 1.0398 & 0 & 0 & -4.2610 & 6.8492 \\ 1.2010 & 0 & 0 & -6.8492 & -4.2610 \\ -0.8251 & 0 & 0 & 3.3878 & -5.3953 \\ -0.9450 & 0 & 0 & 5.3953 & 3.3878 \end{array}\right] \times 10^{-3}\text{m},$$

where the structure of the data matrix $\boldsymbol{A}$ is replicated in the residual matrix $\tilde{\boldsymbol{E}}_A$.

It is worth noting that Algorithm 1 would not converge when using a convergence threshold as small as $\delta = 1 \times 10^{-10}$ or when scaling the identity matrix by a scalar as small as 1×10^2 in the assignment of $\boldsymbol{S}$. Moreover, the model check $\|\boldsymbol{y} - \boldsymbol{A}\hat{\xi} + \tilde{\boldsymbol{E}}_A\hat{\xi} - \tilde{e}_y\|$ turns out to be relatively large when compared to those of other examples. The difference in the values of the model checks from Tab. 5 and 6 are reflected in the corresponding residuals, which differ at the level of 1×10^{-6} m for some elements (well below the precision of the data).

It is conjectured that the relatively large value for the model check and the sensitivity of the algorithm to the convergence criterion are due, at least in part, to the contrived constraints of this experiment. We leave further details regarding the numerical properties of the new algorithm for a future study.

5.2.3 A rigid transformation, weighted case for singular Q_1

In this final example, the similarity transformation is effectively transformed to a rigid transformation by use of a quadratic constraint only in the model. The model is modified by simply removing the linear constraint $\boldsymbol{K}\xi = \kappa_0$ from Equation (47c). Algorithm 1 is modified accordingly by omitting all terms associated with the linear constraint. Finally, the assignment $\alpha_0^2 = 1$ is made, implying $(\omega \cos\alpha)^2 + (\omega \sin\alpha)^2 = 1 \Rightarrow \omega = 1$, which ensures a rigid transformation.

The identity matrix was used for $\boldsymbol{S}$. The solution is presented in Tab. 7, and the residuals follow. Note the relatively large increase in the TSSR value compared to the unconstrained case of chapter 5.2.1. This confirms our expectation that the TSSR should increase when constraints are added to the model, particularly when the residuals prove to be so sensitive to a relatively small change of α_0^2 in the quadratic constraint.

Tab. 6: Solution for the 2-dimensional similarity transformation with linear and quadratic constraints. Linear quantities are shown in units of meters.

Quantity	Estimate/Calculation
$\hat{\xi}_1 := \hat{t}_1$	−69.726354
$\hat{\xi}_2 := \hat{t}_2$	35.078215
$\hat{\xi}_3 := \widehat{\omega\cos\alpha}$	0.98765501
$\hat{\xi}_4 := \widehat{\omega\sin\alpha}$	−0.15642921
$\hat{\xi}_1 / \hat{\xi}_2$	−1.9877395
$\hat{\omega}$	0.99996626
Ω (TSSR)	6.164035
$\hat{\sigma}_0^2$	$(0.877784)^2$
redundancy	8
$\lVert \boldsymbol{y} - \boldsymbol{A}\hat{\xi} + \tilde{\boldsymbol{E}}_A\hat{\xi} - \tilde{e}_y \rVert$	$< 3 \times 10^{-6}$
$\boldsymbol{K}\hat{\xi}$	0
$\hat{\xi}^T \boldsymbol{M}\hat{\xi}$	$(0.99996626)^2$
δ (convergence threshold)	1×10^{-9}
$\boldsymbol{S}$	$1 \times 10^3 \cdot \boldsymbol{I}_4$
Number of iterations	5

Tab. 7: Solution for the 2-dimensional similarity transformation with a quadratic constraint. Linear quantities are shown in units of meters.

Quantity	Estimate/Calculation
$\hat{\xi}_1 := \hat{t}_1$	−69.738828
$\hat{\xi}_2 := \hat{t}_2$	35.070627
$\hat{\xi}_3 := \widehat{\omega\cos\alpha} = \widehat{\cos\alpha}$	0.98768834
$\hat{\xi}_4 := \widehat{\omega\sin\alpha} = \widehat{\sin\alpha}$	−0.15643449
$\hat{\omega}$	1
Ω (TSSR)	15.719221
$\hat{\sigma}_0^2$	$(1.498534)^2$
redundancy	7
$\lVert \boldsymbol{y} - \boldsymbol{A}\hat{\xi} + \tilde{\boldsymbol{E}}_A\hat{\xi} - \tilde{e}_y \rVert$	$< 4 \times 10^{-9}$
$\hat{\xi}^T \boldsymbol{M}\hat{\xi}$	1
δ (convergence threshold)	1×10^{-10}
$\boldsymbol{S}$	$\boldsymbol{I}_4$
Number of iterations	5

Finally, the residuals are given by

$$[\tilde{e}_y \mid \tilde{E}_A] = \\ = \left[\begin{array}{r|cc rr} 0.4653 & 0 & 0 & -0.9131 & 9.3364 \\ 1.7878 & 0 & 0 & -9.3364 & -0.9131 \\ -0.7649 & 0 & 0 & 3.5299 & -2.5409 \\ -0.3854 & 0 & 0 & 2.5409 & 3.5299 \\ -2.1356 & 0 & 0 & 9.3052 & -10.5701 \\ -1.7694 & 0 & 0 & 10.5701 & 9.3052 \\ 2.1500 & 0 & 0 & -10.4468 & 3.8332 \\ 0.4238 & 0 & 0 & -3.8332 & -10.4468 \\ 0.2852 & 0 & 0 & -1.4753 & -0.0585 \\ -0.0568 & 0 & 0 & 0.0585 & -1.4753 \end{array}\right] \times 10^{-3}\text{m},$$

where the structure of the data matrix $\boldsymbol{A}$ is replicated in the residual matrix $\tilde{\boldsymbol{E}}_A$.

6 Conclusions and outlook

In this contribution, we have presented a new algorithm for solving the weighted TLS problem within an EIV-Model having linear and quadratic constraints and (significantly) singular dispersion matrices. The flexibility of the algorithm has been demonstrated through a few numerical examples.

The numerical experiments revealed that, in some cases, the convergence properties of the algorithm were affected by the choice of the matrix $\boldsymbol{S}$ required by the algorithm. Equation (43c) reveals that $\boldsymbol{S}$ plays an apparent balancing role with the matrix $\boldsymbol{Q}_1$. It would be interesting to develop guidelines for the choice of $\boldsymbol{S}$, perhaps based on theoretical considerations or even on empirical studies.

References

Fang, X.: Weighted Total Least Squares Solutions for Applications in Geodesy. PhD thesis, Report. No. 294, Dept. of Geodesy and Geoinformatics, Leibniz University Hannover, Germany, 2011.

Fang, X.: A structured and constrained Total Least-Squares solution for geodetic applications. Studia Geophysica et Geodaetica, in print, 2014.

Felus, Y.; Burtch, R.: On symmetrical three-dimensional datum conversion. GPS Solutions, 13:65–74, 2009.

Felus, Y.; Schaffrin, B.: Fitting variogram models by Total Least-Squares. Poster Paper, pres. at the Heiskanen-Symp., Columbus, Ohio, USA, on CD., 2002.

Felus, Y.; Schaffrin, B.: A Total Least-Squares approach in two stages for semivariogram modeling of aeromagnetic data. In Cheng, Q.; Bonham-Carter, G., editors, GIS and Spatial Analysis, pages 215–220. China University of Geosciences Printing House: Beijing, PR China. Proc. of the IAMG ’05, 2005.

Golub, G.; Van Loan, C.: An analysis of the total least-squares problem. SIAM Journal on Numerical Analysis, 17:883–893, 1980.

Magnus, J.; Neudecker, H.: Matrix Differential Calculus with Applications in Statistics and Econometrics. Wiley: Chichester, England, third edition, 2007.

Mahboub, V.: On weighted total least-squares for geodetic transformations. Journal of Geodesy, 86:359–367, 2012.

Markovsky, I.; Rastello, M.; Premoli, A.; Kukush, A.; Van Huffel, S.: The element-wise weighted total least-squares problem. Computational Statistics & Data Analysis, 50:181–209, 2006.

Neitzel, F.; Schaffrin, B.: On the Gauss-Helmert Model with a singular dispersion matrix, where BQ is of lesser row rank than B. Studia Geophysica et Geodaetica. Revised in April, 2013, 2013.

Schaffrin, B.: A note on Constrained Total Least-Squares estimation. Linear Algebra and its Applications, 417:245–258, 2006.

Schaffrin, B.: Connecting the dots: The straight-line case revisited. zfv – Zeitschrift für Geodaesie, Geoinformation und Landmanagement, 132:385–394, 2007.

Schaffrin, B.; Felus, Y.: On Total Least-Squares Adjustment with Constraints. In Sansò, F., editor, A Window on the Future of Geodesy, pages 417–421. IAG-Symp., Vol. 128, Springer: Berlin/Heidelberg/New York, 2005.

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

Schaffrin, B.; Lee, I.; Felus, Y.; Choi, Y.: Total Least-Squares (TLS) for geodetic straight-line and plane adjustment. Bollettino di Geodesia e Scienze Affini, 65:141–168, 2006.

Schaffrin, B.; Neitzel, F.; Uzun, S.; Mahboub, V.: Modifying Cadzow’s algorithm to generate the optimal TLS-solution for the structured EIV-Model of a similarity transformation. Journal of Geodetic Science, 2:98–106, 2012.

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

Schaffrin, B.; Snow, K.; Neitzel, F.: On the Errors-In-Variables Model with singular covariance matrices. Journal of Geodetic Science, in print, 2014.

Schaffrin, B.; Wieser, A.: On weighted total least-squares adjustment for linear regression. Journal of Geodesy, 82:415–421, 2008.

Snow, K.: Topics in Total Least-Squares Adjustment within the Errors-In-Variables Model: Singular Cofactor Matrices and Prior Information. PhD thesis, Report. No. 502, Div. of Geodetic Science, School of Earth Sciences, The Ohio State University, Columbus, Ohio, USA, 2012.

Van Huffel, S.; Vandewalle, J.: The Total Least-Squares Problem: Computational Aspects and Analysis. Society for Industrial and Applied Mathematics: Philadelphia/PA, USA, 1991.

Authors’ addresses

Dr. Shahram Jazaeri
Department of Surveying and Geomatics Engineering
College of Engineering
University of Tehran, Iran
jazayeri@alumni.ut.ac.ir

Prof. Dr. habil. Burkhard Schaffrin
Geodetic Science Program, School of Earth Sciences
The Ohio State University
275 Mendenhall Laboratory, 125 South Oval Mall, Columbus, Ohio 43210-1308, USA
aschaffrin@earthlink.net

Dr. Kyle Snow
Topcon Positioning Systems, Inc.
4400 Easton Commons, Suite 100, Columbus, Ohio 43219, USA
ksnow6378137@gmail.com

This article also is digitally available under www.geodaesie.info.