Veröffentlichungen der DGK

Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften

Reihe C Dissertationen Heft Nr. 841

Georgios Malissiovas

New nonlinear adjustment approaches for applications in Geodesy and related fields

München 2020

Verlag der Bayerischen Akademie der Wissenschaften

ISSN 0065-5325 ISBN 978-3-7696-5253-6

Diese Arbeit ist gleichzeitig veröffentlicht in:
Wissenschaftliche Arbeiten der Fachrichtung Geodäsie und Geoinformationstechnik
der Technischen Universität Berlin
http://dx.doi.org/10.14279/depositonce-9194, Berlin 2019

Veröffentlichungen der DGK

Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften

Reihe C Dissertationen Heft Nr. 841

New nonlinear adjustment approaches for applications in Geodesy and related fields

Von der Fakultät VI – Planen Bauen Umwelt

der Technischen Universität Berlin

zur Erlangung des Grades

Doktor der Ingenieurwissenschaften (Dr.-Ing.)

genehmigte Dissertation

Vorgelegt von

M.Sc. Georgios Malissiovas

Geboren am 10.11.1984 in Tripoli, Griechenland

München 2020

Verlag der Bayerischen Akademie der Wissenschaften

ISSN 0065-5325 ISBN 978-3-7696-5253-6

Diese Arbeit ist gleichzeitig veröffentlicht in:
Wissenschaftliche Arbeiten der Fachrichtung Geodäsie und Geoinformationstechnik
der Technischen Universität Berlin
http://dx.doi.org/10.14279/depositonce-9194, Berlin 2019

Adresse der DGK:

Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften (DGK)
Alfons-Goppel-Straße 11 ● D – 80 539 München
Telefon +49 – 331 – 288 1685 ● Telefax +49 – 331 – 288 1759
E-Mail post@dgk.badw.de ● http://www.dgk.badw.de

Prüfungskommission:

Vorsitzender: Prof. Dr.-Ing. Martin Kada, Technische Universität Berlin

Referent: Prof. Dr.-Ing. Frank Neitzel, Technische Universität Berlin

Korreferenten: Prof. Dr. techn. Wolf-Dieter Schuh, Universität Bonn
Prof. Dr. Andreas Wieser, ETH Zürich
Priv.-Doz. Dr. techn. habil. Svetozar Petrović, GeoForschungsZentrum Potsdam

Tag der mündlichen Prüfung: 10.05.2019

ISSN 0065-5325 ISBN 978-3-7696-5253-6

Summary

This dissertation deals with a class of nonlinear adjustment problems that has a direct least squares solution for certain weighting cases. In the literature of mathematical statistics these problems are expressed in a nonlinear model called Errors-In-Variables (EIV) and their solution became popular as total least squares (TLS). The TLS solution is direct and involves the use of singular value decomposition (SVD), presented in most cases for adjustment problems with equally weighted and uncorrelated measurements. Additionally, several weighted total least squares (WTLS) algorithms have been published in the last years for deriving iterative solutions, when more general weighting cases have to be taken into account and without linearizing the problem in any step of the solution process.

This research provides firstly a well defined mathematical relationship between TLS and direct least squares solutions. As a by-product, a systematic approach for the direct solution of these adjustments is established, using a consistent and complete mathematical formalization. By transforming the problem to the solution of a quadratic or cubic algebraic equation, which is identical with those resulting from TLS, it will be shown that TLS is an algorithmic approach already known to the geodetic community and not a new method.

A second contribution of this work is the clear overview of weighted least squares solutions for the discussed class of problems, i.e. the WTLS solution in the terminology of the statistical community. It will be shown that for certain weighting cases a direct solution still exists, for which two new solution strategies will be proposed. Further, stochastic models with more general weight matrices are examined, including correlations between the measurements or even singular cofactor matrices. New algorithms are developed and presented, that provide iterative weighted least squares solutions without linearizing the original nonlinear problem.

The aim of this work is the popularization of the TLS approach, by presenting a complete framework for obtaining a (weighted) least squares solution for the investigated class of nonlinear adjustment problems. The proposed approaches and the implemented algorithms can be employed for obtaining direct solutions in engineering tasks for which efficiency is important, while iterative solutions can be derived for stochastic models with more general weights.

Zusammenfassung

Die vorliegende Dissertation beschäftigt sich mit einer Klasse von nichtlinearen Ausgleichungsproblemen, die eine direkte Lösung nach der Methode der kleinsten Quadrate unter spezifischen Gewichtungsfällen aufweisen. In der Literatur der mathematischen Statistik werden derartige Probleme in einem nichtlinearen Modell namens Errors-In-Variables (EIV) ausgedrückt und deren Lösung wurde als Total Least Squares (TLS) populär. In den meisten Fällen lässt sich für gleich gewichtete und unkorrelierte Messungen eine TLS Lösung direkt durch eine Singulärwertzerlegung (SVD) bestimmen. Darüber hinaus wurden in den letzten Jahren mehrere Weighted Total Least Squares (WTLS) Algorithmen zur Herleitung iterativer Lösungen veröffentlicht, bei denen allgemeinere Gewichtungsfälle berücksichtigt werden können, ohne das Problem in jedem Schritt des Lösungsprozesses zu linearisieren.

Zunächst wird in dieser Arbeit eine klar definierte mathematische Beziehung zwischen TLS und direkter Lösungen nach der Methode der kleinsten Quadrate dargestellt. Des Weiteren wird ein systematischer Ansatz zur direkten Lösung derartiger Ausgleichungsproblemen unter Verwendung einer konsistenten und vollständigen mathematischen Formalisierung entwickelt. Durch die Überführung des Problems in die Lösung einer quadratischen oder kubischen algebraischen Gleichung wird gezeigt, dass TLS ein algorithmischer Ansatz ist, der der geodätischen Gemeinschaft bereits bekannt ist und keine neue Methode darstellt.

Ein weiterer Beitrag dieser Arbeit besteht in einer klaren Übersicht von gewichteten Kleinste-Quadrate Lösungen für die hier diskutierte Klasse von Problemen, wie z.B. der WTLS-Lösung aus der Terminologie der statistischen Gemeinschaft. Es wird gezeigt, dass für bestimmte Gewichtungsfälle noch eine direkte Lösung existiert, wofür zwei neue Lösungsstrategien vorgestellt werden. Weiterhin werden stochastische Modelle mit allgemeineren Gewichtsmatrizen untersucht, einschließlich Korrelationen zwischen den Messungen oder sogar singulären Kofaktor-Matrizen. Es werden neue Algorithmen entwickelt und vorgestellt, die gewichtete Kleinste-Quadrate Lösungen iterativ berechnen, ohne das ursprüngliche nichtlineare Problem zu linearisieren.

Das Ziel dieser Arbeit ist die Popularisierung des TLS-Ansatzes, indem eine umfassende Strategie zur Berechnung einer (gewichteten) Kleinste-Quadrate Lösung für die betrachtete Klasse an nichtlinearen Ausgleichungsproblemen bereitgestellt wird. Die vorgeschlagenen Ansätze und implementierten Algorithmen können zur Berechnung direkter Lösungen in vielen Ingenieuraufgaben eingesetzt werden, bei denen Effizienz wichtig ist, während für stochastische Modelle mit allgemeineren Gewichten auf die iterativen Lösungsansätze zurückgegriffen werden kann.

Contents

List of Figures

List of Tables

Abbreviations

CTLS	Constrained Total Least Squares
EIV	Errors In Variables
EVD	Eigen Value Decomposition
GHM	Gauss Helmert Model
GMM	Gauss Markov Model
MCS	Monte Carlo Simulation
NS	Neitzel-Schaffrin
STLS	Structured Total Least Squares
SVD	Singular Value Decomposition
TLS	Total Least Squares
UT	Uncented Transformation
VC	Variance Covariance
WTLS	Weighted Total Least Squares
2D	Two dimensionional
3D	Three dimensional

1 Introduction and Motivation

In geodetic practice, engineers are engaged in performing measurements for the numerical description of reality, including the characteristics of some physical phenomena or geometrical properties of real objects. The desired values often cannot be measured directly, but they are linked to the measured values via a mathematical model. The mathematical modelling of the measurement results, together with the errors that influence them, results in an under-determined algebraic problem. The target is in most cases the "optimal" estimation of some unknown parameters. For more than two centuries mathematicians and geodesists have solved these adjustment problems using the method of least squares, according to the fundamental studies of Gauss (1809) and Gauss (1823). A least squares estimate can be obtained by minimizing a defined objective function (the sum of squared residuals) and thus by solving a system of normal equations, i.e. a system of equations that follow from the partial derivatives of the objective function with respect to all unknowns. Depending on the nature of the problem, a least squares adjustment can be linear or nonlinear. Obviously linear least squares problems can be solved using the rules of linear algebra, whilst the nonlinear cases require most of the time a numerical method for obtaining a solution. This thesis investigates only the second type of problems.

The solution of nonlinear adjustment problems with the use of least squares has a long history and the simplicity of the "recipe" of this method is recognized by its wide application in all scientific fields that deal with redundant observations and seeking for an "optimal" solution. Helmert (1924), for instance, proposed a least squares solution by linearizing the functional model of the nonlinear adjustment problem. Deming (1931) and Deming (1934) tackled the same problem by developing algorithms based on the iterative linearization of the functional model following the Gauss-Newton approach. Pope (1974) pointed out that a least squares solution of a nonlinear adjustment problem can be obtained either by using the Gauss-Newton approach, or by setting up and linearizing the nonlinear normal equations, according to the Newton-Raphson approach. It is often in geodetic literature that the least squares principle is applied in the form of two adjustment models, namely the Gauss-Markov Model (GMM), see (Niemeier 2008, p. 137 ff.), and the Gauss-Helmert Model (GHM), see e.g. (Niemeier 2008, p. 172 ff.). In (Krakiwsky 1975, pp. 7-26) these models can be found under the name parametric (case) adjustment and combined (case) adjustment, as well. A least squares solution from both models is based on the Gauss-Newton approach.

Furthermore, a class of nonlinear least squares problems exists, for which a direct solution is possible for the nonlinear normal equations, especially for such cases where the solving of the normal equations can be converted into an eigenvalue problem. Thus, a solution is obtained by computing the roots of a polynomial (i.e. the characteristic equation of the eigenvalues) and a direct solution can be possible depending on the polynomial's degree. Such adjustment problems have been discussed in geodetic literature since long time. Linkwitz (1960), for instance, presented a least squares solution for two adjustment problems that belong to this class, the fitting of a straight line in two dimensions (2D) and the fitting of a plane in three dimensions

(3D), while the coordinates of the points in all directions are considered as measurements. Another example is the work of Jovičić et al. (1982), who investigated the fitting of a straight line in 3D. Therefore, it can be seen from past publications that the solution of such nonlinear least squares problems, using eigenvalue decomposition (EVD), has been already a standard procedure for the members of the geodetic community. Nevertheless, an alternative approach was developed during the last decades for the solution of this class of nonlinear least squares problems by the mathematical community. This is called total least squares (TLS) and has been firstly defined and presented by Golub and Van Loan (1980). Since then many researchers dealt with the solution of adjustment problems using TLS and developed modern and sophisticated algorithms. As it was defined in (Golub and Van Loan 1980) or (Van Huffel and Vandewalle 1991, p. 33 ff.), the TLS solution is related only with this class of nonlinear adjustment problems which can be expressed within an errors in variables (EIV) model and solved with the use of singular value decomposition (SVD), without involving any kind of linearization of the functional model.

Within literature the work coming from the TLS community is often distinguished from the classical least squares by stating that TLS functions differently. There are expectations that TLS might produce a "more realistic" result than the classic least squares, as indicated for example in (Felus and Schaffrin 2005) or (Schaffrin et al. 2006). Petrovic (2003) has already pointed out that this view has been caused possibly by the work of Golub and Van Loan (1980), where the solution of TLS was compared with that of least squares for fitting a straight line in 2D. In that study, for the least squares solution it was assumed that only the y-coordinates are regarded as observed values and the x-coordinates as error free, which led to the misleading conclusion that TLS functions differently from least squares. For geodesists it has already been clear that the most important steps for the adjustment of observations is to formulate a correct model and minimize the correct objective function. When these requirements are fulfilled, then for a linear problem the solution will be unique, regardless of the solution strategy that has been followed.

Contrary to the belief that TLS is an additional method like least squares (or even a generalisation of it), several scientists have shown that TLS describes just a particular algorithmic approach to find the (weighted) least squares solution. One of the first critical views on TLS can be found in the appendix of (Petrovic 2003). This author concluded that least squares and TLS are not individual methods, but applications of minimizing the sum of squared residuals. Afterwards, Neitzel and Petrovic (2008) and Neitzel (2010) showed on two practical examples that in fact TLS can be regarded as a special case of the least squares method within the GHM. The iterative solution of the GHM has been proven to be numerically equivalent to the TLS solution in both cases. Other contributions that follow this line of thinking are (Reinking 2008) and (Mihajlovic and Cvijetinovic 2016). These studies provided the motivation for the first part of this research, which has been partly published in (Malissiovas et al. 2016). In that article, a clear mathematical relationship between TLS and direct least squares solutions has been presented.

A TLS solution has been mainly investigated when the measurements are equally weighted and uncorrelated. Consequently, the following scientific challenges focused on the algorithmic development of TLS when individual precisions or correlations are postulated for the measurements. The solutions from these algorithms are iterative, they do not include a linearization of the functional model and have been published under the name "weighted TLS" (WTLS). For instance, Schaffrin and Wieser (2008) developed a WTLS algorithm for the weigthed least squares solution of fitting a straight line in 2D, which was the basis for further algorithmic developments from Shen et al. (2011), Fang (2011), Amiri-Simkooei and Jazaeri (2012) and Mahboub (2012). Additionally, various names for the TLS solution emerged due to algorithmic complications caused by the stochastic model of each problem. For instance, the term STLS (structured TLS) was presented in

(Schaffrin et al. 2012) for the solution of a 2D similarity transformation of coordinates, or the term CTLS (constrained TLS) presented in (Abatzoglou et al. 1991) or (Schaffrin 2006). Despite the name TLS, in all above cases the solution has been obtained iteratively and does not follow the definition of TLS that was established by Golub and Van Loan (1980). A clear overview of the latest WTLS algorithms has been presented in the dissertation of Snow (2012), who also covered special cases of singular cofactor matrices being present in the model.

Nevertheless, York (1966) and Williamson (1968) had already presented weighted least squares solutions for the problems that have been tackled in terms of WTLS, i.e. iterative solutions of the nonlinear normal equations without linearizing the functional model. York (1968) and Petrović et al. (1983) developed also algorithms even for the case of correlated observations. These contributions provide the motivation for the next research question of this dissertation. In this second part, the solution of the discussed class of nonlinear least squares problems is investigated for various cases where the observations have different precisions or correlations.

1.1 Research contributions

In the first place a mathematical relationship between direct least squares solutions and TLS is developed by investigating four individual adjustment problems, namely: the fitting of a straight line in 2D, the fitting of a straight line in 3D, the fitting of a plane in 3D and the 2D similarity transformation of coordinates. As a consequence, a systematic approach is established for the direct least squares solution of these nonlinear adjustment problems. This leads to a deeper understanding of the underlying principles of TLS adjustment and to the class of problems that can be solved via SVD, i.e. the solving of nonlinear normal equations can be transformed into an eigenvalue problem.

The second part of this study focuses on the solution of the class of nonlinear least squares problems, when individual precisions and correlations between the measurements have to be taken into account. Novel algorithms are implemented to provide direct and iterative weighted least squares solutions, depending on the stochastic model of each problem. In contrast to the known WTLS algorithms, the presented solutions do not always need iterations but can have in some weighted cases a direct solution. When iterations are necessary, the developed algorithms can provide a weighted least squares solution without performing any linearization of the problem under investigation. The cases of singular cofactor matrices in the stochastic model are also covered (when the criterion of Neitzel and Schaffrin (2016) is fulfilled) by the developed approaches, without the need of any special treatment of the problem.

From an engineering point of view, the proposed solutions and algorithms can be an asset for providing the (weighted) least squares solution for this class of problems without any iteration or starting values for the unknown parameters. Thus, this is an advantageous algorithmic option for researchers and engineers in terms of efficiency. The adjustment approaches that were developed so far have a significant impact on several applications in geodesy. For instance, 3D point clouds obtained by terrestrial laser scanners can be easily handled, depending on the needs of the engineering task. The developed algorithms can handle efficiently the vast amount of data, providing for example direct solutions for fitting planes or estimating the transformation parameters between several data sets. On the other hand, the presented algorithms that provide iterative weighted least squares solutions can be utilized also when correlations between the observations have to be taken into account.

One of the objectives from the contributions of this dissertation is to present a simple explanation of the TLS and WTLS solutions to scientists, researchers and engineers from all scientific fields that deal with adjustment calculus. Thus, it can enable them to use appropriately the TLS/WTLS algorithms for the solution of nonlinear adjustment problems. Furthermore, the proposed solution strategies, accompanied by individual algorithms for each case, can be utilized for the solution of this class of problems in the future without the need of TLS or WTLS theory and practice.

1.2 Organization of this thesis

This thesis is organized into two main parts, the part describing fundamental ideas of adjustment and the methodological contributions, and involves six chapters. The content of each chapter is summarized in the following.

Fundamentals

Chapter 2 describes fundamental requirements for the adjustment of observations and gives a definition of the method of least squares and hence sets the basis for the definition of nonlinear least squares problems in this dissertation. These requirements are deduced from a review of related work on adjustment calculus. Since the mathematical formalization of nonlinear normal equations with a solution using EVD or SVD is one of the main goals of this thesis, basic mathematical notions and concepts which serve as foundation for this purpose are briefly discussed.

Chapter 3 presents a part of existing solution strategies and numerical methods for the solution of nonlinear least squares problems, based on the requirements of chapter 2. In this core chapter of the thesis, both the traditional geodetic and the most modern solution strategies are elaborated in depth. Three fundamental models, namely the GMM, the GHM and the EIV model, are defined and mathematically described. Based on these models, further concepts such as the Gauss-Newton, the TLS and the WTLS approach are defined and formalized. Scientific questions emerge from the developed concepts and are outlined at the end of this chapter.

Methodological contributions

Chapter 4 is dedicated to the development of a well defined mathematical relation between direct least squares solutions and TLS solutions for a class of nonlinear adjustment problems. It involves the elaboration of both solutions for four adjustment problems that often occur in practice. Based on this, a new strategy for the direct solution of such problems is elaborated and presented in detail. Subsequently, common features are identified between the investigated adjustment problems that can serve as classification factors for others that belong to this group. The chapter concludes with a discussion on the nature of TLS adjustment and raises further scientific questions.

Chapter 5 addresses different solutions of the same class of nonlinear least squares problems as in chapter 4, by postulating individual stochastic models for each case. In contrast to the iterative algorithms known as WTLS, various weighting cases are identified in this chapter that can still lead to a direct solution, based on the same concepts as in chapter 4. Additionally, a standard solution strategy is presented for cases where iterations are necessary for a weighted least squares solution, without applying any linearization to the investigated problems. From this, individual algorithms are designed that fit into the framework of the proposed solution approach. The feasibility of the developed algorithms is also demonstrated on cases where the cofactor matrix of the problem is singular.

Chapter 6 illustrates the application of the developed methodologies and algorithms for two adjustment cases. The first case presents the proposed direct and iterative least squares solution for fitting a straight line to a set of measured points in 2D, while postulating various weights and correlations between the observations. In the same line of thinking the 2D similarity transformation of coordinates is examined as a second example.

Chapter 7 summarizes this work and draws conclusions with respect to the stated research goals as well as the scientific questions that emerged in each chapter. It reviews and evaluates the results of this investigation, lists and discusses contributions to the field of adjustment calculus and identifies and outlines scientific problems that could be tackled in a future research.

Part I - Fundamentals

2 Adjustment calculus

This chapter briefly summarizes the concept of adjustment calculus. This is necessary for understanding the fundamentals of the adjustment of observations and the estimation of unknown parameters. Many parts of this chapter have been adopted from the work of Pasioti (2015) and have been extended appropriately to fit the needs of this dissertation.

It can be said that adjustment calculus is a part of geodesy and all scientific fields that deal with redundant measurements. The measurement results are necessary for describing numerically the characteristics of some physical phenomena or geometrical properties of real objects. The target is in most cases the "optimal" estimation of some unknown parameters by means of an under-determined algebraic system that occurs from the mathematical modelling of the measured quantities[1]. Additionally, the statistical properties of the estimated parameters are of great importance for drawing conclusions concerning the precision and the reliability of the adjustment results. Based on the adjustment phases presented by Neitzel and Petrovic (2008), five basic steps can be identified for the gradual process of any adjustment problem. These can be summarized with the following systematic stages:

1. Definition of the problem/task:

   In this step it must be clearly defined what are the measurements that are subject to errors, which parameters are fixed (error free) and which parameters are unknown.

2. Clear description of the mathematical model:

   The mathematical model is the combination of an appropriate functional model, which is selected by the user, with the stochastic model that incorporates the stochastic properties of the measurements.

3. Selection of the method for a solution of the adjustment problem:

   A solution of an adjustment problem can be derived by selecting an appropriate criterion for the unknown errors, depending on their distribution. The choice of such a criterion implies the method that will be followed. For example, in case of normally distributed errors the least squares method should be employed.

4. Calculation of the adjustment results:

   A least squares solution for an adjustment problem can be obtained by solving a system of equations (i.e. the normal equations). Various approaches exist for the solution of the normal equations in numerical mathematics, that depend on the nature (linear or nonlinear) and the well-posedness of the problem.

[1] As it will be explained in subsection 2.1.1, the functional model of an adjustment problem includes unknown parameters to be estimated and residuals, which are also unknown, leading to an under-determined set of equations

5. Computation of stochastics:

   The precision measures associated with the adjustment results can be computed after obtaining the solution of the problem using, for example, the rules of error propagation.

These fundamental stages for the solution of adjustment problems serve as a basis for the explanation of the theory of adjustment calculus that is presented in this chapter.

2.1 Mathematical modelling of adjustment problems

Under the theoretical assumption that measurements are contaminated by errors (Taylor 1982, p. 3), an under-determined algebraic system emerges that is strongly related to the stochastic properties of the measurements. Important for the solution of an adjustment problem is the clear description of the deterministic part (the derived under-determined system of equations) and the stochastic part[2] (the measurement errors and their disribution). The first is described by the *functional model* and the second by the *stochastic model*, which are always combined to form the *mathematical model* of the problem, as it has been defined in (Mikhail and Ackermann 1976, p. 5) or (Perović 2005, p. 72).

For obtaining meaningful adjustment results, a correct formulation of the mathematical model is essential in every case and should clearly answer the following questions:

- What are the observations that are subject to measurement errors?
- What is the precision of the observations?
- What are the unknown parameters to be estimated?
- Which parameters are error-free or fixed/constant?

The two fundamental parts of a mathematical model, i.e. the functional and stochastic model, are explained in the following subsections.

2.1.1 The functional model

Observations l_i (with $i = 1, ..., n$) and unknown parameters x_j ($j = 1, ..., m$) can be mathematically related through a set of functions, which are often referred as the *functional model*. Three main categories of functional relationships are disinguished in the geodetic literature, see for example (Wells and Krakiwsky 1971, pp. 102-104) or (Perović 2005, p. 63):

1. Implicit functional relationships between the observations

$$f(l_i) \approx 0.$$

[2]see (Perović 2005, p. 55) for a definition of the deterministic and stochastic model.

2. Explicit functional relationships between the observations and the unknown parameters

$$l_i \approx f(x_j).$$

3. Explicit or implicit functional relationships between the observations and the unknown parameters

$$f(l_i, x_j) \approx 0.$$

The presented equations are approximately equal, as the observed quantities are subject to errors and have to be corrected, as it is explained below. According to (Wells and Krakiwsky 1971, p. 104), the mathematical relationships of categories 1 and 2 are only special cases of the most general case in 3. This study investigates adjustment problems in which some unknown parameters need to be estimated [3].

Explicit functional relationships between observations and unknown parameters

Let a number of observations $l_1, ..., l_n$ be performed and $x_1, ..., x_m$ unknown parameters to be estimated. The mathematical relationship between the observations and the unknown parameters can be expressed by the equation system

$$\begin{aligned} l_1 &\approx f_1(x_1, ..., x_m), \\ l_2 &\approx f_2(x_1, ..., x_m), \\ &\vdots \\ l_n &\approx f_n(x_1, ..., x_m). \end{aligned} \tag{2.1}$$

Depending on the number of equations in (2.1), three individual cases can emerge (Helmert 1924, p. 47):

1. The number of observations is smaller than the number of the parameters to be estimated ($n < m$). There is no unique solution (there are infinitely many solutions).
2. The number of observations is equal to the number of the parameters to be estimated ($n = m$). There exists a unique solution, but the presence of blunders cannot be identified.
3. The number of observations is larger than the number of the parameters to be estimated ($n > m$). This is an overdetermined system with no unique solution.

A usual geodetic problem consists of repeated observations, seeking for a solution to a set of unknown parameters. Obviously, the third case presented above is what geodesists usually face. Furthermore, in equation (2.1) the measured values l have to be corrected for their random errors. According to (Bjerhammar 1973, p. 1), the true value $\tilde{l}$ can be obtained by excluding the error (e) from the measurement:

$$\tilde{l} = l - e. \tag{2.2}$$

[3] Category 1 leads to the famous adjustment of condition equations (when taking into account the necessary residuals). However, this adjustment problem is out of the scope of this work. More information for this case can be found in most common literature of adjustment calculations in geodetic science, see for example (Helmert 1924, p. 228 ff.).

However, because the true values are only a theoretical concept, the adjusted value of a measurement $\hat{l}$ is usually considered by adding a correction/residual v

$$\hat{l} = l + v. \tag{2.3}$$

Thus, including the necessary residuals in the functional model yields the under-determined[4] system of equations

$$\begin{aligned} l_1 + v_1 &= f_1(x_1, ..., x_m), \\ l_2 + v_2 &= f_2(x_1, ..., x_m), \\ &\vdots \\ l_n + v_n &= f_n(x_1, ..., x_m). \end{aligned} \tag{2.4}$$

The developed equations (with included residuals) are the *observation equations*. As claimed in (Wells and Krakiwsky 1971, p. 102), adjustment problems of this type are widely known as *adjustment of observation equations*, *parametric adjustment* or *adjustment of indirect observations*[5].

Implicit functional relationships between observations and unknown parameters

A number of observations has been collected and some unknown parameters need to be estimated. Assuming that the mathematical relationship between the observations and the unknown parameters can be expressed by

$$\begin{aligned} f_1(l_1, ..., l_n,\ x_1, ..., x_m) &\approx 0, \\ f_2(l_1, ..., l_n,\ x_1, ..., x_m) &\approx 0, \\ &\vdots \\ f_r(l_1, ..., l_n,\ x_1, ..., x_m) &\approx 0. \end{aligned} \tag{2.5}$$

Depending on the number of equations r in (2.5), an adjustment problem exists when the number of equations is larger than the number of the parameters to be estimated ($r > m$). Adding the necessary residuals to the observations yields

$$\begin{aligned} f_1(l_1 + v_1, ..., l_n + v_n,\ x_1, ..., x_m) &= 0, \\ f_2(l_1 + v_1, ..., l_n + v_n,\ x_1, ..., x_m) &= 0, \\ &\vdots \\ f_r(l_1 + v_1, ..., l_n + v_n,\ x_1, ..., x_m) &= 0. \end{aligned} \tag{2.6}$$

This system of equations is known as *condition equations with unknowns*. Adjustment problems of this type can be found under the name *combined adjustment* or *adjustment of condition equations and unknown parameters* (Wells and Krakiwsky 1971, p. 102) [6].

[4]The developed functional model is under-determined, because v_i are also unknown quantities.

[5]More information from the traditional German literature can be found in (Helmert 1924, p. 43) or (Linkwitz 1960, p. 156). It is worth noticing that the German term “Vermittelnde Beobachtungen” for this type of adjustment cannot be translated literally.

[6]In the German literature this type of adjustment problems is known as “Bedingte Beobachtungen mit Unbekannten”, see for example (Helmert 1924, p. 52) or (Linkwitz 1960, p. 192).

Constraints between the unknown parameters

Additionally, part of the functional model can be a number of constraints that have to be enforced on the unknown parameters. For example, such a constraint can be formulated as

$$u(x_1, x_2, ..., x_m) = 0, \tag{2.7}$$

with u denoting here a function of the unknowns. This is just a special case of the condition equations (2.6). Adjustment problems with imposed constraints can be divided into two cases:

- Adjustment of observation equations with constraints between the unknown parameters, see for example (Helmert 1924, p. 262) or (Linkwitz 1960, p. 197).
- Adjustment of condition equations and unknowns with constraints between the unknown parameters, as explaned in (Wells and Krakiwsky 1971, p. 142) and (Mikhail and Ackermann 1976, p. 213 ff.).

2.1.2 The stochastic model

In a complete mathematical model the stochastic properties of the obtained observations must be taken into account. In the literature, the theoretical error that can influence an observation is denoted with the term "standard deviation" of the measurement, as for example in (Niemeier 2008, p. 6). The measurement standard deviations are often based on the physical characteristics of some instrument that has been used for measurements and imply how precise the observations are. This *a priori*[7] information influences the adjustment procedure in a sense of how much an observation contributes to the adjustment (i.e. a precise observation is more valuable than another with a low precision).

Theoretical variances and covariances of a random variable

In statistics the precision of an observation can be expressed in terms of its variance. Based on the theory of errors the *theoretical variance* of a discrete random variable is defined as

$$\sigma^2 = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} e_i^2 = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} (l_i - \tilde{l})^2 = E\{(l_i - \tilde{l})^2\}, \tag{2.8}$$

see, e.g. (Bjerhammar, 1973, p. 22), (Niemeier, 2008, p. 23) or any other textbook on statistics. $\tilde{l}$ denotes the true value of l, e the observation errors and $E\{\}$ the expectation of a variable. N is the total number of random errors with the same probability that belong to a universal set. The standard deviation can be defined as the positive square root of the variance

$$\sigma = +\sqrt{\sigma^2}. \tag{2.9}$$

[7]In the literature the term "*a priori*" is often confused with the term "prior". While the former defines some information that has been obtained after theoretical deduction and has been traditionally used by scientists for expressing such arguments, the latter refers only to some moment in time.

The *theoretical covariance* between two measurements, for example l_1 and l_2, can be written as

$$\sigma_{1,2} = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} (l_{1_i} - \tilde{l}_1)(l_{2_i} - \tilde{l}_2) = E\{(l_{1_i} - \tilde{l}_1)(l_{2_i} - \tilde{l}_2)\}, \tag{2.10}$$

with $\tilde{l}_1$ and $\tilde{l}_2$ denoting the true values of l_1 and l_2, respectively.

Empirical variances and covariances of a random variable

In most cases the true value $\tilde{l}$ is not known and thus it is replaced by the expected value $E\{l\}$ (for least squares problems the mean value $\bar{l}$ is regarded as the expected value $E\{l\}$), see for example (Montgomery and Runger 2010, p. 74) or (Everitt and Skrondal 2010, p. 156) for a definition.

An estimate for the variance of l has been given by (Bjerhammar, 1973, p. 37) or (Dekking et al., 2005, p. 292) as

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} v_i^2 = \frac{1}{n-1} \sum_{i=1}^{n} (l_i - E\{l\})^2 = \frac{1}{n-1} \sum_{i=1}^{n} (l_i - \bar{l})^2, \tag{2.11}$$

with n denoting a finite number of residuals v, that belong to a subset of the universal set containing the random errors of the problem. In (Linnik, 1961, p. 79) and (Bjerhammar, 1973, p. 127) it has been shown that for a linear adjustment problem, the estimated variance s^2 is an unbiased estimate of σ^2, i.e.

$$E\{s^2\} = \sigma^2. \tag{2.12}$$

Equivalently, an estimate for the covariance between two measurements l_1 and l_2 can be derived by

$$s_{1,2} = \frac{1}{n-1} \sum_{i=1}^{n} (l_{1_i} - E\{l_1\})(l_{2_i} - E\{l_2)\} = \frac{1}{n-1} \sum_{i=1}^{n} (l_{1_i} - \bar{l}_1)(l_{2_i} - \bar{l}_2). \tag{2.13}$$

Storing all the observations in one vector

$$\mathbf{L} = [l_1, l_2, \cdots, l_n]^{\mathrm{T}}, \tag{2.14}$$

the variances and covariances of all measurements can be expressed in matrix notation, like in (Wells and Krakiwsky 1971, p. 89) or (Niemeier 2008, p. 29), stored in the variance-covariance (VC) matrix

$$\mathbf{\Sigma}_{\mathrm{LL}} = \frac{1}{n-1} (\mathbf{L} - E\{\mathbf{L}\})(\mathbf{L} - E\{\mathbf{L}\})^{\mathrm{T}} = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} & \cdots & \cdots \\ \sigma_{2,1} & \sigma_2^2 & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \cdots & \sigma_n^2 \end{bmatrix}. \tag{2.15}$$

(Ghilani 2010, p. 166) is one of the many textbooks where it was shown that the cofactor matrix of the observations can be computed by

$$\mathbf{Q}_{\mathrm{LL}} = \frac{1}{\sigma_0^2} \mathbf{\Sigma}_{\mathrm{LL}}, \tag{2.16}$$

which for a regular cofactor matrix results in the weight matrix

$$\mathbf{P} = \mathbf{Q}_{\mathrm{LL}}^{-1}. \tag{2.17}$$

Matrices $\mathbf{\Sigma}_{\mathrm{LL}}$, $\mathbf{Q}_{\mathrm{LL}}$ and $\mathbf{P}$ become diagonal, when the observed quantities are uncorrelated. Nevertheless, it is worth mentioning that for correlated observations the cofactor matrix $\mathbf{Q}_{\mathrm{LL}}$ is not always regular, as it was pointed out by (Ghilani, 2010, p. 160). In such cases the inverse weight matrix cannot be computed and a specific strategy has to be employed for a solution that is able to deal with singular cofactor matrices, like for example in the work of Neitzel and Schaffrin (2016).

The theoretical variance of the universal weight[8] σ_0^2 has been defined by (Mikhail and Ackermann, 1976, p. 65) as an arbitrary scalar value that influences the stochastic model. Following (Linnik 1961, p. 100), in case of large measurement samples it can be estimated by

$$s_0^2 = \frac{1}{r_d} \sum_{i=1}^{n} v_i^2, \quad \text{redundancy}: \;\; r_d = n - m, \tag{2.18}$$

with

$$E\left\{s_0^2\right\} = \sigma_0^2. \tag{2.19}$$

According to (Niemeier 2008, p. 165), listing all residuals in one vector $\mathbf{v}$ and introducing the weight matrix from equation (2.17), an estimate for the theoretical variance of the unit weight can be computed by

$$s_0^2 = \frac{\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v}}{r_d}. \tag{2.20}$$

2.1.3 Criteria for the solution of adjustment problems

Both types of adjustment that were discussed in the previous subsection resulted in an under-determined mathematical problem. Infinite possibilities exist for the solution of an adjustment problem, by choosing a suitable criterion for the residuals. The most common of all is the *least squares criterion*, also known as the *method of least squares* or *the minimization of the L_2-norm.* A least squares estimate $\hat{x}_j$ of the unknown parameters can be computed by minimizing the sum of squared residuals, formulated symbolically for equally weighted and uncorrelated observations by the *objective function*

$$\Omega(v_1, v_2, \ldots, v_n) = \sum_{i=1}^{n} v_i^2 \rightarrow \min. \tag{2.21}$$

Taking into account the individual precisions of the observations, the objective function obtains the form

$$\Omega(v_1, v_2, \ldots, v_n) = \sum_{i=1}^{n} p_i v_i^2 \rightarrow \min, \tag{2.22}$$

[8] Also known as a priori variance of the unit weight or reference variance.

or with correlated observations

$$\Omega(v_1, v_2, \ldots, v_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} p_{i,j}\ v_i\ v_j \rightarrow \min. \tag{2.23}$$

The objective function can be expressed equivalently in matrix notation as

$$\Omega(\mathbf{v}) = \mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v} \rightarrow \min. \tag{2.24}$$

As already mentioned by many authors, for example by Gonin (1989) or Neitzel (2004), different criteria can be employed for a solution of an adjustment problem, other than least squares. An example of such a criterion is the *minimization of L_1 norm*, also known as the *method of least absolute deviations.* A solution for the estimation of the unknown parameters for this case can be obtained by minimizing the sum of absolute residuals, i.e. the objective function

$$\Omega(\mathbf{v}) = \sum_{i=1}^{n} |v_i| \rightarrow \min. \tag{2.25}$$

Algorithmic approaches for the solution of this problem can be distinguished between the rigorous solutions via linear programming or using the Simplex-Algorithm, as presented for example in (Dantzig 1949), (Dantzig 1963) or (Fuchs 1980) and the simulated L_1 solutions via reweighted least squares adjustment as offered for example by (Krarup et al., 1980). In general, the minimization of any norm can be employed in this way. The *minimization of L_{p} norm*, with $\mathrm{p} \in \mathbb{R}$ and $\mathrm{p} \geq 1$, can be formulated similarly to the presented objective functions, see for example (Marx, 2013), by

$$\Omega(v_1, v_2, \ldots, v_n) = \sum_{i=1}^{n} |v_i|^{\mathrm{p}} \rightarrow \min. \tag{2.26}$$

Additionally, the M-estimators are based on the maximum likelihood method and can be employed for a robust solution of an adjustment problem. Such solutions have been presented for example by Huber (1964) or Hampel (1980).

2.2 Adjustment of observations with the method of least squares

Amongst all methods (e.g. L_1, L_2, L_{p} or M-estimators) for the solution of adjustment problems, the method of least squares is the most popular. There is a rich literature that deals with adjustment solutions based on the principles of least squares, for example (Helmert 1924), (Linkwitz 1960), (Linnik 1961), (Deming 1964), (Wells and Krakiwsky 1971), (Bjerhammar 1973), (Krakiwsky 1975), (Meissl 1982), (Perović 2005), (Niemeier 2008) or (Ghilani 2010).

The method of least squares has been utilized for around two centuries to provide "optimal" solutions for under-determined algebraic problems that often occur in the mathematical modelling of measurement results. From a historical perspective, the French mathematician Adrien-Marie Legendre (1752-1833) was the first who applied this method to the astronomic problem of determining the orbits of comets. In his book "Nouvelles méthodes pour la détermination des orbites des comètes", that was published in 1805, he included an appendix under the title “Sur la méthode des moindres quarrés”, which can be translated into English as “On the method of least squares”.

Four years later, the famous German geodesist and mathematician Johann Carl Friedrich Gauss (1777-1855) presented his theory for the calculation of the orbits of celestial bodies in (Gauss 1809). In his work, Gauss made extensive use of the method of least squares and claimed that he has already been using it since 1795. Although the statement of Gauss that he was the first being using least squares can be judged on its merits as contradictory (in 1795 Gauss was 18 years old), he is deservedly acknowledged as the pioneer of the method of least squares, as it has been already discussed by Stigler (1981). Gauss not only used least squares for the solution of his geodetic problem, but also introduced the statistical distribution of the errors, as well as the precision of the observed quantities, as an important parameter for obtaining the most probable estimate of the unknown parameters. In this first work the famous Gaussian (or normal) distribution has been presented for the first time. Nevertheless, the most important outcome is that both Legendre and Gauss contributed greatly to the scientific community, with a method for the adjustment of observations that is widely used in various scientific fields.

A year after Gauss’s first publication, Pierre-Simon Laplace (1749–1827) investigated Gauss’s method of least squares and the errors distribution utilizing the central limit theorem. The development of the method of least squares continues with the second publication of Gauss (also known as the 2nd foundation of least squares), see (Gauss 1823). In his second book Gauss argued that the method of least squares could be employed also in case of not normally distributed errors. However, in this case the solution is not the most probable but can be considered as the most appropriate or the most plausible. After that, the application of least squares met success in various scientific fields and in geodetic science the work of Helmert (1872) became a standard textbook for the application of this method to geodetic problems with redundant observations.

2.2.1 Statistical formulation of least squares problems

Since the work of Gauss, the method of least squares is related to the normal distribution of the measurement errors, as it has been already pointed out in (Bjerhammar 1973, p. 80). For normally distributed observations the method of least squares will result in the most probable solution for the unknown quantities or equivalently in the solution of maximum likelihood. The relationship between least squares and the error distribution can be put forward with the following simple example adopted from (Petrović et al. 1983):

Example 2.2.1. least squares and normally distributed errors:

An adjustment problem is under investigation, where n observations l and their residuals v are related to m unknown parameters x. Suppose that the measurement errors originate from a universal set of normally distributed errors[9], expressed symbolically by

$$e_i \sim N(0, \sigma_i). \tag{2.27}$$

In this line of thinking, the residuals v are also assumed to be normally distributed, $v_i \sim N(0, \sigma_i)$, with known expectations

$$E(v_i) = 0 \tag{2.28}$$

[9] In fact, errors of different measurements (or type of measurements) belong to different normal distributions.

and variances

$$E\left(v_i^2\right) = \sigma_i^2. \tag{2.29}$$

The probability density function (or Likelihood function, see Linnik 1961, p. 321 for a similar example) of each individual residual is then

$$\begin{aligned}
\phi_1(v_1) &= \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left(-\frac{v_1^2}{2\sigma_1^2}\right), \\
\phi_2(v_2) &= \frac{1}{\sigma_2\sqrt{2\pi}} \exp\left(-\frac{v_2^2}{2\sigma_2^2}\right), \\
&\vdots \\
\phi_n(v_n) &= \frac{1}{\sigma_n\sqrt{2\pi}} \exp\left(-\frac{v_n^2}{2\sigma_n^2}\right).
\end{aligned} \tag{2.30}$$

Assuming that the observations are uncorrelated and their random errors independent, the vector of residuals

$$\mathbf{v} = \left[v_1, v_2, \ldots, v_n\right]^{\mathrm{T}}, \tag{2.31}$$

would have the probability density function

$$L(v_1, v_2, \ldots, v_n) = \phi_1\phi_2\ldots\phi_n = \prod_{i=1}^{n} \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left(-\frac{v_i^2}{2\sigma_i^2}\right) = \frac{1}{\prod\limits_{i=1}^{n}\left(\sigma_i\sqrt{2\pi}\right)} \exp\left(-\sum_{i=1}^{n}\frac{v_i^2}{2\sigma_i^2}\right). \tag{2.32}$$

Expressing the weights for the residuals by

$$p_i = \frac{1}{\sigma_i^2}, \tag{2.33}$$

function (2.32) can be reformulated to

$$L(v_1, v_2, \ldots, v_n) = \frac{\prod\limits_{i=1}^{n}\sqrt{p_i}}{\left(\sqrt{2\pi}\right)^n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} p_i v_i^2\right) = K_1 \exp\left(-\frac{1}{2}\sum_{i=1}^{n} p_i v_i^2\right), \tag{2.34}$$

with the introduced parameter K_1 being a positive constant. There are infinite solutions for the residuals v, as well as the unknown parameters x. A solution is required that maximizes the probability density function L (i.e. the maximum likelihood solution for the unknowns). Thus, for $K_1 > 0$ it is obvious that

$$L(v_1, v_2, \ldots, v_n) \to \max \Leftrightarrow \exp\left(-\frac{1}{2}\sum_{i=1}^{n} p_i v_i^2\right) \to \max. \tag{2.35}$$

According to (Bronshtein et al. 2005, pp. 49-50), function $L(v_1, v_2, \ldots, v_n)$ is strictly monotonically decreasing. It follows

$$\exp\left(-\frac{1}{2}\sum_{i=1}^{n} p_i v_i^2\right) \to \max \Leftrightarrow \frac{1}{2}\sum_{i=1}^{n} p_i v_i^2 \to \min, \tag{2.36}$$

i.e.

$$L(v_1, v_2, \ldots, v_n) \to \max \Leftrightarrow \sum_{i=1}^{n} p_i v_i^2 \to \min. \tag{2.37}$$

Thus, it can be said that for uncorrelated observations with normally distributed errors the solution of maximum likelihood (by maximizing the probability density function) is equivalent to the least squares solution (by minimizing the sum of weighted squared residuals). Equivalent explanations can be found in (Merimman 1877, p. 16), (Helmert 1924, pp. 94-98), (Wells and Krakiwsky 1971, p. 93) or (Petrović et al. 1983). The first of these authors states in a few sentences that "*... the most probable values of quantities, which are the object of measurements, are those which render the sum of the squares of the errors a minimum.*".

It should be pointed out that the method of least squares can be applied in cases that the errors are not normally distributed as well. However, the estimated unknown parameters would not be the most probable, but in some cases can be acceptable or appropriate.

2.2.2 Least squares parameter estimation

A least squares adjustment problem can be seen as an optimization problem, due to the fact that always the extreme values of an objective function are requested (i.e., for least squares problems the minimum of the sum of squared residuals). To clearly demonstrate the procedure for obtaining a least squares solution, a simple adjustment problem will be employed. For example, a linear functional model between a set of n observations l, their residuals v and the unknown parameters x, which can be expressed by the observation equations

$$\begin{aligned} l_1 + v_1 &= f_1(x_1, x_2, \ldots, x_m), \\ l_2 + v_2 &= f_2(x_1, x_2, \ldots, x_m), \\ &\vdots \\ l_n + v_n &= f_n(x_1, x_2, \ldots, x_m), \end{aligned} \tag{2.38}$$

with f denoting linear functions of the unknown parameters. Obviously an adjustment problem exists when the number of observations is larger than the number of unknown parameters ($n > m$). Such a functional model occurs in practice, for example when the length of a side of an object has been measured repeatedly and the length needs to be estimated (in this case only one unknown parameter is requested). Another adjustment example of such a linear functional model, often presented in geodetic and statistical literature, assumes that the y-coordinates from a set of n points in 2D have been measured, while the x-coordinates are taken as error-free and the parameters of a straight line that fits best to the points are unknown and need to be estimated. Nevertheless, the adjustment problem will be kept more general here. Solving (2.38) for the residuals, yields the system of equations

$$\begin{aligned} v_1 &= f_1(x_1, x_2, \ldots, x_m) - l_1, \\ v_2 &= f_2(x_1, x_2, \ldots, x_m) - l_2, \\ &\vdots \\ v_n &= f_n(x_1, x_2, \ldots, x_m) - l_n. \end{aligned} \tag{2.39}$$

Postulating uncorrelated observations of equal precision, the least squares method leads to the objective function

$$\Omega(v_1, v_2, \ldots, v_n) = \sum_{i=1}^{n} v_i^2, \tag{2.40}$$

which after substituting the residuals from (2.39) can be written as

$$\Omega(x_1, x_2, \ldots, x_m) = \sum_{i=1}^{n} \left[f_i(x_1, x_2, \ldots, x_m) - l_i \right]^2 . \tag{2.41}$$

Estimates of the unknown parameters $\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_m$ are required, that minimize the objective function (2.41). According to *Fermat's Theorem* (Bronshtein et al. 2005, p. 388), the extreme values of a function can be determined by setting the first derivative with respect to the unknown terms equal to zero. This is a necessary condition to obtain the stationary points of the objective function Ω but it is not sufficient for the least squares solution, in the sense that the derived stationary points can still be a maximum, a minimum or even a saddle point (also called inflection point for cases of one variable). Stationary points of the objective function (2.41) can be computed by

$$\begin{aligned}
\frac{\partial \Omega(x_1, x_2, \ldots, x_m)}{\partial x_1} &= \sum_{i=1}^{n} 2 \left[f_i(x_1, x_2, \ldots, x_m) - l_i \right] \frac{\partial f_i(x_1, x_2, \ldots, x_m)}{\partial x_1} = 0, \\
\frac{\partial \Omega(x_1, x_2, \ldots, x_m)}{\partial x_2} &= \sum_{i=1}^{n} 2 \left[f_i(x_1, x_2, \ldots, x_m) - l_i \right] \frac{\partial f_i(x_1, x_2, \ldots, x_m)}{\partial x_2} = 0, \\
&\vdots \\
\frac{\partial \Omega(x_1, x_2, \ldots, x_m)}{\partial x_m} &= \sum_{i=1}^{n} 2 \left[f_i(x_1, x_2, \ldots, x_m) - l_i \right] \frac{\partial f_i(x_1, x_2, \ldots, x_m)}{\partial x_m} = 0,
\end{aligned} \tag{2.42}$$

which is a system of m equations with m unknown parameters, known as the system of *normal equations*[10]. To ensure that the computed stationary point of an objective function of several variables is minimum, the matrix including the second order partial derivatives should be built as

$$\mathbf{D} = \begin{bmatrix} \frac{\partial^2 \Omega}{\partial x_1^2} & \frac{\partial^2 \Omega}{\partial x_1 \partial x_2} & \cdots & \cdots \\ \frac{\partial^2 \Omega}{\partial x_2 \partial x_1} & \frac{\partial^2 \Omega}{\partial x_2^2} & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \cdots & \frac{\partial^2 \Omega}{\partial x_m^2} \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} & \cdots & \cdots \\ d_{21} & d_{22} & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ldots & d_{mm} \end{bmatrix} . \tag{2.43}$$

According to (Bronshtein et al. 2005, p. 402), if all subdeterminants of matrix $\mathbf{D}$ are positive (i.e. $\mathbf{D}$ is a positive definite matrix):

$$\begin{gathered} d_{11} > 0, \\ d_{11} d_{22} - d_{12} d_{21} > 0, \\ \vdots \end{gathered} \tag{2.44}$$

then it is guaranteed that the computed extreme value of Ω is a minimum. Furthermore, for functions $f_i(x_1, x_2, \ldots, x_m)$ in equation (2.38) being linear, the partial derivatives with respect to the unknown parameters are constants, resulting in linear normal equations in (2.42). The least squares estimates of the

[10]In (Wells and Krakiwsky 1971, p. 87) an equivalent explanation of the normal equations is presented using matrix notation.

unknown parameters $\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_m$ can be computed in this case straightforward. The estimated unknowns can be used in equation (2.39) to determine the residuals

$$\begin{aligned} \hat{v}_1 &= f_1(\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_m) - l_1, \\ \hat{v}_2 &= f_2(\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_m) - l_2, \\ &\vdots \\ \hat{v}_n &= f_n(\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_m) - l_n, \end{aligned} \tag{2.45}$$

leading to the adjusted observations

$$\begin{aligned} \hat{l}_1 &= l_1 + \hat{v}_1, \\ \hat{l}_2 &= l_2 + \hat{v}_2, \\ &\vdots \\ \hat{l}_n &= l_n + \hat{v}_n. \end{aligned} \tag{2.46}$$

Least squares parameter estimation with constraints

In many cases there are additional constraints between the unknowns that have to be taken into account, for example a constraint

$$u(x_1, x_2, \ldots, x_m) = 0, \tag{2.47}$$

has to be enforced to the unknown parameters. The functional model includes not only the observation equations (2.38), but also the constraint (2.47). Following (Bronshtein et al. 2005, p. 403), the *Lagrange multiplier method* can be employed in this case for obtaining a least squares estimate for the unknown parameters, that minimizes the objective function (2.41). Therefore, by combining $\Omega(x_1, x_2, \ldots, x_m)$ with the constraint (2.47), leads to the formulation of the Lagrange function (also known as Lagrangian)

$$\mathrm{K}(x_1, x_2, \ldots, x_m, k) = \sum_{i=1}^{n} \left[(f_i(x_1, x_2, \ldots, x_m) - l_i)\right]^2 - 2k\left(u(x_1, x_2, \ldots, x_m)\right), \tag{2.48}$$

where parameter k denotes the Lagrange multiplier. The stationary points of K can be obtained by taking the partial derivatives with respect to all unknown parameters and setting them to zero, which yields the normal equation system

$$
\begin{aligned}
\frac{\partial \mathrm{K}(x_1, x_2, \ldots, x_m, k)}{\partial x_1} &= \sum_{i=1}^{n} 2\left[f_i(x_1, x_2, \ldots, x_m) - l_i\right] \frac{\partial f_i(x_1, x_2, \ldots, x_m)}{x_1} - 2k\frac{\partial u(x_1, x_2, \ldots, x_m)}{\partial x_1} = 0,\\
\frac{\partial \mathrm{K}(x_1, x_2, \ldots, x_m, k)}{\partial x_2} &= \sum_{i=1}^{n} 2\left[f_i(x_1, x_2, \ldots, x_m) - l_i\right] \frac{\partial f_i(x_1, x_2, \ldots, x_m)}{x_2} - 2k\frac{\partial u(x_1, x_2, \ldots, x_m)}{\partial x_2} = 0,\\
&\vdots\\
\frac{\partial \mathrm{K}(x_1, x_2, \ldots, x_m, k)}{\partial x_m} &= \sum_{i=1}^{n} 2\left[f_i(x_1, x_2, \ldots, x_m) - l_i\right] \frac{\partial f_i(x_1, x_2, \ldots, x_m)}{x_m} - 2k\frac{\partial u(x_1, x_2, \ldots, x_m)}{\partial x_m} = 0,\\
\frac{\partial \mathrm{K}(x_1, x_2, \ldots, x_m, k)}{\partial k} &= -2\left(u(x_1, x_2, \ldots, x_m)\right) = 0.
\end{aligned}
\tag{2.49}
$$

The derived system of normal equations would be linear, as long as the functional model of the problem is also linear (i.e. the observation equations and the constraint are linear). Thus, the least squares solution for the unknown parameters and the Lagrange multiplier can be obtained straightforward by solving the normal equations (2.49). Matrix notation can be easily utilized for such solutions of linear least squares problems. Examples of this procedure using matrices can be found in most common adjustment literature, like (Koch and Pope 1969), (Wolf 1979), (Wells and Krakiwsky 1971, p. 142ff.), (Mikhail and Ackermann 1976, p. 213ff.) or (Perović, 2005, p. 189ff.).

2.2.3 Definition of linear and nonlinear least squares problems

Before discussing any strategy for solving nonlinear least squares problems, it is necessary to define what is a linear and what a nonlinear least squares adjustment problem. A thorough analysis of the adjustment theory can be be found in (Pasioti 2015), which makes clear the different interpretations that exist in geodetic literature concerning the nature of least squares problems. For example, Teunissen and Knickmeyer (1988) described the nature of an adjustment in terms of the solution space curvature. Nevertheless, two clear definitions are given in (Pasioti 2015) regarding linear and nonlinear least squares, which are adopted here and extended by taking into account the special cases with constraints between the unknown parameters. Therefore, the following definitions can be stated:

Definition 2.1. *Linear least squares problems*

A least squares problem is linear, when the observation/condition equations and the additional constraints are linear both with respect to the unknown parameters and the residuals.

Definition 2.2. *Nonlinear least squares problems*

A least squares problem is nonlinear, when the observation/condition equations or the additional constraints are nonlinear with respect to the unknown parameters or the residuals.

Obviously, a linear functional model will lead to linear normal equations, respectively a nonlinear functional model to nonlinear normal equations.

2.3 Error estimation of adjustment results

In addition to the parameter estimation and the adjustment of the observed quantities using least squares or some other criterion, it is important to know the precision of the derived parameters in order to verify the quality of the adjustment results and draw valid conclusions. The target is to calculate how an infinitesimal change of the observed quantities would affect the unknown parameters. In other words, it is investigated how the variances and covariances of the observations $(l_1, l_2, \ldots, l_n)$ propagate to the unknown parameters $(x_1, x_2, \ldots, x_m)$ and the residuals $(v_1, v_2, \ldots, v_n)$. This procedure can be found under the name "error propagation" or "variance and covariance propagation" and has been presented in various publications, for example in (Wells and Krakiwsky 1971, p. 20), (Mikhail and Ackermann 1976, p. 76ff.), (Cross 1994, p. 32) or (Niemeier 2008, p. 51ff.). A brief explanation of the error propagation is presented below for two individual cases, regarding linear and nonlinear functional relationships.

Computation of stochastic parameters in linear functional relationships

Assuming for a moment a linear functional relationship between the unknown parameters and the observations

$$\begin{aligned} x_1 &= f_1(l_1, l_2, \ldots, l_n), \\ x_2 &= f_2(l_1, l_2, \ldots, l_n), \\ &\vdots \\ x_m &= f_m(l_1, l_2, \ldots, l_n) \end{aligned} \tag{2.50}$$

and taking into account that functions $f_1, f_2, \ldots, f_m$ are linear, the latter equation system can be formulated equivalently in matrix notation as

$$\mathbf{X} = \mathbf{F}\mathbf{L}, \tag{2.51}$$

with vector $\mathbf{X} = [x_1, x_2, \ldots, x_m]^{\mathrm{T}}$ listing all the unknown parameters and vector $\mathbf{L} = [l_1, l_2, \ldots, l_n]^{\mathrm{T}}$ the observed quantities. The computation of the standard deviations of the unknown parameters can be geometrically interpreted as a projection of the error distribution from the measurements on the unknown parameters. A naive example of such a projection is depicted in Figure 2.1.

Further, applying the expectation operator to equation (2.51) yields

$$E\{\mathbf{X}\} = E\{\mathbf{F}\mathbf{L}\}. \tag{2.52}$$

As long as matrix $\mathbf{F}$ is deterministic and only $\mathbf{L}$ can be considered as stochastic, the last equation results in

$$E\{\mathbf{X}\} = \mathbf{F}E\{\mathbf{L}\} \Rightarrow E\{\mathbf{X}\} = \mathbf{F}(\mathbf{L} + \mathbf{v}), \tag{2.53}$$

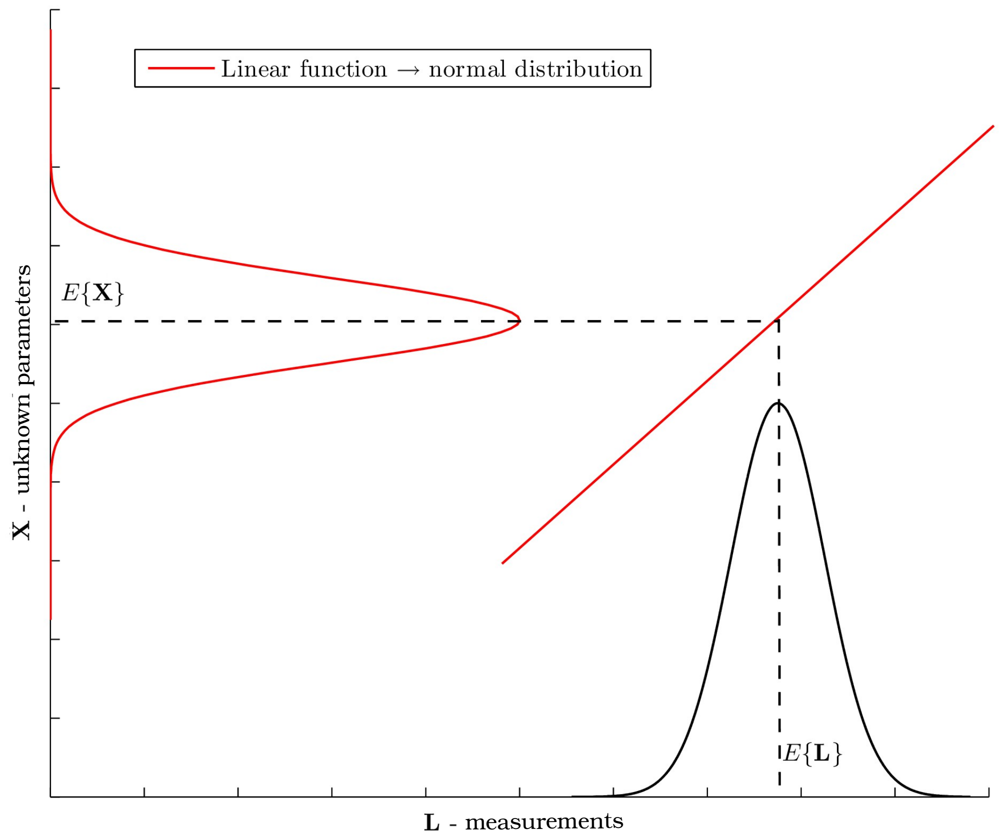

FIGURE 2.1: Simple example of linear variance-covariance propagation

with vector $\mathbf{v}$ holding the residuals of the observations in $\mathbf{L}$. According to Koch and Pope (1969), the theoretical VC matrix of the unknown parameters can be expressed by definition as

$$\begin{aligned}
\mathbf{\Sigma}_{\mathrm{XX}} &= E\left\{[\mathbf{X} - E\{\mathbf{X}\}]\,[\mathbf{X} - E\{\mathbf{X}\}]^{\mathrm{T}}\right\} \\
\Rightarrow \mathbf{\Sigma}_{\mathrm{XX}} &= E\left\{[\mathbf{FL} - (\mathbf{FL} + \mathbf{v})]\,[\mathbf{FL} - \mathbf{F}\,(\mathbf{L} + \mathbf{v})]^{\mathrm{T}}\right\} \\
\Rightarrow \mathbf{\Sigma}_{\mathrm{XX}} &= \mathbf{F}\,\underbrace{E\left\{\mathbf{v}\mathbf{v}^{\mathrm{T}}\right\}}_{\mathbf{\Sigma}_{\mathrm{LL}}}\,\mathbf{F}^{\mathrm{T}} \\
\Rightarrow \mathbf{\Sigma}_{\mathrm{XX}} &= \mathbf{F}\ \mathbf{\Sigma}_{\mathrm{LL}}\ \mathbf{F}^{\mathrm{T}}.
\end{aligned} \tag{2.54}$$

The last equation is known as the *propagation law of variances and covariances* and can be found e.g. in (Wells and Krakiwsky 1971, p. 20) or (Niemeier 2008, p. 56). Introducing the variance of the unit weight σ_0^2, the developed VC matrix can be written equivalently as

$$\mathbf{\Sigma}_{\mathrm{XX}} = \sigma_0^2 \mathbf{F}\ \mathbf{Q}_{\mathrm{LL}}\ \mathbf{F}^{\mathrm{T}}. \tag{2.55}$$

with the respective cofactor matrix being defined as

$$\mathbf{Q}_{\mathrm{XX}} = \mathbf{F}\mathbf{Q}_{\mathrm{LL}}\mathbf{F}^{\mathrm{T}}. \tag{2.56}$$

Analogously, the variances and covariances of the adjustment results (estimated unknown parameters, residuals and adjusted observations) can be derived by applying the law of error propagation that is presented above. A detailed formulation of these VC matrices is given in chapter 3, covering various adjustment cases.

Computation of stochastic parameters in nonlinear functional relationships

In practice, nonlinear functional relationships between the unknown parameters and the observed quantities occur more often than linear ones. For example, assume a nonlinear least squares problem with the estimates of the unknown parameters being expressed by the nonlinear system of equations

$$\begin{aligned} x_1 &= \psi_1(l_1, l_2, \ldots, l_n), \\ x_2 &= \psi_2(l_1, l_2, \ldots, l_n), \\ &\vdots \\ x_m &= \psi_m(l_1, l_2, \ldots, l_n), \end{aligned} \tag{2.57}$$

with $\psi_1, \psi_2, \ldots, \psi_m$ denoting nonlinear functions of the observations. The propagation law of variances and covariances, that was applied in the linear case, cannot be utilized here. Several solution strategies exist for obtaining estimates for the variances and covariances of parameters in nonlinear problems. Following the study of Lösler et al. (2016), four main procedures (but not only these) can be distinguished:

- First order variance-covariance propagation.
- Second order variance-covariance propagation.
- Monte Carlo simulation (MCS).
- Unscented Transformation (UT).

A detailed explanation of the first procedure is presented in chapter 3 for various adjustment examples. The analysis of the remaining procedures is out of the scope of this work. Nevertheless, it is important to discuss some main characteristics of all these cases and point out their main drawbacks that could eventually mislead to wrong interpretations of the adjustment results and unfortunately lead to wrong conclusions.

The *first order variance-covariance propagation* is a procedure based on the linear approximation of the problem and it has been presented e.g. in (Wells and Krakiwsky 1971, p. 21), (Mikhail and Ackermann 1976, pp. 79-81), (Taylor 1982, p. 80), (Jäger et al. 2005, p. 68), (Niemeier 2008, p. 63-64) or (Ghilani 2010, p. 89). The derived linearized expressions can be utilized for obtaining approximate solutions for the variances and covariances of the unknown parameters, based on the law of variance-covariance propagation. Geometrically, the application of the error propagation on a linearized problem can be seen as the projection of the measurements' error distribution on the unknown parameters, by a linear approximation of the "original" nonlinear function. A simple example of this procedure is depicted in Figure 2.2.

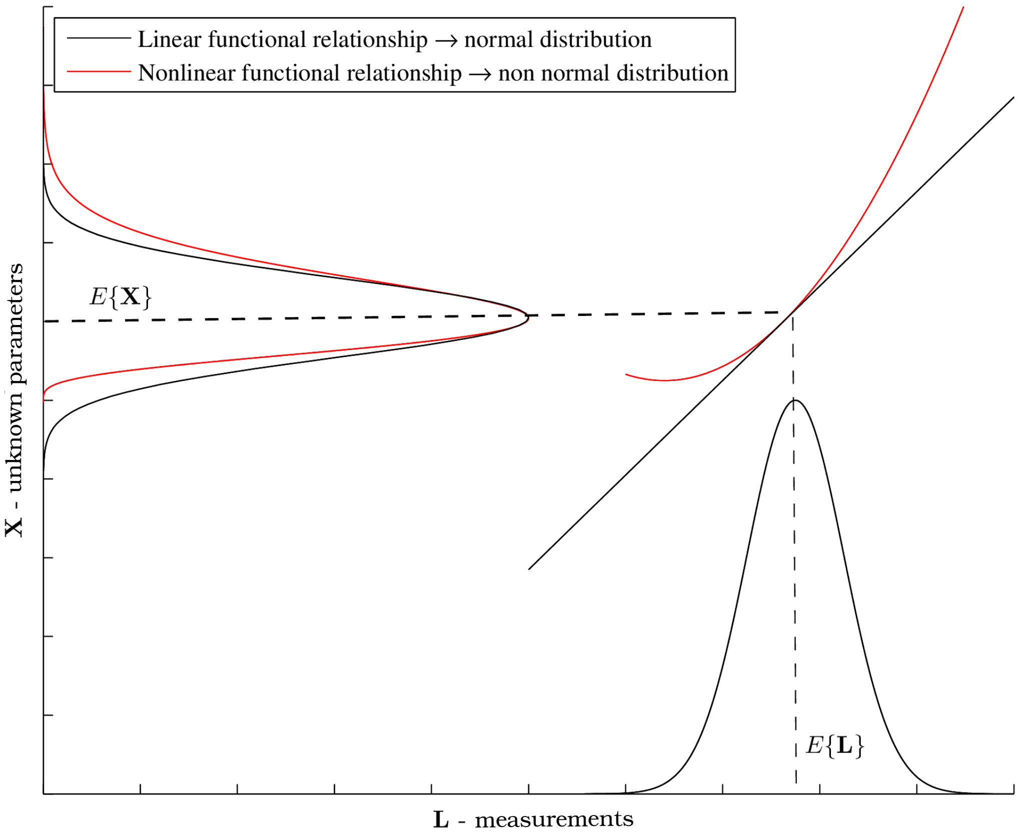

FIGURE 2.2: Simple example of a first order variance-covariance propagation

The solution of the error propagation from a linearized function cannot be always trusted, as it is only a linear approximation of the "real" solution. Therefore, the propagation law of variances and covariances should not be applied to an arbitrary nonlinear problem, but only in these cases where the linearized solution is a "good" approximation of the original nonlinear one. This has been already noticed by (Mikhail and Ackermann, 1976, p. 80), who stated "*Although in practical applications linearized functions are used regularly for the propagation of variances and covariances, it should be pointed out that this is permitted if the range of dispersion in $\tilde{x}_1$, $\tilde{x}_2$ is small when linear approximation is compared to the curvature of the function in the neighborhood of x_1^0, x_2^0. In other words the function should be approximated well by its tangent within the region of interest - that is, the region of dispersion of the random variables.*". In the same line of thinking, Lösler et al. (2016) considered this approximate solution for the variances and covariances as "distorted".

A solution coming from a *second order variance-covariance propagation* can be also seen as an approximate solution for the variances and covariances of the unknown parameters, by taking into account the first and second order terms of the nonlinear functional relationships. The solution coming from this procedure can be seen as a better approximation than the first order variances-covariances, however, it also needs to be verified in the same manner as the first order approximation. A numerical example for this approach can be found in (Lösler et al. 2016).

The *Monte Carlo method* is a famous statistical approach that is based on the sequential generation of statistically random data samples for performing simulations. Some of the authors that employed MCS for estimating the variances and covariances of adjustment results are Alkhatib (2007), Alkhatib and Schuh (2007) or Lösler et al. (2016). The advantage of MCS is that the variances and covariances can be estimated directly by using the nonlinear functional relationships, in contrast to VC propagation that is restricted to approximate functional relationships. To demonstrate a solution for the variances and covariances using MCS, equation (2.57) is expressed in vector form as

$$\mathbf{X} = \Psi(\mathbf{L}), \tag{2.58}$$

with Ψ denoting a vector that lists the nonlinear functions. Following (Lösler et al. 2016), the expectation values and the variances and covariances of the unknown parameters $\mathbf{X}$ can be obtained by means of a MC simulation, using N independent repetitions of a random experiment. The first step involves the generation of N random samples of the measurements

$$\mathbf{L}_j = \mathbf{L} + \Delta\mathbf{L}_j, \quad \text{for } j = 1, \ldots, N \tag{2.59}$$

with the error vector $\Delta\mathbf{L}$ being randomly distributed ($\Delta\mathbf{L} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{\mathrm{LL}})$), that results in

$$\begin{aligned} E\left\{\mathbf{X}\right\} &= \frac{1}{N}\sum_{j=1}^{N}\mathbf{X}_j = \frac{1}{N}\sum_{j=1}^{N}\Psi(\mathbf{L}_j), \\ \mathbf{\Sigma}_{\mathrm{XX}} &= \frac{1}{N}\sum_{j=1}^{N} E\left\{\left[\mathbf{X}_j - E\left\{\mathbf{X}_j\right\}\right]\left[\mathbf{X}_j - E\{\mathbf{X}_j\}\right]^{\mathrm{T}}\right\}. \end{aligned} \tag{2.60}$$

Moreover, the *Unscented Transformation* is a statistical approach that has been developed in the last decades, utilized also for estimating variances and covariances from nonlinear functional relationships. It has been firstly presented by Julier et al. (1995) for filtering nonlinear systems, and later on has been extended in (Julier and Uhlmann 1996) and (Julier and Uhlmann 2000) for the approximation of distribution functions and variances-covariances of unknown parameters. This approach involves a sampling strategy of "sigma points"[11] based on the a priori statistical information of the measurements. The derived sigma points are used in their turn to approximate the distribution functions or variances and covariances of the unknown parameters, even in cases of nonlinear functional relationships. Julier and Uhlmann (2000) state that this approximation is similar to a second order approximation, however, without the need of computing derivatives. A solution from the UT approach has been discussed in the study of Lösler et al. (2016), as well.

2.4 Synopsis of the basics in adjustment calculus

It has been shown that the mathematical modelling of redundant measurement results, as well as the statistical properties of the measurement errors, embody the fundamental parts of every adjustment problem.

[11] "Sigma points" can be seen as synthetic measurements, that have been generated by adding randomly distributed errors to the original measurements.

Only a correct mathematical model can lead to meaningful conclusions regarding estimates of unknown parameters and their standard deviations.

An under-determined system of equations is a consequence of every adjustment of measurements that are contaminated by errors. Depending on the nature of the errors (e.g. random errors) an appropriate method is employed by means of a criterion for the residuals, which can lead to "optimal" adjustment results. A least squares solution for the unknown parameters, the residuals and the adjusted measurements can be obtained by solving a system of normal equations, which is a result of the minimization of a clearly defined objective function. The adjustment results can be evaluated in terms of precision and reliability by computing and interpreting their stochastic parameters. These are statistical measures that can be obtained by using the rules of error propagation or some other approach, depending on the nature of the problem, i.e. the functional relationship between the estimated parameters and the measurements.

A definition of linear and nonlinear least squares problems has been presented in this chapter. Linear functional models lead to linear normal equations and a straightforward solution of the adjustment problem is possible. However, for nonlinear cases the solution can be more complicated and a specific strategy might be necessarily followed. In the next chapter various approaches are discussed for the solution of nonlinear least squares problems. Thus, traditional approaches are covered that have been utilized in geodetic science for many years and can be found in most standard adjustment textbooks, as well as the most modern approaches that have been presented by the mathematical/statistical community in the last decades.

3 Solutions of nonlinear least squares problems

This chapter summarizes two main strategies for solving nonlinear least squares problems. The first has been presented in most common geodetic literature and used extensively by geodesists, while the second has been developed recently by the mathematical and statistical scientific community. Nevertheless, before any discussion about solving nonlinear adjustment problems with the method of least squares, it would be advantageous to make clear the viewpoint on the various models and optimization approaches that are considered in the following.

Nonlinear adjustments occur often not only in geodetic science, but also in geodetic practice. Depending on the functional relationship between the observed quantities, the unknown parameters and the fixed/constant parameters, two nonlinear cases are distinguished here:

- The adjustment of nonlinear observation equations (extended to the case with nonlinear constraints);
- The adjustment of nonlinear condition equations with unknown parameters (extended to the case with nonlinear constraints).

The least squares solution of such adjustment problems can be obtained iteratively involving sometimes a linearization, or in some specific cases directly. Various optimization approaches exist for the solution of such problems and can be classified into *global optimization* or *local optimization*. Some of those approaches have been extensively applied in geodetic science, for instance in (Pope 1974), (Madsen et al. 2004) or (Neitzel and Petrovic 2008), as for example the

- Gauss-Newton,
- Newton-Raphson,
- Levenberg–Marquardt,
- heuristic optimization.

The numerical characteristics of a class of iterative algorithms for the solution of nonlinear least squares problems, with particular focus on the Gauss-Newton approach, have been investigated in (Teunissen 1985) and (Teunissen 1990) in terms of differential geometric concepts.

A comparison between these approaches[1] is out of the scope of this work and only the Gauss-Newton approach will be utilized in the following sections. The rest have been mentioned here for the sake of completeness and will not be analysed further.

An additional algorithmic approach for a class of nonlinear least squares problems has been defined and presented by (Golub and Van Loan 1980) under the name TLS. In the last three decades many authors dealt with this approach and various modern algorithms have been developed since then. A thorough analysis of TLS and its development is provided at a later point in this chapter.

Depending on the nature of the adjustment problem, as well as the chosen optimization strategy, three individual adjustment models are identified and discussed:

- The *Gauss-Markov model* (GMM)
- The *Gauss-Helmert model* (GHM)
- The *Errors in variables model* (EIV)

Figure 3.1 depicts a diagram with the solutions of nonlinear least squares problems that are discussed in this chapter.

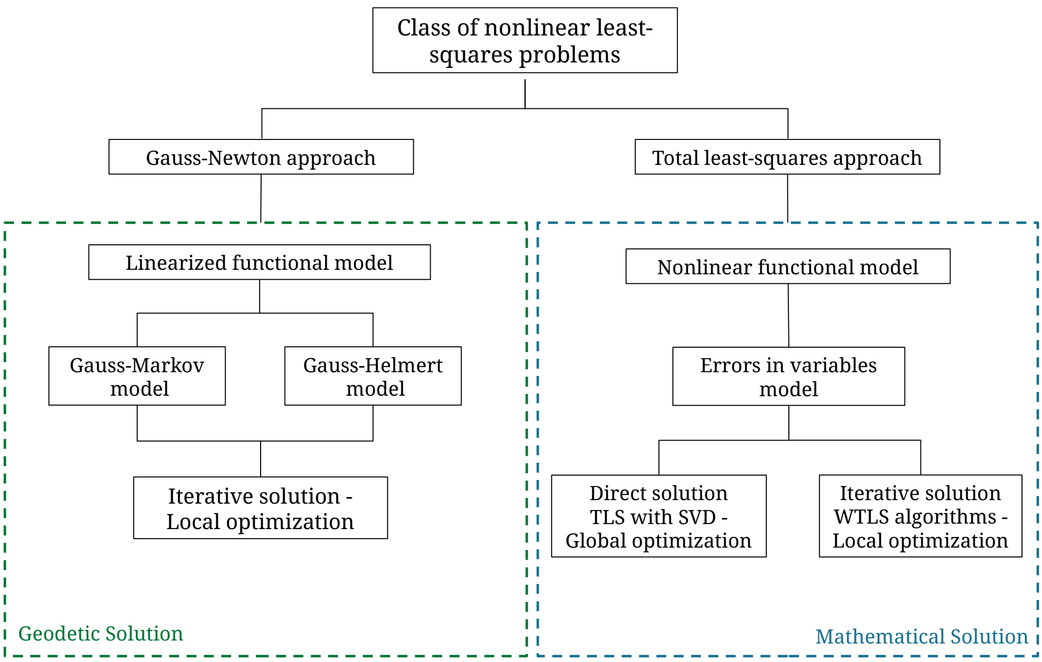

FIGURE 3.1: Two optimization approaches for the solution of a class of nonlinear least squares problems.

It must be clarified that Gauss-Newton is a general optimization approach and can be employed for finding the minimum solution of any nonlinear problem. On the other hand the TLS approach can be utilized only

[1] It is important to mention that all presented approaches are iterative and can be employed for the solution of nonlinear least squares problems. However, without guaranteeing a convergence necessarily. Thus, it can happen that a solution is not always possible by a specific approach, or some can provide a solution and some not, depending on the individual adjustment problem. More information about additional optimization approaches and a thorough analysis of the presented ones can be found in (Madsen et al. 2004).

for a special class of nonlinear least squares problems. It is, thus, necessary for the scope of this work to discuss only adjustment problems that have a solution using TLS.

3.1 Traditional geodetic solutions

Solutions of nonlinear least squares problems with the Gauss-Newton approach have been already discussed in (Wells and Krakiwsky 1971), (Pope 1974), (Niemeier 2008), (Neitzel and Petrovic 2008) or (Neitzel 2010). This involves a linear approximation of the observation/condition equations (as well as the additional constraint between the unknowns in special cases), thus initial values for the unknown parameters are necessary. The derived objective function in this case can have one minimum and after an iterative process reaches a local minimum of the "original" nonlinear problem. The estimated unknown parameters in each iteration serve as corrections for the initial values of the next iteration. This iterative process continues until a predefined threshold (or until a condition is met). Thus, the main characteristics of a least squares solution of an adjustment problem using Gauss-Newton are:

- linear approximation of the nonlinear observation/condition equations,
- linear approximation of the nonlinear constraints between the unknown parameters,
- initial values for the unknown parameters,
- iterative procedure,
- local optimization.

Depending on the individual nonlinear adjustment problem (for example adjustment with observation equations or adjustment with condition equations), a least squares solution can be obtained within the GMM or the GHM. Both adjustment models are thoroughly discussed and analysed in the next subsections.

3.1.1 Adjustment with observation equations and constraints

Point of beginning is the functional model that explicitly relates the observations l_i, the residuals v_i (with $i = 1, 2, ..., n$) and the unknown parameters x_j (with $j = 1, 2, ..., m$) and can be expressed by the nonlinear observation equations [2]

$$\begin{aligned} l_1 + v_1 &= \phi_1(x_1, ..., x_m), \\ l_2 + v_2 &= \phi_2(x_1, ..., x_m), \\ &\vdots \\ l_n + v_n &= \phi_n(x_1, ..., x_m), \end{aligned} \tag{3.1}$$

with ϕ_i denoting nonlinear differentiable functions of the unknown parameters x_j. Storing the observations l_i in a column-vector $\mathbf{L}$, the residuals v_i in vector $\mathbf{v}$ and the nonlinear functions $\phi_i(x_1, ..., x_m)$ in the formal vector $\mathbf{\Phi}(\mathbf{X})$, it is possible to write the system of observation equations (3.1) in vector notation as

$$\mathbf{L} + \mathbf{v} = \mathbf{\Phi}(\mathbf{X}). \tag{3.2}$$

[2] The linear case of this problem has been discussed in section 2.2.2, expressed by the linear system of equations 2.38 with f_i denoting in that case linear functions.

Linearization of the functional model

The first step in the Gauss-Newton approach is the linear approximation of the nonlinear functional model. In case of observation equations, a linearization is performed on the nonlinear differentiable functions $\phi_i(x_1, ..., x_m)$. For instance, the first order Taylor series expansion of the formal vector $\mathbf{\Phi}(\mathbf{X})$ at the point $\mathbf{X}^0$ can be taken, which reads

$$\mathbf{L} + \mathbf{v} = \mathbf{\Phi}(\mathbf{X}) \approx \mathbf{\Phi}\left(\mathbf{X}^0\right) + \left.\frac{\partial \mathbf{\Phi}(\mathbf{X})}{\partial \mathbf{X}}\right|_{\mathbf{X}=\mathbf{X}^0} \left(\mathbf{X} - \mathbf{X}^0\right), \tag{3.3}$$

with $\mathbf{X}^0$ denoting the vector of approximate unknown parameters $x_1^0, x_2^0, ..., x_m^0$ (i.e. initial values are necessary for approximating $\mathbf{X}$). The partial derivatives of the nonlinear functions in $\mathbf{\Phi}(\mathbf{X})$ with respect to the unknown parameters can be expressed equivalently by the Jacobian matrix

$$\mathbf{J_x} = \left.\frac{\partial \mathbf{\Phi}(\mathbf{X})}{\partial \mathbf{X}}\right|_{\mathbf{X}=\mathbf{X}^0} = \begin{bmatrix} \frac{\partial \phi_1}{\partial x_1} & \frac{\partial \phi_1}{\partial x_2} & \cdots & \frac{\partial \phi_1}{\partial x_m} \\ \frac{\partial \phi_2}{\partial x_1} & \frac{\partial \phi_2}{\partial x_2} & \cdots & \frac{\partial \phi_2}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \phi_n}{\partial x_1} & \frac{\partial \phi_n}{\partial x_2} & \cdots & \frac{\partial \phi_n}{\partial x_m} \end{bmatrix}. \tag{3.4}$$

A vector of reduced observations is introduced with

$$\mathbf{l} = \mathbf{L} - \mathbf{\Phi}(\mathbf{X}^0) \tag{3.5}$$

and the vector of corrections, containing the differences between the unknown parameters ($\mathbf{X}$) to be estimated and the approximated ones ($\mathbf{X}^0$), written as

$$\mathbf{x} = \mathbf{X} - \mathbf{X}^0. \tag{3.6}$$

Substituting the Jacobian matrix $\mathbf{J_x}$, the vector of reduced observations $\mathbf{l}$ and the vector of corrections $\mathbf{x}$ in equation (3.3), results in the linearized observation equations

$$\mathbf{l} + \mathbf{v} = \mathbf{J_x}\mathbf{x}. \tag{3.7}$$

Following (Pasioti 2015), the Jacobian matrix $\mathbf{J_x}$ can be represented symbolically as the design matrix $\mathbf{A}$, whose elements are the partial derivatives of the nonlinear functions $\phi_i(x_1, x_2, ..., x_m)$ with respect to all unknown parameters. Thus, the linearized observation equations can be expressed equivalently by

$$\mathbf{l} + \mathbf{v} = \mathbf{A}\mathbf{x}. \tag{3.8}$$

Taking into account the precision of the observed quantities, the stochastic model of the problem can be expressed by the weight matrix $\mathbf{P}$. The mathematical model of this problem represents the well known GMM, as explained in (Niemeier 2008, p. 137). Assuming normally distributed residuals ($\mathbf{v} \sim N(\mathbf{0}, \mathbf{\Sigma}_{\mathrm{LL}})$), the method of least squares will provide the most probable solution for the vector of unknown parameters and the residuals.

3.1.1.1 Least squares parameter estimation within the GMM

For obtaining a least squares solution of an adjustment problem within the GMM, an appropriate objective function has to be built and the sum of squared residuals needs to be minimized. In matrix notation the objective function is

$$\Omega(\mathbf{v}) = \mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v} \rightarrow \min. \tag{3.9}$$

Solving for the residual vector in equation (3.8) yields

$$\mathbf{v} = \mathbf{A}\mathbf{x} - \mathbf{l}, \tag{3.10}$$

which can be used to reformulate the objective function to

$$\Omega(\mathbf{x}) = (\mathbf{A}\mathbf{x} - \mathbf{l})^{\mathrm{T}}\mathbf{P}(\mathbf{A}\mathbf{x} - \mathbf{l}). \tag{3.11}$$

A solution for the unknown correction vector $\mathbf{x}$ is requested that minimizes $\Omega(\mathbf{x})$. Taking the partial derivatives of the objective function with respect to the unknowns and setting the solution to zero yields the normal equation system

$$\frac{\partial\Omega(\mathbf{x})}{\partial\mathbf{x}^{\mathrm{T}}} = \left[\ \frac{\partial\Omega}{\partial x_1}, \ \frac{\partial\Omega}{\partial x_2}, \ \cdots, \ \frac{\partial\Omega}{\partial x_m} \ \right] = 2(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\mathbf{x} - \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l}) = \mathbf{0}. \tag{3.12}$$

Therefore, the least squares estimate for the correction vector is

$$\hat{\mathbf{x}} = (\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A})^{-1}(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l}). \tag{3.13}$$

However, it is worth mentioning that matrix inversion is prone to rounding errors and even the use of double-precision floating-point format for computation is in many cases insufficient. An error analysis in matrix computations and alternative direct and iterative solutions for linear systems can be found e.g. in (Golub and Van Loan 1996) or (Björck 2015).

According to (Perović 2005, p. 84), the second derivative of the objective function with respect to the unknowns is

$$\frac{\partial^2\Omega(\mathbf{x})}{\partial\mathbf{x}^2} = 2(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}). \tag{3.14}$$

As long as the product of matrices $\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}$ is positive semi-definite, the stationary point (3.13) of the objective function will be a minimum.

Iterative solution for the adjustment results

Madsen et al. (2004) explained that a solution for the unknown vector of parameters $\hat{\mathbf{X}}$ can be obtained iteratively, by means of a local minimizer $\hat{\mathbf{x}}$. The initial values given in the first iteration step will define the area where the algorithm starts descending towards the local minimum of the objective function. From all local minima the global minimum is requested, thus "good" initial values are necessary. In every iteration step the estimated vector $\hat{\mathbf{X}}$ is utilized as initial approximation for the next iteration

$$\hat{\mathbf{X}}^0_{i+1} = \hat{\mathbf{X}}_i, \tag{3.15}$$

with i denoting the iteration step. After the termination of the iterative procedure, the final solution for the vector of corrections $\hat{\mathbf{x}}$ is utilized for computing an estimate for the vector of unknown parameters

$$\hat{\mathbf{X}} = \hat{\mathbf{X}}^0_{\text{final}} + \hat{\mathbf{x}}_{\text{final}}, \tag{3.16}$$

with the subscript "final" denoting the last iteration step. The solution for the residuals can be obtained by

$$\hat{\mathbf{v}} = \mathbf{A}\hat{\mathbf{x}}_{\text{final}} - \mathbf{l} \tag{3.17}$$

and the vector of adjusted observations

$$\hat{\mathbf{L}} = \mathbf{L} + \hat{\mathbf{v}}. \tag{3.18}$$

The iterative process can be terminated when adequate *break-off conditions* (stopping criteria) are fulfilled. According to (Pasioti 2015), suitable stopping criteria for the iterations can be the following:

1. Computation error:

   The element of the vector of corrections $\hat{\mathbf{x}}$ with the maximum absolute value, should become smaller or at least equal to a predefined threshold ϵ:

   $$\max|\hat{\mathbf{x}}| \leq \epsilon. \tag{3.19}$$

   In (Pasioti 2015) it is stated "*in practice all computations are performed with software compilers and the usage of decimal places is translated differently to the computer's world. Instead significant places are used. Therefore, a small value for ϵ is highly recommended to be chosen*". Moreover, it must be pointed out that a meaningful choice for the threshold parameter ϵ depends on the significant digits and their position inside the number system (in computer this is a binary system while humans usually think in decimal system).

2. Linearization error:

   The maximum absolute difference between the elements of the estimated vector $\hat{\mathbf{L}}$ (linearized problem) and vector $\mathbf{\Phi}(\hat{\mathbf{X}})$ ("original" nonlinear problem), should become smaller or equal to a predefined value δ:

   $$\max|\hat{\mathbf{L}} - \mathbf{\Phi}(\hat{\mathbf{X}})| \leq \delta. \tag{3.20}$$

For the second criterion, a value for δ should be chosen as close to zero as possible. This threshold value ensures that the linear approximation of the functional model has been performed correctly, as explained in (Pasioti 2015) in a few sentences: "*The linearisation error is a safeguard against wrong linearisation so that the solution of the linearised problem is also the solution of the original nonlinear problem. If the tolerance criterion is not met then the linearisation is inconsistent.*"

3.1.1.2 Error estimation within the GMM

The precision of the estimated unknown parameters in $\hat{\mathbf{x}}$ can be expressed by the VC matrix $\mathbf{\Sigma}_{\hat{\mathrm{X}}\hat{\mathrm{X}}}$, as explained in (Niemeier 2008, p. 272). Approximate solutions for the stochastic properties of the unknown parameters in a nonlinear adjustment problem can be obtained in terms of a first order variance-covariance propagation. This procedure is based on the utilization of the linearized functional model and the employment of the propagation law of variances and covariances that was discussed in section 2.3.

Stochastic properties of the estimated unknown parameters

Point of beginning is the linear functional relationship between the vector of corrections $\mathbf{x}$ and the vector of reduced observations $\mathbf{l}$:

$$\hat{\mathbf{x}} = (\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}\ \mathbf{l}.$$

This equation can be equivalently formulated as

$$\hat{\mathbf{x}} = \mathbf{F}\ \mathbf{l}, \tag{3.21}$$

after introducing matrix

$$\mathbf{F} = (\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}. \tag{3.22}$$

Following (Niemeier 2008, p. 140), only $\mathbf{l}$ is a stochastic parameter, while $\mathbf{F}$ can be taken as "fixed" or as deterministic. Utilizing the propagation law of variances and covariances from section 2.3, the cofactor matrix of $\hat{\mathbf{x}}$ can be computed here by

$$\begin{gathered}
\mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}} = \mathbf{F}\ \mathbf{Q}_{\mathrm{ll}}\ \mathbf{F}^{\mathrm{T}} \\
\Rightarrow \mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}\ \mathbf{Q}_{\mathrm{ll}}\ \left(\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}\right)^{\mathrm{T}} \\
\Rightarrow \mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\underbrace{\mathbf{P}\ \mathbf{Q}_{\mathrm{ll}}}_{\mathbf{I}_n}\ \mathbf{P}\mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1} \\
\Rightarrow \mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1}\underbrace{\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1}}_{\mathbf{I}_m} \\
\Rightarrow \mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}} = (\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A})^{-1}.
\end{gathered} \tag{3.23}$$

In case of large measurement samples the a posteriori variance of the unit weight

$$s_0^2 = \frac{\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v}}{r_d}, \quad \text{with redundancy}: \; r_d = n - m, \tag{3.24}$$

converges stochastically to σ_0^2, with

$$E\left\{s_0^2\right\} = \sigma_0^2 \tag{3.25}$$

and the VC matrix for the estimated unknown parameters can be computed by

$$\mathbf{\Sigma}_{\hat{\mathrm{x}}\hat{\mathrm{x}}} = s_0^2 \; \mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}}. \tag{3.26}$$

Finally, a first order approximate solution for the variances and covariances of the estimated unknown parameters $\hat{\mathbf{X}}$ can be derived as

$$\mathbf{\Sigma}_{\hat{\mathrm{X}}\hat{\mathrm{X}}} = \mathbf{\Sigma}_{\hat{\mathrm{x}}\hat{\mathrm{x}}} \;\; \text{and} \;\; \mathbf{Q}_{\hat{\mathrm{X}}\hat{\mathrm{X}}} = \mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}}. \tag{3.27}$$

Stochastic properties of the residuals and the adjusted observations

Approximate cofactor and VC matrices for the adjusted observations and the computed residuals can be derived following the same line of thinking as in (Niemeier 2008, p. 141). Making use of the linearized functional model (3.8) it is possible to express the adjusted reduced observations by

$$\hat{\mathbf{l}} = \mathbf{l} + \hat{\mathbf{v}} = \mathbf{A}\hat{\mathbf{x}}, \tag{3.28}$$

with the respective cofactor matrix being

$$\mathbf{Q}_{\hat{\mathrm{l}}\hat{\mathrm{l}}} = \mathbf{A}\mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}}\mathbf{A}^{\mathrm{T}}. \tag{3.29}$$

Reformulating appropriately equation (3.28) results in

$$\begin{aligned} \hat{\mathbf{v}} &= \mathbf{A}\hat{\mathbf{x}} - \mathbf{l} \\ \Rightarrow \hat{\mathbf{v}} &= \mathbf{A}\mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l} - \mathbf{l} \\ \Rightarrow \hat{\mathbf{v}} &= \left[\mathbf{A}\mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}}\mathbf{A}^{\mathrm{T}}\mathbf{P} - \mathbf{I}_n\right]\mathbf{l}. \end{aligned} \tag{3.30}$$

Thus, the cofactor matrix of the residuals can be computed by

$$\mathbf{Q}_{\hat{\mathrm{v}}\hat{\mathrm{v}}} = \left[\mathbf{A}\mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}}\mathbf{A}^{\mathrm{T}}\mathbf{P} - \mathbf{I}_n\right]\mathbf{Q}_{\mathrm{ll}}\left[\mathbf{A}\mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}}\mathbf{A}^{\mathrm{T}}\mathbf{P} - \mathbf{I}_n\right]^{\mathrm{T}}, \tag{3.31}$$

which after some examination can be simplified to

$$\mathbf{Q}_{\hat{\mathrm{v}}\hat{\mathrm{v}}} = \mathbf{Q}_{\mathrm{ll}} - \mathbf{Q}_{\hat{\mathrm{l}}\hat{\mathrm{l}}}. \tag{3.32}$$

3.1.1.3 Least squares parameter estimation within the GMM with constraints

In this case a set of constraints between the unknown parameters has to be taken into account in the functional model. If these constraints are represented by nonlinear functional relationships, then they have to be linearized together with the observation equations for a solution using the Gauss-Newton approach. For example, a number of n_c constraints are enforced in the adjustment problem, which can be expressed by the system of nonlinear equations

$$\begin{aligned} \psi_1(x_1, x_2, ..., x_m) &= 0, \\ \psi_2(x_1, x_2, ..., x_m) &= 0, \\ &\vdots \\ \psi_{n_c}(x_1, x_2, ..., x_m) &= 0. \end{aligned} \tag{3.33}$$

Listing the nonlinear functions $\psi_1, \psi_2, \ldots, \psi_{n_c}$ in the formal vector $\mathbf{\Psi}(\mathbf{X})$, it is possible to write the system of constraints in vector notation

$$\mathbf{\Psi}(\mathbf{X}) = \mathbf{0}. \tag{3.34}$$

The first order Taylor series expansion of $\mathbf{\Psi}(\mathbf{X})$ at the point $\mathbf{X}^0$ reads

$$\mathbf{\Psi}(\mathbf{X}) \approx \mathbf{\Psi}(\mathbf{X}^0) + \left.\frac{\partial \mathbf{\Psi}(\mathbf{X})}{\partial \mathbf{X}}\right|_{\mathbf{X}=\mathbf{X}^0} (\mathbf{X} - \mathbf{X}^0). \tag{3.35}$$

The partial derivatives of the constraint functions ψ with respect to the unknown parameters can be represented by a Jacobian matrix

$$\mathbf{J_c} = \left.\frac{\partial \mathbf{\Psi}(\mathbf{X})}{\partial \mathbf{X}}\right|_{\mathbf{X}=\mathbf{X}^0} = \begin{bmatrix} \frac{\partial \psi_1}{\partial x_1} & \frac{\partial \psi_1}{\partial x_2} & \cdots & \frac{\partial \psi_1}{\partial x_m} \\ \frac{\partial \psi_2}{\partial x_1} & \frac{\partial \psi_2}{\partial x_2} & \cdots & \frac{\partial \psi_2}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \psi_{n_c}}{\partial x_1} & \frac{\partial \psi_{n_c}}{\partial x_2} & \cdots & \frac{\partial \psi_{n_c}}{\partial x_m} \end{bmatrix}. \tag{3.36}$$

Introducing the vector of corrections $\mathbf{x} = \mathbf{X} - \mathbf{X}^0$ from equation (3.6), the Jacobian matrix $\mathbf{J_c}$ and the vector of misclosures

$$\mathbf{w} = \mathbf{\Psi}(\mathbf{X}^0) \tag{3.37}$$

into equation (3.34), yields the system of linearized constraints

$$\mathbf{J_c}\,\mathbf{x} + \mathbf{w} = 0. \tag{3.38}$$

The Jacobian $\mathbf{J_c}$ can be regarded as a design matrix, with its elements being the linear approximations of the nonlinear functions in (3.33) with respect to all unknown parameters. Denoting $\mathbf{J_c}$ by $\mathbf{C}$, equation (3.38) can be written as

$$\mathbf{C}\ \mathbf{x} + \mathbf{w} = 0. \tag{3.39}$$

The mathematical model in this case can be regarded as a GMM with constraints between the unknown parameters. A definition can be found also in (Perović 2005, p. 189).

In the special case that constraints are imposed on the unknown parameters, a least squares estimate can be acquired with the method of Lagrange multipliers, as it was explained in subsection 2.2.2. Thus, combining the objective function (3.9) and the constraints (3.39) yields the Lagrangian

$$\mathrm{K}(\mathbf{v}, \mathbf{x}, \mathbf{k}) = \mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v} + 2\mathbf{k}^{\mathrm{T}}(\mathbf{C}\ \mathbf{x} + \mathbf{w}) \rightarrow \min, \tag{3.40}$$

which can be expressed equivalently as

$$\mathrm{K}(\mathbf{x}, \mathbf{k}) = (\mathbf{A}\mathbf{x} - \mathbf{l})^{\mathrm{T}}\mathbf{P}(\mathbf{A}\mathbf{x} - \mathbf{l}) + 2\mathbf{k}^{\mathrm{T}}(\mathbf{C}\ \mathbf{x} + \mathbf{w}). \tag{3.41}$$

The auxiliary vector $\mathbf{k}$ holds the Lagrange multipliers. A least squares solution for vectors $\mathbf{x}$ and $\mathbf{k}$ is required that minimizes the developed Lagrange function. Taking the partial derivatives of $\mathrm{K}(\mathbf{x}, \mathbf{k})$ with respect to all unknowns and setting the solution to zero yields the system of normal equations

$$\begin{gathered}\frac{\partial \mathrm{K}}{\partial \mathbf{x}^{\mathrm{T}}} = 2(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\mathbf{x} - \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l} + \mathbf{C}^{\mathrm{T}}\mathbf{k}) = 0 \\ \Rightarrow \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\mathbf{x} + \mathbf{C}^{\mathrm{T}}\mathbf{k} = \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l},\end{gathered} \tag{3.42}$$

$$\begin{gathered}\frac{\partial \mathrm{K}}{\partial \mathbf{k}^{\mathrm{T}}} = -2(\mathbf{C}\mathbf{x} + \mathbf{w}) = 0 \\ \Rightarrow \mathbf{C}\mathbf{x} = -\mathbf{w}.\end{gathered} \tag{3.43}$$

Equations (3.42) and (3.43) can be combined with the block matrices

$$\begin{bmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A} & \mathbf{C}^{\mathrm{T}} \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{k} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l} \\ -\mathbf{w} \end{bmatrix}. \tag{3.44}$$

The least squares estimate for the vector of corrections and the vector of Lagrange multipliers is

$$\begin{bmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{k}} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A} & \mathbf{C}^{\mathrm{T}} \\ \mathbf{C} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l} \\ -\mathbf{w} \end{bmatrix}. \tag{3.45}$$

According to (Niemeier 2008, p. 265), an equivalent solution can be obtained in case of nonsingular products $[(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A})]$ and $[\mathbf{C}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A})^{-1}\mathbf{C}]$ by

$$\begin{bmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{k}} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l} \\ -\mathbf{w} \end{bmatrix}, \tag{3.46}$$

with the respective quantities

$$
\begin{aligned}
\mathbf{Q}_{22} &= -\left[\mathbf{C}(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A})^{-1}\mathbf{C}^{\mathrm{T}}\right]^{-1}, \\
\mathbf{Q}_{12} &= -\mathbf{Q}_{22}\mathbf{C}(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A})^{-1}, \\
\mathbf{Q}_{21} &= \mathbf{Q}_{12}^{\mathrm{T}}, \\
\mathbf{Q}_{11} &= (\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A})^{-1}(\mathbf{I}_m - \mathbf{C}^{\mathrm{T}}\mathbf{Q}_{12}).
\end{aligned}
\tag{3.47}
$$

$\mathbf{I}_m$ is an identity matrix, with the subscript m denoting the number of unknown parameters and specifies the dimensions of the identity matrix. The least squares estimate for the vector of corrections can be explicitly expressed by

$$\hat{\mathbf{x}} = \mathbf{Q}_{11}(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l}) - \mathbf{Q}_{12}\mathbf{w} \tag{3.48}$$

and the vector of Lagrange multipliers

$$\hat{\mathbf{k}} = \mathbf{Q}_{21}(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l}) - \mathbf{Q}_{22}\mathbf{w}. \tag{3.49}$$

A local minimizer of the Lagrange function (3.41) can be obtained iteratively. Equations (3.16), (3.17) and (3.18) can be further employed to compute the unknown parameters $\hat{\mathbf{X}}$, the residuals $\hat{\mathbf{v}}$ and the vector of adjusted observations $\hat{\mathbf{L}}$.

3.1.1.4 Error estimation within the GMM with constraints

Approximate error estimates for the parameters in $\hat{\mathbf{x}}$ can be derived by making use of the linearized functional relationship

$$\hat{\mathbf{x}} = \mathbf{Q}_{11}(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l}) - \mathbf{Q}_{12}\mathbf{w}. \tag{3.50}$$

Applying the propagation law of variances and covariances to the last equation, the cofactor matrix for the unknown parameters can be found after some investigation in

$$\mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}} = \mathbf{Q}_{11}. \tag{3.51}$$

In this case the number of constraint equations must be taken into account when computing the redundancy of the problem, with the estimated variance of the unit weight

$$s_0^2 = \frac{\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v}}{r_d}, \quad \text{with redundancy :} \quad r_d = n - m + n_c. \tag{3.52}$$

Assuming that s_0^2 converges stochastically to σ_0^2, the VC matrix for $\hat{\mathbf{x}}$ is

$$\mathbf{\Sigma}_{\hat{\mathrm{x}}\hat{\mathrm{x}}} = s_0^2\mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}}, \tag{3.53}$$

Furthermore, the VC and cofactor matrices of the estimated unknown parameters $\hat{\mathbf{X}}$, the adjusted observations $\hat{\mathbf{L}}$ and the residuals $\hat{\mathbf{v}}$ are equivalent to those of section 3.1.1.2.

3.1.2 Adjustment with condition equations and constraints

In this section a nonlinear functional model is under consideration, that implicitly relates the observations l_i, their residuals v_i (with $i = 1, ..., n$) and the unknown parameters x_j ($j = 1, ..., m$) with the condition equations

$$\begin{aligned}
\phi_1(l_1 + v_1, ..., l_n + v_n,\ x_1, ..., x_m) &= 0,\\
\phi_2(l_1 + v_1, ..., l_n + v_n,\ x_1, ..., x_m) &= 0,\\
&\vdots\\
\phi_r(l_1 + v_1, ..., l_n + v_n,\ x_1, ..., x_m) &= 0.
\end{aligned} \tag{3.54}$$

while $\phi_1, \phi_2, ..., \phi_r$ are nonlinear differentiable functions of the unknown parameters and the residuals. This system of condition equations is expressed equivalently in matrix notation as

$$\mathbf{\Phi}(\mathbf{X}, \mathbf{L} + \mathbf{v}) = \mathbf{0}, \tag{3.55}$$

with the formal vector $\mathbf{\Phi}$ holding the nonlinear functional relationship between the vector of observations $\mathbf{L}$, the vector of residuals $\mathbf{v}$ and the vector of unknown parameters $\mathbf{X}$.

Following the Gauss-Newton approach, a linear approximation of the nonlinear condition equations has to be introduced. Here it is important to mention that a correct linearization involves an approximation of both the unknown parameters ${x_j}^0$ and the unknown residuals ${v_i}^0$. This type of linearization leads to the rigorous solution of the nonlinear adjustment problem, as it has been already examined in (Pope 1972), (Lenzmann and Lenzmann 2004) and demonstrated on a practical example by Neitzel and Petrovic (2008). In the latter contributions has been shown very clearly which terms when neglected will produce merely approximate formulas for the linearized problem[3] that yield and unusuable solution. Unfortunately, these approximate formulas can be found in many popular textbooks on adjustment calculus. For more details please refer to (Neitzel 2010).

A rigorous solution within the GHM is presented here for a combined adjustment problem, according to the remarks of (Lenzmann and Lenzmann 2004) and (Neitzel 2010). The first order Taylor series approximation of the formal vector $\mathbf{\Phi}(\mathbf{X}, \mathbf{L} + \mathbf{v})$ at the point $\mathbf{X}^0$ and $\mathbf{v}^0$ reads

$$\begin{aligned}
\mathbf{\Phi}(\mathbf{X}, \mathbf{L} + \mathbf{v}) \approx \mathbf{\Phi}^0(\mathbf{X}^0, \mathbf{L} + \mathbf{v}^0) &+ \left.\frac{\partial \mathbf{\Phi}(\mathbf{X}, \mathbf{L} + \mathbf{v})}{\partial \mathbf{X}}\right|_{\mathbf{X}=\mathbf{X}^0, \mathbf{v}=\mathbf{v}^0} (\mathbf{X} - \mathbf{X}^0)\\
&+ \left.\frac{\partial \mathbf{\Phi}(\mathbf{X}, \mathbf{L} + \mathbf{v})}{\partial \mathbf{v}}\right|_{\mathbf{X}=\mathbf{X}^0, \mathbf{v}=\mathbf{v}^0} (\mathbf{v} - \mathbf{v}^0).
\end{aligned} \tag{3.56}$$

A linear approximation of the condition equations (3.55) can be expressed by

$$\mathbf{\Phi}^0(\mathbf{X}^0, \mathbf{L} + \mathbf{v}^0) + \left.\frac{\partial \mathbf{\Phi}(\mathbf{X}, \mathbf{L} + \mathbf{v})}{\partial \mathbf{X}}\right|_{\mathbf{X}=\mathbf{X}^0, \mathbf{v}=\mathbf{v}^0} (\mathbf{X} - \mathbf{X}^0) + \left.\frac{\partial \mathbf{\Phi}(\mathbf{X}, \mathbf{L} + \mathbf{v})}{\partial \mathbf{v}}\right|_{\mathbf{X}=\mathbf{X}^0, \mathbf{v}=\mathbf{v}^0} (\mathbf{v} - \mathbf{v}^0) = \mathbf{0}. \tag{3.57}$$

[3]This approximate linearization has been introduced in standard adjustment textbooks in the past, for example in (Helmert, 1924, pp. 171-174) and could lead to simpler algebraic equations for the approximate solution of the nonlinear adjustment problem without iterating (i.e. the approximate solution of the problem was obtained after one iteration).

Forming a first Jacobian matrix that contains the partial derivatives of the condition equations with respect to the unknown parameters

$$\mathbf{J_x} = \left.\frac{\partial \mathbf{\Phi}(\mathbf{X}, \mathbf{L} + \mathbf{v})}{\partial \mathbf{X}}\right|_{\mathbf{X}=\mathbf{X}^0, \mathbf{v}=\mathbf{v}^0} = \begin{bmatrix} \frac{\partial \phi_1}{\partial x_1} & \frac{\partial \phi_1}{\partial x_2} & \cdots & \frac{\partial \phi_1}{\partial x_m} \\ \frac{\partial \phi_2}{\partial x_1} & \frac{\partial \phi_2}{\partial x_2} & \cdots & \frac{\partial \phi_2}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \phi_r}{\partial x_1} & \frac{\partial \phi_r}{\partial x_2} & \cdots & \frac{\partial \phi_r}{\partial x_m} \end{bmatrix} \tag{3.58}$$

and a second that contains the partial derivatives of the condition equations with respect to the residuals

$$\mathbf{J_v} = \left.\frac{\partial \mathbf{\Phi}(\mathbf{X}, \mathbf{L} + \mathbf{v})}{\partial \mathbf{v}}\right|_{\mathbf{X}=\mathbf{X}^0, \mathbf{v}=\mathbf{v}^0} = \begin{bmatrix} \frac{\partial \phi_1}{\partial v_1} & \frac{\partial \phi_1}{\partial v_2} & \cdots & \frac{\partial \phi_1}{\partial v_n} \\ \frac{\partial \phi_2}{\partial v_1} & \frac{\partial \phi_2}{\partial v_2} & \cdots & \frac{\partial \phi_2}{\partial v_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \phi_r}{\partial v_1} & \frac{\partial \phi_r}{\partial v_2} & \cdots & \frac{\partial \phi_r}{\partial v_n} \end{bmatrix} \tag{3.59}$$

and introducing them into equation (3.57) results in

$$\mathbf{\Phi}^0(\mathbf{X}^0, \mathbf{L} + \mathbf{v}^0) + \mathbf{J_x}(\mathbf{X} - \mathbf{X}^0) + \mathbf{J_v}(\mathbf{v} - \mathbf{v}^0) = \mathbf{0}. \tag{3.60}$$

The Jacobians $\mathbf{J_x}$ and $\mathbf{J_v}$ can be regarded as design matrices. Denoting $\mathbf{J_x}$ by $\mathbf{A}$ and $\mathbf{J_v}$ by $\mathbf{B}$, this linearized equation system can be equivalently written as

$$\mathbf{\Phi}^0(\mathbf{X}^0, \mathbf{L} + \mathbf{v}^0) + \mathbf{A}(\mathbf{X} - \mathbf{X}^0) + \mathbf{B}(\mathbf{v} - \mathbf{v}^0) = \mathbf{0}. \tag{3.61}$$

Introducing the vector of misclosures

$$\mathbf{w} = \mathbf{\Phi}^0(\mathbf{X}^0, \mathbf{L} + \mathbf{v}^0) - \mathbf{B}\mathbf{v}^0, \tag{3.62}$$

and the vector of corrections $\mathbf{x} = \mathbf{X} - \mathbf{X}^0$ into equation (3.61), yields the linearized functional model

$$\mathbf{B}\mathbf{v} + \mathbf{A}\mathbf{x} + \mathbf{w} = \mathbf{0}. \tag{3.63}$$

The combination of the developed linearized functional model together with the stochastic model for the observed quantities results in the famous *Gauss-Helmert model*. An equivalent definition of this model can be found in various textbooks and publications in geodetic literature, like (Wolf 1978), (Lenzmann and Lenzmann 2004), (Perović 2005, p. 203), (Neitzel and Petrovic 2008) or (Neitzel 2010).

3.1.2.1 Least squares parameter estimation within the GHM

A least squares solution for the unknown corrections $\mathbf{x}$ can be estimated by minimizing the objective function

$$\Omega(\mathbf{v}) = \mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v}. \tag{3.64}$$

Due to the implicit functional relationship between the parameters of this adjustment case, a Lagrangian

$$\mathrm{K}(\mathbf{x}, \mathbf{v}, \mathbf{k}) = \mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v} - 2\mathbf{k}^{\mathrm{T}}(\mathbf{B}\mathbf{v} + \mathbf{A}\mathbf{x} + \mathbf{w}) \tag{3.65}$$

can be formed. Vector $\mathbf{k}$ is the vector of Lagrange multipliers. Computing the partial derivatives of K with respect to all unknown parameters and setting the solution to zero yields the stationary points

$$\begin{aligned}\frac{\partial \mathrm{K}}{\partial \mathbf{v}^{\mathrm{T}}} &= 2\mathbf{P}\mathbf{v} - 2\mathbf{B}^{\mathrm{T}}\mathbf{k} = \mathbf{0}\\ &\Rightarrow \mathbf{v} = \mathbf{Q}_{ll}\mathbf{B}^{\mathrm{T}}\mathbf{k},\end{aligned} \tag{3.66}$$

$$\begin{aligned}\frac{\partial \mathrm{K}}{\partial \mathbf{x}^{\mathrm{T}}} &= -2\mathbf{k}^{\mathrm{T}}\mathbf{A} = \mathbf{0}\\ &\Rightarrow \mathbf{A}^{\mathrm{T}}\mathbf{k} = \mathbf{0},\end{aligned} \tag{3.67}$$

$$\frac{\partial \mathrm{K}}{\partial \mathbf{k}^{\mathrm{T}}} = -2\left(\mathbf{B}\mathbf{v} + \mathbf{A}\mathbf{x} + \mathbf{w}\right) = \mathbf{0}. \tag{3.68}$$

Inserting the residual vector from equation (3.66) into (3.68) gives

$$\mathbf{B}\mathbf{Q}_{ll}\mathbf{B}^{\mathrm{T}}\mathbf{k} + \mathbf{A}\mathbf{x} + \mathbf{w} = \mathbf{0}. \tag{3.69}$$

Combining equations (3.67) and (3.69) with the block matrices

$$\begin{bmatrix} \mathbf{B}\mathbf{Q}_{ll}\mathbf{B}^{\mathrm{T}} & \mathbf{A} \\ \mathbf{A}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{k} \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} -\mathbf{w} \\ \mathbf{0} \end{bmatrix}, \tag{3.70}$$

the solution for the unknown parameters is obtained by

$$\begin{bmatrix} \hat{\mathbf{k}} \\ \hat{\mathbf{x}} \end{bmatrix} = \begin{bmatrix} \mathbf{B}\mathbf{Q}_{ll}\mathbf{B}^{\mathrm{T}} & \mathbf{A} \\ \mathbf{A}^{\mathrm{T}} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} -\mathbf{w} \\ \mathbf{0} \end{bmatrix}. \tag{3.71}$$

Under the condition that the product $[\mathbf{B}\mathbf{Q}_{ll}\mathbf{B}^{\mathrm{T}}]$ is not singular, the last equation can be expressed by

$$\begin{bmatrix} \hat{\mathbf{k}} \\ \\ \hat{\mathbf{x}} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} \begin{bmatrix} -\mathbf{w} \\ \\ \mathbf{0} \end{bmatrix}, \tag{3.72}$$

as it has been presented in (Niemeier 2008, p. 177), with the respective quantities

$$\begin{aligned} \mathbf{Q}_{22} &= -\left[\mathbf{A}^{\mathrm{T}}(\mathbf{B}\mathbf{Q}_{ll}\mathbf{B}^{\mathrm{T}})^{-1}\mathbf{A}\right]^{-1}, \\ \mathbf{Q}_{12} &= -(\mathbf{B}\mathbf{Q}_{ll}\mathbf{B}^{\mathrm{T}})^{-1}\mathbf{A}\mathbf{Q}_{22}, \\ \mathbf{Q}_{21} &= \mathbf{Q}_{12}^{\mathrm{T}}, \\ \mathbf{Q}_{11} &= (\mathbf{B}\mathbf{Q}_{ll}\mathbf{B}^{\mathrm{T}})^{-1}(\mathbf{I}_n - \mathbf{A}\mathbf{Q}_{21}). \end{aligned} \tag{3.73}$$

Explicit expressions for the vector of corrections and the vector of Lagrange multipliers are

$$\hat{\mathbf{x}} = -\mathbf{Q}_{21}\mathbf{w} = -\left[\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{Q}_{ll}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}\right]^{-1}\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{Q}_{ll}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{w} \tag{3.74}$$

and

$$\hat{\mathbf{k}} = -\mathbf{Q}_{11}\mathbf{w} = -\left(\mathbf{B}\mathbf{Q}_{ll}\mathbf{B}^{\mathrm{T}}\right)^{-1}(\mathbf{A}\hat{\mathbf{x}} + \mathbf{w}). \tag{3.75}$$

A least squares solution for the vector of unknown parameters can be computed iteratively by

$$\hat{\mathbf{X}}_i = \hat{\mathbf{x}}_i + \mathbf{X}^0 \tag{3.76}$$

and approximate estimates for the residuals by

$$\hat{\mathbf{v}}_i = \mathbf{Q}_{ll}\mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}}_i, \tag{3.77}$$

with i indicating the iteration step. Due to the iterative procedure of the Gauss-Newton approach, a local minimizer $\hat{\mathbf{x}}$ for the Lagrange function $\mathrm{K}(\mathbf{x}, \mathbf{v}, \mathbf{k})$ will be estimated by a series of adjustments within the GHM. Two vectors have to be updated in each iteration step in this case. The solution for the vector containing the unknown parameters $\hat{\mathbf{X}}$ will be introduced as the vector of initial values for the unknown parameters in the next iteration

$$\hat{\mathbf{X}}_{i+1}^0 = \hat{\mathbf{X}}_i,$$

and the computed residual vector will be used as an initial residual vector

$$\mathbf{v}_{i+1}^0 = \hat{\mathbf{v}}_i. \tag{3.78}$$

A solution can be obtained after the fulfillment of the two stopping criteria (break-off conditions), according to equations (3.19) and (3.20). The final estimated vector of corrections can be utilized for computing the vector of unknown parameters

$$\hat{\mathbf{X}} = \hat{\mathbf{X}}^0_{\text{final}} + \hat{\mathbf{x}}_{\text{final}}, \tag{3.79}$$

the residuals

$$\hat{\mathbf{v}}_{\text{final}} = \mathbf{Q}_{\text{ll}}\mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}}_{\text{final}}, \tag{3.80}$$

and the adjusted observations

$$\hat{\mathbf{L}} = \mathbf{L} + \hat{\mathbf{v}}_{\text{final}}, \tag{3.81}$$

with the subscript "final" denoting the last iteration step.

3.1.2.2 Error estimation within the GHM

Point of beginning for the error estimates in the case of a GHM is the linearized functional relationship between the estimated parameters $\hat{\mathbf{x}}$ and the vector of misclosures $\mathbf{w}$ of equation (3.74), written as

$$\hat{\mathbf{x}} = -\left[\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{Q}_{\text{ll}}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}\right]^{-1}\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{Q}_{\text{ll}}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{w}.$$

Following (Niemeier 2008, p. 178), it is necessary to derive an expression for the vector of misclosures $\mathbf{w}$ as a linear function of the vector of the observations. Therefore, taking a first order Taylor approximation of $\mathbf{w}$ from equation (3.62) with respect to the vector of observations $\mathbf{L}$, results in

$$\mathbf{w} = \underbrace{\mathbf{\Phi}^0(\mathbf{X}^0, \mathbf{L}^0 + \mathbf{v}^0)}_{\mathbf{0}} + \left.\frac{\partial \mathbf{w}}{\partial \mathbf{L}}\right|_{\mathbf{L}=\mathbf{L}^0}(\mathbf{L} - \mathbf{L}^0). \tag{3.82}$$

Forming a Jacobian matrix that contains the partial derivatives of $\mathbf{w}$ with respect to the observations

$$\mathbf{J}_{\mathbf{w}} = \left.\frac{\partial \mathbf{w}}{\partial \mathbf{L}}\right|_{\mathbf{L}=\mathbf{L}^0} \tag{3.83}$$

and by introducing a vector of reduced observations

$$\mathbf{l} = \mathbf{L} - \mathbf{L}^0 \tag{3.84}$$

in equation (3.82), returns

$$\mathbf{w} = \mathbf{J}_{\mathbf{w}}\,\mathbf{l}. \tag{3.85}$$

At the last step of the iterative procedure, i.e. the final results for the unknown parameters $\hat{\mathbf{x}}_{\text{final}}$, it can be shown that the elements of the Jacobian matrix $\mathbf{J}_{\mathbf{w}}$ will be equal to the elements of matrix $\mathbf{B}$ from equation

(3.63). The vector of misclosures can be related linearly with the reduced vector of observations, with

$$\mathbf{w} = \mathbf{B}\ \mathbf{l}, \tag{3.86}$$

which can be introduced in (3.74) to derive

$$\begin{aligned} \hat{\mathbf{x}} &= -\left[\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}\right]^{-1}\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\right)^{-1}\ \mathbf{B}\ \mathbf{l}, \\ \hat{\mathbf{x}} &= \mathbf{F}\ \mathbf{l}. \end{aligned} \tag{3.87}$$

The auxiliary matrix $\mathbf{F}$ can be defined in this case as

$$\mathbf{F} = -\left[\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}\right]^{-1}\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{B}. \tag{3.88}$$

Assuming that $\mathbf{l}$ has the same stochastic properties as $\mathbf{L}$, the former can be regarded as the only stochastic parameter in equation (3.87).The cofactor matrix for the vector of corrections is

$$\begin{aligned} &\Rightarrow \mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}} = \mathbf{F}\ \mathbf{Q}_{\mathrm{ll}}\ \mathbf{F}^{\mathrm{T}} \\ &\Rightarrow \mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}} = \left[\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}\right]^{-1}\mathbf{A}^{\mathrm{T}}\underbrace{\left(\mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{B}\ \mathbf{Q}_{\mathrm{ll}}\ \mathbf{B}^{\mathrm{T}}}_{\mathbf{I}_n}\left(\mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}\left[\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}\right]^{-1} \\ &\Rightarrow \mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}} = \underbrace{\left[\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}\right]^{-1}\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}}_{\mathbf{I}_m}\left[\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}\right]^{-1} \\ &\Rightarrow \mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}} = \left[\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}\right]^{-1} = \mathbf{Q}_{22}. \end{aligned} \tag{3.89}$$

The estimated variance of the unit weight can be computed in this case by

$$s_0^2 = \frac{\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v}}{r_d}, \quad \text{with redundancy}: \ r_d = r - m, \tag{3.90}$$

or using the expression for the residuals from equation (3.66), by

$$s_0^2 = \frac{\mathbf{k}^{\mathrm{T}}\mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{P}\mathbf{v}}{r_d} = \frac{\mathbf{k}^{\mathrm{T}}\mathbf{B}\mathbf{v}}{r_d} = -\frac{\mathbf{k}^{\mathrm{T}}\left(\mathbf{A}\hat{\mathbf{x}} + \mathbf{w}\right)}{r_d}. \tag{3.91}$$

The VC matrix for the vector of corrections is

$$\mathbf{\Sigma}_{\hat{\mathrm{x}}\hat{\mathrm{x}}} = s_0^2\mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}}. \tag{3.92}$$

The residual vector and the vector of adjusted observations can be expressed as functions of the observed quantities. Introducing $\mathbf{k}$ and $\mathbf{w}$ from equations (3.75) and (3.86) into equation (3.66) results in

$$\hat{\mathbf{v}} = \mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\mathbf{k} = -\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\mathbf{Q}_{11}\mathbf{w} = -\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\mathbf{Q}_{11}\mathbf{B}\mathbf{l} \tag{3.93}$$

and

$$\hat{\mathbf{l}} = \mathbf{l} + \hat{\mathbf{v}} = \left(\mathbf{I}_n - \mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\mathbf{Q}_{11}\mathbf{B}\right)\mathbf{l}. \tag{3.94}$$

Following the same line of reasoning as (Niemeier 2008, p. 179), the required cofactor of the residuals can be found from

$$\mathbf{Q}_{\hat{\mathrm{v}}\hat{\mathrm{v}}} = \mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\mathbf{Q}_{11}\mathbf{B}\mathbf{Q}_{\mathrm{ll}} \tag{3.95}$$

and the cofactor for the adjusted observations from

$$\mathbf{Q}_{\hat{\mathrm{l}}\hat{\mathrm{l}}} = \mathbf{Q}_{\mathrm{ll}} - \mathbf{Q}_{\hat{\mathrm{v}}\hat{\mathrm{v}}} = \mathbf{Q}_{\mathrm{ll}}\left(\mathbf{I}_n - \mathbf{B}^{\mathrm{T}}\mathbf{Q}_{11}\mathbf{B}\mathbf{Q}_{\mathrm{ll}}\right). \tag{3.96}$$

Finally, the necessary VC and cofactor matrices of the adjustment results can be computed as

$$\boldsymbol{\Sigma}_{\hat{\mathrm{X}}\hat{\mathrm{X}}} = \boldsymbol{\Sigma}_{\hat{\mathrm{x}}\hat{\mathrm{x}}} \quad , \quad \mathbf{Q}_{\hat{\mathrm{X}}\hat{\mathrm{X}}} = \mathbf{Q}_{\hat{\mathrm{x}}\hat{\mathrm{x}}}$$

and

$$\boldsymbol{\Sigma}_{\hat{\mathrm{L}}\hat{\mathrm{L}}} = \boldsymbol{\Sigma}_{\hat{\mathrm{l}}\hat{\mathrm{l}}} \quad , \quad \mathbf{Q}_{\hat{\mathrm{L}}\hat{\mathrm{L}}} = \mathbf{Q}_{\hat{\mathrm{l}}\hat{\mathrm{l}}}.$$

3.1.2.3 Least squares parameter estimation within the GHM with constraints

An adjustment with condition equations and constraints between the unknown parameters can be treated like the one presented in subsection 3.1.1.4. For constraints that are represented by nonlinear functional relationships, a linear approximation will result in the linearized constraint equations (3.39), which are expressed in this subsection by

$$\mathbf{C}\mathbf{x} + \mathbf{w}_2 = \mathbf{0}, \tag{3.97}$$

with the vector of misclosures $\mathbf{w}_2$ for the constraints. The linearized condition equations (3.63) and the constraints (3.97) represent the linearized functional model of this adjustment problem. Taking into account the stochastic information of the observed quantities, results in the case of a GHM with constraints. A least squares solution can be found by minimizing the Lagrange function

$$\mathrm{K}(\mathbf{x}, \mathbf{v}, \mathbf{k}_1, \mathbf{k}_2) = \mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v} - 2\mathbf{k}_1^{\mathrm{T}}(\mathbf{B}\mathbf{v} + \mathbf{A}\mathbf{x} + \mathbf{w}_1) - 2\mathbf{k}_2^{\mathrm{T}}(\mathbf{C}\mathbf{x} + \mathbf{w}_2) \rightarrow \min, \tag{3.98}$$

with the vectors of Lagrange multipliers $\mathbf{k}_1$ and $\mathbf{k}_2$ and the vector of misclosures $\mathbf{w}_1$ for the linearized condition equations [4]. From the standard procedure for obtaining a least squares estimate for the unknowns,

[4] The subscript "1" is introduced here to diferentiate with the vector of misclosures $\mathbf{w}_2$ for the linearized constraint equations.

the partial derivatives of $\mathrm{K}(\mathbf{x}, \mathbf{v}, \mathbf{k}_1, \mathbf{k}_2)$ with respect to all unknown parameters are computed and set equal to zero:

$$\begin{aligned} \frac{\partial \mathrm{K}}{\partial \mathbf{v}^{\mathrm{T}}} &= 2\left(\mathbf{P}\mathbf{v} - \mathbf{B}^{\mathrm{T}}\mathbf{k}_1\right) = \mathbf{0} \\ &\Rightarrow \mathbf{v} = \mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\mathbf{k}_1, \end{aligned} \tag{3.99}$$

$$\begin{aligned} \frac{\partial \mathrm{K}}{\partial \mathbf{x}^{\mathrm{T}}} &= -2\left({\mathbf{k}_1}^{\mathrm{T}}\mathbf{A} + {\mathbf{k}_2}^{\mathrm{T}}\mathbf{C}\right) = \mathbf{0} \\ &\Rightarrow \mathbf{A}^{\mathrm{T}}\mathbf{k}_1 + \mathbf{C}^{\mathrm{T}}\mathbf{k}_2 = \mathbf{0}, \end{aligned} \tag{3.100}$$

$$\frac{\partial \mathrm{K}}{\partial {\mathbf{k}_1}^{\mathrm{T}}} = -2\left(\mathbf{B}\mathbf{v} + \mathbf{A}\mathbf{x} + \mathbf{w}_1\right) = \mathbf{0}, \tag{3.101}$$

$$\frac{\partial \mathrm{K}}{\partial {\mathbf{k}_2}^{\mathrm{T}}} = -2\left(\mathbf{C}\mathbf{x} + \mathbf{w}_2\right) = \mathbf{0}. \tag{3.102}$$

Introducing the vector of residuals from equation (3.99) into (3.101) yields

$$\mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\mathbf{k}_1 + \mathbf{A}\mathbf{x} + \mathbf{w}_1 = \mathbf{0}. \tag{3.103}$$

Equations (3.100), (3.102) and (3.103) can be expressed as the block matrices

$$\begin{bmatrix} \mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}} & \mathbf{A} & \mathbf{0} \\ \mathbf{A}^{\mathrm{T}} & \mathbf{0} & \mathbf{C}^{\mathrm{T}} \\ \mathbf{0} & \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{k}_1 \\ \mathbf{x} \\ \mathbf{k}_2 \end{bmatrix} = \begin{bmatrix} -\mathbf{w}_1 \\ \mathbf{0} \\ -\mathbf{w}_2 \end{bmatrix}, \tag{3.104}$$

with the least squares solution for the unknown parameters being obtained by

$$\begin{bmatrix} \hat{\mathbf{k}}_1 \\ \hat{\mathbf{x}} \\ \hat{\mathbf{k}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}} & \mathbf{A} & \mathbf{0} \\ \mathbf{A}^{\mathrm{T}} & \mathbf{0} & \mathbf{C}^{\mathrm{T}} \\ \mathbf{0} & \mathbf{C} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} -\mathbf{w}_1 \\ \mathbf{0} \\ -\mathbf{w}_2 \end{bmatrix}. \tag{3.105}$$

For an equivalent solution of the problem, the matrices

$$\mathbf{R} = \mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}} \tag{3.106}$$

and

$$\mathbf{M} = -\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{Q}_{\mathrm{ll}}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A} = -\mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{A} \tag{3.107}$$

can be introduced. If matrix $\mathbf{R}$ is regular, then the vector of Lagrange multipliers $\mathbf{k}_1$ is

$$\mathbf{k}_1 = -\mathbf{R}^{-1}(\mathbf{A}\mathbf{x} + \mathbf{w}_1) \tag{3.108}$$

which can be substituted in equation (3.100), that yields

$$\begin{aligned} &-\mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{A}\mathbf{x} + \mathbf{w}_1) + \mathbf{C}^{\mathrm{T}}\mathbf{k}_2 = \mathbf{0} \\ &\Rightarrow \mathbf{M}\mathbf{x} - \mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{w}_1 + \mathbf{C}^{\mathrm{T}}\mathbf{k}_2 = \mathbf{0}. \end{aligned} \tag{3.109}$$

A least squares solution can be obtained by expressing equations (3.102) and (3.109) as

$$\begin{bmatrix} \mathbf{M} & \mathbf{C}^{\mathrm{T}} \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{k}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{w}_1 \\ -\mathbf{w}_2 \end{bmatrix}, \tag{3.110}$$

resulting in

$$\begin{bmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{k}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{M} & \mathbf{C}^{\mathrm{T}} \\ \mathbf{C} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{w}_1 \\ -\mathbf{w}_2 \end{bmatrix}. \tag{3.111}$$

For a regular $\mathbf{M}$ matrix, the last equation system can be equivalently written as

$$\begin{bmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{k}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{w}_1 \\ -\mathbf{w}_2 \end{bmatrix}, \tag{3.112}$$

with the respective matrices being

$$\begin{aligned} \mathbf{Q}_{22} &= -\left[\mathbf{C}\left(\mathbf{M}^{\mathrm{T}}\mathbf{M}\right)^{-1}\mathbf{M}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\right]^{-1}, \\ \mathbf{Q}_{12} &= \left(\mathbf{M}^{\mathrm{T}}\mathbf{M}\right)^{-1}\mathbf{M}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\mathbf{Q}_{22}, \\ \mathbf{Q}_{21} &= \mathbf{Q}_{12}^{\mathrm{T}}, \\ \mathbf{Q}_{11} &= \left(\mathbf{M}^{\mathrm{T}}\mathbf{M}\right)^{-1}\mathbf{M}^{\mathrm{T}}(\mathbf{I}_m - \mathbf{C}^{\mathrm{T}}\mathbf{Q}_{12}). \end{aligned} \tag{3.113}$$

Thus, the vector of corrections can be expressed explicitly by

$$\hat{\mathbf{x}} = \mathbf{Q}_{11}(\mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{w}_1) - \mathbf{Q}_{12}\mathbf{w}_2 \tag{3.114}$$

and the vector of Lagrange multipliers

$$\hat{\mathbf{k}}_2 = \mathbf{Q}_{21}\left(\mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{w}_1\right) - \mathbf{Q}_{22}\mathbf{w}_2. \tag{3.115}$$

An iterative procedure is necessary also in this adjustment case. A local minimizer of the Lagrange function (3.98) is derived, which results in the least squares solution for the unknown parameters $\hat{\mathbf{X}}$, the residuals $\hat{\mathbf{v}}$ and the vector of adjusted observations $\hat{\mathbf{L}}$, similarly to the discussed adjustment cases of the previous subsections.

3.1.2.4 Error estimation within the GHM with constraints

Starting point is the solution for the vector of corrections within the GHM with constraints

$$\hat{\mathbf{x}} = \mathbf{Q}_{11}(\mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{w}_1) - \mathbf{Q}_{12}\mathbf{w}_2.$$

Introducing the same concept as in section 3.1.2.2, the vector of misclosures $\mathbf{w}_1$ can be related linearly with the vector of observed quantities $\mathbf{l}$ by

$$\mathbf{w}_1 = \mathbf{B}\,\mathbf{l}.$$

Thus, by substituting $\mathbf{w}_1$ into (3.114) gives

$$\hat{\mathbf{x}} = \mathbf{Q}_{11}(\mathbf{A}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{B}\mathbf{l}) - \mathbf{Q}_{12}\mathbf{w}_2. \tag{3.116}$$

Applying the propagation law of vaiances and covariances and after some investigation, the cofactor matrix for $\hat{\mathbf{x}}$ can be computed by

$$\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}} = \mathbf{Q}_{11}. \tag{3.117}$$

Taking into account the number of constraint equations for computing the redundancy of the problem, the estimated variance of the unit weight is

$$s_0^2 = \frac{\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v}}{r_d}, \quad \text{with redundancy}: \;\; r_d = r - m + n_c. \tag{3.118}$$

The VC matrix for the corrections is

$$\mathbf{\Sigma}_{\hat{\mathbf{x}}\hat{\mathbf{x}}} = s_0^2\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}. \tag{3.119}$$

Furthermore, the VC and cofactor matrices of the estimated unknown parameters $\hat{\mathbf{X}}$, the adjusted observations $\hat{\mathbf{L}}$ and the residuals $\hat{\mathbf{v}}$ can be computed as in section 3.1.2.2.

3.2 Total least squares

Modern and sophisticated algorithms have been presented by the mathematical community since the 1980s, for the solution of nonlinear adjustments. These algorithms deal with a class of nonlinear least squares problems, which can be expressed within an EIV model and solved by TLS. A solution coming from TLS does not involve any kind of linearization of the functional model but presupposes the use of SVD, as it was defined in (Golub and Van Loan 1980) or (Van Huffel and Vandewalle 1991, p. 33 ff.). Thus, a TLS solution is obtained by computing the roots of a polynomial (i.e. by solving the characteristic equation of the eigenvalues) and a direct solution can be possible depending on the polynomial's degree. Such solutions have been presented in the literature when postulating, in most cases, equally weighted and uncorrelated measurements.

Various approaches and algorithms have been implemented for the solution of this class of nonlinear least squares problems when different precision is associated with each measurement. The solutions from these algorithms are iterative, they do not include a linearization of the functional model and have been published under the name WTLS. For example, Schaffrin and Wieser (2008) presented a WTLS solution for linear regression, which inspired Shen et al. (2011), Fang (2011), Amiri-Simkooei and Jazaeri (2012) and Mahboub (2012) to present modern WTLS algorithms. Despite the name TLS, in all above cases the solution has been obtained iteratively and does not follow the definition that was established by Golub and Van Loan (1980), i.e. direct solution using SVD. A clear overview of these type of algorithmic solutions has been presented in (Snow 2012), covering also the special cases of cofactor matrices being singular. In that work the term TLS has been used in a more general sense, as it was implied by the following statement"*the terms TLS and TLS solution as used in this dissertation will mean the least squares solution within the EIV model without linearization*".

Two different perspectives can be distinguished for the terms TLS and WTLS solution from the discussion above. The first follows the definition of Golub and Van Loan (1980), and the second that of Snow (2012). In this dissertation the term TLS will refer to the former and the term WTLS to the latter definition. Thus, the main characteristics of a least squares solution of an adjustment problem within the EIV model will be distinguished here by

the TLS approach:

- postulating equally weighted and uncorrelated observations,
- treatment of the nonlinear adjustment problem,
- direct solution derived by SVD,
- global optimization,

the WTLS approach:

- postulating individually weighted and correlated/uncorrelated observations,
- reduction of the derived normal equations,
- iterative solution,
- local optimization.

3.2.1 Nonlinear adjustments within the EIV model

In this section, the modelling of nonlinear least squares problems within the EIV model will be introduced. Therefore, a nonlinear functional model that implicitly relates the observations l_i, the residuals v_i (with $i = 1, 2, ..., n$) and the unknown parameters x_j (with $j = 1, 2, ..., m$) is under consideration. Similarly to equation (3.54), the discussed functional model can be written as the system of nonlinear condition equations

$$
\begin{aligned}
\phi_1(l_1 + v_1, ..., l_n + v_n,\ x_1, ..., x_m) &= 0,\\
\phi_2(l_1 + v_1, ..., l_n + v_n,\ x_1, ..., x_m) &= 0,\\
&\vdots\\
\phi_r(l_1 + v_1, ..., l_n + v_n,\ x_1, ..., x_m) &= 0.
\end{aligned}
\tag{3.120}
$$

with ϕ_i denoting nonlinear differentiable functions of the unknown parameters and the residuals. For a certain class of such nonlinear adjustment problems[5] it is possible to formulate this equation system in matrix notation by the functional model

$$
\begin{aligned}
\mathbf{L} + \mathbf{v}_{\mathrm{L}} &= (\mathbf{A} + \mathbf{V}_{\mathrm{A}})\,\mathbf{X},\\
\dim(\mathbf{A}) &= n \times m,\\
\mathrm{rank}(\mathbf{A}) &= m < n,
\end{aligned}
\tag{3.121}
$$

where $\mathbf{L}$ and $\mathbf{v}_{\mathrm{L}}$ are the vectors of observations and their residuals[6], respectively. Matrix $\mathbf{A}$ contains the coefficients of the functional model with respect to the unknown parameters x_j, except the residuals v_i which are stored in the residual matrix $\mathbf{V}_{\mathrm{A}}$. $\mathbf{X}$ is the vector containing the unknown parameters. Here, it is worth mentioning the differences, by definition, of the observation vector $\mathbf{L}$ with the vector of reduced observations $\mathbf{l}$ and the vector of unknown parameters $\mathbf{X}$ with the vector of corrections $\mathbf{x}$, as it has already been explained in section 3.1.1. The presented functional model in (3.121), accompanied by the stochastic model of the measurements, leads to the nonlinear mathematical model known as "*Errors In Variables*". A definition of the EIV model can be found alternatively in (Golub and Van Loan 1980), (Bickel and Ritov 1987), (Van Huffel and Vandewalle 1989) or (Van Huffel and Vandewalle 1991, p. 5).

In contrast to the classical representation of the functional model of an adjustment problem (see for example the GMM or the GHM), the latter formulation involves a coefficient matrix $\mathbf{A}$ that includes measured quantities that are under the influence of random errors. Therefore, the necessary residuals are added to the measurements, symbolized in $\mathbf{A}$, by means of the residual matrix $\mathbf{V}_{\mathrm{A}}$. It is of course not the elements (or the variables) of matrix $\mathbf{A}$ that are subject to errors, but the measurements that are symbolized by these elements of $\mathbf{A}$. The following simple example illustrates this type of functional modelling:

Example 3.2.1. Assuming that the coordinates of a set of points in 2D have been observed in both directions (i.e. in x and y direction). The question is how to fit a straight line to the observed points. The simplest representation of a straight line in plane is

$$y = a\ x + b. \tag{3.122}$$

[5] These adjustment problems can be solved using a traditional geodetic approach, which involved a linearization of the functional model that resulted in the formulation of the GHM and an iterative procedure for obtaining a least squares solution.

[6] For notation reasons residuals and residual vectors are introduced here and not the notion of errors and error vectors as it is usual in the TLS literature, see e.g. (Golub and Van Loan 1980, Van Huffel and Vandewalle 1989).

Adding the necessary residuals to the measurements yields

$$\begin{aligned} y_1 + v_{y_1} &= a\ (x_1 + v_{x_1}) + b, \\ y_2 + v_{y_2} &= a\ (x_2 + v_{x_2}) + b, \\ &\vdots \\ y_n + v_{y_n} &= a\ (x_n + v_{x_n}) + b. \end{aligned} \tag{3.123}$$

y and x are the observed coordinates of the 2D points, with their corresponding residuals denoted by v_y and v_x. a and b are the unknown line parameters and need to be estimated. This nonlinear equation system can be formulated equivalently by the EIV model (3.121):

$$\mathbf{L} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \mathbf{v}_{\mathrm{L}} = \begin{bmatrix} v_{y_1} \\ v_{y_2} \\ \vdots \\ v_{y_n} \end{bmatrix}, \mathbf{X} = \begin{bmatrix} a \\ b \end{bmatrix}, \mathbf{A} = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \\ x_n & 1 \end{bmatrix}, \mathbf{V}_{\mathrm{A}} = \begin{bmatrix} v_{x_1} & 0 \\ v_{x_2} & 0 \\ \vdots & \\ v_{x_n} & 0 \end{bmatrix}. \tag{3.124}$$

In this adjustment example it can be easily seen that the elements of the "design matrix" $\mathbf{A}$ symbolize some of the measured quantities of the adjustment problem, that are subject to errors and thus the corresponding residuals are listed in matrix $\mathbf{V}_{\mathrm{A}}$. Therefore, the term/name "Errors in Variables" is misleading. Always, the measured quantities are subject to errors and not the variables of a matrix.

It is necessary to point out that the stochastic model within the EIV model, involves both the stochastic properties of the measurements in vector $\mathbf{L}$, as well as in the design matrix $\mathbf{A}$. Thus, an appropriate weight matrix can be described by

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{\mathrm{L}} & \mathbf{P}_{\mathrm{LA}} \\ \mathbf{P}_{\mathrm{AL}} & \mathbf{P}_{\mathrm{A}} \end{bmatrix}, \tag{3.125}$$

with the cofactor matrix

$$\mathbf{Q}_{\mathrm{LL}} = \begin{bmatrix} \mathbf{Q}_{\mathrm{L}} & \mathbf{Q}_{\mathrm{LA}} \\ \mathbf{Q}_{\mathrm{AL}} & \mathbf{Q}_{\mathrm{A}} \end{bmatrix} \tag{3.126}$$

and the variance-covariance matrix

$$\mathbf{\Sigma}_{\mathrm{LL}} = \begin{bmatrix} \mathbf{\Sigma}_{\mathrm{L}} & \mathbf{\Sigma}_{\mathrm{LA}} \\ \mathbf{\Sigma}_{\mathrm{AL}} & \mathbf{\Sigma}_{\mathrm{A}} \end{bmatrix}. \tag{3.127}$$

For uncorrelated observations the terms off the diagonal in matrices $\mathbf{P}$, $\mathbf{Q}_{\mathrm{LL}}$ and $\mathbf{\Sigma}_{\mathrm{LL}}$ become zero. In case of normally distributed errors the most probable solution for the undetermined parameters of this adjustment problem can be obtained by employing the method of least squares. Therefore, two individual solution strategies are presented in the following sections regarding adjustment problems that can be expressed by an EIV model. The first is direct and presumes the use of an orthogonal decomposition (TLS), while the second is iterative but without involving any kind of linearization (WTLS).

3.2.1.1 Least squares parameter estimation using TLS

By the definition of TLS (Golub and Van Loan 1980), (Van Huffel and Vandewalle 1991, p. 33), a solution of an adjustment problem within the EIV model is based on the minimization of the objective function

$$|| \ [\mathbf{V}_{\mathrm{A}}, \mathbf{v_l}] \ ||_F = || \ \hat{\mathbf{A}}, \hat{\mathbf{L}}] - [\mathbf{A}, \mathbf{L}] \ ||_F \rightarrow \min, \tag{3.128}$$

with $|| \ ||_F$ being the Frobenius norm of a matrix, defined in (Van Huffel and Vandewalle 1991, p. 21) or in (Felus and Burtch 2009) with

$$|| \ [\mathbf{V}_{\mathrm{A}}, \mathbf{v_l}] \ ||_F = \sqrt{trace \left([\mathbf{V}_{\mathrm{A}}, \mathbf{v_l}]^{\mathrm{T}} [\mathbf{V}_{\mathrm{A}}, \mathbf{v_l}]\right)}. \tag{3.129}$$

The adjusted matrix $\hat{\mathbf{A}}$ and vector $\hat{\mathbf{L}}$ are

$$[\hat{\mathbf{A}}, \hat{\mathbf{L}}] = [\mathbf{A}, \mathbf{L}] + [\mathbf{V}_{\mathrm{A}}, \mathbf{v}_{\mathrm{L}}]. \tag{3.130}$$

A solution for vector $\mathbf{X}$ has been presented in the TLS literature by decomposing the augmented matrix $[\mathbf{A}, \mathbf{L}]$ (i.e. the matrix containing the coefficient matrix $\mathbf{A}$ and the observation vector $\mathbf{L}$) with the help of SVD, resulting in

$$\mathbf{U} \ \mathbf{\Sigma} \ \mathbf{W}^{\mathrm{T}} = [\mathbf{A}, \mathbf{L}], \tag{3.131}$$

with the following matrices described in (Bronshtein et al. 2005, p. 285) :

- Matrix $\mathbf{U} \in \mathbb{R}^{n \times n}$ is orthogonal and contains the left singular vectors ($\mathbf{u}$) of matrix $[\mathbf{A}, \mathbf{l}]$:

$$\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n], \text{ with } \mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{I}_n \tag{3.132}$$

  and n is the number of observation equations (this is the number of rows of matrix $\mathbf{A}$).

- Matrix $\mathbf{W} \in \mathbb{R}^{(m+1)\times(m+1)}$ is orthogonal and contains the right singular vectors ($\mathbf{w}$) of matrix $[\mathbf{A}, \mathbf{l}]$:

$$\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_{m+1}], \text{ with } \mathbf{W}^{\mathrm{T}}\mathbf{W} = \mathbf{I}_{m+1} \tag{3.133}$$

  and m is the number of unknown parameters (this is the number of columns of matrix $\mathbf{A}$).

- Matrix $\mathbf{\Sigma} \in \mathbb{R}^{n\times(m+1)}$ has the form

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \tag{3.134}$$

  with the diagonal matrix $\mathbf{\Sigma}_1 \in \mathbb{R}^{(m+1)\times(m+1)}$ carrying the singular values (σ) of matrix $[\mathbf{A}, \mathbf{l}]$:

$$\mathbf{\Sigma}_1 = \begin{bmatrix} \sigma_1 & & & \mathbf{0} \\ & \sigma_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \sigma_{m+1} \end{bmatrix}, \tag{3.135}$$

with $\sigma_1 \geqq \ldots \geqq \sigma_{m+1} \geqq 0$.

According to the procedure of (Van Huffel and Vandewalle, 1991, p. 35) or (Felus and Schaffrin 2005), the TLS solution can be derived by scaling appropriately the right singular vector ($\mathbf{w}_{\min}$) of matrix $\mathbf{W}$ that corresponds to the minimum singular value ($\sigma_{\min}$). This is the last column of $\mathbf{W}$ and can be written as

$$\mathbf{w}_{\min} = \mathbf{w}_{m+1} = \left[w_{1,m+1}, \cdots, w_{m,m+1}, w_{m+1,m+1}\right]^{\mathrm{T}}. \tag{3.136}$$

The TLS solution for the vector of unknowns is

$$\hat{\mathbf{X}} = -\frac{1}{w_{m+1,m+1}}\left[w_{1,m+1}, \cdots, w_{m,m+1}\right]^{\mathrm{T}}. \tag{3.137}$$

It must be mentioned that the last equation is usually presented with a negative sign, as in (Felus and Schaffrin 2005), which is caused by the form of the functional model. However, the functional model can always be expressed in such a way that the negative sign is not necessary anymore, see for instance (Malissiovas et al. 2016).

An equivalent formulation of the objective function

To derive the TLS solution, Schaffrin et al. (2012) and (Snow 2012) minimized the sum of the weighted squared residuals

$$\begin{gathered}\Omega(\mathbf{v}_{\mathrm{L}}, \mathbf{v}_{\mathrm{A}}) = \mathbf{v}_{\mathrm{L}}^{\mathrm{T}}\mathbf{P}_{\mathrm{L}}\mathbf{v}_{\mathrm{L}} + \mathbf{v}_{\mathrm{A}}^{\mathrm{T}}\mathbf{P}_{\mathrm{A}}\mathbf{v}_{\mathrm{A}} \rightarrow \min, \\ \text{with} \quad \mathbf{v}_{\mathrm{A}} := \mathrm{vec}(\mathbf{V}_{\mathrm{A}}),\end{gathered} \tag{3.138}$$

for the case of non-singular cofactor matrices $\mathbf{Q}_{\mathrm{L}}$ and $\mathbf{Q}_{\mathrm{A}}$. “vec” implies a function that stacks the columns of the residual matrix $\mathbf{V}_{\mathrm{A}}$ into one vector. Postulating uncorrelated observed quantities of equal precision, the last equation can be formulated equivalently by

$$\Omega(\mathbf{v}_{\mathrm{L}}, \mathbf{v}_{\mathrm{A}}) = {\mathbf{v}_{\mathrm{L}}}^{\mathrm{T}}\mathbf{v}_{\mathrm{L}} + {\mathbf{v}_{\mathrm{A}}}^{\mathrm{T}}\mathbf{v}_{\mathrm{A}} \rightarrow \min, \tag{3.139}$$

which is equal to the objective function (3.128). Thus, it is already visible that least squares is the method being used for the solution of the adjustment problem. Several contributions have already pointed out that TLS can be regarded as an approach/solution strategy for a special class of nonlinear least squares problems and not as a different method than least squares, for example (Neitzel and Petrovic 2008), (Reinking 2008), (Neitzel 2010) or (Malissiovas et al. 2016).

In this dissertation, the terms TLS and TLS approach will refer to the least squares solution of an adjustment problem within the EIV model, with equally weighted and uncorrelated observations. The solution is derived by minimizing the objective function (3.139) through SVD of the augmented matrix $[\mathbf{A}, \mathbf{L}]$.

3.2.1.2 Least squares parameter estimation using WTLS

The least squares solution of adjustment problems within the EIV model can be derived using WTLS, especially for cases of individually weighted or correlated observations. A variety of WTLS algorithms exists in the literature, like for example in (Schaffrin and Wieser 2008), (Fang 2011), (Mahboub 2012), (Snow 2012) or (Schaffrin and Snow 2014).

For obtaining a WTLS solution, the objective function (3.138) is combined with the condition equations in (3.121) to build the Lagrangian

$$\mathrm{K}\left(\mathbf{v}_{\mathrm{L}}, \mathbf{v}_{\mathrm{A}}, \mathbf{k}, \mathbf{X}\right) = \mathbf{v}_{\mathrm{L}}^{\mathrm{T}}\mathbf{P}_{\mathrm{L}}\mathbf{v}_{\mathrm{L}} + \mathbf{v}_{\mathrm{A}}^{\mathrm{T}}\mathbf{P}_{\mathrm{A}}\mathbf{v}_{\mathrm{A}} + 2\mathbf{v}_{\mathrm{L}}^{\mathrm{T}}\mathbf{P}_{\mathrm{LA}}\mathbf{v}_{\mathrm{A}} - 2\mathbf{k}^{\mathrm{T}}\left(\mathbf{L} + \mathbf{v}_{\mathrm{L}} - \left(\mathbf{A} + \mathbf{V}_{\mathrm{A}}\right)\mathbf{X}\right), \tag{3.140}$$

with $\mathbf{k}$ denoting the vector of Lagrange multipliers. The authors dealing with WTLS, e.g. Schaffrin and Wieser (2008), introduced the Kronecker-Zehfuss product (symbolized by $\otimes$) to express the developed Lagrange function equivalently as

$$\mathrm{K}(\mathbf{v}_{\mathrm{L}}, \mathbf{v}_{\mathrm{A}}, \mathbf{k}, \mathbf{X}) = \mathbf{v}_{\mathrm{L}}^{\mathrm{T}}\mathbf{P}_{\mathrm{L}}\mathbf{v}_{\mathrm{L}} + \mathbf{v}_{\mathrm{A}}^{\mathrm{T}}\mathbf{P}_{\mathrm{A}}\mathbf{v}_{\mathrm{A}} + 2\mathbf{v}_{\mathrm{L}}^{\mathrm{T}}\mathbf{P}_{\mathrm{LA}}\mathbf{v}_{\mathrm{A}} - 2\mathbf{k}^{\mathrm{T}}\left(\mathbf{L} + \mathbf{v}_{\mathrm{L}} - \mathbf{A}\mathbf{X} - \left(\mathbf{X}^{\mathrm{T}} \otimes \mathbf{I}_n\right)\mathbf{v}_{\mathrm{A}}\right). \tag{3.141}$$

The necessary stationary points can be obtained by computing the partial derivatives of K with respect to all unknowns and setting the solution to zero, which according to (Schaffrin and Snow 2014) yields the nonlinear normal equation system

$$\frac{\partial \mathrm{K}}{\partial \mathbf{v}_{\mathrm{L}}^{\mathrm{T}}} = 2\left(\mathbf{P}_{\mathrm{L}}\mathbf{v}_{\mathrm{L}} + \mathbf{P}_{\mathrm{LA}}\mathbf{v}_{\mathrm{A}} - \mathbf{k}\right) = \mathbf{0}, \tag{3.142}$$

$$\frac{\partial \mathrm{K}}{\partial \mathbf{v}_{\mathrm{A}}^{\mathrm{T}}} = 2\left(\mathbf{P}_{\mathrm{A}}\mathbf{v}_{\mathrm{A}} + \mathbf{P}_{\mathrm{AL}}\mathbf{v}_{\mathrm{L}} + \left(\mathbf{X} \otimes \mathbf{I}_n\right)\mathbf{k}\right) = \mathbf{0}, \tag{3.143}$$

$$\begin{aligned}\frac{\partial \mathrm{K}}{\partial \mathbf{X}^{\mathrm{T}}} &= 2\left(\mathbf{A}^{\mathrm{T}}\mathbf{k} + \left(\mathbf{I}_m \otimes \mathbf{k}^{\mathrm{T}}\right)\mathbf{v}_{\mathrm{A}}\right) = \mathbf{0} \\ &\Rightarrow \left(\mathbf{A} + \mathbf{V}_{\mathrm{A}}\right)^{\mathrm{T}}\mathbf{k} = \mathbf{0},\end{aligned} \tag{3.144}$$

$$\frac{\partial \mathrm{K}}{\partial \mathbf{k}^{\mathrm{T}}} = -2\left(\mathbf{L} + \mathbf{v}_{\mathrm{L}} - \mathbf{A}\mathbf{X} - (\mathbf{X}^{\mathrm{T}} \otimes \mathbf{I}_n)\mathbf{v}_{\mathrm{A}}\right) = \mathbf{0}. \tag{3.145}$$

Two individual iterative approaches are developed in the following of this section for solving the system of normal equations (3.142)-(3.145).

WTLS - Approach 1

A solution for the unknown parameters can be obtained by reducing appropriately the derived normal equations. According to (Snow 2012), if $\mathbf{Q}_{\mathrm{LL}}$ is regular then equations (3.142) and (3.143) can be rewritten as

$$\mathbf{v}_{\mathrm{L}} = \left[\mathbf{Q}_{\mathrm{L}} - \mathbf{Q}_{\mathrm{LA}}\left(\mathbf{X} \otimes \mathbf{I}_n\right)\right]\mathbf{k}, \tag{3.146}$$

$$\mathbf{v}_{\mathrm{A}} = \left[\mathbf{Q}_{\mathrm{AL}} - \mathbf{Q}_{\mathrm{A}}\left(\mathbf{X} \otimes \mathbf{I}_n\right)\right]\mathbf{k}, \tag{3.147}$$

using a "bidirectional" substitution. Inserting the explicit expressions of the residual vectors into (3.145) yields

$$
\begin{aligned}
&\mathbf{L} + \left[\mathbf{Q}_{\mathrm{L}} - \mathbf{Q}_{\mathrm{LA}}\left(\mathbf{X}\otimes\mathbf{I}_n\right)\right]\mathbf{k} - \mathbf{AX} + \left(\mathbf{X}^{\mathrm{T}}\otimes\mathbf{I}_n\right)\left[\mathbf{Q}_{\mathrm{AL}} - \mathbf{Q}_{\mathrm{A}}\left(\mathbf{X}\otimes\mathbf{I}_n\right)\right]\mathbf{k} = \mathbf{0} \\
\Rightarrow\; &\underbrace{\left[\mathbf{Q}_{\mathrm{L}} - \mathbf{Q}_{\mathrm{LA}}(\mathbf{X}\otimes\mathbf{I}_n) - (\mathbf{X}^{\mathrm{T}}\otimes\mathbf{I}_n)\mathbf{Q}_{\mathrm{AL}} + \left(\mathbf{X}^{\mathrm{T}}\otimes\mathbf{I}_n\right)\mathbf{Q}_{\mathrm{A}}(\mathbf{X}\otimes\mathbf{I}_n)\right]}_{\mathbf{Q}_1}\;\mathbf{k} = (\mathbf{AX}-\mathbf{L})\,.
\end{aligned}
\tag{3.148}
$$

Introducing approximate values for the vector of unknowns $\mathbf{X}^0$ only on the left hand side of the last equation, it is possible to express the vector of Lagrange multipliers by

$$
\mathbf{k} = \mathbf{Q}_1^{-1}\left(\mathbf{AX}-\mathbf{L}\right), \tag{3.149}
$$

with the auxiliary matrix[7]

$$
\mathbf{Q}_1 = \mathbf{Q}_{\mathrm{L}} - \mathbf{Q}_{\mathrm{LA}}(\mathbf{X}^0\otimes\mathbf{I}_n) - (\mathbf{X}^{0^{\mathrm{T}}}\otimes\mathbf{I}_n)\mathbf{Q}_{\mathrm{AL}} + \left(\mathbf{X}^{0^{\mathrm{T}}}\otimes\mathbf{I}_n\right)\mathbf{Q}_{\mathrm{A}}(\mathbf{X}\otimes\mathbf{I}_n). \tag{3.150}
$$

Consequently, substituting $\mathbf{k}$ in equations (3.146) and (3.147) results in the residual vectors

$$
\mathbf{v}_{\mathrm{L}} = \left[\mathbf{Q}_{\mathrm{L}} - \mathbf{Q}_{\mathrm{LA}}\left(\mathbf{X}\otimes\mathbf{I}_n\right)\right]\;\mathbf{Q}_1^{-1}\left(\mathbf{AX}-\mathbf{L}\right), \tag{3.151}
$$

$$
\mathbf{v}_{\mathrm{A}} = \left[\mathbf{Q}_{\mathrm{AL}} - \mathbf{Q}_{\mathrm{A}}\left(\mathbf{X}\otimes\mathbf{I}_n\right)\right]\mathbf{Q}_1^{-1}\left(\mathbf{AX}-\mathbf{L}\right). \tag{3.152}
$$

Substituting $\mathbf{v}_{\mathrm{A}}$ and $\mathbf{k}$ in (3.144) yields

$$
\begin{aligned}
&\mathbf{A}^{\mathrm{T}}\mathbf{k} + \left(\mathbf{I}_m\otimes\mathbf{k}\right)^{\mathrm{T}}\mathbf{v}_{\mathrm{A}} = \mathbf{0} \\
\Rightarrow\; &\mathbf{A}^{\mathrm{T}}\mathbf{Q}_1^{-1}\left(\mathbf{AX}-\mathbf{L}\right) = \underbrace{-\left(\mathbf{I}_m\otimes\mathbf{k}\right)^{\mathrm{T}}\left[\mathbf{Q}_{\mathrm{AL}} - \mathbf{Q}_{\mathrm{A}}\left(\mathbf{X}\otimes\mathbf{I}_n\right)\right]\mathbf{Q}_1^{-1}}_{\mathbf{R}_1}\left(\mathbf{AX}-\mathbf{L}\right) \\
\Rightarrow\; &\mathbf{A}^{\mathrm{T}}\mathbf{Q}_1^{-1}(\mathbf{AX}-\mathbf{L}) = \mathbf{R}_1(\mathbf{AX}-\mathbf{L}),
\end{aligned}
\tag{3.153}
$$

with the auxiliary matrix[8]

$$
\mathbf{R}_1 = -\left(\mathbf{I}_m\otimes\mathbf{k}^0\right)^{\mathrm{T}}\left[\mathbf{Q}_{\mathrm{AL}} - \mathbf{Q}_{\mathrm{A}}\left(\mathbf{X}^0\otimes\mathbf{I}_n\right)\right]\mathbf{Q}_1^{-1}. \tag{3.154}
$$

In this last equation the vector of Lagrange multipliers needs also to be approximated with

$$
\mathbf{k}^0 = \mathbf{Q}_1^{-1}\left(\mathbf{AX}^0-\mathbf{L}\right). \tag{3.155}
$$

[7]The auxiliary matrix $\mathbf{Q_1}$ coincides with that in (Snow 2012, p. 23) presented in equation (2.11) and is identical to the product of matrices $\mathbf{BQ}_{ll}\mathbf{B}^{\mathrm{T}}$ from (Fang 2011, p. 22) from equation (4.16). It is usual in TLS literature that matrix $\mathbf{Q_1}$ is built without introducing approximate values for the vector of unknowns ($\mathbf{X}^0$). However, this can mislead the readers to think that they deal with a linear problem.

[8]It must be pointed out that "*Algorithm 1*" of section 2.1 in (Snow 2012) contains a typo. This is matrix $\mathbf{R}_1$ in "*Algorithm 1*" that differs from the correct definition of $\mathbf{R}_1$ from equation (2.13c) in that dissertation.

Furthermore, rearranging appropriately equation (3.153) gives

$$\left[\left(\mathbf{A}^{\mathrm{T}}\mathbf{Q}_1^{-1} - \mathbf{R}_1\right)\mathbf{A}\right]\ \mathbf{X} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{Q}_1^{-1} - \mathbf{R}_1\right)\mathbf{L}, \tag{3.156}$$

which under a non-singular product of matrices $\left(\mathbf{A}^{\mathrm{T}}\mathbf{Q}_1^{-1} - \mathbf{R}_1\right)\mathbf{A}$, yields the solution for the vector of unknown parameters

$$\hat{\mathbf{X}} = \left[\left(\mathbf{A}^{\mathrm{T}}\mathbf{Q}_1^{-1} - \mathbf{R}_1\right)\mathbf{A}\right]^{-1}\left(\mathbf{A}^{\mathrm{T}}\mathbf{Q}_1^{-1} - \mathbf{R}_1\right)\mathbf{L}. \tag{3.157}$$

Iterative solution for the adjustment results

A least squares solution for the unknown parameters can be obtained iteratively following the WTLS procedure. The auxiliary matrices $\mathbf{Q}_1$ and $\mathbf{R}_1$ are functions of unknown parameters. Thus, an initial approximation $\mathbf{X}^0$ for the vector of unknowns, gives

$$\mathbf{Q}_{1i} = \mathbf{Q}_{\mathrm{L}} - \mathbf{Q}_{\mathrm{LA}}(\mathbf{X}_i^0 \otimes \mathbf{I}_n) - (\mathbf{X}_i^{0\mathrm{T}} \otimes \mathbf{I}_n)\mathbf{Q}_{\mathrm{AL}} + \left(\mathbf{X}_i^{0\mathrm{T}} \otimes \mathbf{I}_n\right)\mathbf{Q}_{\mathrm{A}}(\mathbf{X}_i^0 \otimes \mathbf{I}_n), \tag{3.158}$$

$$\mathbf{k}_i = (\mathbf{Q}_{1i})^{-1}(\mathbf{A}\mathbf{X}_i^0 - \mathbf{L}), \tag{3.159}$$

and

$$\mathbf{R}_{1i} = -\left(\mathbf{I}_m \otimes \mathbf{k}_i\right)^{\mathrm{T}}\left[\mathbf{Q}_{\mathrm{AL}} - \mathbf{Q}_{\mathrm{A}}\left(\mathbf{X}_i^0 \otimes \mathbf{I}_n\right)\right]\mathbf{Q}_{1i}^{-1}, \tag{3.160}$$

with the superscript i implying the iteration step. A solution for the unknown parameters is obtained by

$$\hat{\mathbf{X}} = \left[\left(\mathbf{A}^{\mathrm{T}}(\mathbf{Q}_{1i})^{-1} - \mathbf{R}_{1i}\right)\mathbf{A}\right]^{-1}\left(\mathbf{A}^{\mathrm{T}}(\mathbf{Q}_{1i})^{-1} - \mathbf{R}_{1i}\right)\mathbf{L} \tag{3.161}$$

and is further introduced as an initial approximation for the unknown parameters in the next iteration step

$$\mathbf{X}_{i+1}^0 = \hat{\mathbf{X}}_i.$$

The final solutions can be obtained after a sufficient stopping criterion is fulfilled. Due to the fact that a linearization has not been applied in any step of the adjustment, this iterative procedure will be terminated only after the "computational error" condition has been fulfilled, as it has been defined in subsection 3.1.1. Therefore, let the vector of corrections being computed from the difference between an estimated vector of unknown parameters and its approximation in an iteration step, expressed by

$$\Delta\mathbf{X}_i = \hat{\mathbf{X}}_i - \mathbf{X}_i^0.$$

The necessary condition for the iteration stop is then given by

$$\max|\Delta\mathbf{X}| \leq \epsilon. \tag{3.162}$$

The elements of the vector of corrections $\Delta\mathbf{X}$ with the maximum absolute value should become smaller or at least equal to a predefined threshold ϵ. A similar stopping criterion can be found in (Snow 2012). The developed WTLS procedure is similar to the one presented in (Snow 2012) as "*Algorithm 1*" and has been primarily developed and presented by Fang (2011) as "*Algorithm 2*".

WTLS - Approach 2

Following (Snow 2012), an alternative WTLS approach exists for obtaining the least squares solution for the vector of unknown parameters. The first step is to reformulate equation (3.144) to

$$\begin{aligned} &\mathbf{A}^{\mathrm{T}}\mathbf{k} = -\mathbf{V}_{\mathrm{A}}^{\mathrm{T}}\mathbf{k} \\ \Rightarrow\ &\mathbf{A}^{\mathrm{T}}\ \mathbf{Q}_1^{-1}(\mathbf{A}\mathbf{X}-\mathbf{L}) = -\mathbf{V}_{\mathrm{A}}^{\mathrm{T}}\ \mathbf{Q}_1^{-1}(\mathbf{A}\mathbf{X}-\mathbf{L}) \\ \Rightarrow\ &(\mathbf{A}+\mathbf{V}_{\mathrm{A}})^{\mathrm{T}}\ \mathbf{Q}_1^{-1}\mathbf{A}\ \mathbf{X} = (\mathbf{A}+\mathbf{V}_{\mathrm{A}})^{\mathrm{T}}\ \mathbf{Q}_1^{-1}\ \mathbf{L}\,. \end{aligned} \tag{3.163}$$

Adding the term $\left[(\mathbf{A}+\mathbf{V}_{\mathrm{A}})^{\mathrm{T}}\mathbf{Q}_1^{-1}\mathbf{V}_{\mathrm{A}}\mathbf{X}\right]$ to both sides of the last equation yields

$$\left[(\mathbf{A}+\mathbf{V}_{\mathrm{A}})^{\mathrm{T}}\ \mathbf{Q}_1^{-1}\,(\mathbf{A}+\mathbf{V}_{\mathrm{A}})\right]\mathbf{X} = (\mathbf{A}+\mathbf{V}_{\mathrm{A}})^{\mathrm{T}}\ \mathbf{Q}_1^{-1}\,(\mathbf{L}+\mathbf{V}_{\mathrm{A}}\mathbf{X})\,. \tag{3.164}$$

The vector of unknown parameters reads

$$\hat{\mathbf{X}} = \left[(\mathbf{A}+\mathbf{V}_{\mathrm{A}})^{\mathrm{T}}\ \mathbf{Q}_1^{-1}\,(\mathbf{A}+\mathbf{V}_{\mathrm{A}})\right]^{-1}(\mathbf{A}+\mathbf{V}_{\mathrm{A}})^{\mathrm{T}}\ \mathbf{Q}_1^{-1}\,(\mathbf{L}+\mathbf{V}_{\mathrm{A}}\mathbf{X})\,. \tag{3.165}$$

Iterative solution for the adjustment results

A solution can be achieved, also in this case, iteratively. An initial approximation of the vector of unknowns $\mathbf{X}^0$, as well as the residual matrix $\mathbf{V}_{\mathrm{A}}^0$, are necessary and lead to

$$\mathbf{Q}_{1i} = \mathbf{Q}_{\mathrm{L}} - \mathbf{Q}_{\mathrm{LA}}(\mathbf{X}_i^0\otimes\mathbf{I}_n) - ({\mathbf{X}_i^0}^{\mathrm{T}}\otimes\mathbf{I}_n)\mathbf{Q}_{\mathrm{AL}} + \left({\mathbf{X}_i^0}^{\mathrm{T}}\otimes\mathbf{I}_n\right)\mathbf{Q}_{\mathrm{A}}(\mathbf{X}_i^0\otimes\mathbf{I}_n), \tag{3.166}$$

$$\hat{\mathbf{X}}_i = \left[\left(\mathbf{A}+\mathbf{V}_{\mathrm{A}i}^0\right)^{\mathrm{T}}(\mathbf{Q}_{1i})^{-1}\left(\mathbf{A}+\mathbf{V}_{\mathrm{A}i}^0\right)\right]^{-1}\left(\mathbf{A}+\mathbf{V}_{\mathrm{A}i}^0\right)^{\mathrm{T}}(\mathbf{Q}_{1i})^{-1}\left(\mathbf{L}+\mathbf{V}_{\mathrm{A}i}^0\ \mathbf{X}_i^0\right) \tag{3.167}$$

and

$$\hat{\mathbf{v}}_{\mathrm{A}_i} = \left[\mathbf{Q}_{\mathrm{AL}} - \mathbf{Q}_{\mathrm{A}}\left(\hat{\mathbf{X}}_i\otimes\mathbf{I}_n\right)\right]\mathbf{Q}_{1i}^{-1}\left(\mathbf{A}\hat{\mathbf{X}}_i-\mathbf{L}\right), \tag{3.168}$$

with i denoting the iteration step. The estimates for $\hat{\mathbf{X}}$ and $\hat{\mathbf{v}}_{\mathrm{A}}$ can be used to update the initial values for the next iteration step

$$\mathbf{X}_{i+1}^0 = \hat{\mathbf{X}}_i \quad \text{and} \quad \mathbf{V}_{\mathrm{A}i+1}^0 = \mathrm{invec}(\hat{\mathbf{v}}_{\mathrm{A}_i}),$$

with "invec" implying the inverse operator of "vec", i.e. rearranging a vector back into a matrix. The iterative procedure should continue until an appropriate "break-off" condition is met. The solution for the unknown parameters of equation (3.165) has been firstly presented by Fang (2011) within "*Algorithm 3*" and is identical to the solution of (Snow 2012) presented in "*Algorithm 2*".

3.3 Discussion and open questions

Two existing strategies have been analysed in this chapter for the solution of nonlinear adjustment problems with the method of least squares. The first has been traditionally used in geodesy and is based on the Gauss-Newton approach. It involves a linearization of the nonlinear functional model, which allows the representation of the mathematical model within a GMM or a GHM. This is a local optimization approach, as the solution from the iterative procedure converges to a local minimum of the objective function. After the assumption of "good" initial values for the unknown parameters, the estimated least squares solution would correspond to the global minimum.

The second approach that has been discussed is TLS, as proposed by Golub and Van Loan (1980) for the direct solution of a class of nonlinear least squares problems using SVD. In that work the solutions of two individual adjustment problems were presented for fitting a straight line in 2D. The least squares solution was estimated when only the y-coordinates of the points were regarded as measurements and the x-coordinates as error free (called ordinary least squares), in contrast to the TLS solution where both coordinates of the points were measurements. Petrovic (2003) has already pointed out that this comparison caused a confusion regarding the method of least squares, as many investigations draw the conclusion that TLS is a different method than least squares, or as stated by Groen (1996) that TLS is a generalization of the least squares method. For geodesists it has been already clear that the most important steps for the adjustment of observations is to build a correct mathematical model and minimize the objective function composed of correct residuals.

The relationship between nonlinear least squares problems and TLS was first placed under scrutiny by Neitzel and Petrovic (2008) and Neitzel (2010) for two individual nonlinear adjustment problems, this of fitting a straight line to equally weighted 2D data and for the 2D similarity transformation of coordinates. It has been shown that the TLS solution is identical to the least squares solution within the GHM, concluding that TLS can be regarded as a special case of least squares within the GHM. Additionally, Reinking (2008) showed that the TLS solution can be obtained using the traditional geodetic approaches. From (Neitzel and Petrovic 2008) and (Neitzel 2010) it can be seen that TLS *is not a new method*, but a new strategy or an approach for the solution of a class of nonlinear least squares problems. Therefore, the investigations in the next chapter try to answer the following arising questions:

- If it is possible to solve an adjustment problem with TLS and SVD, is it also possible to obtain the same eigenvalue problem from a classical least squares approach and solve the problem directly?
- Are there additional nonlinear least squares problems (besides the generally well-known case of the straight line fitting to equally weighted 2D data) which can be solved directly?
- Is it possible to classify those nonlinear least squares problems with a direct solution and solve them by using a systematic approach?

Furthermore, the cases of weighted nonlinear least squares problems, as well as the solution by using WTLS will be examined in a later chapter.

Part II - Methodological contributions

4 Direct solutions of nonlinear least squares problems with equal weights

The current chapter is based on the study of Malissiovas et al. (2016). It is an extended version of this article and includes the most important facts for the solution of adjustment problems with TLS. A clear mathematical relationship is presented between TLS and direct least squares solutions. Additionally, a systematic approach has been developed as a by-product of this investigation, for the direct solution of a class of nonlinear adjustment problems.

4.1 Basic idea and general methodology

The centre of interest is a class of nonlinear least squares problems that can be transformed into solving a polynomial equation (or the characteristic equation of an eigenvalue problem) and have a direct solution, depending on the degree of the resulting polynomial. Solutions for these adjustment problems have been presented in the TLS literature by using SVD. Therefore, the mathematical relationship is examined between direct solutions of nonlinear least squares and solutions coming from TLS for the following four adjustment cases:

1. Fitting of a straight line in 2D;
2. Fitting of a straight line in 3D;
3. Fitting of a plane in 3D;
4. 2D similarity transformation of coordinates.

In all four cases under investigation the coordinates in all directions are regarded as measurements. In TLS literature these problems are often distinguished as EIV model. Moreover, a regular adjustment model is always postulated here with the observed quantities being equally weighted and uncorrelated.

The concept of solving nonlinear least squares problems applied here is based directly on (Jovičić et al. 1982), where the adjustment problem of fitting a straight line to a set of points in 3D space was examined. In that work, the least squares estimate has been obtained by solving an eigenvalue problem, which is one of the key elements of TLS as well. Following the solution strategy from (Jovičić et al. 1982), a systematic approach for solving the four investigated adjustment problems has been established by Malissiovas et al. (2016). The proposed mathematical approach involves a sophisticated parametrization of the problem which can be always solved by building a Lagrange function that results in a quadratic or cubic algebraic equation.

In the following sections the least squares solution of the proposed approach from (Malissiovas et al. 2016) is derived and compared with the TLS solution for the four problems under investigation. It is shown that a clearly defined mathematical model of the adjustment problem, leading to an objective function based on the principle of least squares, is the Occam's razor[1] for TLS. A flowchart presenting both ways of solving directly the discussed nonlinear least squares problems is depicted in Figure 4.1.

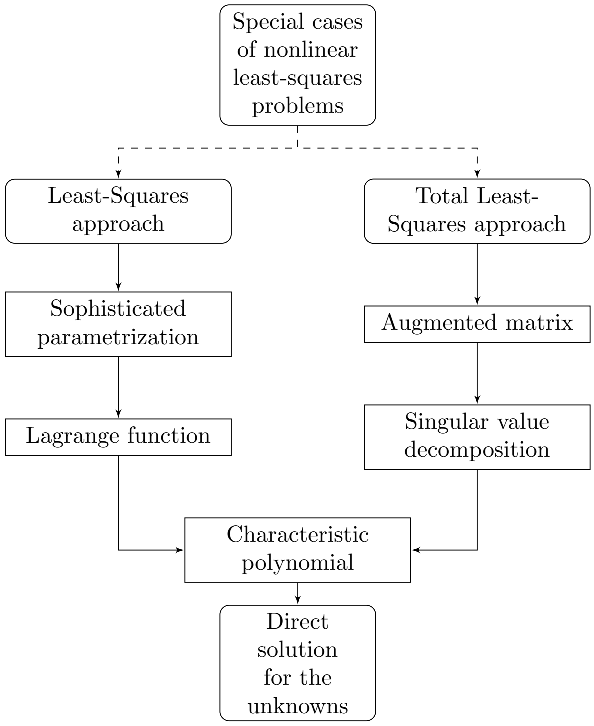

FIGURE 4.1: Flowchart for two possible direct solutions of a class of nonlinear least squares problems.

4.2 Fitting of a straight line in 2D

One of the first attempts to solve the nonlinear problem of least squares for fitting a straight line to a set of points in plane (i.e. in the 2D space) non-iteratively was done by Adcock (1878), who provided an elegant way of finding the direct solution to the problem. Pearson (1901) investigated the same problem by minimizing the sum of the squared orthogonal distances of every point to the requested line and he extended his study to fitting a plane to a set of points in the 3D space as well. On the other hand, the work of Golub and Van Loan (1980) provided an analysis of the TLS solution followed by the contributions of Groen (1996), Van Huffel (2004), Markovsky and Van Huffel (2007) and Schaffrin (2007). These authors always comprised the example of the straight line fit as the most appropriate example for illustrating the idea of TLS.

At the beginning an amount of 2D data is observed, e.g. a set of points with coordinates in x and in y direction. The question is how to fit a straight line to the measured points. The general form of a straight

[1]The Occam's razor can be defined as "the principle (attributed to William of Occam) that in explaining a thing no more assumptions should be made than are necessary. The principle is often invoked to defend reductionism or nominalism.", see for exampe the Oxford dictionary.

line in 2D[2] is (Bronshtein et al. 2005, p. 194)

$$ax + by + c = 0, \tag{4.1}$$

with the constant line parameters a, b, c. The first two parameters denote the components of a vector normal to the requested straight line, which intersepts the x-axis at $\frac{c}{a}$ and the y-axis at $\frac{c}{b}$, as it is depicted in Figure 4.2.

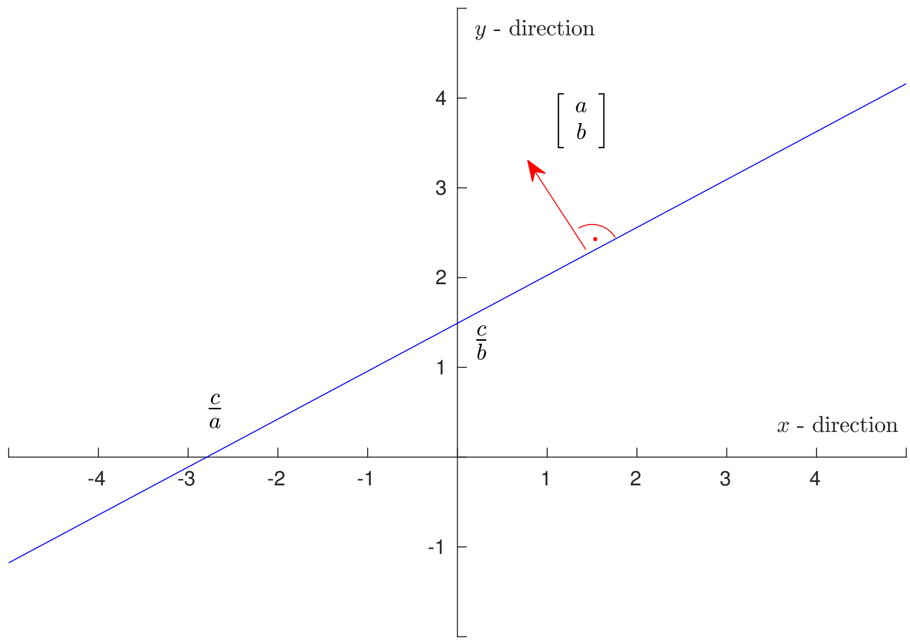

FIGURE 4.2: Representation of a straight line in 2D using equation (4.1).

The Hessian normal form of the straight line can be derived by multiplying equation (4.1) with the normalization parameter [3]

$$f = \pm\frac{1}{\sqrt{a^2 + b^2}}. \tag{4.2}$$

Introducing residuals v_x for the coordinates in the x direction and v_y for the coordinates in the y direction results in the nonlinear system of condition equations

$$a(x_i + v_{x_i}) + b(y_i + v_{y_i}) + c = 0, \tag{4.3}$$

with $i = 1, \ldots, n$, n denoting the number of observed points. Since the system of equations (4.3) is underdetermined, the least squares criterion can be used for an "optimal" solution by minimizing the sum of the squared residuals

$$\sum_{i=1}^{n} v_{x_i}^2 + v_{y_i}^2 \to \min. \tag{4.4}$$

[2]The same problem has been investigated in (Malissiovas et al., 2016), where the straight line in 2D has been represented in coordinate form.

[3]The sign of the scaling factor f is opposite to the sign of parameter c, as it is explained in (Bronshtein et al. 2005, p. 195).

4.2.1 Least squares adjustment with a direct solution

In this section, a direct least squares solution is developed for fitting a straight line in 2D when both coordinates are measurements and subject to errors. The unknown line parameters can be estimated directly by constructing and minimizing an appropriate Lagrange function and by solving a system of homogeneous normal equations. The goal is to show that the proposed approach leads, according to the chosen technique, to the solution of such algebraic equations that are equivalent to TLS. The solution for fitting a straight line in 2D by TLS is presented and analysed in the following subsection.

4.2.1.1 Definition of the problem

For solving an adjustment problem, it is important to clarify which quantities are observations and hence subject to random errors. This is necessary in order to define the objective function of the problem in an appropriate way. In this investigation both coordinates (in the direction of x and y) are subject to measurement errors. Furthermore, let all measurements be uncorrelated and of the same precision. Therefore, the aim is to find the shortest distance of each "measured" point to the requested straight line. As noticed already by Adcock (1878) the same precision of all coordinate measurements corresponds to the normal distances

$$D_i^2 = v_{x_i}^2 + v_{y_i}^2, \tag{4.5}$$

as measure of deviations, with $i = 1, ..., n$ (n is the number of observed points). This problem is depicted in Figure 4.3. Moreover, the normal distance of every point to the requested line can be expressed by (Bronshtein et al. 2005, p. 195)

$$D_i = \frac{a\ x_i + b\ y_i + c}{\sqrt{a^2 + b^2}}. \tag{4.6}$$

There are infinitely many choices for a condition that connects the three unknown parameters a, b and c for the general equation of the straight line. It is possible to restrict the problem to the usual $a = 1$ or $b = 1$, but in these cases some lines in plane are excluded[4]. From all remaining restrictions, the most appropriate for this study is

$$a^2 + b^2 = 1, \tag{4.7}$$

as it allows all lines in the plane to be calculated. Geometrically this restriction can be seen as a normalization of the orthogonal distances from every point to the requested line (i.e. the denominator of the orthogonal distance of equation (4.6) becomes 1), which results in

$$D_i = a\ x_i + b\ y_i + c. \tag{4.8}$$

The developed expressions for the orthogonal distances can serve as observation equations and be utilized as an alternative to the nonlinear condition equations (4.3) for estimating the unknown line parameters. An important remark is that the point coordinates x_i and y_i in this transformed functional model can be treated as fixed parameters. The orthogonal distances D_i are serving as random deviations, thus the observed quantities are zero pseudo-observations denoting the Euclidean distances of the points to the requested line.

[4] choosing $a = 1$, then there is no solution for lines parallel to the x direction and for $b = 1$ no solution for lines parallel to the y direction.

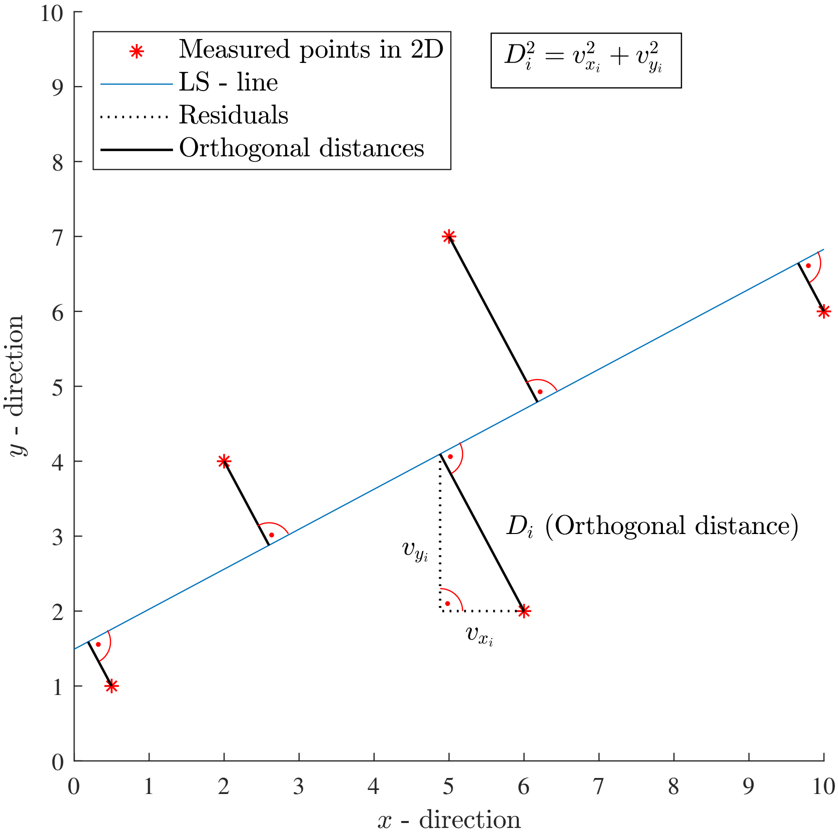

FIGURE 4.3: Example of fitting a straight line to points in 2D with both x and y coordinates subject to measurement errors.

This leads to the linear observation equations[5] for the distances

$$0_i + D_i = a\ x_i + b\ y_i + c. \tag{4.9}$$

Equivalently to the objective function (4.4), the least squares criterion can be applied to obtain the minimum normal distances from a set of points to the fitted line by minimizing the objective function

$$\Omega(a, b, c) = \sum_{i=1}^{n} v_{x_i}^2 + v_{y_i}^2 = \sum_{i=1}^{n} D_i^2 = \sum_{i=1}^{n} (a\ x_i + b\ y_i + c)^2. \tag{4.10}$$

We seek for a least squares solution for the unknown line parameters a, b and c that minimizes (4.10), subject to the restriction (4.7). Consequently, the Lagrangian

$$\mathrm{K}(a, b, c, \lambda) = \Omega(a, b, c) - k(a^2 + b^2 - 1), \tag{4.11}$$

[5]Although the observation equations (4.9) are linear, the least squares problem is nonlinear, due to the nonlinear constraint (4.7).

can be introduced, with k denoting the Lagrange multiplier. Differentiating function K with respect to all unknown parameters and setting the result to zero, yields the system of normal equations

$$\frac{\partial \mathrm{K}}{\partial a} = 2\left(a\left(\sum_{i=1}^{n} x_i^2 - k\right) + b\sum_{i=1}^{n} y_i x_i + c\sum_{i=1}^{n} x_i\right) = 0, \tag{4.12}$$

$$\frac{\partial \mathrm{K}}{\partial b} = 2\left(a\sum_{i=1}^{n} y_i x_i + b\left(\sum_{i=1}^{n} y_i^2 - k\right) + c\sum_{i=1}^{n} y_i\right) = 0, \tag{4.13}$$

$$\frac{\partial \mathrm{K}}{\partial c} = 2\left(a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} y_i + c\,n\right) = 0 \tag{4.14}$$

and

$$\frac{\partial \mathrm{K}}{\partial k} = -\left(a^2 + b^2 - 1\right) = 0. \tag{4.15}$$

Rearranging equation (4.14), yields a solution for parameter

$$c = -a\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) - b\left(\frac{1}{n}\sum_{i=1}^{n} y_i\right), \tag{4.16}$$

in terms of a and b. Introducing c into the normal equations (4.12) and (4.13) results in the reduced system of normal equations

$$a\left[\sum_{i=1}^{n} {x_i}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2 - k\right] + b\left[\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i\right)\right] = 0, \tag{4.17}$$

$$a\left[\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i\right)\right] + b\left[\sum_{i=1}^{n} {y_i}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2 - k\right] = 0. \tag{4.18}$$

If the Lagrange multiplier k were known, then equations (4.17) and (4.18) would form a homogeneous system of linear equations in a and b. Thus, the determinant of the equation system is equal to zero for a nontrivial solution

$$\begin{vmatrix} \left[\sum_{i=1}^{n} {x_i}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2 - k\right] & \left[\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i\right)\right] \\ \left[\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i\right)\right] & \left[\sum_{i=1}^{n} {y_i}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2 - k\right] \end{vmatrix} = 0, \tag{4.19}$$

which leads to a quadratic characteristic equation with two real and positive solutions for the unknown parameter k. The minimum solution, denoted by k_{min}, corresponds to the minimum of the Lagrange function (4.11). The solution for the unknown line parameters a and b can be computed by substituting the Lagrangian factor $\hat{k}_{min}$ into equations (4.17)-(4.18) subject to the chosen restriction (4.7). An equivalent solution can be obtained by transforming the equation system (4.17)-(4.18) into an eigenvalue problem.

4.2.1.2 Simplification of the problem by substituting one unknown parameter

A simplification of this adjustment problem can be easily achieved by replacing the unknown parameter c in the functional model (4.3), which leads to more elegant expressions for the orthogonal distances than equation (4.8). However, such a simplification makes sense only if the requested line will pass through the center of mass of the measured points, located at

$$y_c = \frac{1}{n}\sum_{i=1}^{n} y_i \quad \text{and} \quad x_c = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{4.20}$$

This has been already shown by the developed expression for parameter c in equation (4.16). A similar proof has been also presented by Jovičić et al. (1982) for the 3D line case and by Adcock (1878) and Malissiovas et al. (2016) for the 2D line as well. Therefore, introducing parameter c from equation (4.16) into (4.1), yields

$$a\ (x - x_c) + b\ (y - y_c) = 0. \tag{4.21}$$

The last equation can be further simplified by reducing the coordinates to a coordinate system with its origin located at the centre of mass of the given points. Geometrically, this is equivalent to shifting the coordinate system to a point that coincides numerically with the center of mass of the measured set of points, as it is depicted in Figure 4.4.

Thus, denoting the reduced coordinates of a point by

$$y' = y - y_c \quad \text{and} \quad x' = x - x_c, \tag{4.22}$$

leads to a system of simplified condition equations

$$a\ (x_i' + v_{x_i}) + b\ (y_i' + v_{y_i}) = 0 \tag{4.23}$$

and to the simplified expression for the normal distances

$$D_i = a\ x_i' + b\ y_i'. \tag{4.24}$$

Consequently, the objective function (4.10) can be rewritten as

$$\Omega(a,b) = \sum_{i=1}^{n} D_i^2 = \sum_{i=1}^{n}(a\ x_i' + b\ y_i')^2 = \ a^2\sum_{i=1}^{n} {x_i'}^2 + b^2\sum_{i=1}^{n} {y_i'}^2 + 2ab\sum_{i=1}^{n} y_i' x_i'. \tag{4.25}$$

We seek for a least squares solution for the unknown line parameters a and b that minimizes equation (4.25) subject to the restriction (4.7). The Lagrangian

$$\mathrm{K}(a,b,k) = \Omega(a,b) - k(a^2 + b^2 - 1), \tag{4.26}$$

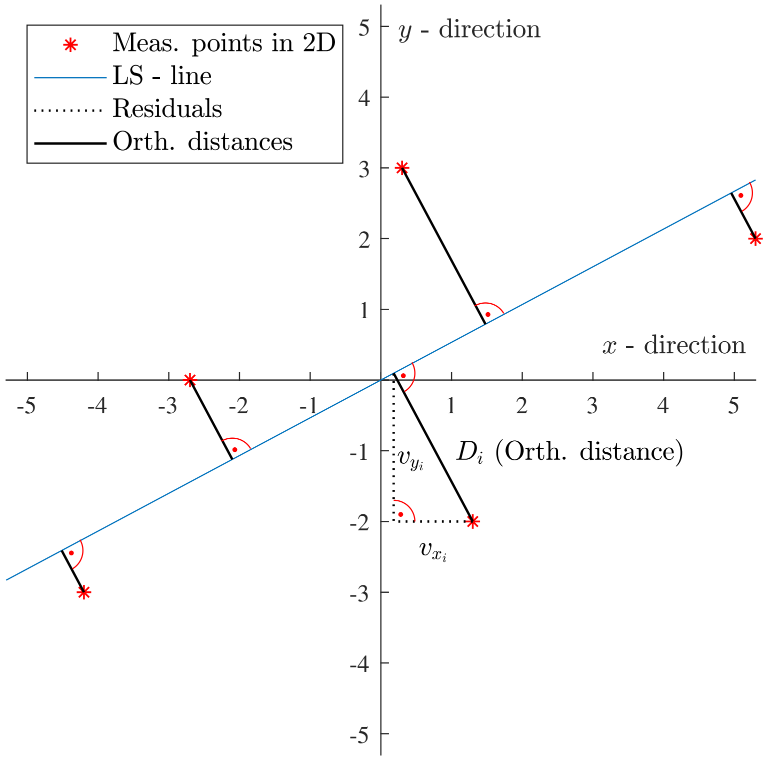

FIGURE 4.4: Example of fitting a straight line to points in 2D with coordinates reduced to the centre of mass of the measured points.

is introduced, where k is the Lagrange multiplier. Differentiating K with respect to all unknowns and setting the partial derivatives to zero, yields the normal equations

$$\frac{\partial \mathrm{K}}{\partial a} = 2a\left(\sum_{i=1}^{n} {x'_i}^2 - k\right) + 2b\left(\sum_{i=1}^{n} y'_i x'_i\right) = 0, \tag{4.27}$$

$$\frac{\partial \mathrm{K}}{\partial b} = 2a\left(\sum_{i=1}^{n} y'_i x'_i\right) + 2b\left(\sum_{i=1}^{n} {y'_i}^2 - k\right) = 0 \tag{4.28}$$

and

$$\frac{\partial \mathrm{K}}{\partial k} = -\left(a^2 + b^2 - 1\right) = 0. \tag{4.29}$$

If the Lagrange multiplier were known, then equations (4.27)-(4.28) would represent a homogeneous system of equations which is linear in the unknown line parameters a and b. A nontrivial solution can be obtained by setting the determinant of the equation system equal to zero:

$$\begin{vmatrix} \left(\sum_{i=1}^{n} {x'_i}^2 - k\right) & \sum_{i=1}^{n} y'_i x'_i \\ \sum_{i=1}^{n} y'_i x'_i & \left(\sum_{i=1}^{n} {y'_i}^2 - k\right) \end{vmatrix} = 0, \tag{4.30}$$

which leads to the quadratic characteristic equation

$$\left(\sum_{i=1}^{n} {x'_i}^2 - k\right)\left(\sum_{i=1}^{n} {y'_i}^2 - k\right) - \left(\sum_{i=1}^{n} y'_i x'_i\right)^2 = 0, \tag{4.31}$$

with one unknown parameter k and two real and positive solutions k_{min} and k_{max}. It can be shown that the smaller of the two solutions for k, denoted by k_{min}, corresponds to the minimum of the Lagrange function (4.26). There are two possibilities to determine a solution for the unknown parameters a and b, either by substituting the Lagrangian factor k_{min} into equations (4.27)-(4.28) or by solving an eigenvalue problem. It can be easily seen that both equations (4.19) and (4.30) result in quadratic characteristic equations that produce identical results for the unknown Lagrange multipliers k and the requested line parameters.

4.2.2 TLS solution with SVD

An alternative solution for finding the line that fits best to a set of points in 2D can be provided by TLS (Golub and Van Loan 1980, Groen 1996). According to (Golub and Van Loan, 1980) this solution can be represented geometrically by minimizing the orthogonal distances, as it is depicted in Figure 4.3. It is noteworthy in that contribution, that the least squares problem of fitting a straight line in 2D was regarded only when the y-coordinates are observations, in contrast to the definition and solution of the least squares problem that was presented in the previous section. Thus, the target is to provide an insight into the TLS approach and show that the TLS solution is equivalent to the one from the developed direct least squares approach of section 4.2.1.

Equation (4.1) for the straight line can be rewritten[6] as

$$y = \beta\ x + \gamma, \tag{4.32}$$

with

$$\beta = -\frac{b}{a}\ ,\ \gamma = -\frac{c}{a}. \tag{4.33}$$

Such a formulation for the straight line implies that the restriction $a = 1$ is taken into account. Therefore, it is not possible to describe all straight lines in plane, however, these are limited cases (in this problem all lines that are parallel to the y direction). Rearranging the functional model of equation (4.23), which already contains the reduced coordinates of the measured points to the centre of mass, yields the system of nonlinear equations

$$(y'_i + v_{y_i}) = \beta\ (x'_i + v_{x_i}), \tag{4.34}$$

[6] Greek letters are chosen to describe the parameters of the straight line in the TLS approach just for readability reasons.

leading to the EIV model of equation (3.121) with the respective quantities

$$\mathbf{L} = \begin{bmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_n \end{bmatrix}, \mathbf{v}_{\mathrm{L}} = \begin{bmatrix} v_{y_1} \\ v_{y_2} \\ \vdots \\ v_{y_n} \end{bmatrix}, \mathbf{X} = \begin{bmatrix} \beta \end{bmatrix}, \mathbf{A} = \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{bmatrix}, \mathbf{V}_{\mathrm{A}} = \begin{bmatrix} v_{x_1} \\ v_{x_2} \\ \vdots \\ v_{x_n} \end{bmatrix}. \tag{4.35}$$

Matrix $\mathbf{A}$ contains the coefficients of equation (4.34) with respect to the unknown parameter β, except of the residuals that are introduced into matrix $\mathbf{V}_{\mathrm{A}}$.

4.2.2.1 TLS solution based on the minimum eigenvalue principle

The TLS solution has been presented amongst others by Felus and Schaffrin (2005), as it has been already explained in subsection 3.2.1.1. The first step is to construct the augmented matrix

$$[\mathbf{A}, \mathbf{L}] = \begin{bmatrix} x'_1 & y'_1 \\ x'_2 & y'_2 \\ \vdots & \vdots \\ x'_n & y'_n \end{bmatrix} \tag{4.36}$$

and decompose it with the help of SVD, which yields

$$\mathbf{U\Sigma W}^{\mathrm{T}} = [\mathbf{A}, \mathbf{L}], \tag{4.37}$$

where the matrices $\mathbf{U}$ and $\mathbf{W}$ contain the left and right singular vectors of the augmented matrix respectively and matrix $\mathbf{\Sigma}$ is diagonal carrying the singular values. The TLS solution is in the right singular vector of matrix $\mathbf{W}$ that corresponds to the minimum singular value:

$$\mathbf{w}_{\min} = \mathbf{w}_{m+1} = [w_{1,m+1}, \cdots, w_{m,m+1}, w_{m+1,m+1}]^{\mathrm{T}}, \tag{4.38}$$

with the vector of unknowns computed by

$$\hat{\mathbf{X}} = -\frac{1}{w_{m+1,m+1}} [w_{1,m+1} : w_{m,m+1}]^{\mathrm{T}}. \tag{4.39}$$

4.2.2.2 Solution by the eigenvalue/eigenvector decomposition

To understand deeper the operation of SVD and the derivation of the adjusted unknowns of equation (4.39) it is important to explain SVD as the solution of the eigenproblem of the symmetric non-negative definite matrices $([\mathbf{A}, \mathbf{L}]^{\mathrm{T}}[\mathbf{A}, \mathbf{L}])$ and $([\mathbf{A}, \mathbf{L}][\mathbf{A}, \mathbf{L}]^{\mathrm{T}})$. According to (Lawson and Hanson 1974, p. 18) matrix $\mathbf{W}$ containing the right singular vectors of $[\mathbf{A}, \mathbf{L}]$ can be also estimated by the eigenvalue decomposition (EVD) of the squared matrix $([\mathbf{A}, \mathbf{L}]^{\mathrm{T}}[\mathbf{A}, \mathbf{L}])$:

$$\mathbf{W\Lambda W}^{\mathrm{T}} = [\mathbf{A}, \mathbf{L}]^{\mathrm{T}}[\mathbf{A}, \mathbf{L}], \tag{4.40}$$

where matrix $\mathbf{\Lambda}$ is a diagonal matrix carrying the eigenvalues of $[\mathbf{A}, \mathbf{L}]$. A relationship between eigenvalues and singular values can be found in (Golub and Van Loan 1989, p. 427), expressed as

$$\lambda_i = {\sigma_i}^2, \tag{4.41}$$

with λ and σ being the eigenvalues and singular values, respectively. For an explicit solution of this eigenproblem, matrix

$$[\mathbf{A}, \mathbf{L}]^{\mathrm{T}}[\mathbf{A}, \mathbf{L}] = \begin{bmatrix} x'_1 & x'_2 & \cdots & x'_n \\ y'_1 & y'_2 & \cdots & y'_n \end{bmatrix} \begin{bmatrix} x'_1 & y'_1 \\ x'_2 & y'_2 \\ \vdots & \vdots \\ x'_n & y'_n \end{bmatrix} = \mathbf{G}, \tag{4.42}$$

can be introduced, which can be rewritten in a more compact form as

$$\mathbf{G} = \begin{bmatrix} \sum_{i=1}^{n} {x'_i}^2 & \sum_{i=1}^{n} y'_i x'_i \\ \sum_{i=1}^{n} y'_i x'_i & \sum_{i=1}^{n} {y'_i}^2 \end{bmatrix}. \tag{4.43}$$

The eigenvalues and eigenvectors of matrix $\mathbf{G}$ can be computed from the eigenvalue problem

$$\mathbf{G}\mathbf{y} = \lambda\mathbf{y} \Rightarrow (\mathbf{G} - \lambda\mathbf{I})\mathbf{y} = \mathbf{0}, \tag{4.44}$$

as explained in (Bronshtein et al. 2005, p. 278). $\mathbf{I}$ is an identity matrix and $\mathbf{y}$ an eigenvector of $\mathbf{G}$. The eigenvalues of $\mathbf{G}$ can be determined by searching for non-trivial solutions $\mathbf{y} \neq 0$, i.e. by solving the characteristic equation of the eigenvalues

$$\begin{vmatrix} \sum_{i=1}^{n} {x'_i}^2 - \lambda & \sum_{i=1}^{n} y'_i x'_i \\ \sum_{i=1}^{n} y'_i x'_i & \sum_{i=1}^{n} {y'_i}^2 - \lambda \end{vmatrix} = 0, \tag{4.45}$$

or equivalently

$$\left(\sum_{i=1}^{n} {x'_i}^2 - \lambda\right)\left(\sum_{i=1}^{n} {y'_i}^2 - \lambda\right) - \left(\sum_{i=1}^{n} y'_i x'_i\right)^2 = 0. \tag{4.46}$$

This quadratic equation has two solutions for the unknown eigenvalues, λ_{min} and λ_{max}. By rearranging the eigenvalues and eigenvectors appropriately, the TLS solution for the line parameter β can be found from equation (4.39). Thus, by normalizing the eigenvector that corresponds to the smallest eigenvalue yields

$$\beta = \frac{\sum_{i=1}^{n} y'_i x'_i}{\sum_{i=1}^{n} {x'_i}^2 - \lambda_{min}}. \tag{4.47}$$

As expected, the TLS solution is identical with the developed direct least squares solution. This can be seen by simply comparing the developed characteristic equation of the eigenvalues (4.46), which corresponds to the quadratic equation (4.31) from the direct least squares solution. The conclusion is that the presented direct least squares solution for the nonlinear straight line fit in 2D already provides the exact result for TLS.

4.3 Fitting of a straight line in 3D

The problem of fitting a straight line to points in 3D space has been examined e.g. by Kampmann and Renner (2004), Kupferer (2004) or Späth (2004). The last from these authors developed an iterative algorithm for minimizing the sum of the squared orthogonal distances of the measured points to the fitted line and thus obtaining a least squares estimate for the unknown line parameters. A similar iterative solution can be found in the investigations of Snow and Schaffrin (2016), who have solved the problem using several models and always obtained identical results. Non-iterative adjustment solutions for the straight line fit in space can be found in the studies of (Jovičić et al. 1982) or (Drixler 1993, p. 46). Both can be transformed into an eigenvalue problem which gives the motivation for investigating the relationship with the TLS solution.

A representation of a straight line in 3D is given in (Bronshtein et al. 2005, p. 217) , expressed as

$$\frac{y-y_0}{a} = \frac{x-x_0}{b} = \frac{z-z_0}{c}, \tag{4.48}$$

for a line that passes through a point with coordinates x_0, y_0 and z_0 and is parallel to a direction vector with components a, b and c. The target is to minimize the errors in all x, y and z coordinates, which implies the nonlinearity of the system of condition equations

$$\begin{aligned} a(x_i + v_{x_i} - x_0) - b(y_i + v_{y_i} - y_0) &= 0, \\ b(z_i + v_{z_i} - z_0) - c(x_i + v_{x_i} - x_0) &= 0, \\ c(y_i + v_{y_i} - y_0) - a(z_i + v_{z_i} - z_0) &= 0, \end{aligned} \tag{4.49}$$

with $i = 1, \ldots, n$, n being the number of observed points in 3D space. The best line passing through the 3D point cloud can be obtained by minimizing the sum of squared residuals from all coordinates

$$\sum_{i=1}^{n} v_{x_i}^2 + v_{y_i}^2 + v_{z_i}^2 \rightarrow \min. \tag{4.50}$$

In order to solve the problem stated above, two additional constraints (or restrictions between the unknown parameters) have to be taken into account, as it has been already discussed by Snow and Schaffrin (2016). However, the selection of a proper restriction is avoided at this point. An appropriate parametrization of the problem is attempted that involves a substitution of some unknown parameters with known, following the same procedure presented in (Malissiovas et al. 2016).

4.3.1 Direct least squares solution for fitting a straight line in 3D

Analogously to the investigated case of the previous section (fitting of a straight line in 2D), the same precision of all coordinate measurements would correspond to the normal distances

$$D_i^2 = v_{x_i}^2 + v_{y_i}^2 + v_{z_i}^2, \tag{4.51}$$

as measures of deviations between the observed points and the requested line. As explained in (Bronshtein et al. 2005, p. 218), the squared normal distance between a point and a line in space is

$$D^2 = \frac{[a(x-x_0) - b(y-y_0)]^2 + [b(z-z_0) - c(x-x_0)]^2 + [c(y-y_0) - a(z-z_0)]^2}{a^2+b^2+c^2}. \tag{4.52}$$

Furthermore, it is possible to reduce the number of the unknowns of the model by replacing the parameters x_0, y_0 and z_0 with the coordinates of the centre of mass[7]

$$y_c = \frac{1}{n}\sum_{i=1}^{n} y_i \ , \quad x_c = \frac{1}{n}\sum_{i=1}^{n} x_i \ , \quad z_c = \frac{1}{n}\sum_{i=1}^{n} z_i, \tag{4.53}$$

of the n 3D points. Therefore, equation (4.48) can be rewritten as

$$\frac{y-y_c}{a} = \frac{x-x_c}{b} = \frac{z-z_c}{c}. \tag{4.54}$$

Solution with coordinates reduced to the centre of mass

A reduction of all coordinates to the centre of mass leads to the simplified functional model

$$\frac{y'}{a} = \frac{x'}{b} = \frac{z'}{c}, \tag{4.55}$$

and the condition equations

$$\begin{aligned} a(x_i' + v_{x_i}) - b(y_i' + v_{y_i}) &= 0, \\ b(z_i' + v_{z_i}) - c(x_i' + v_{x_i}) &= 0, \\ c(y_i' + v_{y_i}) - a(z_i' + v_{z_i}) &= 0, \end{aligned} \tag{4.56}$$

with x', y' and z' being coordinates reduced to the centre of mass of the 3D point cloud. The squared normal distances of the reduced points can be formulated as

$$D_i^2 = \frac{(a\ x_i' - b\ y_i')^2 + (b\ z_i' - c\ x_i')^2 + (c\ y_i' - a\ z_i')^2}{a^2+b^2+c^2}. \tag{4.57}$$

In this case the most appropriate restriction between the unknown parameters can be selected as

$$a^2 + b^2 + c^2 = 1 \tag{4.58}$$

[7] The proof that this parameter replacement is allowed can be found in (Jovičić et al. 1982).

and makes possible to derive a simplified expression for the squared orthogonal distances

$$D_i^2 = (a\ x_i' - b\ y_i')^2 + (b\ z_i' - c\ x_i')^2 + (c\ y_i' - a\ z_i')^2. \tag{4.59}$$

Therefore, the best line can be estimated by minimizing the objective function

$$\begin{aligned} \Omega(a,b,c) = \sum_{i=1}^{n} v_{x_i}^2 + v_{y_i}^2 + v_{z_i}^2 = \sum_{i=1}^{n} D_i^2 \\ = \sum_{i=1}^{n}[(a\ x_i' - b\ y_i')^2 + (b\ z_i' - c\ x_i')^2 + (c\ y_i' - a\ z_i')^2]. \end{aligned} \tag{4.60}$$

In order to derive the minimum of the objective function under the restriction of equation (4.58), the Lagrangian

$$\mathrm{K}(a,b,c,k) = \Omega(a,b,c) - k(a^2 + b^2 + c^2 - 1) \tag{4.61}$$

can be built. A differentiation of function K with respect to all unknowns and setting the resulting partial derivatives to zero, leads to the system of normal equations

$$\frac{\partial \mathrm{K}}{\partial a} = 2a\left(\sum_{i=1}^{n} x_i'^2 + \sum_{i=1}^{n} z_i'^2 - k\right) - 2b\sum_{i=1}^{n} y_i'x_i' - 2c\sum_{i=1}^{n} y_i'z_i' = 0, \tag{4.62}$$

$$\frac{\partial \mathrm{K}}{\partial b} = -2a\sum_{i=1}^{n} y_i'x_i' + 2b\left(\sum_{i=1}^{n} y_i'^2 + \sum_{i=1}^{n} z_i'^2 - k\right) - 2c\sum_{i=1}^{n} x_i'z_i' = 0, \tag{4.63}$$

$$\frac{\partial \mathrm{K}}{\partial c} = -2a\sum_{i=1}^{n} y_i'z_i' - 2b\sum_{i=1}^{n} x_i'z_i' + 2c\left(\sum_{i=1}^{n} y_i'^2 + \sum_{i=1}^{n} x_i'^2 - k\right) = 0, \tag{4.64}$$

and

$$\frac{\partial \mathrm{K}}{\partial k} = -\left(a^2 + b^2 + c^2 - 1\right) = 0. \tag{4.65}$$

Equations (4.62) to (4.64) can be interpreted as a homogeneous system of equations, with the solution for parameter k obtained from

$$\begin{vmatrix} (p_1 - k) & q_1 & q_2 \\ q_1 & (p_2 - k) & q_3 \\ q_2 & q_3 & (p_3 - k) \end{vmatrix} = 0, \tag{4.66}$$

with the respective elements

$$\begin{aligned} p_1 = \sum_{i=1}^{n} x_i'^2 + \sum_{i=1}^{n} z_i'^2 \quad , \quad p_2 = \sum_{i=1}^{n} y_i'^2 + \sum_{i=1}^{n} z_i'^2 \quad , \quad p_3 = \sum_{i=1}^{n} y_i'^2 + \sum_{i=1}^{n} x_i'^2 \\ q_1 = -\sum_{i=1}^{n} y_i'x_i' \quad , \quad q_2 = -\sum_{i=1}^{n} y_i'z_i' \ \text{ and } \ q_3 = -\sum_{i=1}^{n} x_i'z_i'. \end{aligned} \tag{4.67}$$

Equation (4.66) is a cubic characteristic equation[8] with the unknown parameter k. The adjusted line parameters a, b and c can be estimated by substituting k_{min} into equations (4.62) - (4.64) under the

[8] see for example (Bronshtein et al. 2005, p. 261) for the calculation of the value of a determinant of third order.

specified restriction or by transforming the equation system into an eigenvalue problem.

4.3.2 TLS fitting of a straight line in 3D

A TLS solution for fitting a straight line in 3D using SVD has been presented for the first time in (Malissiovas et al. 2016). In this section an equivalent solution is presented using a slightly modified functional model.

In order to build the adjustment model of equation (3.121) it is necessary to derive an appropriate functional model. Thus, rearranging appropriately equation (4.56) yields

$$\begin{aligned} &\alpha(x' + v_{x_i}) - \beta(y' + v_{y_i}) = 0, \\ &0\ \alpha - \beta(z' + v_{z_i}) = x' + v_{x_i}, \\ &-\alpha(z' + v_{z_i}) + 0\ \beta = y' + v_{y_i}, \end{aligned} \tag{4.68}$$

with

$$\alpha = -\frac{a}{c} \quad \text{and} \quad \beta = -\frac{b}{c}. \tag{4.69}$$

The derived system of nonlinear equations can be expressed within an EIV model, with the respective quantities being

$$\mathbf{L} = \begin{bmatrix} 0 \\ x'_1 \\ y'_1 \\ \vdots \\ 0 \\ x'_n \\ y'_n \end{bmatrix}, \mathbf{v}_\mathrm{L} = \begin{bmatrix} 0 \\ v_{x_1} \\ v_{y_2} \\ \vdots \\ 0 \\ v_{x_n} \\ v_{y_n} \end{bmatrix}, \mathbf{X} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \mathbf{A} = \begin{bmatrix} x'_1 & -y'_1 \\ 0 & -z'_1 \\ -z'_1 & 0 \\ \vdots & \vdots \\ x'_n & -y'_n \\ 0 & -z'_n \\ -z'_n & 0 \end{bmatrix}, \mathbf{V}_\mathrm{A} = \begin{bmatrix} v_{x_1} & -v_{y_1} \\ 0 & -v_{z_1} \\ -v_{z_1} & 0 \\ \vdots & \vdots \\ v_{x_n} & -v_{y_n} \\ 0 & -v_{z_n} \\ -v_{z_n} & 0 \end{bmatrix}. \tag{4.70}$$

The first column of matrix $\mathbf{A}$ contains the coefficients of the functional model (4.68) with respect to the unknown parameter α, whilst in the second column are the coefficients with respect to the unknown parameter β.

The augmented matrix is

$$[\mathbf{A}, \mathbf{L}] = \begin{bmatrix} x'_1 & -y'_1 & 0 \\ 0 & -z'_1 & x'_1 \\ -z'_1 & 0 & y'_1 \\ \vdots & \vdots & \vdots \\ x'_n & -y'_n & 0 \\ 0 & -z'_n & x'_n \\ -z'_n & 0 & y'_n \end{bmatrix}. \tag{4.71}$$

The right singular vectors of the augmented matrix can be estimated by the eigenvalue/eigenvector decomposition. Thus, the squared augmented matrix is

$$[\mathbf{A},\mathbf{L}]^{\mathrm{T}}[\mathbf{A},\mathbf{L}] = \begin{bmatrix} p_1 & q_1 & q_2 \\ q_1 & p_2 & q_3 \\ q_2 & q_3 & p_3 \end{bmatrix} = \mathbf{G} \ , \tag{4.72}$$

with the respective elements p and q corresponding to those of equation (4.67). The eigenvalues and eigenvectors of matrix $\mathbf{G}$ can be found by employing the generalised eigenvalue problem, which results in

$$\begin{vmatrix} (p_1-\lambda) & q_1 & q_2 \\ q_1 & (p_2-\lambda) & q_3 \\ q_2 & q_3 & (p_3-\lambda) \end{vmatrix} = 0. \tag{4.73}$$

The derived determinant provides the characteristic equation of the eigenvalues. In this case this is a cubic characteristic equation with three solutions for the unknown eigenvalues λ. The adjusted line parameters $\hat{\alpha}$ and $\hat{\beta}$ can be found by employing the minimum eigenvalue principle of equation (4.38). The right eigenvector corresponding to the smallest eigenvalue of matrix $\mathbf{G}$ holds the TLS solution for the 3D line parameters.

Obviously the elements of matrix $\mathbf{G}$ coincide with those from the direct least squares solution. The determinants (4.66) and (4.73) are equal, leading to identical characteristic equations. Therefore, the TLS solution for the nonlinear problem of the straight line fit in 3D space is identical with the presented direct least squares solution.

4.4 Fitting of a plane in 3D

The third case under investigation is the nonlinear problem of fitting a plane to a 3D point cloud with all coordinates being subject to measurement errors. Also for this case all coordinates of the points are regarded as uncorrelated observations of equal precision. Several results from various TLS algorithms were presented for this problem in (Schaffrin et al. 2006), which resulted in a slight deviation from the least squares solution. Therefore, a mathematical relation between the TLS and least squares solution is built for fitting a plane in 3D, following the same line of thinking as in the previous application cases.

The general equation of a plane in 3D can be found in (Bronshtein et al. 2005, p. 214), which reads

$$ax + by + cz + d = 0, \tag{4.74}$$

with x, y and z being the 3D coordinates of a point that lies in the plane. a, b, c and d are the plane parameters. Assuming that the coordinates in all directions are observed quantities, a system of nonlinear condition equations emerges.

$$a(x_i + v_{x_i}) + b(y_i + v_{y_i}) + c(z_i + v_{z_i}) + d = 0 \tag{4.75}$$

Applying the least squares criterion, the plane that fits best to the observed point cloud can be estimated by minimizing the sum of squared residuals

$$\sum_{i=1}^{n} v_{x_i}^2 + v_{y_i}^2 + v_{z_i}^2 \to \text{min.} \tag{4.76}$$

4.4.1 Direct least squares solution for fitting a plane in 3D

Fitting a plane to points in 3D, with all coordinates being subject to measurement errors, is similar to the case of fitting a straight line in plane. Therefore, the objective function (4.76) is equal to the sum of squared normal distances of the points to the requested plane

$$\sum_{i=1}^{n} v_{x_i}^2 + v_{y_i}^2 + v_{z_i}^2 = \sum_{i=1}^{n} D_i^2, \tag{4.77}$$

with the normal distances being expressed by

$$D_i = \frac{ax_i + by_i + cz_i + d}{\sqrt{a^2 + b^2 + c^2}}, \tag{4.78}$$

as it is shown in (Bronshtein et al. 2005, p. 214). A simplification of the problem is also in this case possible by replacing one unknown parameter. Thus, reducing the coordinates of the point cloud to the centre of mass (see equation 4.53) results in the substitution of parameter d and the simplified functional model

$$ax' + by' + cz' = 0, \tag{4.79}$$

with x', y' and z' denoting the coordinates of a point reduced to the centre of mass of the 3D point cloud.

Solution with coordinates reduced to the centre of mass

The developed functional model (4.79) leads to the system of nonlinear condition equations

$$a(x_i' + v_{x_i}) + b(y_i' + v_{y_i}) + c(z_i' + v_{z_i}) = 0 \tag{4.80}$$

and the orthogonal distances

$$D_i = \frac{ax_i' + by_i' + cz_i'}{\sqrt{a^2 + b^2 + c^2}}. \tag{4.81}$$

Since equation (4.81) can be scaled by an arbitrary factor, which means that only two out of the three parameters a, b or c are independent, an appropriate constraint would be

$$a^2 + b^2 + c^2 = 1. \tag{4.82}$$

Thus, the expression for the normal distances can be rewritten as

$$D_i = ax_i' + by_i' + cz_i' \tag{4.83}$$

and the objective function under minimization is

$$\Omega(a,b,c) = \sum_{i=1}^{n} D_i^2 = \sum_{i=1}^{n} (ax'_i + by'_i + cz'_i)^2. \tag{4.84}$$

A least squares solution for the unknown parameters a, b and c is required that minimizes $\Omega(a,b,c)$, subject to the constraint (4.82). The Lagrangian

$$\mathrm{K}(a,b,c,k) = \Omega(a,b,c) - k(a^2+b^2+c^2-1), \tag{4.85}$$

can be built. The differentiation of K with respect to the unknown plane parameters leads, after setting the partial derivatives to zero, to the system of normal equations

$$\frac{\partial \mathrm{K}}{\partial a} = 2\left[a\left(\sum_{i=1}^{n} x_i'^2 - k\right) + b\left(\sum_{i=1}^{n} y'_i x'_i\right) + c\left(\sum_{i=1}^{n} x'_i z'_i\right)\right] = 0, \tag{4.86}$$

$$\frac{\partial \mathrm{K}}{\partial b} = 2\left[a\left(\sum_{i=1}^{n} y'_i x'_i\right) + b\left(\sum_{i=1}^{n} y_i'^2 - k\right) + c\left(\sum_{i=1}^{n} y'_i z'_i\right)\right] = 0, \tag{4.87}$$

$$\frac{\partial \mathrm{K}}{\partial c} = 2\left[a\left(\sum_{i=1}^{n} x'_i z'_i\right) + b\left(\sum_{i=1}^{n} y'_i z'_i\right) + c\left(\sum_{i=1}^{n} z_i'^2 - k\right)\right] = 0 \tag{4.88}$$

and

$$\frac{\partial \mathrm{K}}{\partial k} = -\left(a^2+b^2+c^2-1\right) = 0. \tag{4.89}$$

The solution for the Lagrange multiplier can be derived from

$$\begin{vmatrix} (r_1-k) & s_1 & s_2 \\ s_1 & (r_2-k) & s_3 \\ s_2 & s_3 & (r_3-k) \end{vmatrix} = 0, \tag{4.90}$$

with the quantities

$$r_1 = \sum_{i=1}^{n} x_i'^2, \quad r_2 = \sum_{i=1}^{n} y_i'^2, \quad r_3 = \sum_{i=1}^{n} z_i'^2,$$
$$s_1 = \sum_{i=1}^{n} y'_i x'_i, \quad s_2 = \sum_{i=1}^{n} x'_i z'_i \;\text{ and }\; s_3 = \sum_{i=1}^{n} y'_i z'_i. \tag{4.91}$$

This is a cubic characteristic equation and has three solutions for k. The unknown plane parameters a, b and c can be estimated, either by inserting $\hat{k}_{min}$ in equations (4.86) - (4.88), under the restriction (4.82), or by transforming the equation system into an eigenvalue problem. The presented direct solution for fitting a plane in 3D coincides with that of Linkwitz (1976).

4.4.2 TLS fitting of a plane in 3D

The TLS solution for fitting a plane in 3D can be derived analogously to the investigations of Schaffrin et al. (2006), following however a different functional model. Based on the presented approach for obtaining a TLS estimate, the functional model of equation (4.79) can be rewritten as

$$z' = -\frac{a}{c}x' - \frac{b}{c}y' \Rightarrow \; z'_i = \alpha x' + \beta y', \tag{4.92}$$

with

$$\alpha = -\frac{a}{c} \; \text{ and } \; \beta = -\frac{b}{c}. \tag{4.93}$$

Therefore, the system of condition equations (4.80) becomes

$$z' + v_{z_i} = \alpha(x' + v_{x_i}) + \beta(y' + v_{y_i}), \tag{4.94}$$

and can be expressed by an EIV model, after introducing the following matrices:

$$\mathbf{L} = \begin{bmatrix} z'_1 \\ z'_2 \\ \vdots \\ z'_n \end{bmatrix}, \mathbf{v}_\mathrm{L} = \begin{bmatrix} v_{z_1} \\ v_{z_2} \\ \vdots \\ v_{z_n} \end{bmatrix}, \mathbf{X} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \mathbf{A} = \begin{bmatrix} x'_1 & y'_1 \\ x'_2 & y'_2 \\ \vdots & \vdots \\ x'_n & y'_n \end{bmatrix}, \mathbf{V}_\mathrm{A} = \begin{bmatrix} v_{x_1} & v_{y_1} \\ v_{x_2} & v_{y_2} \\ \vdots & \vdots \\ v_{x_n} & v_{y_n} \end{bmatrix}. \tag{4.95}$$

The first column of the coefficient matrix $\mathbf{A}$ contains the coefficients of the condition equations (4.94) with respect to the unknown parameter α while in the second column are the coefficients with respect to β. Furthermore, it is possible to build the augmented matrix

$$[\mathbf{A}, \mathbf{L}] = \begin{bmatrix} x'_1 & y'_1 & z'_1 \\ x'_2 & y'_2 & z'_2 \\ \vdots & \vdots & \vdots \\ x'_n & y'_n & z'_n \end{bmatrix} \tag{4.96}$$

and the square matrix

$$[\mathbf{A}, \mathbf{L}]^\mathrm{T}[\mathbf{A}, \mathbf{L}] = \begin{bmatrix} x'_1 & x'_2 & \dots & x'_n \\ y'_1 & y'_2 & \dots & y'_n \\ z'_1 & z'_2 & \dots & z'_n \end{bmatrix} \begin{bmatrix} x'_1 & y'_1 & z'_1 \\ x'_2 & y'_2 & z'_2 \\ \vdots & \vdots & \vdots \\ x'_n & y'_n & z'_n \end{bmatrix} = \mathbf{G}. \tag{4.97}$$

This is equivalent to

$$\mathbf{G} = \begin{bmatrix} r_1 & s_1 & s_2 \\ s_1 & r_2 & s_3 \\ s_2 & s_3 & r_3 \end{bmatrix}, \tag{4.98}$$

with the respective elements being identical to those of equation (4.91). The eigenvalues and eigenvectors of matrix **G** can be computed from the generalised eigenvalue problem, by solving

$$\begin{vmatrix} (r_1-\lambda) & s_1 & s_2 \\ s_1 & (r_2-\lambda) & s_3 \\ s_2 & s_3 & (r_3-\lambda) \end{vmatrix} = 0. \tag{4.99}$$

This characteristic cubic equation has three solutions for the eigenvalue λ. The unknown parameters α and β can be estimated using the minimum eigenvalue principle. The presented least squares solution for fitting a plane in 3D coincides perfectly with the TLS solution. Equations (4.90) and (4.99) are identical and only the name of the unknown (k or λ) is different.

4.5 2D similarity transformation of coordinates

The 2D similarity transformation of coordinates is one of the most frequent geodetic and photogrammetric applications. A first attempt to estimate the TLS solution of the problem using SVD was that of Felus and Schaffrin (2005) by presenting a Strucured TLS (STLS) algorithm for solving the problem. Neitzel (2010) has shown that this algorithm needs to be modified for estimating the correct solution. For this reason, the same problem has been examined again by Schaffrin et al. (2012). Their modified solution is iterative, however, they state that a TLS solution using SVD could be possible. Here, a new approach is presented for a direct solution of the problem (least squares and also TLS solution via SVD).

The well-known equation for the planar coordinate transformation is

$$\begin{bmatrix} X_i \\ Y_i \end{bmatrix} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} \mu & 0 \\ 0 & \mu \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}, \tag{4.100}$$

see for example (Felus and Schaffrin 2005). This can be written equivalently as

$$\begin{aligned} X_i &= (\mu\ \cos\phi)x_i - (\mu\ \sin\phi)y_i + t_x \\ Y_i &= (\mu\ \sin\phi)x_i + (\mu\ \cos\phi)y_i + t_y, \end{aligned} \tag{4.101}$$

with $i = 1, ..., n$, where n is the number of observed homologous points in the source xy and target XY system. The unknown transformation parameters between the two systems are:

- ϕ = rotation angle
- μ = scale factor
- t_x = translation in x direction
- t_y = translation in y direction

Introducing the parameters

$$\xi_1 = \mu\ \cos\phi \quad \text{and} \quad \xi_2 = \mu\ \sin\phi, \tag{4.102}$$

it is possible to obtain the simplified equation system

$$\begin{aligned} X_i &= \xi_1\, x_i - \xi_2\, y_i + t_x, \\ Y_i &= \xi_2\, x_i + \xi_1\, y_i + t_y. \end{aligned} \tag{4.103}$$

If all point coordinates are considered as measured quantities, the necessary residuals are introduced in the functional model resulting in the nonlinear system of condition equations

$$\begin{aligned} X_i + v_{X_i} - \xi_1\,(x_i + v_{x_i}) + \xi_2\,(y_i + v_{y_i}) - t_x &= 0, \\ Y_i + v_{Y_i} - \xi_2\,(x_i + v_{x_i}) - \xi_1\,(y_i + v_{y_i}) - t_y &= 0. \end{aligned} \tag{4.104}$$

The least squares criterion can be employed for an "optimal" solution by minimizing the sum of the squared residuals in both coordinate systems:

$$\sum_{i=1}^{n} v_{x_i}^2 + v_{y_i}^2 + v_{X_i}^2 + v_{Y_i}^2 \to \min. \tag{4.105}$$

4.5.1 Direct least squares solution for the 2D similarity transformation

For a realistic functional model the translation vector has to be present. However, the substitution of the translations from the functional model is possible and this can be proven in the same way as for the previous investigated cases, by showing that

$$\begin{aligned} t_x &= X_c - \xi_1\, x_c + \xi_2\, y_c, \\ t_y &= Y_c - \xi_2\, x_c - \xi_1\, y_c, \end{aligned} \tag{4.106}$$

with x_c and y_c denoting the coordinates of the centre of mass of the points in the source system and X_c and Y_c in the target system, computed by

$$x_c = \frac{1}{n}\sum_{i=1}^{n} x_i\ ,\quad y_c = \frac{1}{n}\sum_{i=1}^{n} y_i\ ,\quad X_c = \frac{1}{n}\sum_{i=1}^{n} X_i\ ,\quad Y_c = \frac{1}{n}\sum_{i=1}^{n} Y_i. \tag{4.107}$$

Therefore, a reduction of all coordinates to their centre of mass leads to a simplified functional model

$$\begin{aligned} X_i' + v_{X_i} - \xi_1\,(x_i' + v_{x_i}) + \xi_2\,(y_i' + v_{y_i}) &= 0, \\ Y_i' + v_{Y_i} - \xi_2\,(x_i' + v_{x_i}) - \xi_1\,(y_i' + v_{y_i}) &= 0. \end{aligned} \tag{4.108}$$

Appropriate parametrization of the problem

In order to obtain a direct solution in the same manner as in the previous sections, an additional unknown parameter has to be taken into consideration [9]. For this reason the functional model (4.108) can be rewritten

[9] Here the problem is overparametrized and a meaningful constraint between the unknown parameters is chosen. However, this is not necessary for obtaining a solution but for being consistent with the solution strategy that was followed in the adjustment problems of the previous sections. Especially for showing that the TLS solution is identical with the solution from the proposed direct least squares approach.

as

$$\begin{aligned}\gamma\,(X'_i + v_{X_i}) + \alpha\,(x'_i + v_{x_i}) - \beta\,(y'_i + v_{y_i}) = 0,\\ \gamma\,(Y'_i + v_{Y_i}) + \beta\,(x'_i + v_{x_i}) + \alpha\,(y'_i + v_{y_i}) = 0,\end{aligned} \tag{4.109}$$

with

$$\xi_1 = -\frac{\alpha}{\gamma} \quad \text{and} \quad \xi_2 = -\frac{\beta}{\gamma}. \tag{4.110}$$

The enforced additional unknown parameter (γ can be seen as an additional parameter) requires a restriction between the unknowns. For the purposes of this research, a "meaningful" constraint is chosen as

$$\alpha^2 + \beta^2 + \gamma^2 = 1. \tag{4.111}$$

Solution with coordinates reduced to the centre of mass

The coordinates of the points in both coordinate systems are subject to measurement errors. By employing the least squares criterion, the goal is to minimize the errors in all homologous points and in both directions. This is equivalent to the minimization of the Euclidean distances between the points in the target system and the transformed homologous points from the source system

$$\sum_{i=1}^{n} v_{x_i}^2 + v_{y_i}^2 + v_{X_i}^2 + v_{Y_i}^2 = \sum_{i=1}^{n} D_i^2 \rightarrow \min, \tag{4.112}$$

with the squared distances between two homologous points expressed as

$$D_i^2 = (\gamma\,X'_i + \alpha\,x'_i - \beta\,y'_i)^2 + (\gamma\,Y'_i + \beta\,x'_i + \alpha\,y'_i)^2. \tag{4.113}$$

Therefore, the objective function under minimization becomes

$$\Omega(\alpha,\beta,\gamma) = \sum_{i=1}^{n} D_i^2 = \sum_{i=1}^{n} \Big[(\gamma\,X'_i + \alpha\,x'_i - \beta\,y'_i)^2 + (\gamma\,Y'_i + \beta\,x'_i + \alpha\,y'_i)^2\Big]. \tag{4.114}$$

A least squares solution for the unknown transformation parameters α, β and γ is desired, that minimizes Ω under the restriction (4.111). Thus, the Lagrange function can be built as

$$\mathrm{K}(\alpha,\beta,\gamma,k) = \Omega(\alpha,\beta,\gamma) - k(\alpha^2 + \beta^2 + \gamma^2 - 1). \tag{4.115}$$

Differentiating the Lagrangian K with respect to all unknown parameters and setting the partial derivatives to zero, yields the system of normal equations

$$\frac{\partial \mathrm{K}}{\partial \alpha} = 2\left[\alpha\left(\sum_{i=1}^{n} {x'_i}^2 + \sum_{i=1}^{n} {y'_i}^2 - k\right) + \gamma\left(\sum_{i=1}^{n} x'_i X'_i + \sum_{i=1}^{n} y'_i Y'_i\right)\right] = 0, \tag{4.116}$$

$$\frac{\partial \mathrm{K}}{\partial \beta} = 2\left[\beta\left(\sum_{i=1}^{n} {x'_i}^2 + \sum_{i=1}^{n} {y'_i}^2 - k\right) + \gamma\left(\sum_{i=1}^{n} x'_i Y'_i - \sum_{i=1}^{n} y'_i X'_i\right)\right] = 0, \tag{4.117}$$

$$\frac{\partial \mathrm{K}}{\partial \gamma} = 2\left[\alpha\left(\sum_{i=1}^{n} x_i' X_i' + \sum_{i=1}^{n} y_i' Y_i'\right) + \beta\left(\sum_{i=1}^{n} x_i' Y_i' - \sum_{i=1}^{n} y_i' X_i'\right) + \gamma\left(\sum_{i=1}^{n} {x_i'}^2 + \sum_{i=1}^{n} {y_i'}^2 - k\right)\right] = 0 \quad (4.118)$$

and

$$\frac{\partial \mathrm{K}}{\partial k} = -\left(\alpha^2 + \beta^2 + \gamma^2 - 1\right) = 0. \quad (4.119)$$

Similarly to the previous cases it is possible to estimate k by solving

$$\begin{vmatrix} (v_1 - k) & w_1 & w_2 \\ w_1 & (v_1 - k) & w_3 \\ w_2 & w_3 & (v_2 - k) \end{vmatrix} = 0, \quad (4.120)$$

which leads to a cubic equation with one unknown parameter. The respective elements are

$$\begin{aligned} v_1 = \sum_{i=1}^{n} {x_i'}^2 + \sum_{i=1}^{n} {y_i'}^2 \quad , \quad v_2 = \sum_{i=1}^{n} {X_i'}^2 + \sum_{i=1}^{n} {Y_i'}^2, \\ w_1 = 0, \quad w_2 = \sum_{i=1}^{n} x_i' X_i' + \sum_{i=1}^{n} y_i' Y_i' \ \text{ and } \ w_3 = \sum_{i=1}^{n} x_i' Y_i' - \sum_{i=1}^{n} y_i' X_i'. \end{aligned} \quad (4.121)$$

The unknown transformation parameters α, β and γ can be estimated either by substituting parameter k_{min} into equations (4.116) - (4.118), under the condition (4.119), or by transforming the equation system into an eigenvalue problem.

4.5.2 TLS 2D similarity transformation

In this section the TLS solution of the 2D similarity transformation is presented. By utilizing the functional model of equation (4.108) and following the same approach as in the presented TLS solutions (subsections 4.2.2, 4.3.2, 4.4.2), the EIV model is introduced with the relevant matrices

$$\mathbf{A} = \begin{bmatrix} x_1' & -y_1' \\ y_1' & x_1' \\ \vdots & \vdots \\ x_n' & -y_n' \\ y_n' & x_n' \end{bmatrix}, \mathbf{V}_{\mathrm{A}} = \begin{bmatrix} v_{x_1} & -v_{y_1} \\ \\ v_{y_1} & v_{x_1} \\ \vdots & \vdots \\ v_{x_n} & -v_{y_n} \\ \\ v_{y_n} & v_{x_n} \end{bmatrix}, \mathbf{L} = \begin{bmatrix} X_1' \\ Y_1' \\ \vdots \\ X_n' \\ Y_n' \end{bmatrix}, \mathbf{v}_{\mathrm{L}} = \begin{bmatrix} v_{X_1} \\ v_{Y_1} \\ \vdots \\ v_{X_n} \\ v_{Y_n} \end{bmatrix}, \hat{\mathbf{X}} = \begin{bmatrix} \hat{\xi}_1 \\ \hat{\xi}_2 \end{bmatrix}. \quad (4.122)$$

The augmented matrix $[\mathbf{A}, \mathbf{L}]$ can be described in this case by

$$[\mathbf{A}, \mathbf{L}] = \begin{bmatrix} x_1' & -y_1' & X_1' \\ y_1' & x_1' & Y_1' \\ \vdots & \vdots & \vdots \\ x_n' & -y_n' & X_n' \\ y_n' & x_n' & Y_n' \end{bmatrix}. \tag{4.123}$$

The right eigenvectors of the augmented matrix can be derived by the eigenvalue/eigenvector decomposition of the squared matrix

$$[\mathbf{A}, \mathbf{L}]^{\mathrm{T}}[\mathbf{A}, \mathbf{L}] = \begin{bmatrix} v_1 & w_1 & w_2 \\ w_1 & v_1 & w_3 \\ w_2 & w_3 & v_2 \end{bmatrix} = \mathbf{G}, \tag{4.124}$$

with the respective elements being equal to those of equation (4.121). The solution for the transformation parameters can be determined from the generalised eigenvalue problem by solving

$$\begin{vmatrix} (v_1 - \lambda) & w_1 & w_2 \\ w_1 & (v_1 - \lambda) & w_3 \\ w_2 & w_3 & (v_2 - \lambda) \end{vmatrix} = 0. \tag{4.125}$$

As expected, equation (4.125) is the same as equation (4.120) and the resulting cubic polynomial equation for the solution of k is identical to the characteristic equation of the eigenvalues λ. The translation terms t_x and t_y can be computed by substituting the estimated parameters into equation (4.106).

4.6 General formulation and classification

The normal equations of the discussed nonlinear least squares problems in this chapter can be transformed into an eigenvalue problem and be solved directly when the characteristic equation is a polynomial of degree four or less. Such adjustment cases are the fitting of a straight line in 2D and 3D, the fitting of a plane in 3D and the 2D similarity transformation of coordinates. The following common features have been identified for the direct solution of these problems:

1. The measured quantities in all adjustment cases were equally weighted and uncorrelated.
2. In the beginning a nonlinear and over-parametrised functional model was used to express each individual problem, see for example equation (4.3) for fitting of a straight line in 2D.
3. Choosing an appropriate restriction between the unknown parameters for the adjustment of each investigated problem, it was possible to obtain an apparently linear relationship between the observations and the unknowns.
4. A reduction of the observed coordinates to the centre of mass was in any case proven to be admissible. This reduction leads everytime to the substitution of some unknown parameters with known ones. However, it must be mentioned that this parameter substitution is not necessary and is only performed for simplifying the problem. Thus, an equivalent solution can be obtained from the respective normal equations, including all unknown parameters.

5. The developed objective function for minimising the sum of squared residuals leads to a homogeneous system of normal equations which is linear with respect to the unknown parameters and has a direct solution (in case that the derived characteristic equation is a polynomial of degree four or less).
6. The derived direct solutions have been proven to be identical with the TLS solutions obtained by using SVD.

These features can be used in the future as criteria for identifying easier and quicker those nonlinear least squares problems that belong to this class. A general formulation of these adjustment problems can be considered, based on the replacement of the "original" nonlinear functional model with an apparently linear one, see features (2.) and (3.) above.

General formulation in matrix notation

All discussed adjustments can be solved in their "original" nonlinear form iteratively by linearizing the condition equations and expressing the problem within a GHM. However, in all cases the problem could be "transformed" in such a way, so that it could be expressed in matrix notation by the system of observation equations

$$\mathbf{L} + \mathbf{v} = \mathbf{A}\ \mathbf{X}, \tag{4.126}$$

with the nonlinear constraint between the unknown parameters being described by the quadratic function [10]

$$\mathbf{X}^{\mathrm{T}}\mathbf{X} = 1. \tag{4.127}$$

Vector $\mathbf{L}$ contains pseudo-observations (zero elements). Vector $\mathbf{v}$ includes the residuals of the pseudo-observations, which are the orthogonal distances. The design matrix $\mathbf{A}$ contains the coefficients of the linear observation equations with respect to the unknown parameters in each problem and vector $\mathbf{X}$ holds the unknown parameters to be estimated.

In order to illustrate this type of functional modeling, the example of fitting of a straight line in 2D can be considered. For instance, n orthogonal distances of the measured points to the straight line are describing the functional model of the problem, with the observation equations

$$\begin{aligned} 0_1 + D_1 &= a\ x'_1 + b\ y'_1, \\ 0_2 + D_2 &= a\ x'_2 + b\ y'_2, \\ &\ \vdots \\ 0_n + D_n &= a\ x'_n + b\ y'_n, \end{aligned} \tag{4.128}$$

[10] In strict notation the product of matrices will also be a matrix, which in this case will have only one element.

under the constraint $a^2 + b^2 = 1$. This system of observation equations can be expressed in matrix notation, as in equation (4.126), with the respective matrices being defined by

$$\mathbf{L} = \begin{bmatrix} 0_1 \\ 0_2 \\ \vdots \\ 0_n \end{bmatrix}, \mathbf{v} = \begin{bmatrix} D_1 \\ D_2 \\ \vdots \\ D_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} a \\ b \end{bmatrix}, \mathbf{A} = \begin{bmatrix} x'_1 & y'_1 \\ x'_2 & y'_2 \\ \vdots & \\ x'_n & y'_n \end{bmatrix}. \tag{4.129}$$

Taking into account the stochastic information of the measured quantities, it can be seen that the problem can be expressed within a GMM with a quadratic constraint (as in the presented solution of section 3.1). Avoiding any kind of linearization, a least squares solution can be obtained by minimizing the objective function

$$\Omega(\mathbf{v}, \mathbf{X}) = \mathbf{v}^{\mathrm{T}}\mathbf{v} \rightarrow \min, \tag{4.130}$$

or taking into account the constraint (4.127), by finding the minimum of the Lagrange function

$$\mathrm{K}(\mathbf{v}, \mathbf{X}, \mathbf{k}) = \mathbf{v}^{\mathrm{T}}\mathbf{v} - k(\mathbf{X}^{\mathrm{T}}\mathbf{X} - 1) \rightarrow \min, \tag{4.131}$$

with k denoting the Lagrange multiplier. Rearranging the observation equations (4.126) and substituting the solution for the residuals in the Lagrangian yields

$$\begin{aligned} &\mathrm{K}(\mathbf{X}, \mathbf{k}) = (\mathbf{AX} - \mathbf{L})^{\mathrm{T}}(\mathbf{AX} - \mathbf{L}) - k(\mathbf{X}^{\mathrm{T}}\mathbf{X} - 1) \\ \Rightarrow\ &\mathrm{K}(\mathbf{X}, \mathbf{k}) = \mathbf{X}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{AX} - 2\mathbf{L}^{\mathrm{T}}\mathbf{AX} + \mathbf{L}^{\mathrm{T}}\mathbf{L} - k(\mathbf{X}^{\mathrm{T}}\mathbf{X} - 1). \end{aligned} \tag{4.132}$$

Taking into consideration that the vector of pseudo-observations $\mathbf{L}$ contains only zero values, the last function can be equivalently written as

$$\mathrm{K}(\mathbf{X}, \mathbf{k}) = \mathbf{X}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{AX} - k(\mathbf{X}^{\mathrm{T}}\mathbf{X} - 1). \tag{4.133}$$

The minimization of the derived Lagrangian can be found already in (Perović 2005, p. 33) as the solution of mathematical problems in quadratic forms, with their extrema being derived using an EVD. This can be proven by taking the partial derivatives of K with respect to the unknowns and setting the solution to zero, which yields the normal equation system

$$\begin{aligned} \frac{\partial \mathrm{K}}{\partial \mathbf{X}^{\mathrm{T}}} &= 2\mathbf{A}^{\mathrm{T}}\mathbf{AX} - 2k\ \mathbf{X} = \mathbf{0} \\ &\Rightarrow \mathbf{A}^{\mathrm{T}}\mathbf{AX} - k\ \mathbf{X} = \mathbf{0}, \end{aligned} \tag{4.134}$$

$$\frac{\partial \mathrm{K}}{\partial k} = \mathbf{X}^{\mathrm{T}}\mathbf{X} - 1 = \mathbf{0}. \tag{4.135}$$

In equation (4.134), parameter k can be seen as an eigenvalue and $\mathbf{X}$ as an eigenvector of the squared matrix $\mathbf{A}^\mathrm{T}\mathbf{A}$. The solution can be computed from the generalised eigenvalue problem

$$\mathbf{A}^\mathrm{T}\mathbf{A}\mathbf{X} - k\ \mathbf{X} = \mathbf{0} \Rightarrow (\mathbf{A}^\mathrm{T}\mathbf{A} - k\mathbf{I})\mathbf{X} = \mathbf{0}, \tag{4.136}$$

with $\mathbf{I}$ denoting an identity matrix. The eigenvalues of matrix $\mathbf{A}^\mathrm{T}\mathbf{A}$ can be determined by searching for non-trivial solutions $\mathbf{X} \neq 0$, i.e. by solving the characteristic equation of the eigenvalues, or equivalently by

$$\det(\mathbf{A}^\mathrm{T}\mathbf{A} - k\mathbf{I}) = \mathbf{0}. \tag{4.137}$$

From the latter developments it can be seen that all discussed problems can be expressed within a GMM, while finding the minimum of the objective function is equivalent to finding the minimum of a quadratic function by employing EVD.

4.7 Discussion and open questions

In this chapter two individual solution strategies have been examined, the systematic approach for direct least squares solutions that has been established already in (Malissiovas et al. 2016) and TLS. A mathematical relationship between the two approaches has been presented by comparing their solutions for four nonlinear least squares problems, the fitting of a straight line in 2D and 3D, the fitting of a plane in 3D, as well as the 2D similarity transformation of coordinates. The discussed adjustment problems have been identified as such, that belong to a certain class of nonlinear least squares and can be transformed into solving a polynomial equation. Thus, depending on the polynomial's degree a direct solution can be possible[11].

An "optimal" estimate for the unknowns is derived by employing the method of least squares and minimizing a well-defined Lagrange function. In all discussed cases the normal equations can be solved with various techniques, for example SVD or EVD and by solving a characteristic equation, which was always identical to the characteristic equation of the eigenvalues from the corresponding TLS solution. The developed approach provides a deep understanding of the concept of TLS for the solution of nonlinear least squares problems.

It has been already known from (Neitzel and Petrovic 2008), (Neitzel 2010), as well as (Reinking 2008), that TLS *is not a new method* per se but a solution strategy for a class of nonlinear least squares problems. In addition, the presented direct solutions of this chapter reveal that TLS is an algorithmic approach for the solution of a class of nonlinear least squares problems using SVD.

Nevertheless, in order to obtain a direct solution for the discussed adjustment problems, either using the proposed systematic approach or with TLS and SVD, it is always assumed that the observations are uncorrelated and have been obtained with equal precision. When postulating a different precision for each observation, then different solution strategies can be utilized. A weighted least squares solution can be obtained in this case using, for example, the Gauss-Newton approach from the traditional geodetic solutions of section 3.1, or by employing one of the WTLS algorithms from section 3.2.1.2.

[11] Direct least squares solution have been presented for the investigated adjustment problems, as the roots of the derived polynomials could be computed directly using known formulas from mathematics, see for more information (Bronshtein et al. 2005, p. 62 ff.)

The following questions arise out of the findings from this chapter:

- If it is possible to obtain directly a solution for this class of nonlinear least squares problems by transforming them into solving a polynomial equation, is it also possible to obtain a similar solution by transforming the weighted least squares problem?
- Are there specific weighted cases of nonlinear least squares problems (besides the generally well-known case of equally weighted observations) which can be solved directly?
- Is it possible to detect those weighted nonlinear least squares problems with a direct solution and solve them by using a systematic approach?
- In cases where a direct weighted least squares solution is not possible, what are the alternative ways? Is it possible to obtain an iterative solution without making any use of linearization of the problem?

Therefore, possible direct solutions are investigated in the next chapter for the discussed class of adjustment problems by postulating different weighting cases for the measured quantities. In these scenarios where a direct solution is not possible, an iterative approach is examined that does not involve a linearization of the problem at any step of the procedure.

5 Direct and iterative solutions of weighted nonlinear least squares problems

5.1 Basic idea and general methodology

Direct solutions have been presented in the previous chapter for a special class of nonlinear least squares problems. The established solution strategy involved the transformation of the normal equations into the solution of a quadratic or cubic algebraic equation (characteristic equation). The mathematical derivations of those solutions were based on the fact that uncorrelated observations have been obtained with equal precision. Thus, the postulated weights of the adjustment were in all cases equal.

The investigations in this chapter focus on the aforementioned questions from section 4.7. Three of the adjustment problems that belong to this class with a direct solution are examined here[1], namely:

- Fitting of a straight line in 2D;
- Fitting of a plane in 3D;
- 2D similarity transformation of coordinates,

and for four individual weighting cases for each problem:

1. Same precision for the coordinates in each direction;
2. Individual precision for the coordinates of each point;
3. Individual precision for each coordinate;
4. Individual precision and correlations between the observations (covering also the cases of singular cofactor matrices).

Direct weighted least squares solutions are proposed in this chapter for the first time for the discussed class of problems. The general idea of these solutions corresponds to the methodology suggested in (Malissiovas et al. 2016). It involves a parameterization of the problem and the formation of a Lagrange function that results in a quadratic or cubic equation. In cases where the problem cannot be transformed to such algebraic equations, modern iterative algorithms are clearly presented that do not require a linearization of the original

[1]The case of fitting a straight line in 3D will not discussed further in this chapter in order to avoid repetition, as it has similarities to the problems of fitting a straight line in 2D and fitting a plane in 3D and therefore could be easily solved following the same strategy that is presented for the latter problems.

problem and are based mainly on the approaches from (York 1966) and (Petrović et al. 1983). The iterative solutions of the developed algorithms can be compared with those from WTLS. Figure 5.1 depicts a flowchart with the proposed solutions for this class of weighted nonlinear least squares problems.

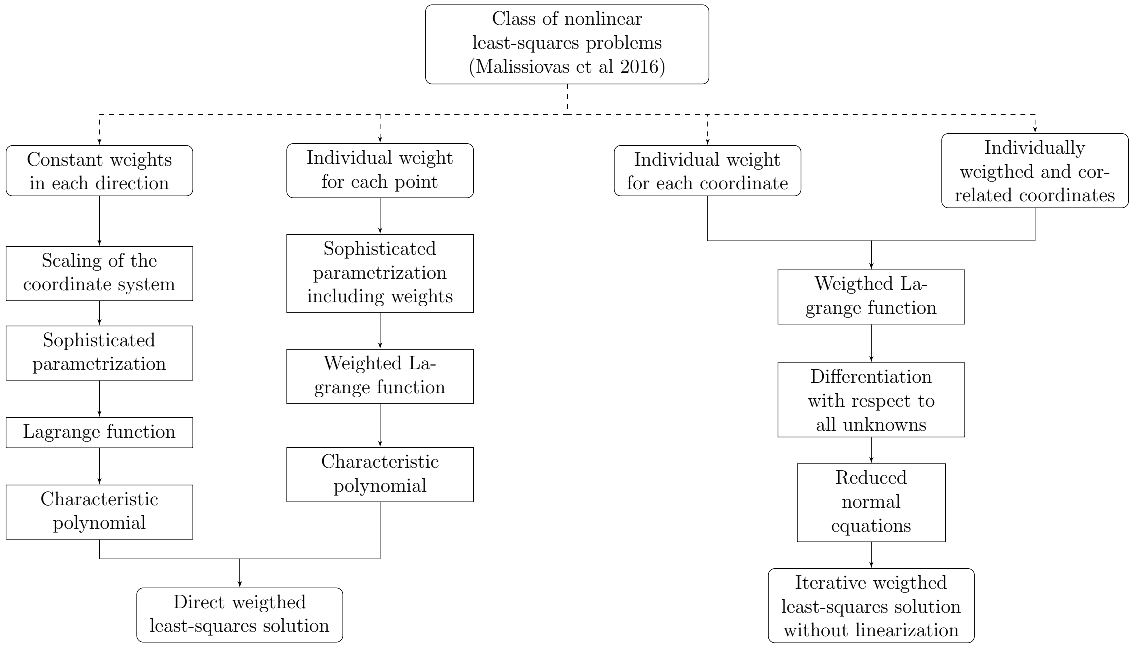

FIGURE 5.1: Flowchart for possible direct and iterative solutions of a class of nonlinear weighted least squares problems.

5.2 Fitting of a straight line in 2D

One of the first attempts to solve the nonlinear least squares problem of fitting a straight line to a set of points in 2D, that have been observed with different precisions, goes back to (York, 1966). In that article the problem has been expressed by a pseudo-cubic equation and a solution has been obtained iteratively and without any kind of linearization. It was further Williamson (1968) who pointed out that the same adjustment problem can be formulated either as a pseudo-quadratic or even as a pseudo-linear equation. A year later York (1968) published an iterative least squares solution including this time correlations between the observations. A thorough investigation and a detailed explanation of the same problem can be found in (Petrović et al., 1983) as well.

Schaffrin and Wieser (2008) have presented a weighted TLS algorithm for linear regression, that inspired Shen et al. (2011) and Amiri-Simkooei and Jazaeri (2012) to develope modern TLS algorithms for solving the same problem. On the other hand, Neitzel and Petrovic (2008) presented a solution within the linearized GHM, following the traditional geodetic procedure for solving nonlinear least squares problems. This includes a linearization of the condition equations and an iterative process that stops after a predefined threshold, as it has been discussed in chapter 2.

Point of beginning in this investigation are the coordinates of a set of points in 2D, assuming that they have been observed with different precisions. The functional model of this problem can be expressed by the general form of a straight line in 2D, presented already by equation (4.1) in section 4.2. Including the necessary residuals in the measured quantities results in the nonlinear condition equations

$$a(x_i + v_{x_i}) + b(y_i + v_{y_i}) + c = 0, \tag{5.1}$$

with $i = 1, \ldots, n$, where n is the number of measured points. The selection of an appropriate restriction between the unknown parameters will be considered at a later point in this section. In the case of measurements that have been obtained with different precisions, σ_{y_i} for the y-coordinates and σ_{x_i} for the x-coordinates, a least squares solution for the unknown line parameters could be based on the minimization of the sum of weighted squared residuals

$$\Omega(v_{x_i}, v_{y_i}) = \sum_{i=1}^{n} p_{x_i}\ v_{x_i}^2 + p_{y_i}\ v_{y_i}^2 \rightarrow \min, \tag{5.2}$$

where p_{x_i} is a weight for the residual v_{x_i} and p_{y_i} for v_{y_i}. The respective weights have been defined in (Helmert 1924, p. 81) as

$$p_{x_i} = \frac{1}{\sigma_{x_i}^2} \quad \text{and} \quad p_{y_i} = \frac{1}{\sigma_{y_i}^2}. \tag{5.3}$$

Therefore, a least squares solution for fitting a straight line in 2D is investigated for four individual weighting cases that often occur in practice:

1. Same precision σ_x for the coordinates in x direction and σ_y in y direction.
2. Individual precision for each point: $\sigma_{x_i} = \sigma_{y_i} \forall i$.
3. Individual precision for each measured coordinate.
4. Individual precision and correlations between the measured 2D coordinates.

5.2.1 Weighting case 1 - Equally weighted observations in each direction

For the first weighting case the coordinates in x direction have been observed with the same precision σ_x, respectively in y direction with σ_y. The weights p_x and p_y can be computed from equation (5.3) and the objective function under minimization becomes

$$\Omega(v_{x_i}, v_{y_i}) = \sum_{i=1}^{n} p_x\ v_{x_i}^2 + p_y\ v_{y_i}^2 \rightarrow \min, \tag{5.4}$$

with

$$\frac{p_x}{p_y} = m,\ m = \text{constant}. \tag{5.5}$$

This problem will have the geometry that is portrayed in Figure 5.2. Thus, it results in the minimization of slanted distances from each observed point to the requested line. However, it can be observed that the ratio between the slanted and the orthogonal distances (i.e. the angles between the orthogonal and the slanted distances) will be constant for every point. From a geometric perspective, the postulated weights can be seen as a homogeneous scale of the coordinate system in both x and y direction. A direct approach

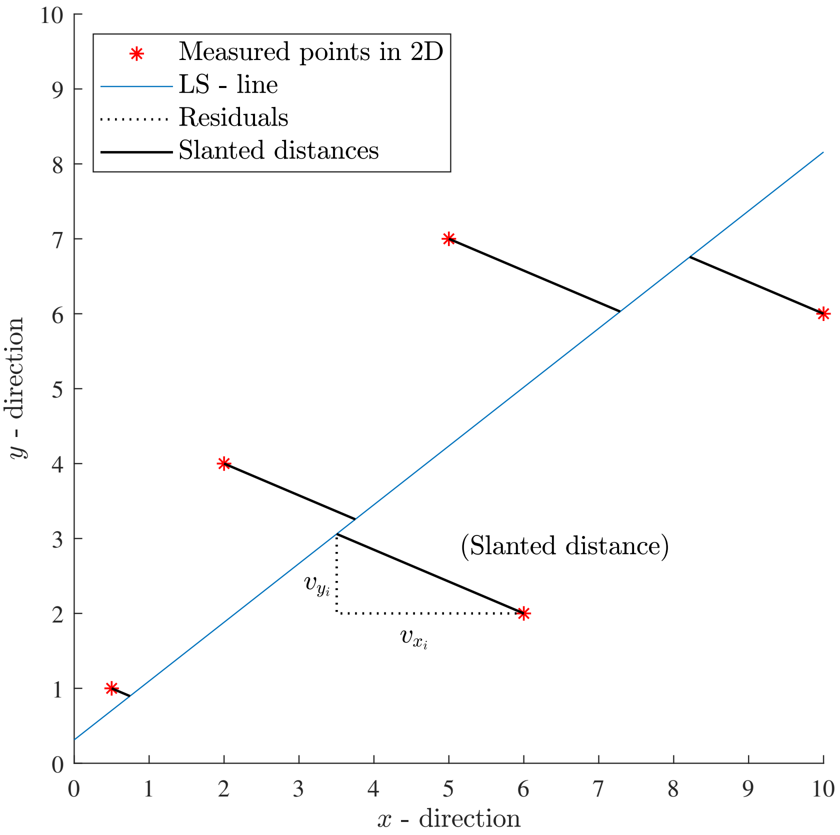

FIGURE 5.2: Example of fitting a straight line to points in 2D, with observed x and y coordinates and p_x, p_y individual constant weights for each coordinate axis.

is presented here, for solving the discussed nonlinear weighted least squares problem. The strategy that will be followed involves the scaling of the observed coordinates beforehand with

$$x_i^s = x_i\sqrt{p_x} \quad \text{and} \quad y_i^s = y_i\sqrt{p_y}, \tag{5.6}$$

with the superscript "s" indicating scale. The scaled coordinates x_i^s and y_i^s can be utilized to derive the requested line in a different coordinate system, where the weights of the observations are equal. In this line of thinking the residuals of the point coordinates will also be scaled accordingly, with

$$v_{x_i}^s = v_{x_i}\sqrt{p_x} \quad \text{and} \quad v_{y_i}^s = v_{y_i}\sqrt{p_y}. \tag{5.7}$$

Substituting the scaled coordinates and their residuals from equations (5.6) and (5.7) into the condition equations (5.1) yields

$$a\frac{1}{\sqrt{p_x}}\left(x_i^s + v_{x_i}^s\right) + b\frac{1}{\sqrt{p_y}}\left(y_i^s + v_{y_i}^s\right) + c = 0. \tag{5.8}$$

Introducing the auxiliary scaled line parameters

$$\begin{aligned} a^s &= a\frac{1}{\sqrt{p_x}}, \\ b^s &= b\frac{1}{\sqrt{p_y}}, \\ c^s &= c, \end{aligned} \tag{5.9}$$

into the condition equations (5.8), yields an alternative functional model to equation (5.1), expressed by

$$a^s\left(x_i^s + v_{x_i}^s\right) + b^s\left(y_i^s + v_{y_i}^s\right) + c^s = 0. \tag{5.10}$$

A meaningful constraint[2] for the solution of this adjustment problem can been chosen here as

$${a^s}^2 + {b^s}^2 = 1. \tag{5.11}$$

The least squares criterion is employed for a solution of the unknown line parameters by minimizing the sum of scaled squared residuals

$$\sum_{i=1}^{n} {v_{x_i}^s}^2 + {v_{y_i}^s}^2 \rightarrow \min. \tag{5.12}$$

Thus, it has been shown that the discussed weighted least squares problem can be transformed into a problem with equal weights in the scaled coordinates, as it is depicted in Figure 5.3.

5.2.1.1 Direct least squares solution in a scaled coordinate system

The sum of squared scaled residuals can be replaced by the sum of squared orthogonal distances

$$\Omega(v_{x_i}^s, v_{y_i}^s) = \sum_{i=1}^{n} {v_{x_i}^s}^2 + {v_{y_i}^s}^2 = \sum_{i=1}^{n} D_i^2. \tag{5.13}$$

The orthogonal distances of the points to the straight line, c.f. equation (4.6), can be expressed for this problem by

$$D_i = \frac{a^s x_i^s + b^s y_i^s + c^s}{\sqrt{{a^s}^2 + {b^s}^2}}, \tag{5.14}$$

which under the constraint (5.11) become

$$D_i = a^s x_i^s + b^s y_i^s + c^s. \tag{5.15}$$

[2] A "meaningful" constraint is chosen here, in the sense that it will lead to simpler equations for the orthogonal distances of the 2D points to the requested line, as it has been discussed in subsection 4.2.

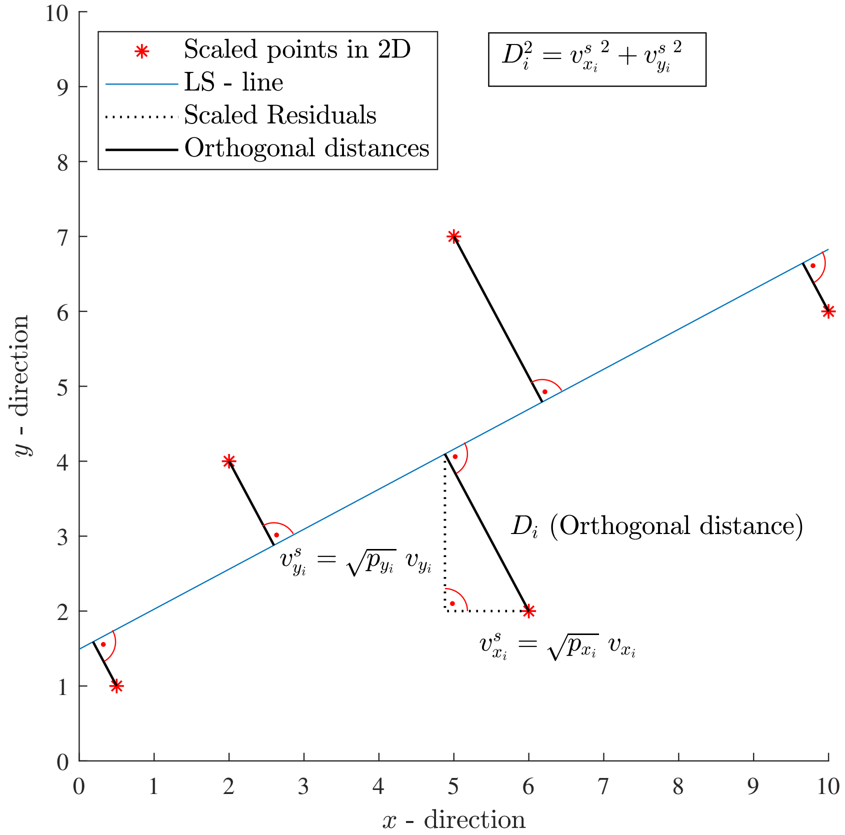

FIGURE 5.3: Example of fitting a straight line to the scaled points in 2D.

Therefore, using the scaled coordinates it is possible to obtain a direct least squares solution for the unknown line parameters, following the same procedure as the one presented in section 4.2.1.

Computation of the line parameters in the original coordinate system

The original line parameters can be computed by substituting the estimated "scaled" line parameters $\hat{a}^s$, $\hat{b}^s$ and $\hat{c}^s$ into equation (5.9):

$$\begin{aligned} \hat{a} &= \hat{a}^s \sqrt{p_x}, \\ \hat{b} &= \hat{b}^s \sqrt{p_y}, \\ \hat{c} &= \hat{c}^s. \end{aligned} \tag{5.16}$$

However, this solution has been restricted to

$$a^{s2} + b^{s2} = \left(a \frac{1}{\sqrt{p_x}}\right)^2 + \left(b \frac{1}{\sqrt{p_y}}\right)^2 = 1. \tag{5.17}$$

The least squares solution which restricts the line parameters to $a^2 + b^2 = 1$ can be easily derived by multiplying the original condition equations with the term

$$\frac{1}{\sqrt{a^2 + b^2}}, \tag{5.18}$$

which yields

$$\frac{a}{\sqrt{a^2 + b^2}} (x_i + v_{x_i}) + \frac{b}{\sqrt{a^2 + b^2}} (y_i + v_{y_i}) + \frac{c}{\sqrt{a^2 + b^2}} = 0. \tag{5.19}$$

Using the information of equation (5.16), the line parameters which are restricted to $a^2 + b^2 = 1$ can be computed by

$$\begin{aligned}
\hat{a} &= \frac{\hat{a}^s \sqrt{p_x}}{\sqrt{\left(\hat{a}^s \sqrt{p_x}\right)^2 + \left(\hat{b}^s \sqrt{p_y}\right)^2}}, \\
\hat{b} &= \frac{\hat{b}^s \sqrt{p_x}}{\sqrt{\left(\hat{a}^s \sqrt{p_x}\right)^2 + \left(\hat{b}^s \sqrt{p_y}\right)^2}}, \\
\hat{c} &= \frac{\hat{c}^s}{\sqrt{\left(\hat{a}^s \sqrt{p_x}\right)^2 + \left(\hat{b}^s \sqrt{p_y}\right)^2}}.
\end{aligned} \tag{5.20}$$

5.2.2 Weighting case 2 - Individually weighted points in 2D

In the second weighting case under investigation, each measured point has been obtained with individual precision

$$\sigma_{x_i} = \sigma_{y_i} \quad \text{and} \quad p_{x_i} = p_{y_i} = p_i \;\; \forall i. \tag{5.21}$$

The ratio between the weights for each point is constant with

$$\frac{p_{x_i}}{p_{y_i}} = 1 \;\; \forall i. \tag{5.22}$$

Taking into consideration (5.21), the objective function (5.2) can be reformulated to

$$\sum_{i=1}^{n} p_{x_i} v_{x_i}^2 + p_{y_i} v_{y_i}^2 = \sum_{i=1}^{n} p_i (v_{x_i}^2 + v_{y_i}^2) \rightarrow \min. \tag{5.23}$$

5.2.2.1 Direct weighted least squares solution

In case of individually weighted points, the sum of weigthed squared residuals can be expressed equivalently by the weighted squared orthogonal distances of each point to the requested line:

$$p_{x_i} v_{x_i}^2 + p_{y_i} v_{y_i}^2 = p_i (v_{x_i}^2 + v_{y_i}^2) = p_i (D_i^2). \tag{5.24}$$

The orthogonal distances are

$$D_i = \frac{ax_i + by_i + c}{\sqrt{a^2 + b^2}}, \tag{5.25}$$

which after taking into account the restriction between the unknown line parameters

$$a^2 + b^2 = 1, \tag{5.26}$$

can be simplified to

$$D_i = ax_i + by_i + c. \tag{5.27}$$

Thus, the objective function (5.23) can be equivalently written as

$$\Omega(a, b, c) = \sum_{i=1}^{n} p_i(v_{x_i}^2 + v_{y_i}^2) = \sum_{i=1}^{n} p_i D_i^2 = \sum_{i=1}^{n} p_i \left(ax_i + by_i + c\right)^2 . \tag{5.28}$$

This adjustment problem is depicted in Figure 5.4.

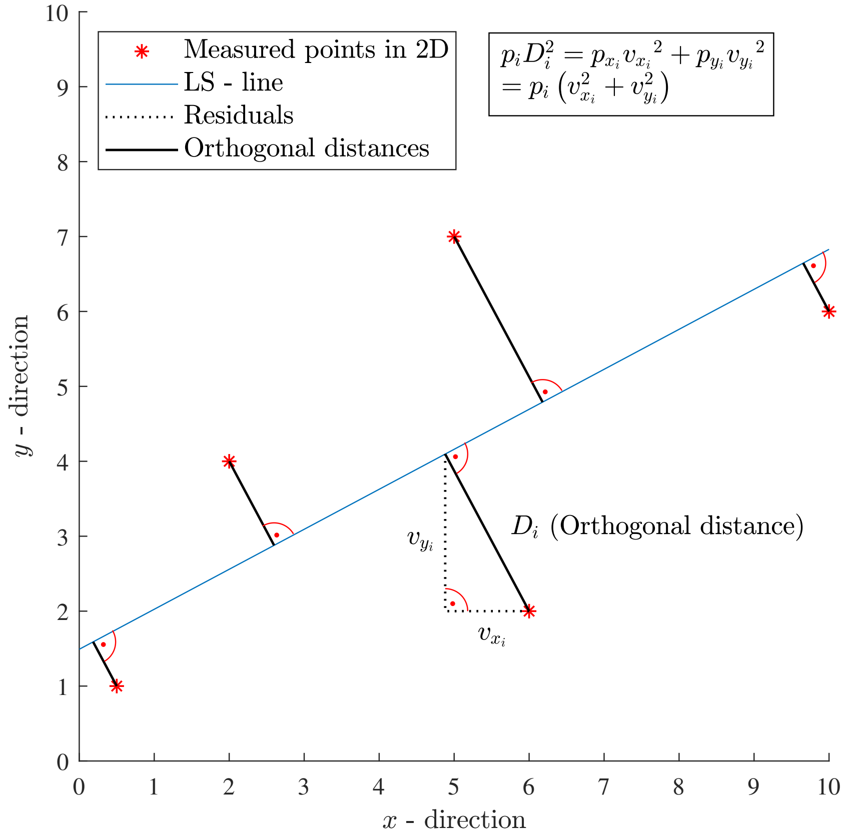

FIGURE 5.4: Example of fitting a straight line to points in 2D with x and y measured coordinates and p_{x_i}, p_{y_i} being equal weights for each point.

We seek for a least squares solution for the unknown line parameters a, b and c that minimizes (5.28), subject to the restriction (5.26). Consequently, the Lagrangian

$$\mathrm{K}(a,b,c,\lambda) = \Omega(a,b,c) - k(a^2+b^2-1), \tag{5.29}$$

can be written, with k denoting the Lagrange multiplier. By differentiating function K with respect to the unknown parameters and setting the partial derivatives to zero results in the system of normal equations

$$\frac{\partial \mathrm{K}}{\partial a} = 2\left[a\Big(\sum_{i=1}^{n} p_i {x_i}^2 - k\Big) + b\Big(\sum_{i=1}^{n} p_i y_i x_i\Big) + c\Big(\sum_{i=1}^{n} p_i x_i\Big)\right] = 0, \tag{5.30}$$

$$\frac{\partial \mathrm{K}}{\partial b} = 2\left[a\Big(\sum_{i=1}^{n} p_i y_i x_i\Big) + b\Big(\sum_{i=1}^{n} p_i {y_i}^2 - k\Big) + c\Big(\sum_{i=1}^{n} p_i y_i\Big)\right] = 0, \tag{5.31}$$

$$\frac{\partial \mathrm{K}}{\partial c} = 2\left[a\sum_{i=1}^{n} p_i x_i + b\sum_{i=1}^{n} p_i y_i + c\sum_{i=1}^{n} p_i\right] = 0 \tag{5.32}$$

and

$$\frac{\partial \mathrm{K}}{\partial k} = -\big(a^2+b^2-1\big) = 0. \tag{5.33}$$

Rearranging equation (5.32) leads to

$$c = -a\,\frac{\sum_{i=1}^{n} p_i x_i}{\sum_{i=1}^{n} p_i} - b\,\frac{\sum_{i=1}^{n} p_i y_i}{\sum_{i=1}^{n} p_i}. \tag{5.34}$$

Introducing the expression for c into the normal equations (5.30) and (5.31), yields the reduced normal equations

$$a\left[\sum_{i=1}^{n} p_i {x_i}^2 - \frac{1}{\sum_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i x_i\right)^2 - k\right] + b\left[\sum_{i=1}^{n} p_i y_i x_i - \frac{1}{\sum_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i x_i \sum_{i=1}^{n} p_i y_i\right)\right] = 0 \tag{5.35}$$

and

$$a\left[\sum_{i=1}^{n} p_i y_i x_i - \frac{1}{\sum_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i x_i \sum_{i=1}^{n} p_i y_i\right)\right] + b\left[\sum_{i=1}^{n} p_i {y_i}^2 - \frac{1}{\sum_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i y_i\right)^2 - k\right] = 0. \tag{5.36}$$

Equations (5.35) and (5.36) form a homogeneous system of linear equations with respect to the unknown line parameters. The determinant of the equation system is equal to zero for a nontrivial solution

$$\begin{vmatrix} \left[\sum_{i=1}^{n} p_i {x_i}^2 - \frac{1}{\sum_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i x_i\right)^2 - k\right] & \left[\sum_{i=1}^{n} p_i y_i x_i - \frac{1}{\sum_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i x_i \sum_{i=1}^{n} p_i y_i\right)\right] \\ \left[\sum_{i=1}^{n} p_i y_i x_i - \frac{1}{\sum_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i x_i \sum_{i=1}^{n} p_i y_i\right)\right] & \left[\sum_{i=1}^{n} p_i {y_i}^2 - \frac{1}{\sum_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i y_i\right)^2 - k\right] \end{vmatrix} = 0, \quad (5.37)$$

which leads to a quadratic characteristic equation with two real and positive solutions for k. The minimum solution, denoted by k_{min}, corresponds to the minimum of the Lagrange function (5.29). The solution for the unknown line parameters a and b can be computed by substituting the Lagrangian factor k_{min} into equations (5.35)-(5.36), subject to the chosen restriction, or by transforming the equation system into an eigenvalue problem.

5.2.3 Weighting case 3 - Individually weighted 2D coordinates

The third weighting case under investigation is more general than the previous two, as the measured coordinates have been observed with individual precisions. As far as known, a direct least squares solution for this nonlinear problem has not been found. Iterative solutions, however, have been presented by various authors in the past. Some of those do not make use of any linearization of the original problem, such as for example York (1966), York (1968) or Petrović et al. (1983). In the next subsection it is shown that a direct solution for the discussed adjustment problem is not possible. However, an iterative procedure is presented that is based to a large extent on the derivations of (Petrović et al. 1983).

Minimizing the sum of weighted squared residuals

Starting from the condition equations (5.1) and the objective function (5.2), it is possible to build the Lagrangian

$$\mathrm{K}(v_{x_i}, v_{y_i}, a, b, c, k_i) = \Omega(v_{x_i}, v_{y_i}) - 2\sum_{i=1}^{n} k_i(a(x_i + v_{x_i}) + b(y_i + v_{y_i}) + c), \quad (5.38)$$

with k_i denoting the Lagrange multipliers.The normal equation system of this adjustment problem can be described by

$$\frac{\partial \mathrm{K}}{\partial v_{x_i}} = 2p_{x_i}v_{x_i} - 2ak_i = 0,$$

$$\Rightarrow v_{x_i} = \frac{ak_i}{p_{x_i}}, \quad (5.39)$$

$$\frac{\partial \mathrm{K}}{\partial v_{y_i}} = 2p_{y_i} v_{y_i} - 2bk_i = 0,$$
$$\Rightarrow v_{y_i} = \frac{bk_i}{p_{y_i}}, \tag{5.40}$$

$$\frac{\partial \mathrm{K}}{\partial k_i} = -2\left[a(x_i + v_{x_i}) + b(y_i + v_{y_i}) + c\right] = 0, \tag{5.41}$$

$$\frac{\partial \mathrm{K}}{\partial a} = -2\sum_{i=1}^{n} k_i(x_i + v_{x_i}) = 0, \tag{5.42}$$

$$\frac{\partial \mathrm{K}}{\partial b} = -2\sum_{i=1}^{n} k_i(y_i + vy_i) = 0, \tag{5.43}$$

$$\frac{\partial \mathrm{K}}{\partial c} = -2\sum_{i=1}^{n} k_i = 0. \tag{5.44}$$

Equations (5.39) to (5.44) form a nonlinear system of $3n+3$ equations. Introducing the residuals from equations (5.39) and (5.40) into (5.41), yields the expression for the Lagrange multipliers

$$k_i = w_i(ax_i + by_i + c), \tag{5.45}$$

with the auxiliary weighting factors[3]

$$w_i = -\left(\frac{a^2}{p_{x_i}} + \frac{b^2}{p_{y_i}}\right)^{-1}. \tag{5.46}$$

Introducing k_i into equation (5.44), results in the expression for parameter

$$c = -\frac{a\sum_{i=1}^{n} w_i x_i + b\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i}. \tag{5.47}$$

Additionally, substituting k_i into equations (5.39) and (5.40) yields explicit expressions for the residual vectors

$$v_{x_i} = -\frac{w_i}{p_x}\, a(ax_i + by_i + c) \tag{5.48}$$

and

$$v_{y_i} = -\frac{w_i}{p_y}\, b(ax_i + by_i + c). \tag{5.49}$$

[3] It is interesting to note the relationship between the developed weighting factors w_i, the coefficients "L_i" and the weighting factors "W_i" from (Deming 1964, p. 134,181) and (York 1966).

Utilizing the developed expressions for v_{y_i} and v_{x_i}, a minimum of the original objective function Ω can be found, instead of minimizing the Lagrangian K. This approach gives the possibility to show why a direct least squares is not possible for this weighted case. Therefore, substituting the developed residuals directly in the objective function (5.2) yields

$$\begin{aligned}
\Omega(v_{x_i}, v_{y_i}) &= \sum_{i=1}^{n} p_{x_i}\ v_{x_i}^2 + p_{y_i}\ v_{y_i}^2 \\
&= \sum_{i=1}^{n} p_{x_i} \left[-\frac{w_i}{p_{x_i}}\ a\,(ax_i + by_i + c) \right]^2 + p_{y_i} \left[-\frac{w_i}{p_{y_i}}\ b(ax_i + by_i + c) \right]^2 \\
&= \sum_{i=1}^{n} - \left(\frac{a^2}{p_{x_i}} + \frac{b^2}{p_{y_i}} \right) \frac{(ax_i + by_i + c)^2}{\left(\frac{a^2}{p_{x_i}} + \frac{b^2}{p_{y_i}} \right)^2} \\
&= \sum_{i=1}^{n} - \frac{1}{\left(\frac{a^2}{p_{x_i}} + \frac{b^2}{p_{y_i}} \right)} (ax_i + by_i + c)^2 \\
&= \sum_{i=1}^{n} w_i (ax_i + by_i + c)^2 \\
&= \sum_{i=1}^{n} D_i^2.
\end{aligned} \tag{5.50}$$

From the last equation it can be seen that the problem of minimizing the sum of weighted squared residuals can be transformed into the minimization of the slanted distances

$$D_i = -\frac{1}{\sqrt{\dfrac{a^2}{p_{x_i}} + \dfrac{b^2}{p_{y_i}}}}\ (ax_i + by_i + c) = \sqrt{w_i}\ (ax_i + by_i + c). \tag{5.51}$$

A direct solution for this weighted case is not possible, as there is no restriction for the unknown parameters that could lead to a linear formulation of the distances in equation (5.51). From a different perspective it could be said that a direct least squares solution is possible, when the auxiliary weighting factors w_i in equation (5.46) can be set equal to a constant value, by selecting a meaningful restriction between the unknown parameters. The geometry of this problem is depicted in Figure (5.5).

5.2.3.1 Iterative least squares solution without linearization

An iterative solution for the discussed adjustment problem can be obtained without performing any kind of linearization, similar to (York 1966) and (Petrović et al. 1983). The unknown line parameters a and b can be estimated by introducing the residuals from equations (5.48) - (5.49) and k_i from (5.45), into equations

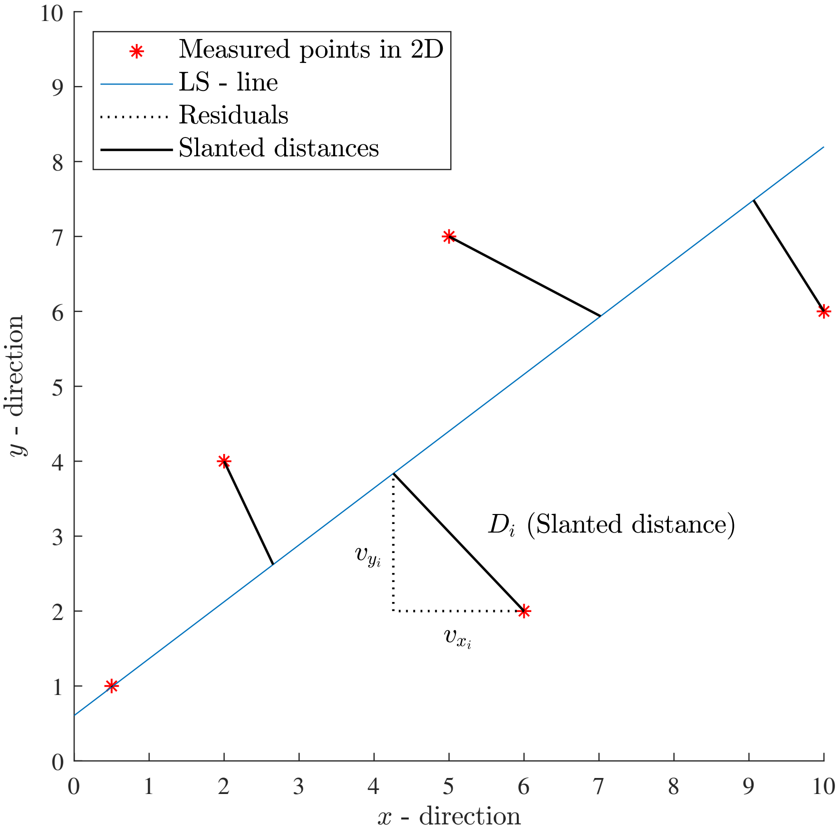

FIGURE 5.5: Example of fitting a straight line to points in 2D with observed x_i and y_i coordinates and p_{x_i}, p_{y_i} individual weights for the coordinates.

(5.42) and (5.43). This yields the reduced normal equations

$$a \sum_{i=1}^{n} \frac{k_i^2}{p_{x_i}} = - \sum_{i=1}^{n} w_i (a x_i + b y_i + c) x_i \tag{5.52}$$

and

$$b \sum_{i=1}^{n} \frac{k_i^2}{p_{y_i}} = - \sum_{i=1}^{n} w_i (a x_i + b y_i + c) y_i. \tag{5.53}$$

Rearranging appropriately equations (5.52) and (5.53) gives

$$a f_1 = b f_2 \tag{5.54}$$

and

$$b f_3 = a f_2, \tag{5.55}$$

with the respective quantities being

$$
\begin{aligned}
f_1 &= \sum_{i=1}^{n} \frac{k_i^2}{p_{x_i}} - \sum_{i=1}^{n} w_i\ x_i^2 + \frac{1}{\sum_{i=1}^{n} w_i} \left(\sum_{i=1}^{n} w_i\ x_i \right)^2 , \\
f_2 &= \sum_{i=1}^{n} (w_i\ x_i\ y_i) - \frac{1}{\sum_{i=1}^{n} w_i} \left(\sum_{i=1}^{n} w_i\ y_i \sum_{i=1}^{n} w_i\ x_i \right) , \\
f_3 &= \sum_{i=1}^{n} \frac{k_i^2}{p_{y_i}} - \sum_{i=1}^{n} w_i\ y_i^2 + \frac{1}{\sum_{i=1}^{n} w_i} \left(\sum_{i=1}^{n} w_i\ y_i \right)^2 .
\end{aligned}
\tag{5.56}
$$

Furthermore, a simplification of the problem is feasible by setting a restriction between the unknown line parameters a and b. It is possible to take into account the general restriction $a^2 + b^2 = 1$, however, for convenience the problem is restricted here to $b = 1$. Thus, a solution for the remaining unknown is

$$
a = \frac{f_2}{f_1}, \tag{5.57}
$$

with f_1 and f_2 containing some unknown parameters, according to their definition in (5.56). Thus, equation (5.57) becomes pseudo-linear after selecting an approximate value a^0. The estimated line parameters can be utilized as new starting values in each iteration step, until a break-off condition is met. As a linearization has not been applied in any step of the adjustment, this iterative procedure can be terminated after the condition of the "computational error" is fulfilled, as it was presented in chapter 3. Algorithm 1 has been developed for estimating the weighted least squares solution for fitting a line to a set of points in 2D, based on the presented procedure of this subsection.

Algorithm 1 Least squares fitting of a straight line in 2D with general weights

```
1: Choose approximate value for a^0.
2: Define parameter b = 1.
3: Set threshold ε for the break-off condition of the iteration process.
4: Set parameter d_a = |â − a^0| = ∞, for entering the iteration process.
5: while d_a > ε do
6:     Compute parameters w_i, k_i and estimate ĉ.
7:     Compute the coefficients f_1 and f_2.
8:     Estimate parameter â.
9:     Compute parameter d_a = |â − a^0|.
10:    Update the approximate values with the estimated ones (a^0 = â).
11: end while
12: return â and ĉ, with b = 1.
```

Iterative procedure of pseudo-quadratic and pseudo-cubic equations

For the sake of a complete overview of the iterative algorithms that can produce the weighted least squares solution for fitting a straight line in 2D, two alternatives to the developed pseudo-linear equations are presented here. A thorough analysis of (5.52) and (5.53) can lead to a pseudo-cubic equation instead, as it has been shown by York (1966), which reads

$$a^3 \sum_{i=1}^{n} \frac{w_i^2 \; x_i'^2}{p_{x_i}} - 2a^2 \sum_{i=1}^{n} \frac{w_i^2 \; x_i' y_i'}{p_{x_i}} - a\left[\sum_{i=1}^{n} w_i x_i'^2 - \sum_{i=1}^{n} \frac{w_i^2 \; y_i'^2}{p_{x_i}}\right] + \sum_{i=1}^{n} w_i x_i' y_i', \tag{5.58}$$

restricted to $b = 1$. The coordinates x' and y' are the reduced coordinates to the pseudo-centre of the mass and can be computed by

$$x_i' = x_i - \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} \quad \text{and} \quad y_i' = y_i - \frac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i}. \tag{5.59}$$

The term "pseudo-centre" originates from (Deming 1964, p. 134,181) and owes its name to the auxiliary parameters w_i, which will change their values in each iteration step. Taking into account the interesting comments of Williamson (1968), this pseudo-cubic equation can be alternatively expressed by a pseudo-quadratic or even as a pseudo-linear one. Such a pseudo-quadratic equation has been presented in (York, 1968) including correlations between the observations, which is equivalent to the development of (Petrović et al., 1983) when setting the correlations equal to zero and is expressed by

$$a^2 \sum_{i=1}^{n} \frac{w_i^2}{p_{x_i}} (p_{y_i} \; x_i' \; y_i') + a\left(\sum_{i=1}^{n} \frac{w_i^2}{p_{x_i}} y_i'^2 - \sum_{i=1}^{n} \frac{w_i^2}{p_{y_i}} x_i'^2\right) - \sum_{i=1}^{n} \frac{w_i^2}{p_{y_i}} x_i'^2 y_i'^2, \tag{5.60}$$

under the restriction of $b = 1$. Iterative algorithms can be easily built for the least squares solution of the line parameters, by making use either of the pseudo-cubic (5.58) or the pseudo-quadratic equation (5.60).

5.2.4 Weighting case 4 - Individually weighted and correlated 2D coordinates

The developed iterative procedure of the previous subsection can be extended to include correlations between the observed quantities. Correlations are often considered in geodetic applications, as it is typical that the 2D coordinates of points are not the original measured quantities, but they have been obtained for example by polar measurements. In another example, these coordinates are the outcome of some previous adjustment, for instance of a 2D network. The precision of the adjusted 2D coordinates is obtained in most cases from a linear error propagation that sometimes results in a cofactor matrix that includes correlations between the observations, depending on the adjustment problem.

It is very important to point out here that the precisions of the coordinates coming from a linearized error propagation is just an approximation, as it has been discussed in section 2.3. Therefore, when the original measurements are polar coordinates, then these should be utilized for obtaining a rigorous least squares solution. However, the adjusted 2D Cartesian coordinates of points and their approximated stochastic model from the error propagation are often used in practice. Therefore, this subsection is dedicated to those practical least squares solutions for fitting straight lines in 2D, taking into account the approximated variances and covariances of the 2D point coordinates.

It would be beneficial at this point to introduce vector/matrix notation, in order to derive simpler equations for the solution of the adjustment problem. Firstly, the cofactor matrix $\mathbf{Q}_{\mathrm{LL}}$ is given or obtained from a previous adjustment and can be written as

$$\mathbf{Q}_{\mathrm{LL}} = \begin{bmatrix} \mathbf{Q}_{\mathrm{xx}} & \mathbf{Q}_{\mathrm{xy}} \\ \mathbf{Q}_{\mathrm{yx}} & \mathbf{Q}_{\mathrm{yy}} \end{bmatrix}, \text{ with } \mathbf{Q}_{\mathrm{xy}} = \mathbf{Q}_{\mathrm{yx}}^{\mathrm{T}}. \tag{5.61}$$

$\mathbf{Q}_{\mathrm{xx}}$ and $\mathbf{Q}_{\mathrm{yy}}$ are the cofactor matrices of the measured 2D coordinates and matrices $\mathbf{Q}_{\mathrm{xy}}$, $\mathbf{Q}_{\mathrm{yx}}$ hold the correlations between the coordinates. The respective weight matrices

$$\mathbf{P} = \mathbf{Q}_{\mathrm{LL}}^{-1} = \begin{bmatrix} \mathbf{P}_{\mathrm{xx}} & \mathbf{P}_{\mathrm{xy}} \\ \mathbf{P}_{\mathrm{yx}} & \mathbf{P}_{\mathrm{yy}} \end{bmatrix}, \text{ with } \mathbf{P}_{\mathrm{xy}} = \mathbf{P}_{\mathrm{yx}}^{\mathrm{T}}. \tag{5.62}$$

The nonlinear condition equations (5.1) can be expressed equivalently in vector notation by

$$a(\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}) + b(\mathbf{y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{y}}) + c\ \mathbf{e} = \mathbf{0}, \tag{5.63}$$

with vectors $\mathbf{x}_{\mathrm{c}}$ and $\mathbf{y}_{\mathrm{c}}$ listing the coordinates of the measured points

$$\mathbf{x}_{\mathrm{c}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \ \mathbf{y}_{\mathrm{c}} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \tag{5.64}$$

and vectors $\mathbf{v}_{\mathrm{x}}$ and $\mathbf{v}_{\mathrm{y}}$ carrying the corresponding residuals

$$\mathbf{v}_{\mathrm{x}} = \begin{bmatrix} v_{x_1} \\ v_{x_2} \\ \vdots \\ v_{x_n} \end{bmatrix}, \ \mathbf{v}_{\mathrm{y}} = \begin{bmatrix} v_{y_1} \\ v_{y_2} \\ \vdots \\ v_{y_n} \end{bmatrix}. \tag{5.65}$$

Vector $\mathbf{e}$ is a vector of ones

$$\mathbf{e} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \tag{5.66}$$

with dimension being equal to the number of n measured points. A least squares solution of the problem can be derived by minimizing an objective function expressed in matrix notation by

$$\Omega(\mathbf{v}_{\mathrm{x}}, \mathbf{v}_{\mathrm{y}}) = \mathbf{v}_{\mathrm{x}}^{\mathrm{T}}\mathbf{P}_{\mathrm{xx}}\mathbf{v}_{\mathrm{x}} + \mathbf{v}_{\mathrm{y}}^{\mathrm{T}}\mathbf{P}_{\mathrm{yy}}\mathbf{v}_{\mathrm{y}} + 2\ \mathbf{v}_{\mathrm{x}}^{\mathrm{T}}\mathbf{P}_{\mathrm{xy}}\mathbf{v}_{\mathrm{y}}. \tag{5.67}$$

5.2.4.1 Iterative least squares solution without linearization

Combining the developed condition equations (5.63) and the objective function (5.67), leads to the Lagrange function

$$\mathrm{K}(a, b, c, \mathbf{v}_\mathrm{x}, \mathbf{v}_\mathrm{y}, \mathbf{k}) = \Omega(\mathbf{v}_\mathrm{x}, \mathbf{v}_\mathrm{y}) - 2\mathbf{k}^\mathrm{T}\left[a(\mathbf{x}_\mathrm{c} + \mathbf{v}_\mathrm{x}) + b(\mathbf{y}_\mathrm{c} + \mathbf{v}_\mathrm{y}) + c\ \mathbf{e}\right], \tag{5.68}$$

with $\mathbf{k}$ denoting the vector of Lagrange multipliers. Following the same procedure as in the previous sections, a minimum for K is obtained by differentiation with respect to all unknown parameters and setting the partial derivatives to zero, resulting in the system of nonlinear normal equations

$$\frac{\partial \mathrm{K}}{\partial \mathbf{v}_\mathrm{x}^\mathrm{T}} = 2\left(\mathbf{P}_\mathrm{xx}\mathbf{v}_\mathrm{x} + \mathbf{P}_\mathrm{xy}\mathbf{v}_\mathrm{y} - a\mathbf{k}\right) = \mathbf{0}, \tag{5.69}$$

$$\frac{\partial \mathrm{K}}{\partial \mathbf{v}_\mathrm{y}^\mathrm{T}} = 2\left(\mathbf{P}_\mathrm{yy}\mathbf{v}_\mathrm{y} + \mathbf{P}_\mathrm{yx}\mathbf{v}_\mathrm{x} - b\mathbf{k}\right) = \mathbf{0}, \tag{5.70}$$

$$\frac{\partial \mathrm{K}}{\partial \mathbf{k}^\mathrm{T}} = -2\left[a(\mathbf{x}_\mathrm{c} + \mathbf{v}_\mathrm{x}) + b(\mathbf{y}_\mathrm{c} + \mathbf{v}_\mathrm{y}) + c\ \mathbf{e}\right] = \mathbf{0}, \tag{5.71}$$

$$\frac{\partial \mathrm{K}}{\partial a} = -2\mathbf{k}^\mathrm{T}\left(\mathbf{x}_\mathrm{c} + \mathbf{v}_\mathrm{x}\right) = 0, \tag{5.72}$$

$$\frac{\partial \mathrm{K}}{\partial b} = -2\mathbf{k}^\mathrm{T}\left(\mathbf{y}_\mathrm{c} + \mathbf{v}_\mathrm{y}\right) = 0, \tag{5.73}$$

$$\frac{\partial \mathrm{K}}{\partial c} = -2\mathbf{k}^\mathrm{T}\mathbf{e} = 0. \tag{5.74}$$

A linearisation or approximation of the original problem is avoided here. Explicit expressions for the residuals can be obtained by expressing equations (5.69) and (5.70) using block matrices:

$$\begin{bmatrix} \mathbf{P}_\mathrm{xx} & \mathbf{P}_\mathrm{xy} \\ \mathbf{P}_\mathrm{yx} & \mathbf{P}_\mathrm{yy} \end{bmatrix} \begin{bmatrix} \mathbf{v}_\mathrm{x} \\ \mathbf{v}_\mathrm{y} \end{bmatrix} = \begin{bmatrix} a\mathbf{k} \\ b\mathbf{k} \end{bmatrix}. \tag{5.75}$$

The residual vectors can be computed by

$$\begin{bmatrix} \mathbf{v}_\mathrm{x} \\ \mathbf{v}_\mathrm{y} \end{bmatrix} = \begin{bmatrix} \mathbf{P}_\mathrm{xx} & \mathbf{P}_\mathrm{xy} \\ \mathbf{P}_\mathrm{yx} & \mathbf{P}_\mathrm{yy} \end{bmatrix}^{-1} \begin{bmatrix} a\mathbf{k} \\ b\mathbf{k} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_\mathrm{xx} & \mathbf{Q}_\mathrm{xy} \\ \mathbf{Q}_\mathrm{yx} & \mathbf{Q}_\mathrm{yy} \end{bmatrix} \begin{bmatrix} a\mathbf{k} \\ b\mathbf{k} \end{bmatrix}, \tag{5.76}$$

which yields

$$\mathbf{v}_\mathrm{x} = \left(a\mathbf{Q}_\mathrm{xx} + b\mathbf{Q}_\mathrm{xy}\right)\mathbf{k} \tag{5.77}$$

and

$$\mathbf{v}_\mathrm{y} = \left(a\mathbf{Q}_\mathrm{yx} + b\mathbf{Q}_\mathrm{yy}\right)\mathbf{k}. \tag{5.78}$$

Equivalently, the residual vectors can be computed by using the properties of inverting block matrices, as it has been discussed in (Snow 2012, p. 22). Moreover, introducing the expressions for $\mathbf{v}_\mathrm{x}$ and $\mathbf{v}_\mathrm{y}$ into equation (5.71) yields

$$\left(a^2\mathbf{Q}_\mathrm{xx} + b^2\mathbf{Q}_\mathrm{yy} + ab\mathbf{Q}_\mathrm{xy} + ab\mathbf{Q}_\mathrm{yx}\right)\mathbf{k} = -\left(a\mathbf{x}_\mathrm{c} + b\mathbf{y}_\mathrm{c} + c\mathbf{e}\right). \tag{5.79}$$

Introducing appropriate approximate values a^0 and b^0 in the left hand side of the last equation, allows us to build the auxiliary matrix

$$\mathbf{W} = {a^0}^2\mathbf{Q}_\mathrm{xx} + {b^0}^2\mathbf{Q}_\mathrm{yy} + a^0b^0\mathbf{Q}_\mathrm{xy} + a^0b^0\mathbf{Q}_\mathrm{yx} \tag{5.80}$$

and write equation (5.79) as

$$\mathbf{W}\mathbf{k} = -\left(a\mathbf{x}_\mathrm{c} + b\mathbf{y}_\mathrm{c} + c\mathbf{e}\right). \tag{5.81}$$

In case of regular cofactor matrices, matrix $\mathbf{W}$ is also regular and invertible (the case of singular cofactor matrices in $\mathbf{W}$ is discussed in the next subsection). This leads to the vector of Lagrange multipliers

$$\mathbf{k} = -\mathbf{W}^{-1}\left(a\mathbf{x}_\mathrm{c} + b\mathbf{y}_\mathrm{c} + c\mathbf{e}\right). \tag{5.82}$$

Furthermore, substituting vector $\mathbf{k}$ into the normal equation (5.74) results in

$$\begin{gathered}\mathbf{e}^\mathrm{T}\mathbf{k} = 0\\ \Rightarrow -\mathbf{e}^\mathrm{T}\mathbf{W}^{-1}\left(a\mathbf{x}_\mathrm{c} + b\mathbf{y}_\mathrm{c} + c\mathbf{e}\right) = 0\\ \Rightarrow c = -a\left[\left(\mathbf{e}^\mathrm{T}\mathbf{W}^{-1}\mathbf{e}\right)^{-1}\mathbf{e}^\mathrm{T}\mathbf{W}^{-1}\mathbf{x}_\mathrm{c}\right] - b\left[\left(\mathbf{e}^\mathrm{T}\mathbf{W}^{-1}\mathbf{e}\right)^{-1}\mathbf{e}^\mathrm{T}\mathbf{W}^{-1}\mathbf{y}_\mathrm{c}\right].\end{gathered} \tag{5.83}$$

The solution for the unknown line parameters a and b can be obtained by analyzing further equations (5.72) and (5.73). Taking into account the solution for the vector of Lagrange multipliers from equation (5.82) and for the residual vectors from (5.77) and (5.78), yields the reduced system of equations

$$-\mathbf{x}_\mathrm{c}^\mathrm{T}\mathbf{W}^{-1}\left(a\mathbf{x}_\mathrm{c} + b\mathbf{y}_\mathrm{c} + c\mathbf{e}\right) + \mathbf{k}^\mathrm{T}\left(a\mathbf{Q}_\mathrm{xx}\mathbf{k} + b\mathbf{Q}_\mathrm{xy}\mathbf{k}\right) = 0 \tag{5.84}$$

and

$$-\mathbf{y}_\mathrm{c}^\mathrm{T}\mathbf{W}^{-1}\left(a\mathbf{x}_\mathrm{c} + b\mathbf{y}_\mathrm{c} + c\mathbf{e}\right) + \mathbf{k}^\mathrm{T}\left(a\mathbf{Q}_\mathrm{xy}\mathbf{k} + b\mathbf{Q}_\mathrm{yy}\mathbf{k}\right) = 0. \tag{5.85}$$

Substituting also parameter c from (5.83), the last two equations can be equivalently written as

$$a\ f_1 + b\ f_2 = 0 \tag{5.86}$$

and

$$b\ f_3 + a\ f_2 = 0, \tag{5.87}$$

with the respective quantities being

$$\begin{aligned} f_1 &= \mathbf{k}^{\mathrm{T}}\mathbf{Q}_{\mathrm{xx}}\mathbf{k} - \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{x}_{\mathrm{c}} + \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\left(\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\right)^{-1}\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{x}_{\mathrm{c}}, \\ f_2 &= \mathbf{k}^{\mathrm{T}}\mathbf{Q}_{\mathrm{xy}}\mathbf{k} - \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{y}_{\mathrm{c}} + \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\left(\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\right)^{-1}\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{x}_{\mathrm{c}}, \\ f_3 &= \mathbf{k}^{\mathrm{T}}\mathbf{Q}_{\mathrm{yy}}\mathbf{k} - \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{y}_{\mathrm{c}} + \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\left(\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\right)^{-1}\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{y}_{\mathrm{c}}. \end{aligned} \tag{5.88}$$

Equations (5.86) and (5.87) form a system of pseudo-linear equations with two unknown parameters (i.e. the line parameters a and b). The term "pseudo-linear" comes from the fact that parameters f_1, f_2 and f_3 contain matrix $\mathbf{W}$, which has been built using the approximations a^0 and b^0. Furthermore, a restriction of the problem to $b = 1$ leads to the solution for the remaining unknown parameter

$$a = -\frac{f_2}{f_1}. \tag{5.89}$$

An iterative procedure for estimating the unknown line parameters can be found in Algorithm 2. A similar iterative solution for this weighting case has been presented in (York 1968).

Algorithm 2 Least squares fitting of a straight line in 2D with general weights and correlations

1: Choose approximate value for a^0.
2: Define parameter $b = 1$.
3: Set threshold ϵ for the break-off condition of the iteration process.
4: Set parameter $d_a = |\hat{a} - a^0| = \infty$, for entering the iteration process.
5: **while** $d_a > \epsilon$ **do**
6: Compute the auxiliary matrix $\mathbf{W}$, and the vector of Lagrange multipliers $\mathbf{k}$.
7: Estimate parameter $\hat{c}$.
8: Compute the coefficients f_1 and f_2.
9: Estimate parameter $\hat{a}$.
10: Compute parameter $d_a = |\hat{a} - a^0|$.
11: Update the approximated parameter with the estimated ($a^0 = \hat{a}$).
12: **end while**
13: **return** $\hat{a}$ and $\hat{c}$, with $b = 1$.

5.2.4.2 Solution for singular cofactor matrices

Postulating regular cofactor matrices in equation (5.81) permitted the inversion of matrix $\mathbf{W}$. This led to the reduction of the normal equations and the solution for the unknown line parameters. However, if the given cofactor matrices are singular, then a solution for the vector of Lagrange multipliers cannot be obtained using equation (5.82), as long as matrix $\mathbf{W}$ is not invertible anymore. However, a solution with iterations is possible also for the case of singular cofactor matrices following a rather different procedure.

Taking into account equation (5.81), together with the normal equations (5.72)-(5.74), the following system of equations emerges:

$$\begin{aligned}
\mathbf{W}\mathbf{k} &= -\left(a\mathbf{x}_{\mathrm{c}} + b\mathbf{y}_{\mathrm{c}} + c\mathbf{e}\right), \\
\mathbf{k}^{\mathrm{T}}\left(\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}\right) &= 0, \\
\mathbf{k}^{\mathrm{T}}\left(\mathbf{y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{y}}\right) &= 0, \\
\mathbf{k}^{\mathrm{T}}\mathbf{e} &= 0.
\end{aligned} \tag{5.90}$$

If the chosen restriction between the unknown parameters is $b = 1$, then the developed normal equations can be written as[4]

$$\begin{aligned}
\mathbf{W}\mathbf{k} &= -\left(a\mathbf{x}_{\mathrm{c}} + \mathbf{y}_{\mathrm{c}} + c\mathbf{e}\right), \\
\mathbf{k}^{\mathrm{T}}\left(\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}\right) &= 0, \\
\mathbf{k}^{\mathrm{T}}\mathbf{e} &= 0.
\end{aligned} \tag{5.91}$$

This equation system can be expressed in a block matrix form, after introducing approximate values for the residual vector $\mathbf{v}_{\mathrm{x}}^{0}$

$$\begin{bmatrix} \mathbf{W} & \mathbf{x}_{\mathrm{c}} & \mathbf{e} \\ \left(\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}^{0}\right)^{\mathrm{T}} & 0 & 0 \\ \mathbf{e}^{\mathrm{T}} & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{k} \\ a \\ c \end{bmatrix} = \begin{bmatrix} -\mathbf{y}_{\mathrm{c}} \\ 0 \\ 0 \end{bmatrix}. \tag{5.92}$$

This can be equivalently written as

$$\mathbf{N} \begin{bmatrix} \mathbf{k} \\ \mathbf{X} \end{bmatrix} = \mathbf{n}, \tag{5.93}$$

with matrices

$$\mathbf{N} = \begin{bmatrix} \mathbf{W} & \mathbf{x}_{\mathrm{c}} & \mathbf{e} \\ \left(\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}^{0}\right)^{\mathrm{T}} & 0 & 0 \\ \mathbf{e}^{\mathrm{T}} & 0 & 0 \end{bmatrix}, \quad \mathbf{n} = \begin{bmatrix} -\mathbf{y}_{\mathrm{c}} \\ 0 \\ 0 \end{bmatrix} \tag{5.94}$$

and the vector of unknown parameters

$$\mathbf{X} = \begin{bmatrix} a \\ c \end{bmatrix}. \tag{5.95}$$

[4] The third equation, $\mathbf{k}^{\mathrm{T}}\left(\mathbf{y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{y}}\right) = 0$, is not taken into account as we set b as a fixed parameter.

A least squares solution for this adjustment problem can be computed by

$$\begin{bmatrix} \hat{\mathbf{k}} \\ \hat{\mathbf{X}} \end{bmatrix} = \mathbf{N}^{-1}\mathbf{n}, \tag{5.96}$$

without applying any linearization to the original problem.

Solution with a symmetric normal matrix N

The developed normal matrix $\mathbf{N}$ from equation (5.94) is nonsymmetric. An equivalent solution of the problem using, however, a symmetric matrix $\mathbf{N}$ can be obtained by adding the term $a\mathbf{v}_{\mathrm{x}}$ to both sides of equation (5.81)[5]. This leads to the equation system

$$\begin{aligned} \mathbf{W}\mathbf{k} + a\left(\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}\right) + c\mathbf{e} &= -\mathbf{y}_{\mathrm{c}} + a\mathbf{v}_{\mathrm{x}}, \\ \mathbf{k}^{\mathrm{T}}\left(\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}\right) &= 0, \\ \mathbf{k}^{\mathrm{T}}\mathbf{e} &= 0 \end{aligned} \tag{5.97}$$

or written in block matrix form

$$\begin{bmatrix} \mathbf{W} & \mathbf{A} \\ \mathbf{A}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{k} \\ \mathbf{X} \end{bmatrix} = \begin{bmatrix} \mathbf{w} \\ \mathbf{0} \end{bmatrix}, \tag{5.98}$$

with matrix

$$\mathbf{A} = \left[\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}^{0}\ ,\ \mathbf{e}\right] \tag{5.99}$$

and vector

$$\mathbf{w} = -\mathbf{y}_{\mathrm{c}} + a^{0}\mathbf{v}_{\mathrm{x}}^{0}. \tag{5.100}$$

It is worth noticing the similarity of matrix $\mathbf{A}$ and vector $\mathbf{w}$ with those from the GHM, as explained in section 3.1.2, without however applying a linearization to the functional model. Therefore, a solution of the adjustment problem can be obtained by equation (5.96), after introducing

$$\mathbf{N} = \begin{bmatrix} \mathbf{W} & \mathbf{A} \\ \mathbf{A}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \quad , \quad \mathbf{n} = \begin{bmatrix} \mathbf{w} \\ \mathbf{0} \end{bmatrix}, \tag{5.101}$$

where $\mathbf{N}$ is in this case symmetric.

The inversion of matrix $\mathbf{N}$ depends on the rank deficiency of matrix $\mathbf{W}$, which is not invertible as long as the cofactor matrices of the adjustment problem are singular. An elegant way to ensure that a unique solution exists, even when singular cofactor matrices must be employed, is the fullfilment of the Neitzel-Schaffrin

[5] A similar procedure for deriving a symetric normal matrix $\mathbf{N}$ is already known in the literature dealing with WTLS algorithms. An example can be found in (Snow 2012, pp. 25-26).

(NS) criterion that was proposed by Neitzel and Schaffrin (2016). As a linearization of the problem has been avoided in the presented solution strategy, a similar criterion is presented in the following that will ensure a unique solution for the unknown parameters. Starting with the equation system from equation (5.98), with

- rank of $\mathbf{W} \leq n$, with n = number of condition equations;
- rank of $\mathbf{A} = m$, with m = number of unknown parameters;
- redundancy : $r_d = n - m$;

while the rank of matrix $\mathbf{W}$ will be smaller than n in cases of singular cofactor matrices. Similar to the NS criterion for the GHM, a unique solution will exist if the rank of the augmented matrix $[\mathbf{W} \mid \mathbf{A}]$ is equal to the number of condition equations n of the problem (which is equal to the rank of matrix $\mathbf{B}$ in the GHM). A criterion that would ensure a unique solution of the problem can be described in this case by

$$\operatorname{rank}\left([\mathbf{W} \mid \mathbf{A}]\right) = n. \tag{5.102}$$

An iterative procedure is presented in Algorithm 3, for the special cases of fitting straight lines in 2D with singular cofactor matrices. It must be pointed out that this is a general procedure and can be used to derive a solution also for the previous examined weighting cases for fitting a straight line in 2D. However, this iteration process involves more complicated normal equations and thus must be preferred when a singular cofactor matrix is given.

Algorithm 3 Least squares fitting of a straight line in 2D with singular cofactor matrices

Choose approximate values for a^0, $\mathbf{v}_\mathrm{x}^0$ and $\mathbf{v}_\mathrm{y}^0$.
Define parameter $b = 1$.
Set threshold ϵ for the break-off condition of the iteration process.
Set parameter $d_a = |\hat{a} - a^0| = \infty$, for entering the iteration process.
while $d_a > \epsilon$ **do**
    Compute matrices $\mathbf{W}$, $\mathbf{A}$ and vector $\mathbf{w}$.
    Build matrix $\mathbf{N}$ and vector $\mathbf{n}$.
    Estimate the vector of unknowns $\begin{bmatrix} \hat{\mathbf{k}} \\ \hat{\mathbf{X}} \end{bmatrix}$.
    Compute the residual vectors $\mathbf{v}_\mathrm{x}$ and $\mathbf{v}_\mathrm{y}$.
    Compute parameter $d_a = |\hat{a} - a^0|$.
    Update the approximate values with the estimated ones, with $a^0 = \hat{a}$, $\mathbf{v}_\mathrm{x}^0 = \mathbf{v}_\mathrm{x}$ and $\mathbf{v}_\mathrm{y}^0 = \mathbf{v}_\mathrm{y}$.
end while
return $\hat{a}$ and $\hat{c}$, with $b = 1$.

5.3 Fitting of a plane in 3D

The problem of fitting a plane to an observed point cloud in 3D is investigated here, with the observed coordinates being obtained with different precisions. The general form of a plane in 3D has been already presented in subsection 4.4:

$$ax + by + cz + d = 0,$$

with x, y and z denoting the 3D coordinates of a point that lies in the plane. a, b, c and d are the unknown plane parameters. Random errors will influence the observed quantities, thus individual residuals can be introduced in the functional model, leading to the condition equations

$$a(x_i + v_{x_i}) + b(y_i + v_{y_i}) + c(z_i + v_{z_i}) + d = 0. \tag{5.103}$$

An "optimal" solution for the unknown plane parameters is possible by minimizing the sum of weighted squared residuals.The least squares solution of the best plane is investigated, for four different weighting cases:

1. Same precision σ_x for the coordinates in x direction, σ_y in y direction and σ_z in z direction.
2. Individual precision for each point: $\sigma_{x_i} = \sigma_{y_i} = \sigma_{z_i} \forall i$.
3. Individual precision for each coordinate.
4. Individual precision and correlations between the measured 3D coordinates.

5.3.1 Weighting case 1 - Equally weighted observations in each direction

The fitting a plane to a 3D point cloud is examined in this subsection, with the measured point coordinates being observed with the same precision in each direction. The estimation of the unknown plane parameters is based on the minimization of the sum of weighted squared residuals

$$\sum_{i=1}^{n} p_x\ v_{x_i}^2 + p_y\ v_{y_i}^2 + p_z\ v_{z_i}^2 \rightarrow \min, \tag{5.104}$$

with the constant weights p_x, p_y and p_z being computed by

$$p_x = \frac{1}{\sigma_x^2} \quad , \quad p_y = \frac{1}{\sigma_y^2} \quad \text{and} \quad p_z = \frac{1}{\sigma_z^2}. \tag{5.105}$$

Following the same solution strategy as in subsection 5.2.1, the observed 3D coordinates are multiplied beforehand with the respective weights leading to the scaled coordinates

$$x_i^s = x_i \sqrt{p_x}\ ,\ y_i^s = y_i \sqrt{p_y} \quad \text{and} \quad z_i^s = z_i \sqrt{p_z} \tag{5.106}$$

and the respective residuals

$$v_{x_i}^s = v_{x_i} \sqrt{p_x}\ ,\ v_{y_i}^s = v_{y_i} \sqrt{p_y} \quad \text{and} \quad v_{z_i}^s = v_{z_i} \sqrt{p_z}. \tag{5.107}$$

Substituting the scaled coordinates and their residuals into the condition equations (5.103) yields

$$a\frac{1}{\sqrt{p_x}}\left(x_i^s+v_{x_i}^s\right)+b\frac{1}{\sqrt{p_y}}\left(y_i^s+v_{y_i}^s\right)+c\frac{1}{\sqrt{p_z}}\left(z_i^s+v_{z_i}^s\right)+d=0. \tag{5.108}$$

Furthermore, introducing the auxiliary scaled plane parameters

$$\begin{aligned} a^s &= a\frac{1}{\sqrt{p_x}}, \\ b^s &= b\frac{1}{\sqrt{p_y}}, \\ c^s &= c\frac{1}{\sqrt{p_z}}, \\ d^s &= d, \end{aligned} \tag{5.109}$$

in equation (5.108), results in the functional model

$$a^s\left(x_i^s+v_{x_i}^s\right)+b^s\left(y_i^s+v_{y_i}^s\right)+c^s\left(z_i^s+v_{z_i}^s\right)+d^s=0. \tag{5.110}$$

Similar to the procedure of subsection 4.4.1, a "meaningful" restriction between the unknown plane parameters is

$${a^s}^2+{b^s}^2+{c^s}^2=1. \tag{5.111}$$

The least squares solution for the unknown plane parameters can be derived by minimizing the sum of squared scaled residuals

$$\sum_{i=1}^{n} {v_{x_i}^s}^2+{v_{y_i}^s}^2+{v_{z_i}^s}^2 \rightarrow \min. \tag{5.112}$$

Direct least squares solution in a scaled coordinate system

Utilizing the scaled 3D coordinates, the original problem is transformed to a problem with equal weights. Thus, the sum of squared residuals is equal to the sum of squared orthogonal distances

$$\Omega(v_{x_i}^s,v_{y_i}^s)=\sum_{i=1}^{n} {v_{x_i}^s}^2+{v_{y_i}^s}^2+{v_{z_i}^s}^2=\sum_{i=1}^{n} D_i^2, \tag{5.113}$$

with the distances of the measured points to the requested plane

$$D_i=\frac{a^s x_i^s+b^s y_i^s+c^s z_i^s+d^s}{\sqrt{{a^s}^2+{b^s}^2+{c^s}^2}}. \tag{5.114}$$

Taking into account the selected restriction from equation (5.111), the expressions for the orthogonal distances become

$$D_i=a^s x_i^s+b^s y_i^s+c^s z_i^s+d^s. \tag{5.115}$$

A least squares solution for the unknown plane parameters can be derived in this case directly, following the same procedure as the one presented in subsection 4.4.

Computation of the plane parameters in the original coordinate system

Introducing the estimated plane parameters $\hat{a}^s$, $\hat{b}^s$, $\hat{c}^s$ and $\hat{d}^s$ into equation (5.109), yields the plane parameters in the original coordinate system

$$\begin{aligned} \hat{a} &= \hat{a}^s \sqrt{p_x}, \\ \hat{b} &= \hat{b}^s \sqrt{p_y}, \\ \hat{c} &= \hat{c}^s \sqrt{p_z}, \\ \hat{d} &= d^s, \end{aligned} \tag{5.116}$$

however, being restricted to

$$a^{s2} + b^{s2} + c^{s2} = \left(\frac{a}{\sqrt{p_x}}\right)^2 + \left(\frac{b}{\sqrt{p_y}}\right)^2 + \left(\frac{c}{\sqrt{p_z}}\right)^2 = 1. \tag{5.117}$$

Therefore, a least squares solution that is restricted to $a^2 + b^2 + c^2 = 1$ can be computed by

$$\begin{aligned} \hat{a} &= \frac{\hat{a}^s \sqrt{p_x}}{\sqrt{\left(\hat{a}^s \sqrt{p_x}\right)^2 + \left(\hat{b}^s \sqrt{p_y}\right)^2 + \left(\hat{c}^s \sqrt{p_z}\right)^2}}, \\ \hat{b} &= \frac{\hat{b}^s \sqrt{p_x}}{\sqrt{\left(\hat{a}^s \sqrt{p_x}\right)^2 + \left(\hat{b}^s \sqrt{p_y}\right)^2 + \left(\hat{c}^s \sqrt{p_z}\right)^2}}, \\ \hat{c} &= \frac{\hat{c}^s \sqrt{p_z}}{\sqrt{\left(\hat{a}^s \sqrt{p_x}\right)^2 + \left(\hat{b}^s \sqrt{p_y}\right)^2 + \left(\hat{c}^s \sqrt{p_z}\right)^2}}, \\ \hat{d} &= \frac{\hat{d}^s}{\sqrt{\left(\hat{a}^s \sqrt{p_x}\right)^2 + \left(\hat{b}^s \sqrt{p_y}\right)^2 + \left(\hat{c}^s \sqrt{p_z}\right)^2}}. \end{aligned} \tag{5.118}$$

5.3.2 Weighting case 2 - Individually weighted points in 3D

In the second weighting scenario for fitting a plane in 3D, every point has been observed with individual precision:

$$\begin{gathered} \sigma_{x_i} = \sigma_{y_i} = \sigma_{z_i} \forall i \\ \Rightarrow p_{x_i} = p_{y_i} = p_{z_i} = p_i \forall i. \end{gathered} \tag{5.119}$$

Also in this case the ratio between the weights of each point is constant:

$$\frac{p_{x_i}}{p_{y_i}} = \text{constant} \quad , \quad \frac{p_{y_i}}{p_{z_i}} = \text{constant} \quad \text{and} \quad \frac{p_{x_i}}{p_{z_i}} = \text{constant}. \tag{5.120}$$

A least squares solution for the unknown plane parameters can be found by minimizing

$$\Omega(v_{x_i}, v_{y_i}, v_{z_i}) = \sum_{i=1}^{n} p_{x_i} v_{x_i}^2 + p_{y_i} v_{y_i}^2 + p_{z_i} v_{z_i}^2, \tag{5.121}$$

which after taking into account the stochastic model of equation (5.119), becomes

$$\Omega(v_{x_i}, v_{y_i}, v_{z_i}) = \sum_{i=1}^{n} p_i (v_{x_i}^2 + v_{y_i}^2 + v_{z_i}^2) \rightarrow \min. \tag{5.122}$$

Direct weighted least squares solution

The orthogonal distances of the measured points to the requested plane are

$$D_i = \frac{a x_i + b y_i + c z_i + d}{\sqrt{a^2 + b^2 + c^2}} \tag{5.123}$$

and after selecting the restriction

$$a^2 + b^2 + c^2 = 1, \tag{5.124}$$

can be equivalently written as

$$D_i = a x_i + b y_i + c z_i + d. \tag{5.125}$$

The sum of weighted squared residuals of this adjustment problem is equal to the sum of weighted squared orthogonal distances

$$p_i D_i^2 = p_i \left(v_{x_i}^2 + v_{y_i}^2 + v_{z_i}^2 \right). \tag{5.126}$$

Thus, the objective function is

$$\Omega(a, b, c, d) = \sum_{i=1}^{n} p_i (v_{x_i}^2 + v_{y_i}^2 + v_{z_i}^2) = p_i D_i^2 = p_i \left(a x_i + b y_i + c z_i + d \right)^2. \tag{5.127}$$

To obtain a solution for the unknown plane parameters which minimizes equation (5.127) under the chosen restriction, the Lagrangian

$$\mathrm{K}(a, b, c, k) = \sum_{i=1}^{n} p_i (a x_i + b y_i + c z_i + d)^2 - k(a^2 + b^2 + c^2 - 1), \tag{5.128}$$

can be built. The normal equation system for this problem is

$$\frac{\partial \mathrm{K}}{\partial a} = 2 \left[a \left(\sum_{i=1}^{n} p_i {x_i}^2 - k \right) + b \left(\sum_{i=1}^{n} p_i y_i x_i \right) + c \left(\sum_{i=1}^{n} p_i x_i z_i \right) + d \left(\sum_{i=1}^{n} p_i x_i \right) \right] = 0, \tag{5.129}$$

$$\frac{\partial \mathrm{K}}{\partial b} = 2 \left[a \left(\sum_{i=1}^{n} p_i y_i x_i \right) + b \left(\sum_{i=1}^{n} p_i {y_i}^2 - k \right) + c \left(\sum_{i=1}^{n} p_i y_i z_i \right) + d \left(\sum_{i=1}^{n} p_i y_i \right) \right] = 0, \tag{5.130}$$

$$\frac{\partial \mathrm{K}}{\partial c} = 2\left[a\left(\sum_{i=1}^{n} p_i x_i z_i\right) + b\left(\sum_{i=1}^{n} p_i y_i z_i\right) + c\left(\sum_{i=1}^{n} p_i {z_i}^2 - k\right) + d\left(\sum_{i=1}^{n} p_i z_i\right)\right] = 0, \tag{5.131}$$

$$\frac{\partial \mathrm{K}}{\partial d} = 2\left[a\sum_{i=1}^{n} p_i x_i + b\sum_{i=1}^{n} p_i y_i + c\sum_{i=1}^{n} p_i z_i + d\sum_{i=1}^{n} p_i\right] = 0, \tag{5.132}$$

$$\frac{\partial \mathrm{K}}{\partial k} = -\left(a^2 + b^2 + c^2 - 1\right) = 0. \tag{5.133}$$

Equation (5.132) can be rearranged to

$$d = -\frac{1}{\sum\limits_{i=1}^{n} p_i}\left(a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} y_i + c\sum_{i=1}^{n} z_i\right). \tag{5.134}$$

The expression for d can be subsequently introduced into equations (5.129) - (5.131). This yields the reduced normal equations

$$\begin{aligned} &a\left[\sum_{i=1}^{n} p_i {x_i}^2 - \frac{1}{\sum\limits_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i x_i\right)^2 - k\right] + b\left[\sum_{i=1}^{n} p_i x_i y_i - \frac{1}{\sum\limits_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i x_i \sum_{i=1}^{n} p_i y_i\right)\right] + \\ &c\left[\sum_{i=1}^{n} p_i x_i z_i - \frac{1}{\sum\limits_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i x_i \sum_{i=1}^{n} p_i z_i\right)\right] = 0, \end{aligned} \tag{5.135}$$

$$\begin{aligned} &a\left[\sum_{i=1}^{n} p_i x_i y_i - \frac{1}{\sum\limits_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i x_i \sum_{i=1}^{n} p_i y_i\right)\right] + b\left[\sum_{i=1}^{n} p_i {y_i}^2 - \frac{1}{\sum\limits_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i y_i\right)^2 - k\right] + \\ &c\left[\sum_{i=1}^{n} p_i y_i z_i - \frac{1}{\sum\limits_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i y_i \sum_{i=1}^{n} p_i z_i\right)\right] = 0 \end{aligned} \tag{5.136}$$

and

$$
\begin{aligned}
&a\left[\sum_{i=1}^{n} p_i x_i z_i - \frac{1}{\sum_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i x_i \sum_{i=1}^{n} p_i z_i\right)\right] + b\left[\sum_{i=1}^{n} p_i y_i z_i - \frac{1}{\sum_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i y_i \sum_{i=1}^{n} p_i z_i\right)\right] + \\
&\qquad c\left[\sum_{i=1}^{n} p_i {z_i}^2 - \frac{1}{\sum_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i z_i\right)^2 - k\right] = 0,
\end{aligned}
\tag{5.137}
$$

which represent a homogeneous system of equations. A nontrivial solution is possible by setting the determinant equal to zero:

$$
\begin{vmatrix} (f_1 - k) & g_1 & g_2 \\ g_1 & (f_2 - k) & g_3 \\ g_2 & g_3 & (f_3 - k) \end{vmatrix} = 0,
\tag{5.138}
$$

with the respective quantities being

$$
\begin{aligned}
f_1 &= \sum_{i=1}^{n} p_i {x_i}^2 - \frac{1}{\sum_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i x_i\right)^2, \quad f_2 = \sum_{i=1}^{n} p_i {y_i}^2 - \frac{1}{\sum_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i y_i\right)^2, \\
f_3 &= \sum_{i=1}^{n} p_i {z_i}^2 - \frac{1}{\sum_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i z_i\right)^2, \\
g_1 &= \sum_{i=1}^{n} p_i x_i y_i - \frac{1}{\sum_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i x_i \sum_{i=1}^{n} p_i y_i\right), \quad g_2 = \sum_{i=1}^{n} p_i x_i z_i - \frac{1}{\sum_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i x_i \sum_{i=1}^{n} p_i z_i\right), \\
g_3 &= \sum_{i=1}^{n} p_i y_i z_i - \frac{1}{\sum_{i=1}^{n} p_i}\left(\sum_{i=1}^{n} p_i y_i \sum_{i=1}^{n} p_i z_i\right).
\end{aligned}
\tag{5.139}
$$

Equation (5.138) is a cubic characteristic equation with three solutions for k. The unknown plane parameters can be estimated either by substituting k_{min} into equations (5.135)-(5.137) or by solving an eigenvalue problem.

5.3.3 Weighting case 3 - Individually weighted 3D coordinates

For the third weighting case of fitting a plane to a 3D point cloud, the observed 3D coordinates of each point have been obtained with individual precisions. The objective function under minimization can be expressed in this case as

$$
\Omega(v_{x_i}, v_{y_i}, v_{z_i}) = \sum_{i=1}^{n} p_{x_i} v_{x_i}^2 + p_{y_i} v_{y_i}^2 + p_{z_i} v_{z_i}^2 .
\tag{5.140}
$$

Iterative least squares solution without linearization

A least squares solution can be computed for this nonlinear adjustment by minimizing the Lagrange function

$$\mathrm{K}(a, b, c, v_{x_i}, v_{y_i}, v_{zi}, k_i) = \Omega(v_{x_i}, v_{y_i}, v_{z_i}) - 2\sum_{i=1}^{n} k_i \left[a(x_i + v_{x_i}) + b(y_i + v_{y_i}) + c(z_i + v_{z_i}) + d\right]. \quad (5.141)$$

Avoiding any kind of linearization of the problem, a differentiation of function K with respect to all unknown parameters and setting the partial derivatives equal to zero yields the system of normal equations

$$\begin{aligned} \frac{\partial \mathrm{K}}{\partial v_{x_i}} &= 2p_{x_i} v_{x_i} - 2ak_i = 0 \\ &\Rightarrow v_{x_i} = \frac{ak_i}{p_{x_i}}, \end{aligned} \quad (5.142)$$

$$\begin{aligned} \frac{\partial \mathrm{K}}{\partial v_{y_i}} &= 2p_{y_i}\ v_{y_i} - 2bk_i = 0 \\ &\Rightarrow v_{y_i} = \frac{bk_i}{p_{y_i}}, \end{aligned} \quad (5.143)$$

$$\begin{aligned} \frac{\partial \mathrm{K}}{\partial v_{z_i}} &= 2p_{z_i} v_{z_i} - 2ck_i = 0 \\ &\Rightarrow v_{z_i} = \frac{ck_i}{p_{z_i}}, \end{aligned} \quad (5.144)$$

$$\frac{\partial \mathrm{K}}{\partial k} = -2\left[a(x_i + v_{x_i}) + b(y_i + v_{y_i}) + c(z_i + v_{zi}) + d\right] = 0, \quad (5.145)$$

$$\frac{\partial \mathrm{K}}{\partial a} = -2\sum_{i=1}^{n} k_i (x_i + v_{x_i}) = 0, \quad (5.146)$$

$$\frac{\partial \mathrm{K}}{\partial b} = -2\sum_{i=1}^{n} k_i (y_i + v_{y_i}) = 0, \quad (5.147)$$

$$\frac{\partial \mathrm{K}}{\partial c} = -2\sum_{i=1}^{n} k_i (z_i + v_{z_i}) = 0, \quad (5.148)$$

$$\frac{\partial \mathrm{K}}{\partial d} = -2\sum_{i=1}^{n} k_i = 0. \quad (5.149)$$

Equations (5.142)-(5.149) represent a nonlinear system of $4n+4$ equations. The residuals from (5.142)-(5.144) are introduced into equation (5.145), that yields the expression for the Lagrange multipliers

$$k_i = w_i(ax_i + by_i + cz_i + d), \quad (5.150)$$

with the auxiliary weighting factors

$$w_i = -\left(\frac{a^2}{p_{x_i}} + \frac{b^2}{p_{y_i}} + \frac{c^2}{p_{zi}}\right)^{-1}. \tag{5.151}$$

Introducing k_i into (5.149) returns

$$d = -\frac{a\sum_{i=1}^{n} w_i\ x_i + b\sum_{i=1}^{n} w_i\ y_i + c\sum_{i=1}^{n} w_i\ z_i}{\sum_{i=1}^{n} w_i}. \tag{5.152}$$

Substituting the residuals v_{x_i}, v_{y_i}, v_{z_i} and the Lagrange multipliers k_i into equations (5.146), (5.147) and (5.148), results in the system of reduced equations

$$a\sum_{i=1}^{n} \frac{k_i^2}{p_{x_i}} = -\sum_{i=1}^{n} w_i\ (ax_i + by_i + cz_i + d)x_i, \tag{5.153}$$

$$b\sum_{i=1}^{n} \frac{k_i^2}{p_{y_i}} = -\sum_{i=1}^{n} w_i(ax_i + by_i + cz_i + d)y_i \tag{5.154}$$

and

$$c\sum_{i=1}^{n} \frac{k_i^2}{p_{z_i}} = -\sum_{i=1}^{n} w_i(ax_i + by_i + cz_i + d)z_i. \tag{5.155}$$

Taking into account parameter d yields

$$af_1 = bf_2 + cf_3, \tag{5.156}$$

$$bf_4 = af_2 + cf_5 \tag{5.157}$$

and

$$cf_6 = af_3 + bf_5, \tag{5.158}$$

with the quantities

$$
\begin{aligned}
f_1 &= \sum_{i=1}^{n} \frac{k_i^2}{p_{x_i}} - \sum_{i=1}^{n} w_i\ x_i^2 + \frac{1}{\sum\limits_{i=1}^{n} w_i} \left(\sum_{i=1}^{n} w_i\ x_i \right), \\
f_2 &= \sum_{i=1}^{n} (w_i\ x_i\ y_i) - \frac{1}{\sum\limits_{i=1}^{n} w_i} \left(\sum_{i=1}^{n} w_i\ x_i \sum_{i=1}^{n} w_i\ y_i \right), \\
f_3 &= \sum_{i=1}^{n} (w_i\ x_i\ z_i) - \frac{1}{\sum\limits_{i=1}^{n} w_i} \left(\sum_{i=1}^{n} w_i\ x_i \sum_{i=1}^{n} w_i\ z_i \right), \\
f_4 &= \sum_{i=1}^{n} \frac{k_i^2}{p_{y_i}} - \sum_{i=1}^{n} w_i\ y_i^2 + \frac{1}{\sum\limits_{i=1}^{n} w_i} \left(\sum_{i=1}^{n} w_i\ y_i \right), \\
f_5 &= \sum_{i=1}^{n} (w_i\ y_i\ z_i) - \frac{1}{\sum\limits_{i=1}^{n} w_i} \left(\sum_{i=1}^{n} w_i\ y_i \sum_{i=1}^{n} w_i\ z_i \right), \\
f_6 &= \sum_{i=1}^{n} \frac{k_i^2}{p_{z_i}} - \sum_{i=1}^{n} w_i\ z_i^2 + \frac{1}{\sum\limits_{i=1}^{n} w_i} \left(\sum_{i=1}^{n} w_i\ z_i \right).
\end{aligned}
\tag{5.159}
$$

At this stage a meaningful restriction between the unknown parameters can be selected. For the sake of convenience the restriction $c = 1$ is chosen. Therefore, solving equation (5.157) for b and introducing it in (5.156) gives

$$\hat{a} = \left(f_1 - \frac{f_2^2}{f_4} \right)^{-1} \left(\frac{f_2\ f_5}{f_4} + f_3 \right) \tag{5.160}$$

and

$$\hat{b} = \hat{a} \left(\frac{f_2}{f_4} \right) + \frac{f_5}{f_4}. \tag{5.161}$$

Equations (5.160) and (5.161) become pseudo-linear after approximating functions $f_1, f_2, \ldots, f_6$. Therefore, initial values for a^0 and b^0 are necessary for computing the auxiliary weighting factors w_i, the Lagrange multipliers k_i, as well as functions $f_1, f_2, \ldots, f_6$. The estimated plane parameters can be utilized as new starting values in each iteration step, until a break-off condition is met. Based on the presented procedure, Algorithm 4 has been developed for estimating the weighted least squares solution of fitting a plane to a 3D point-cloud.

Algorithm 4 Least squares fitting of a plane to points in 3D with general weights

```
Choose approximate values for a^0, b^0.
Define parameter c = 1.
Set threshold ε for the break-off condition of the iteration process.
Set parameter d_a = |â − a^0| = ∞ and d_b = |b̂ − b^0| = ∞, for entering the iteration process.
while d_a > ε or d_b > ε do
    Compute parameters k_i, p_i and estimate d̂.
    Compute the coefficients f_1, f_2, ..., f_5.
    Estimate parameters â and b̂.
    Compute parameter d_a = |â − a^0| and d_b = |b̂ − b^0|.
    Update the approximate values with the estimated ones (a^0 = â and b^0 = b̂).
end while
return â, b̂ and d̂, with c = 1.
```

5.3.4 Weighting case 4 - Individually weighted and correlated 3D coordinates

For the fourth weighted adjustment problem under investigation, correlations are introduced between the measurements. Therefore, the cofactor matrix of the 3D point coordinates is given or obtained from a previous adjustment and is expressed by

$$\mathbf{Q}_{\mathrm{LL}} = \begin{bmatrix} \mathbf{Q}_{\mathrm{xx}} & \mathbf{Q}_{\mathrm{xy}} & \mathbf{Q}_{\mathrm{xz}} \\ \mathbf{Q}_{\mathrm{yx}} & \mathbf{Q}_{\mathrm{yy}} & \mathbf{Q}_{\mathrm{yz}} \\ \mathbf{Q}_{\mathrm{zx}} & \mathbf{Q}_{\mathrm{zy}} & \mathbf{Q}_{\mathrm{zz}} \end{bmatrix}, \text{ with } \mathbf{Q}_{\mathrm{xy}} = \mathbf{Q}_{\mathrm{yx}}^{\mathrm{T}}, \ \mathbf{Q}_{\mathrm{xz}} = \mathbf{Q}_{\mathrm{zx}}^{\mathrm{T}} \text{ and } \mathbf{Q}_{\mathrm{yz}} = \mathbf{Q}_{\mathrm{zy}}^{\mathrm{T}}. \tag{5.162}$$

$\mathbf{Q}_{\mathrm{xx}}$, $\mathbf{Q}_{\mathrm{yy}}$ and $\mathbf{Q}_{\mathrm{zz}}$ represent the cofactor matrices for the x, y and z coordinates, respectively. Matrices $\mathbf{Q}_{\mathrm{xy}}$, $\mathbf{Q}_{\mathrm{xz}}$, $\mathbf{Q}_{\mathrm{yz}}$, $\mathbf{Q}_{\mathrm{yx}}$, $\mathbf{Q}_{\mathrm{zx}}$ and $\mathbf{Q}_{\mathrm{zy}}$ are holding the correlations between the 3D point coordinates. The respective weight matrices are

$$\mathbf{P} = \mathbf{Q}_{\mathrm{LL}}^{-1} = \begin{bmatrix} \mathbf{P}_{\mathrm{xx}} & \mathbf{P}_{\mathrm{xy}} & \mathbf{P}_{\mathrm{xz}} \\ \mathbf{P}_{\mathrm{yx}} & \mathbf{P}_{\mathrm{yy}} & \mathbf{P}_{\mathrm{yz}} \\ \mathbf{P}_{\mathrm{zx}} & \mathbf{P}_{\mathrm{zy}} & \mathbf{P}_{\mathrm{zz}} \end{bmatrix}. \tag{5.163}$$

The nonlinear condition equations (5.103) can be expressed equivalently in vector notation by

$$a(\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}) + b(\mathbf{y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{y}}) + c(\mathbf{z}_{\mathrm{c}} + \mathbf{v}_{\mathrm{z}}) + d\,\mathbf{e} = \mathbf{0}. \tag{5.164}$$

Vectors $\mathbf{x}_{\mathrm{c}}$, $\mathbf{y}_{\mathrm{c}}$ and $\mathbf{z}_{\mathrm{c}}$ include the 3D point coordinates

$$\mathbf{x}_{\mathrm{c}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \ \mathbf{y}_{\mathrm{c}} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \text{ and } \mathbf{z}_{\mathrm{c}} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \tag{5.165}$$

and the residual vectors $\mathbf{v}_\mathrm{x}$, $\mathbf{v}_\mathrm{y}$ and $\mathbf{v}_\mathrm{z}$ list the corresponding residuals

$$\mathbf{v}_\mathrm{x} = \begin{bmatrix} v_{x_1} \\ v_{x_2} \\ \vdots \\ v_{x_n} \end{bmatrix}, \ \mathbf{v}_\mathrm{y} = \begin{bmatrix} v_{y_1} \\ v_{y_2} \\ \vdots \\ v_{y_n} \end{bmatrix} \text{ and } \mathbf{v}_\mathrm{z} = \begin{bmatrix} v_{z_1} \\ v_{z_2} \\ \vdots \\ v_{z_n} \end{bmatrix}. \tag{5.166}$$

$\mathbf{e}$ is a vector of ones with length being equal to the number of measured points. A solution for this weighted least squares problem can be derived by minimizing the objective function

$$\Omega(\mathbf{v}_\mathrm{x}, \mathbf{v}_\mathrm{y}, \mathbf{v}_\mathrm{z}) = \mathbf{v}_\mathrm{x}^\mathrm{T}\mathbf{P}_\mathrm{xx}\mathbf{v}_\mathrm{x} + \mathbf{v}_\mathrm{y}^\mathrm{T}\mathbf{P}_\mathrm{yy}\mathbf{v}_\mathrm{y} + \mathbf{v}_\mathrm{z}^\mathrm{T}\mathbf{P}_\mathrm{zz}\mathbf{v}_\mathrm{z} + 2\ \mathbf{v}_\mathrm{x}^\mathrm{T}\mathbf{P}_\mathrm{xy}\mathbf{v}_\mathrm{y} + 2\ \mathbf{v}_\mathrm{x}^\mathrm{T}\mathbf{P}_\mathrm{xz}\mathbf{v}_\mathrm{z} + 2\ \mathbf{v}_\mathrm{y}^\mathrm{T}\mathbf{P}_\mathrm{yz}\mathbf{v}_\mathrm{z}. \tag{5.167}$$

Iterative least squares solution without linearization

The objective function (5.167) can be combined with the nonlinear condition equations (5.164) to build the Lagrangian

$$\mathrm{K}(a, b, c, d, \mathbf{v}_\mathrm{x}, \mathbf{v}_\mathrm{y}, \mathbf{v}_\mathrm{z}, \mathbf{k}) = \Omega(\mathbf{v}_\mathrm{x}, \mathbf{v}_\mathrm{y}, \mathbf{v}_\mathrm{z}) - 2\mathbf{k}^\mathrm{T}\left[a(\mathbf{x}_\mathrm{c} + \mathbf{v}_\mathrm{x}) + b(\mathbf{y}_\mathrm{c} + \mathbf{v}_\mathrm{y}) + c(\mathbf{z}_\mathrm{c} + \mathbf{v}_\mathrm{z}) + d\ \mathbf{e}\right]. \tag{5.168}$$

$\mathbf{k}$ is the vector of Lagrange multipliers. The resulting normal equations are

$$\frac{\partial \mathrm{K}}{\partial \mathbf{v}_\mathrm{x}^\mathrm{T}} = 2\left(\mathbf{P}_\mathrm{xx}\mathbf{v}_\mathrm{x} + \mathbf{P}_\mathrm{xy}\mathbf{v}_\mathrm{y} + \mathbf{P}_\mathrm{xz}\mathbf{v}_\mathrm{z} - a\mathbf{k}\right) = \mathbf{0}, \tag{5.169}$$

$$\frac{\partial \mathrm{K}}{\partial \mathbf{v}_\mathrm{y}^\mathrm{T}} = 2\left(\mathbf{P}_\mathrm{yy}\mathbf{v}_\mathrm{y} + \mathbf{P}_\mathrm{yx}\mathbf{v}_\mathrm{x} + \mathbf{P}_\mathrm{yz}\mathbf{v}_\mathrm{z} - b\mathbf{k}\right) = \mathbf{0}, \tag{5.170}$$

$$\frac{\partial \mathrm{K}}{\partial \mathbf{v}_\mathrm{z}^\mathrm{T}} = 2\left(\mathbf{P}_\mathrm{yy}\mathbf{v}_\mathrm{y} + \mathbf{P}_\mathrm{zx}\mathbf{v}_\mathrm{x} + \mathbf{P}_\mathrm{zy}\mathbf{v}_\mathrm{y} - c\mathbf{k}\right) = \mathbf{0}, \tag{5.171}$$

$$\frac{\partial \mathrm{K}}{\partial \mathbf{k}^\mathrm{T}} = -2\left[a(\mathbf{x}_\mathrm{c} + \mathbf{v}_\mathrm{x}) + b(\mathbf{y}_\mathrm{c} + \mathbf{v}_\mathrm{y}) + c(\mathbf{z}_\mathrm{c} + \mathbf{v}_\mathrm{z}) + d\ \mathbf{e}\right] = \mathbf{0}, \tag{5.172}$$

$$\frac{\partial \mathrm{K}}{\partial a} = -2\mathbf{k}^\mathrm{T}\left(\mathbf{x}_\mathrm{c} + \mathbf{v}_\mathrm{x}\right) = 0, \tag{5.173}$$

$$\frac{\partial \mathrm{K}}{\partial b} = -2\mathbf{k}^\mathrm{T}\left(\mathbf{y}_\mathrm{c} + \mathbf{v}_\mathrm{y}\right) = 0, \tag{5.174}$$

$$\frac{\partial \mathrm{K}}{\partial c} = -2\mathbf{k}^\mathrm{T}\left(\mathbf{z}_\mathrm{c} + \mathbf{v}_\mathrm{z}\right) = 0, \tag{5.175}$$

$$\frac{\partial \mathrm{K}}{\partial d} = -2\mathbf{k}^\mathrm{T}\mathbf{e} = 0. \tag{5.176}$$

A solution for the unknown plane parameters can be obtained by analyzing the derived normal equations. Expressing equations (5.169)-(5.171) with block matrices

$$\begin{bmatrix} \mathbf{P}_{\mathrm{xx}} & \mathbf{P}_{\mathrm{xy}} & \mathbf{P}_{\mathrm{xz}} \\ \mathbf{P}_{\mathrm{yx}} & \mathbf{P}_{\mathrm{yy}} & \mathbf{P}_{\mathrm{yz}} \\ \mathbf{P}_{\mathrm{zx}} & \mathbf{P}_{\mathrm{zy}} & \mathbf{P}_{\mathrm{zz}} \end{bmatrix} \begin{bmatrix} \mathbf{v}_{\mathrm{x}} \\ \mathbf{v}_{\mathrm{y}} \\ \mathbf{v}_{\mathrm{z}} \end{bmatrix} = \begin{bmatrix} a\mathbf{k} \\ b\mathbf{k} \\ c\mathbf{k} \end{bmatrix}, \tag{5.177}$$

it is possible to derive the residual vectors

$$\begin{bmatrix} \mathbf{v}_{\mathrm{x}} \\ \mathbf{v}_{\mathrm{y}} \\ \mathbf{v}_{\mathrm{z}} \end{bmatrix} = \begin{bmatrix} \mathbf{P}_{\mathrm{xx}} & \mathbf{P}_{\mathrm{xy}} & \mathbf{P}_{\mathrm{xz}} \\ \mathbf{P}_{\mathrm{yx}} & \mathbf{P}_{\mathrm{yy}} & \mathbf{P}_{\mathrm{yz}} \\ \mathbf{P}_{\mathrm{zx}} & \mathbf{P}_{\mathrm{zy}} & \mathbf{P}_{\mathrm{zz}} \end{bmatrix}^{-1} \begin{bmatrix} a\mathbf{k} \\ b\mathbf{k} \\ c\mathbf{k} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_{\mathrm{xx}} & \mathbf{Q}_{\mathrm{xy}} & \mathbf{Q}_{\mathrm{xz}} \\ \mathbf{Q}_{\mathrm{yx}} & \mathbf{Q}_{\mathrm{yy}} & \mathbf{Q}_{\mathrm{yz}} \\ \mathbf{Q}_{\mathrm{zx}} & \mathbf{Q}_{\mathrm{zy}} & \mathbf{Q}_{\mathrm{zz}} \end{bmatrix} \begin{bmatrix} a\mathbf{k} \\ b\mathbf{k} \\ c\mathbf{k} \end{bmatrix}, \tag{5.178}$$

or equivalently

$$\mathbf{v}_{\mathrm{x}} = \left(a\mathbf{Q}_{\mathrm{xx}} + b\mathbf{Q}_{\mathrm{xy}} + c\mathbf{Q}_{\mathrm{xz}}\right)\mathbf{k}, \tag{5.179}$$

$$\mathbf{v}_{\mathrm{y}} = \left(a\mathbf{Q}_{\mathrm{yx}} + b\mathbf{Q}_{\mathrm{yy}} + c\mathbf{Q}_{\mathrm{yz}}\right)\mathbf{k}, \tag{5.180}$$

and

$$\mathbf{v}_{\mathrm{z}} = \left(a\mathbf{Q}_{\mathrm{zx}} + b\mathbf{Q}_{\mathrm{zy}} + c\mathbf{Q}_{\mathrm{zz}}\right)\mathbf{k}. \tag{5.181}$$

Substituting further $\mathbf{v}_{\mathrm{x}}$, $\mathbf{v}_{\mathrm{y}}$ and $\mathbf{v}_{\mathrm{z}}$ into equation (5.172) gives

$$\mathbf{W}\mathbf{k} = -\left(a\mathbf{x}_{\mathrm{c}} + b\mathbf{y}_{\mathrm{c}} + c\mathbf{z}_{\mathrm{c}} + d\mathbf{e}\right). \tag{5.182}$$

The auxiliary matrix

$$\mathbf{W} = {a^0}^2\mathbf{Q}_{\mathrm{xx}} + {b^0}^2\mathbf{Q}_{\mathrm{yy}} + {c^0}^2\mathbf{Q}_{\mathrm{zz}} + a^0 b^0\left(\mathbf{Q}_{\mathrm{xy}} + \mathbf{Q}_{\mathrm{yx}}\right) + a^0 c^0\left(\mathbf{Q}_{\mathrm{xz}} + \mathbf{Q}_{\mathrm{zx}}\right) + b^0 c^0\left(\mathbf{Q}_{\mathrm{yz}} + \mathbf{Q}_{\mathrm{zy}}\right) \tag{5.183}$$

can be constructed after introducing approximate values for parameters a^0, b^0 and c^0. In case of regular cofactor matrices, matrix $\mathbf{W}$ is also regular and invertible. Thus, a solution for the vector of Lagrange multipliers can be obtained by

$$\mathbf{k} = -\mathbf{W}^{-1}\left(a\mathbf{x}_{\mathrm{c}} + b\mathbf{y}_{\mathrm{c}} + c\mathbf{z}_{\mathrm{c}} + d\mathbf{e}\right). \tag{5.184}$$

Introducing vector $\mathbf{k}$ into the normal equation (5.176) yields the expression

$$\mathbf{e}^{\mathrm{T}}\mathbf{k} = 0$$

$$\Rightarrow -\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\left(a\mathbf{x}_{\mathrm{c}} + b\mathbf{y}_{\mathrm{c}} + c\mathbf{z}_{\mathrm{c}} + d\mathbf{e}\right) = 0 \tag{5.185}$$

$$\Rightarrow d = -a\left[\left(\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\right)^{-1}\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{x}_{\mathrm{c}}\right] - b\left[\left(\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\right)^{-1}\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{y}_{\mathrm{c}}\right] - c\left[\left(\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\right)^{-1}\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{z}_{\mathrm{c}}\right].$$

The solution for the unknown plane parameters a, b and c can be computed by analyzing further the normal equations (5.173)-(5.175). Taking into account vector $\mathbf{k}$, as well as the residual vectors $\mathbf{v}_{\mathrm{x}}$, $\mathbf{v}_{\mathrm{y}}$ and $\mathbf{v}_{\mathrm{z}}$, yields the reduced system of equations

$$-\mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\left(a\mathbf{x}_{\mathrm{c}} + b\mathbf{y}_{\mathrm{c}} + c\mathbf{y}_{\mathrm{c}} + d\mathbf{e}\right) + \mathbf{k}^{\mathrm{T}}\left(a\mathbf{Q}_{\mathrm{xx}} + b\mathbf{Q}_{\mathrm{xy}} + c\mathbf{Q}_{\mathrm{xz}}\right)\mathbf{k} = 0, \tag{5.186}$$

$$-\mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\left(a\mathbf{x}_{\mathrm{c}} + b\mathbf{y}_{\mathrm{c}} + c\mathbf{y}_{\mathrm{c}} + d\mathbf{e}\right) + \mathbf{k}^{\mathrm{T}}\left(a\mathbf{Q}_{\mathrm{xx}} + b\mathbf{Q}_{\mathrm{xy}} + c\mathbf{Q}_{\mathrm{xz}}\right)\mathbf{k} = 0, \tag{5.187}$$

and

$$-\mathbf{z}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\left(a\mathbf{x}_{\mathrm{c}} + b\mathbf{y}_{\mathrm{c}} + c\mathbf{y}_{\mathrm{c}} + d\mathbf{e}\right) + \mathbf{k}^{\mathrm{T}}\left(a\mathbf{Q}_{\mathrm{zx}} + b\mathbf{Q}_{\mathrm{zy}} + c\mathbf{Q}_{\mathrm{zz}}\right)\mathbf{k} = 0. \tag{5.188}$$

Introducing parameter d from (5.185) into these three equations results in

$$af_1 + bf_2 + cf_3 = 0, \tag{5.189}$$

$$bf_4 + af_2 + cf_5 = 0 \tag{5.190}$$

$$cf_6 + af_3 + bf_5 = 0, \tag{5.191}$$

with the defined parameters

$$\begin{aligned}
f_1 &= \mathbf{k}^{\mathrm{T}}\mathbf{Q}_{\mathrm{xx}}\mathbf{k} - \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{x}_{\mathrm{c}} + \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\left(\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\right)^{-1}\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{x}_{\mathrm{c}},\\
f_2 &= \mathbf{k}^{\mathrm{T}}\mathbf{Q}_{\mathrm{xy}}\mathbf{k} - \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{y}_{\mathrm{c}} + \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\left(\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\right)^{-1}\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{x}_{\mathrm{c}},\\
f_3 &= \mathbf{k}^{\mathrm{T}}\mathbf{Q}_{\mathrm{xz}}\mathbf{k} - \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{z}_{\mathrm{c}} + \mathbf{z}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\left(\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\right)^{-1}\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{x}_{\mathrm{c}},\\
f_4 &= \mathbf{k}^{\mathrm{T}}\mathbf{Q}_{\mathrm{yy}}\mathbf{k} - \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{y}_{\mathrm{c}} + \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\left(\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\right)^{-1}\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{y}_{\mathrm{c}},\\
f_5 &= \mathbf{k}^{\mathrm{T}}\mathbf{Q}_{\mathrm{yz}}\mathbf{k} - \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{z}_{\mathrm{c}} + \mathbf{z}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\left(\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\right)^{-1}\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{y}_{\mathrm{c}},\\
f_6 &= \mathbf{k}^{\mathrm{T}}\mathbf{Q}_{\mathrm{zz}}\mathbf{k} - \mathbf{z}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{z}_{\mathrm{c}} + \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\left(\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{e}\right)^{-1}\mathbf{e}^{\mathrm{T}}\mathbf{W}^{-1}\mathbf{z}_{\mathrm{c}}.
\end{aligned} \tag{5.192}$$

Equations (5.189) - (5.191) form a system of pseudo-linear equations with three unknown parameters. A restriction between the plane parameters can be chosen at this point. Selecting the restriction of $c = 1$, the remaining unknown plane parameters are

$$\hat{a} = \frac{f_2 f_5 - f_3 f_4}{f_1 f_4 - f_2^2} \tag{5.193}$$

and

$$\hat{b} = -\hat{a}\frac{f_2}{f_4} - \frac{f_5}{f_4}. \tag{5.194}$$

An iterative procedure for estimating the unknown plane parameters, based on the presented approach, can be found in Algorithm 5.

Algorithm 5 Least squares fitting of a plane in 3D with general weights and correlations

Choose approximate values for a^0 and b^0.
Define parameter $c = 1$.
Set threshold ϵ for the break-off condition of the iteration process.
Set parameters $d_a = |\hat{a} - a^0| = \infty$ and $d_b = |\hat{b} - b^0| = \infty$, for entering the iteration process.
while $d_a > \epsilon$ or $d_b > \epsilon$ **do**
    Compute the auxiliary matrix $\mathbf{W}$, and the vector of Lagrange multipliers $\mathbf{k}$.
    Estimate parameter $\hat{d}$.
    Compute the coefficients $f_1, f_2, \ldots, f_5$.
    Estimate parameters $\hat{a}$ and $\hat{b}$.
    Compute parameters $d_a = |\hat{a} - a^0|$ and $d_b = |\hat{b} - b^0|$.
    Update the approximate values with the estimated ones, $a^0 = \hat{a}$ and $b^0 = \hat{b}$.
end while
return $\hat{a}$, $\hat{b}$ and $\hat{d}$, with $c = 1$.

Solution for singular cofactor matrices

An iterative solution for the case of singular cofactor matrices is possible also for this adjustment problem, following the same procedure that has been presented in subsection 5.2.4.2. Therefore, putting together equation (5.182) with the normal equations (5.173)-(5.176), results in the system of nonlinear equations

$$\begin{gathered}
\mathbf{W}\mathbf{k} = -\left(a\mathbf{x}_\mathrm{c} + b\mathbf{y}_\mathrm{c} + c\mathbf{z}_\mathrm{c} + d\mathbf{e}\right), \\
\mathbf{k}^\mathrm{T}\left(\mathbf{x}_\mathrm{c} + \mathbf{v}_\mathrm{x}\right) = 0, \\
\mathbf{k}^\mathrm{T}\left(\mathbf{y}_\mathrm{c} + \mathbf{v}_\mathrm{y}\right) = 0, \\
\mathbf{k}^\mathrm{T}\left(\mathbf{z}_\mathrm{c} + \mathbf{v}_\mathrm{z}\right) = 0, \\
\mathbf{k}^\mathrm{T}\mathbf{e} = 0.
\end{gathered} \tag{5.195}$$

Furthermore, selecting $c = 1$ as a meaningful restriction between the unknown parameters and introducing approximate values for the residual vectors $\mathbf{v}_{\mathrm{x}}^0$ and $\mathbf{v}_{\mathrm{y}}^0$, the last equation system can be expressed by[6]

$$\begin{bmatrix} \mathbf{W} & \mathbf{x}_{\mathrm{c}} & \mathbf{y}_{\mathrm{c}} & \mathbf{e} \\ \left(\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}^0\right)^{\mathrm{T}} & 0 & 0 & 0 \\ \left(\mathbf{y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{y}}^0\right)^{\mathrm{T}} & 0 & 0 & 0 \\ \mathbf{e}^{\mathrm{T}} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{k} \\ a \\ b \\ d \end{bmatrix} = \begin{bmatrix} -\mathbf{z}_{\mathrm{c}} \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{5.196}$$

This can be equivalently written as

$$\mathbf{N} \begin{bmatrix} \mathbf{k} \\ \mathbf{X} \end{bmatrix} = \mathbf{n}, \tag{5.197}$$

after introducing matrices

$$\mathbf{N} = \begin{bmatrix} \mathbf{W} & \mathbf{x}_{\mathrm{c}} & \mathbf{y}_{\mathrm{c}} & \mathbf{e} \\ \left(\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}^0\right)^{\mathrm{T}} & 0 & 0 & 0 \\ \left(\mathbf{y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{y}}^0\right)^{\mathrm{T}} & 0 & 0 & 0 \\ \mathbf{e}^{\mathrm{T}} & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{n} = \begin{bmatrix} -\mathbf{z}_{\mathrm{c}} \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{5.198}$$

and the vector of unknown parameters

$$\mathbf{X} = \begin{bmatrix} a \\ b \\ d \end{bmatrix}. \tag{5.199}$$

A solution of this adjustment problem can be derived by

$$\begin{bmatrix} \hat{\mathbf{k}} \\ \hat{\mathbf{X}} \end{bmatrix} = \mathbf{N}^{-1}\mathbf{n}. \tag{5.200}$$

Solution with a symmetric normal matrix N

Similarly to the procedure of subsection 5.2.4.2, an equivalent solution of the problem using a symmetric matrix $\mathbf{N}$ can be obtained by adding the term $a\mathbf{v}_{\mathrm{x}} + b\mathbf{v}_{\mathrm{y}}$ to both sides of (5.182). Therefore, the equation system (5.196) can be expressed as

$$\begin{bmatrix} \mathbf{W} & \mathbf{A} \\ \mathbf{A}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{k} \\ \mathbf{X} \end{bmatrix} = \begin{bmatrix} \mathbf{w} \\ \mathbf{0} \end{bmatrix}, \tag{5.201}$$

with matrix

$$\mathbf{A} = \left[\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}^0 \;, \mathbf{y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{y}}^0 \;, \; \mathbf{e}\right] \tag{5.202}$$

[6]The fourth equation, $\mathbf{k}^{\mathrm{T}}\left(\mathbf{z}_{\mathrm{c}} + \mathbf{v}_{\mathrm{z}}\right) = 0$, is not taken into account as parameter c is treated as known.

and vector

$$\mathbf{w} = -\mathbf{z}_{\mathrm{c}} + a^0\mathbf{v}_{\mathrm{x}}^0 + b^0\mathbf{v}_{\mathrm{y}}^0. \tag{5.203}$$

A solution of the adjustment problem can be obtained by equation (5.200), after introducing

$$\mathbf{N} = \begin{bmatrix} \mathbf{W} & \mathbf{A} \\ \mathbf{A}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \quad , \quad \mathbf{n} = \begin{bmatrix} \mathbf{w} \\ \mathbf{0} \end{bmatrix}, \tag{5.204}$$

with $\mathbf{N}$ being symmetric.

The inversion of matrix $\mathbf{N}$ depends on the rank deficiency of matrix $\mathbf{W}$, that will be faced when employing a singular cofactor matrix. Similar to the adjustment problem from subsection 5.2.4.2, the presented criterion (5.102) will ensure the existence of a unique solution.

An iterative procedure is presented in Algorithm 6 for estimating the plane parameters when singular cofactor matrices are given.

Algorithm 6 Least squares fitting of a plane in 3D with singular cofactor matrices.

1: Choose approximate values for a^0, b^0, $\mathbf{v}_{\mathrm{x}}^0$, $\mathbf{v}_{\mathrm{y}}^0$ and $\mathbf{v}_{\mathrm{z}}^0$.
2: Define parameter $c = 1$.
3: Set threshold ϵ for the break-off condition of the iteration process.
4: Define parameters $d_a = |\hat{a} - a^0| = \infty$ and $d_b = |\hat{b} - b^0| = \infty$, for entering the iteration process.
5: **while** $d_a > \epsilon$ or $d_b > \epsilon$ **do**
6: Compute matrices $\mathbf{W}$, $\mathbf{A}$ and vector $\mathbf{w}$.
7: Build matrix $\mathbf{N}$ and vector $\mathbf{n}$.
8: Estimate the unknown vector $\begin{bmatrix} \hat{\mathbf{k}} \\ \hat{\mathbf{X}} \end{bmatrix}$.
9: Compute the residual vectors $\mathbf{v}_{\mathrm{x}}$, $\mathbf{v}_{\mathrm{y}}$ and $\mathbf{v}_{\mathrm{z}}$.
10: Compute parameters $d_a = |\hat{a} - a^0|$ and $d_b = |\hat{b} - b^0|$.
11: Update the approximate values with the estimated ones, with $a^0 = \hat{a}$, $b^0 = \hat{b}$, $\mathbf{v}_{\mathrm{x}}^0 = \mathbf{v}_{\mathrm{x}}$, $\mathbf{v}_{\mathrm{y}}^0 = \mathbf{v}_{\mathrm{y}}$ and $\mathbf{v}_{\mathrm{z}}^0 = \mathbf{v}_{\mathrm{z}}$.
12: **end while**
13: **return** $\hat{a}$, $\hat{b}$ and $\hat{d}$, with $c = 1$.

5.4 2D similarity transformation of coordinates

This subsection deals with the least squares solution of the 2D similarity transformation of coordinates, where homologous points in two coordinate systems have been measured with different precisions. This problem has been treated partly in (Teunissen, 1985, p. 148 ff.) as "symmetric Helmert transformation". A direct solution was obtained in terms of an eigenvalue problem, when cofactor matrices with special block diagonal strucure were applied to the coordinates of the points in the two coordinate systems. Additionaly, Marx (2017) derived direct solutions for "point-wise" and proportional weight matrices in the two coordinate systems.

The functional model of this problem has been already presented in subsection 4.5, by the system of equations

$$X_i = \xi_1 \, x_i - \xi_2 \, y_i + t_x,$$
$$Y_i = \xi_2 \, x_i + \xi_1 \, y_i + t_y.$$

Assuming that all point coordinates are measured quantities the necessary residuals are introduced in the functional model, resulting in the system of nonlinear condition equations

$$\begin{aligned} X_i + v_{X_i} - \xi_1 \, (x_i + v_{x_i}) + \xi_2 \, (y_i + v_{y_i}) - t_x &= 0, \\ Y_i + v_{Y_i} - \xi_2 \, (x_i + v_{x_i}) - \xi_1 \, (y_i + v_{y_i}) - t_y &= 0, \end{aligned} \tag{5.205}$$

with $i = 1, \ldots, n$ indicating the number of observed homologous points. This functional model can be expressed equivalently by the overparameterized system

$$\begin{aligned} \gamma \, (X_i + v_{X_i}) + \alpha \, (x_i + v_{x_i}) - \beta \, (y_i + v_{y_i}) + t_x &= 0, \\ \gamma \, (Y_i + v_{Y_i}) + \beta \, (x_i + v_{x_i}) + \alpha \, (y_i + v_{y_i}) + t_y &= 0, \end{aligned} \tag{5.206}$$

with

$$\xi_1 = -\frac{\alpha}{\gamma} \quad \text{and} \quad \xi_2 = -\frac{\beta}{\gamma}. \tag{5.207}$$

Following the procedure of subsection 4.5.1, a constraint can been chosen as

$$\alpha^2 + \beta^2 + \gamma^2 = 1. \tag{5.208}$$

The least squares criterion is utilized for an "optimal" estimate of the unknown transformation parameters, by minimizing the sum of weighted squared residuals

$$\sum_{i=1}^{n} \left(p_{X_i} {v_{X_i}}^2 + p_{Y_i} {v_{Y_i}}^2 + p_{x_i} {v_{x_i}}^2 + p_{y_i} {v_{y_i}}^2 \right) \rightarrow \min, \tag{5.209}$$

where p_{X_i}, p_{Y_i} are the weights influencing the residuals of the coordinates in the target system, respectively p_{x_i}, p_{y_i} the weights in the source system. This adjustment problem is investigated in the next sections for four different weighting cases:

1. Same precision $\sigma_X = \sigma_Y$, for the coordinates in X and Y direction of the points in the target coordinate system. Same precision $\sigma_x = \sigma_y$, for the coordinates in the directions of x and y of the points in the source coordinate system.
2. Individual precision for each pair of homologous points in the two systems: $\sigma_{x_i} = \sigma_{y_i} = \sigma_{X_i} = \sigma_{Y_i} \forall i$.
3. Individual precision for each coordinate.
4. Individual precisions and correlations between the measured 2D coordinates in each system.

5.4.1 Weighting case 1 - Equally weighted observations in each coordinate system

For the first weighting case the measured points in the target system have been observed with the same precision $\sigma_X = \sigma_Y$ and the points in the source system with $\sigma_x = \sigma_y$. Thus, the objective function under minimization becomes

$$\Omega(v_{X_i}, v_{Y_i}, v_{x_i}, v_{y_i}) = \sum_{i=1}^{n} p_{XY} \left({v_{X_i}}^2 + {v_{Y_i}}^2\right) + p_{xy} \left({v_{x_i}}^2 + {v_{y_i}}^2\right) \to \min. \tag{5.210}$$

with

$$\begin{aligned} \frac{1}{{\sigma_X}_i^2} &= \frac{1}{{\sigma_Y}_i^2} = p_{XY} = constant \quad \forall i, \\ \frac{1}{{\sigma_x}_i^2} &= \frac{1}{{\sigma_y}_i^2} = p_{xy} = constant \quad \forall i. \end{aligned} \tag{5.211}$$

From a geometric perspective the postulated weights can be seen as a homogeneous scale in each coordinate system, according to the respective weight. For obtaining a direct solution, the coordinates are transformed linearly by multiplying them with the respective weights

$$X_i^s = X_i \sqrt{p_{XY}}\ ,\ Y_i^s = Y_i \sqrt{p_{XY}} \quad \text{and} \quad x_i^s = x_i \sqrt{p_{xy}}\ ,\ y_i^s = y_i \sqrt{p_{xy}}. \tag{5.212}$$

The scaled coordinates X_i^s, Y_i^s from the target coordinate system and x_i^s, y_i^s from the source coordinate system can be used to derive the requested transformation parameters with equally weighted observations. In this line of thinking the residuals are

$$v_{X_i}^s = v_{X_i} \sqrt{p_{XY}}\ ,\ v_{Y_i}^s = v_{Y_i} \sqrt{p_{XY}} \quad \text{and} \quad v_{x_i}^s = v_{x_i} \sqrt{p_{xy}}\ ,\ v_{y_i}^s = v_{y_i} \sqrt{p_{xy}}. \tag{5.213}$$

Substituting the scaled coordinates and their residuals from equations (5.212) and (5.213) into the condition equations (5.206) yields

$$\begin{aligned} \gamma \frac{1}{\sqrt{p_{XY}}}(X_i^s + v_{X_i}^s) + \alpha \frac{1}{\sqrt{p_{xy}}}(x_i^s + v_{x_i}^s) - \beta \frac{1}{\sqrt{p_{xy}}}(y_i^s + v_{y_i}^s) + t_x = 0, \\ \gamma \frac{1}{\sqrt{p_{XY}}}(Y_i^s + v_{Y_i}^s) + \beta \frac{1}{\sqrt{p_{xy}}}(x_i^s + v_{x_i}^s) + \alpha \frac{1}{\sqrt{p_{xy}}}(y_i^s + v_{y_i}^s) + t_y = 0. \end{aligned} \tag{5.214}$$

Introducing the scaled transformation parameters

$$\begin{aligned} \alpha^s &= \alpha \frac{1}{\sqrt{p_{xy}}}, \\ \beta^s &= \beta \frac{1}{\sqrt{p_{xy}}}, \\ \gamma^s &= \gamma \frac{1}{\sqrt{p_{XY}}}, \\ t_x^s &= t_x, \\ t_y^s &= t_y, \end{aligned} \tag{5.215}$$

into equation (5.214), results in

$$\begin{aligned}\gamma^s \left(X_i^s + v_{X_i}^s\right) + \alpha^s \left(x_i^s + v_{x_i}^s\right) - \beta^s \left(y_i^s + v_{y_i}^s\right) + t_x^s = 0,\\ \gamma^s \left(Y_i^s + v_{Y_i}^s\right) + \beta^s \left(x_i^s + v_{x_i}^s\right) + \alpha^s \left(y_i^s + v_{y_i}^s\right) + t_y^s = 0.\end{aligned} \tag{5.216}$$

A meaningful constraint between the transformation parameters is

$$\alpha^{s2} + \beta^{s2} + \gamma^{s2} = 1 \tag{5.217}$$

and an “optimal” solution is possible by minimizing the sum of scaled squared residuals

$$\sum_{i=1}^{n} {v_{X_i}^s}^2 + {v_{Y_i}^s}^2 + {v_{x_i}^s}^2 + {v_{y_i}^s}^2 \to \min. \tag{5.218}$$

Direct least squares solution

The scaling of the measured 3D coordinates leads to the transformation of the original problem into a problem with equal weights. This means that the sum of squared residuals is equal to the sum of squared Euclidean distances between the points in the target system and the transformed points from the source system:

$$\sum_{i=1}^{n} {v_{X_i}^s}^2 + {v_{Y_i}^s}^2 + {v_{x_i}^s}^2 + {v_{y_i}^s}^2 = \sum_{i=1}^{n} D_i^2 \to \min. \tag{5.219}$$

The squared distances between the homologous points are

$$D_i^2 = (\gamma^s X_i^s + \alpha^s x_i^s - \beta^s y_i^s + t_x^s)^2 + (\gamma^s Y_i^s + \beta^s x_i^s + \alpha^s y_i^s + t_y^s)^2. \tag{5.220}$$

A least squares solution for the transformation parameters can be obtained directly, following the procedure of subsection 4.5.

Computation of the transformation parameters in the original coordinate system

The original trasnformation parameters can be computed by substituting the estimated scaled tranformation parameters into equation (5.215):

$$\begin{aligned}\hat{\alpha} &= \hat{\alpha}^s \sqrt{p_{xy}},\\ \hat{\beta} &= \hat{\beta}^s \sqrt{p_{xy}},\\ \hat{\gamma} &= \hat{\gamma}^s \sqrt{p_{XY}},\\ \hat{t}_x &= \hat{t}_x^s,\\ \hat{t}_y &= \hat{t}_y^s.\end{aligned} \tag{5.221}$$

However, the developed solution is restricted to

$$\alpha^{s2} + \beta^{s2} + \gamma^{s2} = \left(\alpha \frac{1}{\sqrt{p_{xy}}}\right)^2 + \left(\beta \frac{1}{\sqrt{p_{xy}}}\right)^2 + \left(\gamma \frac{1}{\sqrt{p_{XY}}}\right)^2 = 1. \tag{5.222}$$

The least squares solution which restricts the transformation parameters to $\alpha^2+\beta^2+\gamma^2=1$ can be found by

$$\begin{aligned}
\hat{\alpha} &= \frac{\hat{\alpha}^s\sqrt{p_{xy}}}{\sqrt{\left(\hat{\alpha}^s\sqrt{p_{xy}}\right)^2+\left(\hat{\beta}^s\sqrt{p_{xy}}\right)^2+\left(\hat{\gamma}^s\sqrt{p_{XY}}\right)^2}},\\
\hat{\beta} &= \frac{\hat{\beta}^s\sqrt{p_{xy}}}{\sqrt{\left(\hat{\alpha}^s\sqrt{p_{xy}}\right)^2+\left(\hat{\beta}^s\sqrt{p_{xy}}\right)^2+\left(\hat{\gamma}^s\sqrt{p_{XY}}\right)^2}},\\
\hat{\gamma} &= \frac{\hat{\gamma}^s\sqrt{p_{XY}}}{\sqrt{\left(\hat{\alpha}^s\sqrt{p_{xy}}\right)^2+\left(\hat{\beta}^s\sqrt{p_{xy}}\right)^2+\left(\hat{\gamma}^s\sqrt{p_{XY}}\right)^2}},\\
\hat{t}_x &= \frac{\hat{t}_x^s}{\sqrt{\left(\hat{\alpha}^s\sqrt{p_{xy}}\right)^2+\left(\hat{\beta}^s\sqrt{p_{xy}}\right)^2+\left(\hat{\gamma}^s\sqrt{p_{XY}}\right)^2}},\\
\hat{t}_y &= \frac{\hat{t}_y^s}{\sqrt{\left(\hat{\alpha}^s\sqrt{p_{xy}}\right)^2+\left(\hat{\beta}^s\sqrt{p_{xy}}\right)^2+\left(\hat{\gamma}^s\sqrt{p_{XY}}\right)^2}}.
\end{aligned} \tag{5.223}$$

5.4.2 Weighting case 2 - Individual weight for each pair of homologous points in both systems

In the second weighting case for the 2D similarity transformation the homologous points in the source and target system have been measured with individual precisions

$$\sigma_{X_i}=\sigma_{Y_i}=\sigma_{x_i}=\sigma_{y_i}\forall i, \tag{5.224}$$

with the respective weights

$$p_{X_i}=p_{Y_i}=p_{x_i}=p_{y_i}=p_i. \tag{5.225}$$

A least squares solution invlolves the minimization of the objective function (5.209), which taking into account the postulated weights becomes

$$\sum_{i=1}^{n} p_{X_i}{v_{X_i}}^2+p_{Y_i}{v_{Y_i}}^2+p_{x_i}{v_{x_i}}^2+p_{y_i}{v_{y_i}}^2=\sum_{i=1}^{n} p_i\left({v_{X_i}}^2+{v_{Y_i}}^2+{v_{x_i}}^2+{v_{y_i}}^2\right)\rightarrow \min. \tag{5.226}$$

Direct weighted least squares solution

It has been already shown in section 4.5 that the sum of squared residuals of this problem is equal to the sum of squared Euclidean distances

$$\sum_{i=1}^{n} D_i^2=\sum_{i=1}^{n}(\gamma X_i+\alpha x_i-\beta y_i+t_x)^2+(\gamma Y_i+\beta x_i+\alpha y_i+t_y)^2. \tag{5.227}$$

For equally weighted homologous points, it is also true that

$$\sum_{i=1}^{n} p_i \left({v_{X_i}}^2 + {v_{Y_i}}^2 + {v_{x_i}}^2 + {v_{y_i}}^2 \right) = \sum_{i=1}^{n} p_i D_i^2. \tag{5.228}$$

Therefore, the objective function under minimization can be written as

$$\Omega(\alpha, \beta, \gamma, t_x, t_y) = \sum_{i=1}^{n} p_i D_i^2 = \sum_{i=1}^{n} p_i \left[(\gamma X_i + \alpha x_i - \beta y_i + t_x)^2 + (\gamma Y_i + \beta x_i + \alpha y_i + t_y)^2 \right]. \tag{5.229}$$

A meaningful restriction for the unknown parameters is chosen also for this weighting case as

$$\alpha^2 + \beta^2 + \gamma^2 = 1. \tag{5.230}$$

We attempt to obtain the least squares solution for the unknown transformation parameters, that minimizes equation (5.229) under the chosen restriction. Thus, the Lagrangian can be written as

$$\mathrm{K}(\alpha, \beta, \gamma, t_x, t_y, k) = \Omega(\alpha, \beta, \gamma, t_x, t_y) - k(\alpha^2 + \beta^2 + \gamma^2 - 1), \tag{5.231}$$

with the normal equations

$$\begin{aligned} \frac{\partial \mathrm{K}}{\partial \alpha} = 2 \Bigg[\alpha \left(\sum_{i=1}^{n} p_i {x_i}^2 + \sum_{i=1}^{n} p_i {y_i}^2 - k \right) + \gamma \left(\sum_{i=1}^{n} p_i x_i X_i + \sum_{i=1}^{n} p_i y_i Y_i \right) \\ + t_x \sum_{i=1}^{n} p_i x_i + t_y \sum_{i=1}^{n} p_i y_i \Bigg] = 0, \end{aligned} \tag{5.232}$$

$$\begin{aligned} \frac{\partial \mathrm{K}}{\partial \beta} = 2 \Bigg[\beta \left(\sum_{i=1}^{n} p_i {x_i}^2 + \sum_{i=1}^{n} p_i {y_i}^2 - k \right) + \gamma \left(\sum_{i=1}^{n} p_i x_i Y_i - \sum_{i=1}^{n} p_i y_i X_i \right) \\ - t_x \sum_{i=1}^{n} p_i y_i + t_y \sum_{i=1}^{n} p_i x_i \Bigg] = 0, \end{aligned} \tag{5.233}$$

$$\begin{aligned} \frac{\partial \mathrm{K}}{\partial \gamma} = 2 \Bigg[\alpha \left(\sum_{i=1}^{n} p_i x_i X_i + \sum_{i=1}^{n} p_i y_i Y_i \right) + \beta \left(\sum_{i=1}^{n} p_i x_i Y_i - \sum_{i=1}^{n} p_i y_i X_i \right) + \gamma \left(\sum_{i=1}^{n} p_i {x_i}^2 + \sum_{i=1}^{n} p_i {y_i}^2 - k \right) \\ + t_x \sum_{i=1}^{n} p_i X_i + t_y \sum_{i=1}^{n} p_i Y_i \Bigg] = 0, \end{aligned} \tag{5.234}$$

$$\frac{\partial \mathrm{K}}{\partial t_x} = 2 \left(t_x \sum_{i=1}^{n} p_i + \alpha \sum_{i=1}^{n} p_i x_i - \beta \sum_{i=1}^{n} p_i y_i + \gamma \sum_{i=1}^{n} p_i X_i \right) = 0, \tag{5.235}$$

$$\frac{\partial \mathrm{K}}{\partial t_y} = 2 \left(t_y \sum_{i=1}^{n} p_i + \alpha \sum_{i=1}^{n} p_i y_i + \beta \sum_{i=1}^{n} p_i x_i + \gamma \sum_{i=1}^{n} p_i Y_i \right) = 0 \tag{5.236}$$

and

$$\frac{\partial \mathrm{K}}{\partial k} = -\left(\alpha^2 + \beta^2 + \gamma^2 - 1\right) = 0. \tag{5.237}$$

A solution for the trasnlation parameters can be derived by rearanging equations (5.235) and (5.236), which yields

$$t_x = -\alpha \frac{\sum_{i=1}^{n} p_i x_i}{\sum_{i=1}^{n} p_i} + \beta \frac{\sum_{i=1}^{n} p_i y_i}{\sum_{i=1}^{n} p_i} - \gamma \frac{\sum_{i=1}^{n} p_i X_i}{\sum_{i=1}^{n} p_i} \tag{5.238}$$

and

$$t_y = -\alpha \frac{\sum_{i=1}^{n} p_i y_i}{\sum_{i=1}^{n} p_i} - \beta \frac{\sum_{i=1}^{n} p_i x_i}{\sum_{i=1}^{n} p_i} - \gamma \frac{\sum_{i=1}^{n} p_i Y_i}{\sum_{i=1}^{n} p_i}. \tag{5.239}$$

Substituting t_x and t_y into equations (5.232) - (5.234), results in the system of reduced normal equations

$$\begin{gathered}\alpha \left[\sum_{i=1}^{n} p_i {x_i}^2 + \sum_{i=1}^{n} p_i {y_i}^2 - \frac{\left(\sum_{i=1}^{n} p_i x_i\right)^2}{\sum_{i=1}^{n} p_i} - \frac{\left(\sum_{i=1}^{n} p_i y_i\right)^2}{\sum_{i=1}^{n} p_i} - k \right] + \\ \gamma \left[\sum_{i=1}^{n} p_i x_i X_i + \sum_{i=1}^{n} p_i y_i Y_i - \frac{\sum_{i=1}^{n} p_i x_i \sum_{i=1}^{n} p_i X_i}{\sum_{i=1}^{n} p_i} - \frac{\sum_{i=1}^{n} p_i y_i \sum_{i=1}^{n} p_i Y_i}{\sum_{i=1}^{n} p_i} \right] = 0,\end{gathered} \tag{5.240}$$

$$\begin{gathered}\beta \left[\sum_{i=1}^{n} p_i {x_i}^2 + \sum_{i=1}^{n} p_i {y_i}^2 - \frac{\left(\sum_{i=1}^{n} p_i x_i\right)^2}{\sum_{i=1}^{n} p_i} - \frac{\left(\sum_{i=1}^{n} p_i y_i\right)^2}{\sum_{i=1}^{n} p_i} - k \right] + \\ \gamma \left[\sum_{i=1}^{n} p_i x_i Y_i - \sum_{i=1}^{n} p_i y_i X_i + \frac{\sum_{i=1}^{n} p_i y_i \sum_{i=1}^{n} p_i X_i}{\sum_{i=1}^{n} p_i} - \frac{\sum_{i=1}^{n} p_i x_i \sum_{i=1}^{n} p_i Y_i}{\sum_{i=1}^{n} p_i} \right] = 0,\end{gathered} \tag{5.241}$$

and

$$\alpha\left[\sum_{i=1}^{n}p_ix_iX_i+\sum_{i=1}^{n}p_iy_iY_i-\frac{\sum_{i=1}^{n}p_ix_i\sum_{i=1}^{n}p_iX_i}{\sum_{i=1}^{n}p_i}-\frac{\sum_{i=1}^{n}p_iy_i\sum_{i=1}^{n}p_iY_i}{\sum_{i=1}^{n}p_i}\right]+$$
$$\beta\left[\sum_{i=1}^{n}p_ix_iY_i-\sum_{i=1}^{n}p_iy_iX_i+\frac{\sum_{i=1}^{n}p_iy_i\sum_{i=1}^{n}p_iX_i}{\sum_{i=1}^{n}p_i}-\frac{\sum_{i=1}^{n}p_ix_i\sum_{i=1}^{n}p_iY_i}{\sum_{i=1}^{n}p_i}\right]+ \tag{5.242}$$
$$\gamma\left[\sum_{i=1}^{n}p_i{X_i}^2+\sum_{i=1}^{n}p_i{Y_i}^2-\frac{\left(\sum_{i=1}^{n}p_iX_i\right)^2}{\sum_{i=1}^{n}p_i}-\frac{\left(\sum_{i=1}^{n}p_iY_i\right)^2}{\sum_{i=1}^{n}p_i}-k\right]=0.$$

The Lagrange multiplier k can be calculated by solving

$$\begin{vmatrix}(v_1-k) & w_1 & w_2\\ w_1 & (v_1-k) & w_3\\ w_2 & w_3 & (v_2-k)\end{vmatrix}=0, \tag{5.243}$$

which is a cubic characteristic equation with the unknown parameter k.

The respective elements are

$$v_1=\sum_{i=1}^{n}p_i{x_i}^2+\sum_{i=1}^{n}p_i{y_i}^2-\frac{\left(\sum_{i=1}^{n}p_ix_i\right)^2}{\sum_{i=1}^{n}p_i}-\frac{\left(\sum_{i=1}^{n}p_iy_i\right)^2}{\sum_{i=1}^{n}p_i},$$
$$v_2=\sum_{i=1}^{n}p_i{X_i}^2+\sum_{i=1}^{n}p_i{Y_i}^2-\frac{\left(\sum_{i=1}^{n}p_iX_i\right)^2}{\sum_{i=1}^{n}p_i}-\frac{\left(\sum_{i=1}^{n}p_iY_i\right)^2}{\sum_{i=1}^{n}p_i},$$

$$w_1=0, \tag{5.244}$$

$$w_2=\sum_{i=1}^{n}p_ix_iX_i+\sum_{i=1}^{n}p_iy_iY_i-\frac{\sum_{i=1}^{n}p_ix_i\sum_{i=1}^{n}p_iX_i}{\sum_{i=1}^{n}p_i}-\frac{\sum_{i=1}^{n}p_iy_i\sum_{i=1}^{n}p_iY_i}{\sum_{i=1}^{n}p_i},$$
$$w_3=\sum_{i=1}^{n}p_ix_iY_i-\sum_{i=1}^{n}p_iy_iX_i+\frac{\sum_{i=1}^{n}p_iy_i\sum_{i=1}^{n}p_iX_i}{\sum_{i=1}^{n}p_i}-\frac{\sum_{i=1}^{n}p_ix_i\sum_{i=1}^{n}p_iY_i}{\sum_{i=1}^{n}p_i}.$$

The solution for the transformation parameters α, β and γ can be estimated either by substituting parameter k_{min} into the reduced normal equations (5.240) - (5.242) or by transforming them and solve an eigenvalue problem.

5.4.3 Weighting case 3 - Individually weighted coordinates

In the third investigated weighting case for the 2D similarity transformation the coordinates of the points have been measured with different precisions, leading to individual weights for the residuals of each coordinate. The variances of the observed coordinates in the source system can be stored in the variance-covariance matrix

$$\mathbf{\Sigma}_{\mathrm{LL}1} = \begin{bmatrix} \mathbf{\Sigma}_{\mathrm{xx}} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_{\mathrm{yy}} \end{bmatrix} = \mathbf{Q}_{\mathrm{LL}1} \ , \ \text{for} \ \ \sigma_0^2 = 1, \tag{5.245}$$

with the sub-matrices

$$\mathbf{\Sigma}_{\mathrm{xx}} = \begin{bmatrix} \sigma_{x_1}^2 & & & \mathbf{0} \\ & \sigma_{x_2}^2 & & \\ & & \ddots & \\ \mathbf{0} & & & \sigma_{x_n}^2 \end{bmatrix} \ \text{and} \ \ \mathbf{\Sigma}_{\mathrm{yy}} = \begin{bmatrix} \sigma_{y_1}^2 & & & \mathbf{0} \\ & \sigma_{y_2}^2 & & \\ & & \ddots & \\ \mathbf{0} & & & \sigma_{y_n}^2 \end{bmatrix}. \tag{5.246}$$

Selecting the variance of the unit weight being equal to one, the cofactor of the coordinates in the source system is equal to the variance-covariance matrix ($\mathbf{Q}_{\mathrm{LL}1} = \mathbf{\Sigma}_{\mathrm{LL}1}$) and the respective weight matrix is

$$\mathbf{P}_1 = \mathbf{Q}_{\mathrm{LL}1}^{-1} = \begin{bmatrix} \mathbf{P}_{\mathrm{xx}} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{\mathrm{yy}} \end{bmatrix}, \tag{5.247}$$

with the weight sub-matrices being expressed by

$$\mathbf{P}_{\mathrm{xx}} = \begin{bmatrix} p_{x_1} & & & \mathbf{0} \\ & p_{x_2} & & \\ & & \ddots & \\ \mathbf{0} & & & p_{x_n} \end{bmatrix} \ \text{and} \ \ \mathbf{P}_{\mathrm{yy}} = \begin{bmatrix} p_{y_1} & & & \mathbf{0} \\ & p_{y_2} & & \\ & & \ddots & \\ \mathbf{0} & & & p_{y_n} \end{bmatrix}. \tag{5.248}$$

Analogously, the variance-covariance matrix of the observed coordinates in the target system can be written as

$$\mathbf{\Sigma}_{\mathrm{LL}2} = \begin{bmatrix} \mathbf{\Sigma}_{\mathrm{XX}} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_{\mathrm{YY}} \end{bmatrix} = \mathbf{Q}_{\mathrm{LL}2} \ , \ \text{for} \ \ \sigma_0^2 = 1, \tag{5.249}$$

with

$$\mathbf{\Sigma}_{\mathrm{XX}} = \begin{bmatrix} \sigma_{X_1}^2 & & & \mathbf{0} \\ & \sigma_{X_2}^2 & & \\ & & \ddots & \\ \mathbf{0} & & & \sigma_{X_n}^2 \end{bmatrix} \ \text{and} \ \ \mathbf{\Sigma}_{\mathrm{YY}} = \begin{bmatrix} \sigma_{Y_1}^2 & & & \mathbf{0} \\ & \sigma_{Y_2}^2 & & \\ & & \ddots & \\ \mathbf{0} & & & \sigma_{Y_n}^2 \end{bmatrix}. \tag{5.250}$$

Thus, the weight matrix for the coordinates in the target system is

$$\mathbf{P}_2 = \mathbf{Q}_{\mathrm{LL}2}^{-1} = \begin{bmatrix} \mathbf{P}_{\mathrm{XX}} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{\mathrm{YY}} \end{bmatrix}, \tag{5.251}$$

with the weight sub-matrices

$$\mathbf{P}_{\mathrm{XX}} = \begin{bmatrix} p_{X_1} & & & \mathbf{0} \\ & p_{X_2} & & \\ & & \ddots & \\ \mathbf{0} & & & p_{X_n} \end{bmatrix} \text{ and } \mathbf{P}_{\mathrm{YY}} = \begin{bmatrix} p_{Y_1} & & & \mathbf{0} \\ & p_{Y_2} & & \\ & & \ddots & \\ \mathbf{0} & & & p_{Y_n} \end{bmatrix}. \tag{5.252}$$

The nonlinear condition equations (5.205) can be equivalently written in vector notation as

$$\begin{aligned} \mathbf{X}_{\mathrm{c}} + \mathbf{v}_{\mathrm{X}} - \xi_1(\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}) + \xi_2(\mathbf{y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{y}}) - t_x\mathbf{e} = \mathbf{0}, \\ \mathbf{Y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{Y}} - \xi_2(\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}) - \xi_1(\mathbf{y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{y}}) - t_y\mathbf{e} = \mathbf{0}, \end{aligned} \tag{5.253}$$

with vectors $\mathbf{X}_{\mathrm{c}}$, $\mathbf{Y}_{\mathrm{c}}$ listing the coordinates of the points in the target system and $\mathbf{x}_{\mathrm{c}}$, $\mathbf{y}_{\mathrm{c}}$ the coordinates in the source system:

$$\mathbf{X}_{\mathrm{c}} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}, \ \mathbf{Y}_{\mathrm{c}} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \ \mathbf{x}_{\mathrm{c}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \ \mathbf{y}_{\mathrm{c}} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \tag{5.254}$$

while vectors $\mathbf{v}_{\mathrm{X}}$, $\mathbf{v}_{\mathrm{Y}}$, $\mathbf{v}_{\mathrm{x}}$ and $\mathbf{v}_{\mathrm{y}}$ contain the residuals of the corresponding coordinates

$$\mathbf{v}_{\mathrm{X}} = \begin{bmatrix} v_{X_1} \\ v_{X_2} \\ \vdots \\ v_{X_n} \end{bmatrix}, \ \mathbf{v}_{\mathrm{Y}} = \begin{bmatrix} v_{Y_1} \\ v_{Y_2} \\ \vdots \\ v_{Y_n} \end{bmatrix}, \ \mathbf{v}_{\mathrm{x}} = \begin{bmatrix} v_{x_1} \\ v_{x_2} \\ \vdots \\ v_{x_n} \end{bmatrix}, \ \mathbf{v}_{\mathrm{y}} = \begin{bmatrix} v_{y_1} \\ v_{y_2} \\ \vdots \\ v_{y_n} \end{bmatrix}. \tag{5.255}$$

$\mathbf{e}$ is a vector of ones, with length being equal to the number of homologous points. A solution based on the least squares principle can be obtained by minimizing the objective function (5.209), written in matrix notation as

$$\Omega(\mathbf{v}_{\mathrm{X}}, \mathbf{v}_{\mathrm{Y}}, \mathbf{v}_{\mathrm{x}}, \mathbf{v}_{\mathrm{y}}) = \mathbf{v}_{\mathrm{X}}^{\mathrm{T}}\mathbf{P}_{\mathrm{XX}}\mathbf{v}_{\mathrm{X}} + \mathbf{v}_{\mathrm{Y}}^{\mathrm{T}}\mathbf{P}_{\mathrm{YY}}\mathbf{v}_{\mathrm{Y}} + \mathbf{v}_{\mathrm{x}}^{\mathrm{T}}\mathbf{P}_{\mathrm{xx}}\mathbf{v}_{\mathrm{x}} + \mathbf{v}_{\mathrm{y}}^{\mathrm{T}}\mathbf{P}_{\mathrm{yy}}\mathbf{v}_{\mathrm{y}}. \tag{5.256}$$

Iterative least squares solution without linearization

Utilizing the nonlinear functional model of equation (5.253) and the objective function (5.256), a least squares solution can be derived by minimizing the Lagrange function

$$\begin{aligned}
&\mathrm{K}(\xi_1, \xi_2, t_x, t_y, \mathbf{v}_\mathrm{X}, \mathbf{v}_\mathrm{Y}, \mathbf{v}_\mathrm{x}, \mathbf{v}_\mathrm{y}, \mathbf{k}_1, \mathbf{k}_2) = \Omega(\mathbf{v}_\mathrm{X}, \mathbf{v}_\mathrm{Y}, \mathbf{v}_\mathrm{x}, \mathbf{v}_\mathrm{y}) \\
&-2\mathbf{k}_1^\mathrm{T}\left[-(\mathbf{X}_\mathrm{c} + \mathbf{v}_\mathrm{X}) + \xi_1(\mathbf{x}_\mathrm{c} + \mathbf{v}_\mathrm{x}) - \xi_2(\mathbf{y}_\mathrm{c} + \mathbf{v}_\mathrm{y}) + t_x\mathbf{e}\right] \\
&-2\mathbf{k}_2^\mathrm{T}\left[-(\mathbf{Y}_\mathrm{c} + \mathbf{v}_\mathrm{Y}) + \xi_2(\mathbf{x}_\mathrm{c} + \mathbf{v}_\mathrm{x}) + \xi_1(\mathbf{y}_\mathrm{c} + \mathbf{v}_\mathrm{y}) + t_y\mathbf{e}\right],
\end{aligned} \tag{5.257}$$

with $\mathbf{k}_1$ and $\mathbf{k}_2$ denoting the vectors of Lagrange multipliers. A linearization of the problem is avoided also here. Differentiating the Lagrangian with respect to all unknowns and setting the partial derivatives to zero yields

$$\frac{\partial \mathrm{K}}{\partial \mathbf{v}_\mathrm{X}^\mathrm{T}} = 2\left(\mathbf{P}_\mathrm{XX}\mathbf{v}_\mathrm{X} + \mathbf{k}_1\right) = \mathbf{0}, \tag{5.258}$$

$$\frac{\partial \mathrm{K}}{\partial \mathbf{v}_\mathrm{Y}^\mathrm{T}} = 2\left(\mathbf{P}_\mathrm{YY}\mathbf{v}_\mathrm{Y} + \mathbf{k}_2\right) = \mathbf{0}, \tag{5.259}$$

$$\frac{\partial \mathrm{K}}{\partial \mathbf{v}_\mathrm{x}^\mathrm{T}} = 2\left(\mathbf{P}_\mathrm{xx}\mathbf{v}_\mathrm{x} - \xi_1\mathbf{k}_1 - \xi_2\mathbf{k}_2\right) = \mathbf{0}, \tag{5.260}$$

$$\frac{\partial \mathrm{K}}{\partial \mathbf{v}_\mathrm{y}^\mathrm{T}} = 2\left(\mathbf{P}_\mathrm{yy}\mathbf{v}_\mathrm{y} + \xi_2\mathbf{k}_1 - \xi_1\mathbf{k}_2\right) = \mathbf{0}, \tag{5.261}$$

$$\frac{\partial \mathrm{K}}{\partial \mathbf{k}_1^\mathrm{T}} = -2\left[-(\mathbf{X}_\mathrm{c} + \mathbf{v}_\mathrm{X}) + \xi_1(\mathbf{x}_\mathrm{c} + \mathbf{v}_\mathrm{x}) - \xi_2(\mathbf{y}_\mathrm{c} + \mathbf{v}_\mathrm{y}) + t_x\mathbf{e}\right] = \mathbf{0}, \tag{5.262}$$

$$\frac{\partial \mathrm{K}}{\partial \mathbf{k}_2^\mathrm{T}} = -2\left[-(\mathbf{Y}_\mathrm{c} + \mathbf{v}_\mathrm{Y}) + \xi_2(\mathbf{x}_\mathrm{c} + \mathbf{v}_\mathrm{x}) + \xi_1(\mathbf{y}_\mathrm{c} + \mathbf{v}_\mathrm{y}) + t_y\mathbf{e}\right] = \mathbf{0}, \tag{5.263}$$

$$\frac{\partial \mathrm{K}}{\partial \xi_1} = -2\left[\mathbf{k}_1^\mathrm{T}\left(\mathbf{x}_\mathrm{c} + \mathbf{v}_\mathrm{x}\right) + \mathbf{k}_2^\mathrm{T}\left(\mathbf{y}_\mathrm{c} + \mathbf{v}_\mathrm{y}\right)\right] = 0, \tag{5.264}$$

$$\frac{\partial \mathrm{K}}{\partial \xi_2} = -2\left[-\mathbf{k}_1^\mathrm{T}\left(\mathbf{y}_\mathrm{c} + \mathbf{v}_\mathrm{y}\right) + \mathbf{k}_2^\mathrm{T}\left(\mathbf{x}_\mathrm{c} + \mathbf{v}_\mathrm{x}\right)\right] = 0, \tag{5.265}$$

$$\frac{\partial \mathrm{K}}{\partial t_x} = -2\mathbf{k}_1^\mathrm{T}\mathbf{e} = 0 \tag{5.266}$$

and

$$\frac{\partial \mathrm{K}}{\partial t_y} = -2\mathbf{k}_2^{\mathrm{T}}\mathbf{e} = 0. \tag{5.267}$$

Equations (5.258)-(5.267) represent a nonlinear system of $6n+4$ equations. Substituting the residuals from (5.258)-(5.261) into equations (5.262) and (5.263) yields

$$\left(\xi_1^2\mathbf{Q}_{\mathrm{xx}} + \xi_2^2\mathbf{Q}_{\mathrm{yy}} + \mathbf{Q}_{\mathrm{XX}}\right)\mathbf{k}_1 + \left(\xi_1\xi_2\mathbf{Q}_{\mathrm{xx}} - \xi_1\xi_2\mathbf{Q}_{\mathrm{yy}}\right)\mathbf{k}_2 = -\left(\xi_1\mathbf{x}_{\mathrm{c}} - \xi_2\mathbf{y}_{\mathrm{c}} + t_x\mathbf{e} - \mathbf{X}_{\mathrm{c}}\right) \tag{5.268}$$

and

$$\left(\xi_1\xi_2\mathbf{Q}_{\mathrm{xx}} - \xi_1\xi_2\mathbf{Q}_{\mathrm{yy}}\right)\mathbf{k}_1 + \left(\xi_2^2\mathbf{Q}_{\mathrm{xx}} + \xi_1^2\mathbf{Q}_{\mathrm{yy}} + \mathbf{Q}_{\mathrm{YY}}\right)\mathbf{k}_2 = -\left(\xi_2\mathbf{x}_{\mathrm{c}} + \xi_1\mathbf{y}_{\mathrm{c}} + t_y\mathbf{e} - \mathbf{Y}_{\mathrm{c}}\right). \tag{5.269}$$

Introducing approximate values for the unknown transformation parameters only in the left-hand side of the last two equations, it is possible to write

$$\mathbf{W}_1\mathbf{k}_1 + \mathbf{W}_2\mathbf{k}_2 = -\left(\xi_1\mathbf{x}_{\mathrm{c}} - \xi_2\mathbf{y}_{\mathrm{c}} + t_x\mathbf{e} - \mathbf{X}_{\mathrm{c}}\right) \tag{5.270}$$

and

$$\mathbf{W}_2\mathbf{k}_1 + \mathbf{W}_3\mathbf{k}_2 = -\left(\xi_2\mathbf{x}_{\mathrm{c}} + \xi_1\mathbf{y}_{\mathrm{c}} + t_y\mathbf{e} - \mathbf{Y}_{\mathrm{c}}\right), \tag{5.271}$$

with the auxiliary matrices defined as

$$\begin{aligned}\mathbf{W}_1 &= {\xi_1^0}^2\mathbf{Q}_{\mathrm{xx}} + {\xi_2^0}^2\mathbf{Q}_{\mathrm{yy}} + \mathbf{Q}_{\mathrm{XX}},\\ \mathbf{W}_2 &= \xi_1^0\xi_2^0\mathbf{Q}_{\mathrm{xx}} - \xi_1^0\xi_2^0\mathbf{Q}_{\mathrm{yy}},\\ \mathbf{W}_3 &= {\xi_2^0}^2\mathbf{Q}_{\mathrm{xx}} + {\xi_1^0}^2\mathbf{Q}_{\mathrm{yy}} + \mathbf{Q}_{\mathrm{YY}}.\end{aligned} \tag{5.272}$$

Since the introduced cofactor matrices from equations (5.245) and (5.249) are diagonal and therefore regular, then matrices $\mathbf{W}_1$, $\mathbf{W}_2$ and $\mathbf{W}_3$ are also regular and invertible. In this case a solution for the vectors of Lagrange multipliers can be found by

$$\mathbf{k}_1 = \mathbf{W}_5\left(\xi_2\mathbf{x}_{\mathrm{c}} + \xi_1\mathbf{y}_{\mathrm{c}} + t_y\mathbf{e} - \mathbf{Y}_{\mathrm{c}}\right) - \mathbf{W}_4\left(\xi_1\mathbf{x}_{\mathrm{c}} - \xi_2\mathbf{y}_{\mathrm{c}} + t_x\mathbf{e} - \mathbf{X}_{\mathrm{c}}\right) \tag{5.273}$$

and

$$\mathbf{k}_2 = \mathbf{W}_5\left(\xi_1\mathbf{x}_{\mathrm{c}} - \xi_2\mathbf{y}_{\mathrm{c}} + t_x\mathbf{e} - \mathbf{X}_{\mathrm{c}}\right) - \mathbf{W}_6\left(\xi_2\mathbf{x}_{\mathrm{c}} + \xi_1\mathbf{y}_{\mathrm{c}} + t_y\mathbf{e} - \mathbf{Y}_{\mathrm{c}}\right). \tag{5.274}$$

The respective matrices are

$$\begin{aligned}&\mathbf{W}_4 = \left(\mathbf{W}_1 - \mathbf{W}_2\mathbf{W}_3^{-1}\mathbf{W}_2\right)^{-1},\\ &\mathbf{W}_5 = \left(\mathbf{W}_1 - \mathbf{W}_2\mathbf{W}_3^{-1}\mathbf{W}_2\right)^{-1}\mathbf{W}_2\mathbf{W}_3^{-1}, \quad \mathbf{W}_6 = \left(\mathbf{W}_3 - \mathbf{W}_2\mathbf{W}_3^{-1}\mathbf{W}_2\right)^{-1}\end{aligned}. \tag{5.275}$$

Inserting the expressions for $\mathbf{k}_1$ and $\mathbf{k}_2$ into equations (5.266) and (5.267) gives the solution for the translation parameters

$$t_x\left[\mathbf{e}^{\mathrm{T}}\mathbf{W}_4\mathbf{e}\right] = t_y\left[\mathbf{e}^{\mathrm{T}}\mathbf{W}_5\mathbf{e}\right] + \left[\mathbf{e}^{\mathrm{T}}\mathbf{W}_5\left(\xi_2\mathbf{x}_{\mathrm{c}} + \xi_1\mathbf{y}_{\mathrm{c}} - \mathbf{Y}_{\mathrm{c}}\right) - \mathbf{e}^{\mathrm{T}}\mathbf{W}_4\left(\xi_1\mathbf{x}_{\mathrm{c}} - \xi_2\mathbf{y}_{\mathrm{c}} - \mathrm{X}\right)\right], \tag{5.276}$$

$$t_y\left[\mathbf{e}^{\mathrm{T}}\mathbf{W}_6\mathbf{e}\right] = t_x\left[\mathbf{e}^{\mathrm{T}}\mathbf{W}_5\mathbf{e}\right] + \left[\mathbf{e}^{\mathrm{T}}\mathbf{W}_5\left(\xi_1\mathbf{x}_{\mathrm{c}} - \xi_2\mathbf{y}_{\mathrm{c}} - \mathbf{X}_{\mathrm{c}}\right) - \mathbf{e}^{\mathrm{T}}\mathbf{W}_6\left(\xi_2\mathbf{x}_{\mathrm{c}} + \xi_1\mathbf{y}_{\mathrm{c}} - \mathbf{Y}_{\mathrm{c}}\right)\right]. \tag{5.277}$$

Estimates for the unknown transformation parameters ξ_1 and ξ_2 can be computed by substituting the derived residual vectors from equations (5.258) - (5.261), the vectors of Lagrange multipliers $\mathbf{k}_1$, $\mathbf{k}_2$ from (5.273) and (5.274), as well as the translation parameters (5.276) and (5.277) into the normal equations (5.264) and (5.265). This results in the reduced system of equations

$$\xi_1 f_1 + \xi_2 f_2 + f_3 = 0 \tag{5.278}$$

and

$$\xi_2 f_4 + \xi_1 f_2 + f_5 = 0, \tag{5.279}$$

with the respective quantities

$$\begin{aligned}
f_1 &= \mathbf{k}_1^{\mathrm{T}}\mathbf{Q}_{\mathrm{xx}}\mathbf{k}_1 + \mathbf{k}_2^{\mathrm{T}}\mathbf{Q}_{\mathrm{yy}}\mathbf{k}_2 + \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_5\mathbf{y}_{\mathrm{c}} - \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_4\mathbf{x}_{\mathrm{c}} + \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_5\mathbf{x}_{\mathrm{c}} - \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_6\mathbf{y}_{\mathrm{c}},\\
f_2 &= \mathbf{k}_1^{\mathrm{T}}\mathbf{Q}_{\mathrm{xx}}\mathbf{k}_2 - \mathbf{k}_2^{\mathrm{T}}\mathbf{Q}_{\mathrm{yy}}\mathbf{k}_1 + \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_5\mathbf{x}_{\mathrm{c}} + \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_4\mathbf{y}_{\mathrm{c}} - \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_5\mathbf{y}_{\mathrm{c}} - \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_6\mathbf{x}_{\mathrm{c}},\\
f_3 &= \left(\mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_5 - \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_4\right)\left(t_x\mathbf{e} - \mathbf{X}_{\mathrm{c}}\right) + \left(\mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_5 - \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_6\right)\left(t_y\mathbf{e} - \mathbf{Y}_{\mathrm{c}}\right),\\
f_4 &= \mathbf{k}_2^{\mathrm{T}}\mathbf{Q}_{\mathrm{xx}}\mathbf{k}_2 + \mathbf{k}_1^{\mathrm{T}}\mathbf{Q}_{\mathrm{yy}}\mathbf{k}_1 - \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_5\mathbf{y}_{\mathrm{c}} - \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_6\mathbf{x}_{\mathrm{c}} - \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_4\mathbf{y}_{\mathrm{c}} - \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_5\mathbf{x}_{\mathrm{c}},\\
f_5 &= \left(\mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_5 + \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_4\right)\left(t_x\mathbf{e} - \mathbf{X}_{\mathrm{c}}\right) - \left(\mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_6 + \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_5\right)\left(t_y\mathbf{e} - \mathbf{Y}_{\mathrm{c}}\right).
\end{aligned} \tag{5.280}$$

Last but not least, solving (5.279) for ξ_2 and introducing it in (5.278) leads to the least squares estimate for

$$\hat{\xi}_1 = \left(f_1 f_4 - f_2^2\right)^{-1}\left(f_2 f_5 - f_3 f_4\right) \tag{5.281}$$

and

$$\hat{\xi}_2 = -\xi_1\frac{f_2}{f_4} - \frac{f_5}{f_4}. \tag{5.282}$$

An iterative solution is possible by choosing meaningful approximate values for the unknown transformation parameters. Thus, the last two equations become pseudo-linear with the functions $f_1, f_2, \ldots, f_5$ being approximated. An iterative procedure can be found in Algorithm 7 for the weighted least squares solution of the 2D similarity transformation of coordinates.

Algorithm 7 Least squares 2D similarity transformation of coordinates with general weights

Choose approximate values for ξ_1^0 and ξ_2^0.
Set threshold ϵ for the break-off condition of the iteration process.
Set parameters $d_{\xi_1} = |\hat{\xi}_1 - \xi_1^0| = \infty$ and $d_{\xi_2} = |\hat{\xi}_2 - \xi_2^0| = \infty$, for entering the iteration process.
while $d_{\xi_1} = |\hat{\xi}_1 - \xi_1^0| > \epsilon$ or $d_{\xi_2} = |\hat{\xi}_2 - \xi_2^0| > \epsilon$ **do**
    Compute the auxiliary matrices $\mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_6$.
    Estimate the translation parameters $\hat{t}_x$ and $\hat{t}_y$.
    Compute the vectors of Lagrange multipliers $\mathbf{k}_1$ and $\mathbf{k}_2$.
    Compute the coefficients $f_1, f_2, \ldots, f_5$.
    Estimate parameters $\hat{\xi}_1$ and $\hat{\xi}_2$.
    Compute parameter $d_{\xi_1} = |\hat{\xi}_1 - \xi_1^0|$ and $d_{\xi_2} = |\hat{\xi}_2 - \xi_2^0|$.
    Update the approximate values with the estimated ones ($\xi_1^0 = \hat{\xi}_1$, $\xi_2^0 = \hat{\xi}_2$) .
end while
return $\hat{\xi}_1$, $\hat{\xi}_1$, $\hat{t}_x$ and $\hat{t}_y$.

5.4.4 Weighting case 4 - Individually weighted and correlated coordinates in each coordinate system

In the last investigated weighting case for the 2D similarity transformation, correlations are introduced between the measured coordinates of the points in each coordinate system. Therefore, two cofactor matrices are given. The first is related to the coordinates in the source system

$$\mathbf{Q}_{\mathrm{LL}1} = \begin{bmatrix} \mathbf{Q}_{\mathrm{xx}} & \mathbf{Q}_{\mathrm{xy}} \\ \mathbf{Q}_{\mathrm{yx}} & \mathbf{Q}_{\mathrm{yy}} \end{bmatrix}, \text{ with } \mathbf{Q}_{\mathrm{xy}} = \mathbf{Q}_{\mathrm{yx}}^{\mathrm{T}}, \tag{5.283}$$

while the second concerns the coordinates in the target system

$$\mathbf{Q}_{\mathrm{LL}2} = \begin{bmatrix} \mathbf{Q}_{\mathrm{XX}} & \mathbf{Q}_{\mathrm{XY}} \\ \mathbf{Q}_{\mathrm{YX}} & \mathbf{Q}_{\mathrm{YY}} \end{bmatrix}, \text{ with } \mathbf{Q}_{\mathrm{XY}} = \mathbf{Q}_{\mathrm{YX}}^{\mathrm{T}}. \tag{5.284}$$

The respective weights of this problem can be computed by

$$\mathbf{P}_1 = \mathbf{Q}_{\mathrm{LL}1}^{-1} = \begin{bmatrix} \mathbf{P}_{\mathrm{xx}} & \mathbf{P}_{\mathrm{xy}} \\ \mathbf{P}_{\mathrm{yx}} & \mathbf{P}_{\mathrm{yy}} \end{bmatrix} \tag{5.285}$$

and

$$\mathbf{P}_2 = \mathbf{Q}_{\mathrm{LL}2}^{-1} = \begin{bmatrix} \mathbf{P}_{\mathrm{XX}} & \mathbf{P}_{\mathrm{XY}} \\ \mathbf{P}_{\mathrm{YX}} & \mathbf{P}_{\mathrm{YY}} \end{bmatrix}. \tag{5.286}$$

Taking into account the stochastic model described above, the objective function (5.256) can be extended to

$$\Omega(\mathbf{v}_{\mathrm{X}}, \mathbf{v}_{\mathrm{Y}}, \mathbf{v}_{\mathrm{x}}, \mathbf{v}_{\mathrm{y}}) = \mathbf{v}_{\mathrm{X}}^{\mathrm{T}}\mathbf{P}_{\mathrm{XX}}\mathbf{v}_{\mathrm{X}} + \mathbf{v}_{\mathrm{Y}}^{\mathrm{T}}\mathbf{P}_{\mathrm{YY}}\mathbf{v}_{\mathrm{Y}} + \mathbf{v}_{\mathrm{x}}^{\mathrm{T}}\mathbf{P}_{\mathrm{xx}}\mathbf{v}_{\mathrm{x}} + \mathbf{v}_{\mathrm{y}}^{\mathrm{T}}\mathbf{P}_{\mathrm{yy}}\mathbf{v}_{\mathrm{y}} + 2\ \mathbf{v}_{\mathrm{X}}^{\mathrm{T}}\mathbf{P}_{\mathrm{XY}}\mathbf{v}_{\mathrm{Y}} + 2\ \mathbf{v}_{\mathrm{x}}^{\mathrm{T}}\mathbf{P}_{\mathrm{xy}}\mathbf{v}_{\mathrm{y}}. \quad (5.287)$$

Iterative least squares solution without linearization

The developed objective function can be combined with the nonlinear condition equations (5.253) to form the Lagrange function

$$\begin{aligned}
&\mathrm{K}(\xi_1, \xi_2, t_x, t_y, \mathbf{v}_{\mathrm{X}}, \mathbf{v}_{\mathrm{Y}}, \mathbf{v}_{\mathrm{x}}, \mathbf{v}_{\mathrm{y}}, \mathbf{k}_1, \mathbf{k}_2) = \Omega(\mathbf{v}_{\mathrm{X}}, \mathbf{v}_{\mathrm{Y}}, \mathbf{v}_{\mathrm{x}}, \mathbf{v}_{\mathrm{y}}) \\
&-2\mathbf{k}_1^{\mathrm{T}}\left[-(\mathbf{X}_{\mathrm{c}} + \mathbf{v}_{\mathrm{X}}) + \xi_1(\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}) - \xi_2(\mathbf{y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{y}}) + t_x\mathbf{e}\right] \qquad (5.288) \\
&-2\mathbf{k}_2^{\mathrm{T}}\left[-(\mathbf{Y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{Y}}) + \xi_2(\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}) + \xi_1(\mathbf{y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{y}}) + t_y\mathbf{e}\right],
\end{aligned}$$

with $\mathbf{k}_1$ and $\mathbf{k}_2$ denoting vectors of Lagrange multipliers. Differentiating the Lagrangian with respect to all unknowns and setting the result to zero, yields the system of normal equations

$$\frac{\partial \mathrm{K}}{\partial \mathbf{v}_{\mathrm{X}}^{\mathrm{T}}} = 2\left(\mathbf{P}_{\mathrm{XX}}\mathbf{v}_{\mathrm{X}} + \mathbf{P}_{\mathrm{XY}}\mathbf{v}_{\mathrm{Y}} + \mathbf{k}_1\right) = \mathbf{0}, \qquad (5.289)$$

$$\frac{\partial \mathrm{K}}{\partial \mathbf{v}_{\mathrm{Y}}^{\mathrm{T}}} = 2\left(\mathbf{P}_{\mathrm{YY}}\mathbf{v}_{\mathrm{Y}} + \mathbf{P}_{\mathrm{YX}}\mathbf{v}_{\mathrm{X}} + \mathbf{k}_2\right) = \mathbf{0}, \qquad (5.290)$$

$$\frac{\partial \mathrm{K}}{\partial \mathbf{v}_{\mathrm{x}}^{\mathrm{T}}} = 2\left(\mathbf{P}_{\mathrm{xx}}\mathbf{v}_{\mathrm{x}} + \mathbf{P}_{\mathrm{xy}}\mathbf{v}_{\mathrm{y}} - \xi_1\mathbf{k}_1 - \xi_2\mathbf{k}_2\right) = \mathbf{0}, \qquad (5.291)$$

$$\frac{\partial \mathrm{K}}{\partial \mathbf{v}_{\mathrm{y}}^{\mathrm{T}}} = 2\left(\mathbf{P}_{\mathrm{yy}}\mathbf{v}_{\mathrm{y}} + \mathbf{P}_{\mathrm{yx}}\mathbf{v}_{\mathrm{x}} + \xi_2\mathbf{k}_1 - \xi_1\mathbf{k}_2\right) = \mathbf{0}, \qquad (5.292)$$

$$\frac{\partial \mathrm{K}}{\partial \mathbf{k}_1^{\mathrm{T}}} = -2\left[-(\mathbf{X}_{\mathrm{c}} + \mathbf{v}_{\mathrm{X}}) + \xi_1(\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}) - \xi_2(\mathbf{y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{y}}) + t_x\mathbf{e}\right] = \mathbf{0}, \qquad (5.293)$$

$$\frac{\partial \mathrm{K}}{\partial \mathbf{k}_2^{\mathrm{T}}} = -2\left[-(\mathbf{Y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{Y}}) + \xi_2(\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}) + \xi_1(\mathbf{y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{y}}) + t_y\mathbf{e}\right] = \mathbf{0}, \qquad (5.294)$$

$$\frac{\partial \mathrm{K}}{\partial \xi_1} = -2\left[\mathbf{k}_1^{\mathrm{T}}(\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}) + \mathbf{k}_2^{\mathrm{T}}(\mathbf{y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{y}})\right] = 0, \qquad (5.295)$$

$$\frac{\partial \mathrm{K}}{\partial \xi_2} = -2\left[-\mathbf{k}_1^{\mathrm{T}}(\mathbf{y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{y}}) + \mathbf{k}_2^{\mathrm{T}}(\mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}})\right] = 0, \qquad (5.296)$$

$$\frac{\partial \mathrm{K}}{\partial t_x} = -2\mathbf{k}_1^{\mathrm{T}}\mathbf{e} = 0 \tag{5.297}$$

and

$$\frac{\partial \mathrm{K}}{\partial t_y} = -2\mathbf{k}_2^{\mathrm{T}}\mathbf{e} = 0. \tag{5.298}$$

The first two of the developed normal equations can be expressed using block matrices:

$$\begin{bmatrix} \mathbf{P}_{\mathrm{XX}} & \mathbf{P}_{\mathrm{XY}} \\ \mathbf{P}_{\mathrm{YX}} & \mathbf{P}_{\mathrm{YY}} \end{bmatrix} \begin{bmatrix} \mathbf{v}_{\mathrm{X}} \\ \mathbf{v}_{\mathrm{Y}} \end{bmatrix} = - \begin{bmatrix} \mathbf{k}_1 \\ \mathbf{k}_2 \end{bmatrix}. \tag{5.299}$$

Thus, a solution for the residual vectors in the target system can be computed by

$$\begin{bmatrix} \mathbf{v}_{\mathrm{X}} \\ \mathbf{v}_{\mathrm{Y}} \end{bmatrix} = - \begin{bmatrix} \mathbf{P}_{\mathrm{XX}} & \mathbf{P}_{\mathrm{XY}} \\ \mathbf{P}_{\mathrm{YX}} & \mathbf{P}_{\mathrm{YY}} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{k}_1 \\ \mathbf{k}_2 \end{bmatrix} = - \begin{bmatrix} \mathbf{Q}_{\mathrm{XX}} & \mathbf{Q}_{\mathrm{XY}} \\ \mathbf{Q}_{\mathrm{YX}} & \mathbf{Q}_{\mathrm{YY}} \end{bmatrix} \begin{bmatrix} \mathbf{k}_1 \\ \mathbf{k}_2 \end{bmatrix}, \tag{5.300}$$

or equivalently by

$$\mathbf{v}_{\mathrm{X}} = -\mathbf{Q}_{\mathrm{XX}}\mathbf{k}_1 - \mathbf{Q}_{\mathrm{XY}}\mathbf{k}_2 \tag{5.301}$$

and

$$\mathbf{v}_{\mathrm{Y}} = -\mathbf{Q}_{\mathrm{YX}}\mathbf{k}_1 - \mathbf{Q}_{\mathrm{YY}}\mathbf{k}_2. \tag{5.302}$$

In the same manner, explicit expressions for the residual vectors of the coordinates in the source system can be obtained by utilizing equations (5.291) and (5.292), which yields

$$\mathbf{v}_{\mathrm{x}} = \xi_1 \left(\mathbf{Q}_{\mathrm{xx}}\mathbf{k}_1 + \mathbf{Q}_{\mathrm{xy}}\mathbf{k}_2\right) + \xi_2 \left(-\mathbf{Q}_{\mathrm{xy}}\mathbf{k}_1 + \mathbf{Q}_{\mathrm{xx}}\mathbf{k}_2\right) \tag{5.303}$$

and

$$\mathbf{v}_{\mathrm{y}} = \xi_1 \left(\mathbf{Q}_{\mathrm{yx}}\mathbf{k}_1 + \mathbf{Q}_{\mathrm{yy}}\mathbf{k}_2\right) + \xi_2 \left(-\mathbf{Q}_{\mathrm{yy}}\mathbf{k}_1 + \mathbf{Q}_{\mathrm{yx}}\mathbf{k}_2\right). \tag{5.304}$$

For a solution of $\mathbf{k}_1$ and $\mathbf{k}_2$, all residual vectors are introduced into equations (5.293) and (5.294), resulting in

$$\left(\xi_1^2\mathbf{Q}_{\mathrm{xx}} - \xi_1\xi_2\mathbf{Q}_{\mathrm{xy}} - \xi_1\xi_2\mathbf{Q}_{\mathrm{yx}} + \xi_2^2\mathbf{Q}_{\mathrm{yy}} + \mathbf{Q}_{\mathrm{XX}}\right)\mathbf{k}_1 + \left(\xi_1\xi_2\mathbf{Q}_{\mathrm{xx}} + \xi_1^2\mathbf{Q}_{\mathrm{xy}} - \xi_2^2\mathbf{Q}_{\mathrm{yx}} - \xi_1\xi_2\mathbf{Q}_{\mathrm{yy}} + \mathbf{Q}_{\mathrm{XY}}\right)\mathbf{k}_2 = \\ -\left(\xi_1\mathbf{x}_{\mathrm{c}} - \xi_2\mathbf{y}_{\mathrm{c}} + t_x\mathbf{e} - \mathbf{X}_{\mathrm{c}}\right) \tag{5.305}$$

and

$$\left(\xi_1\xi_2\mathbf{Q}_{\mathrm{xx}} + \xi_1^2\mathbf{Q}_{\mathrm{yx}} - \xi_2^2\mathbf{Q}_{\mathrm{xy}} - \xi_1\xi_2\mathbf{Q}_{\mathrm{yy}} + \mathbf{Q}_{\mathrm{YX}}\right)\mathbf{k}_1 + \left(\xi_2^2\mathbf{Q}_{\mathrm{xx}} + \xi_1\xi_2\mathbf{Q}_{\mathrm{xy}} + \xi_1\xi_2\mathbf{Q}_{\mathrm{yx}} + \xi_1^2\mathbf{Q}_{\mathrm{yy}} + \mathbf{Q}_{\mathrm{YY}}\right)\mathbf{k}_2 = \\ -\left(\xi_2\mathbf{x}_{\mathrm{c}} + \xi_1\mathbf{y}_{\mathrm{c}} + t_y\mathbf{e} - \mathbf{Y}_{\mathrm{c}}\right). \tag{5.306}$$

Introducing approximate values for the unknown transformation parameters only in the left-hand side of the last two equations, it is possible to rewrite them as

$$\mathbf{W}_1\mathbf{k}_1 + \mathbf{W}_2\mathbf{k}_2 = -\left(\xi_1\mathbf{x}_{\mathrm{c}} - \xi_2\mathbf{y}_{\mathrm{c}} + t_x\mathbf{e} - \mathbf{X}_{\mathrm{c}}\right) \tag{5.307}$$

and

$$\mathbf{W}_3\mathbf{k}_2 + \mathbf{W}_4\mathbf{k}_1 = -\left(\xi_2\mathbf{x}_{\mathrm{c}} + \xi_1\mathbf{y}_{\mathrm{c}} + t_y\mathbf{e} - \mathbf{Y}_{\mathrm{c}}\right), \tag{5.308}$$

with the auxiliary matrices

$$\begin{aligned}
\mathbf{W}_1 &= {\xi_1^0}^2\mathbf{Q}_{\mathrm{xx}} - \xi_1^0\xi_2^0\mathbf{Q}_{\mathrm{xy}} - \xi_1^0\xi_2^0\mathbf{Q}_{\mathrm{yx}} + {\xi_2^0}^2\mathbf{Q}_{\mathrm{yy}} + \mathbf{Q}_{\mathrm{XX}},\\
\mathbf{W}_2 &= \xi_1^0\xi_2^0\mathbf{Q}_{\mathrm{xx}} + {\xi_1^0}^2\mathbf{Q}_{\mathrm{xy}} - {\xi_2^0}^2\mathbf{Q}_{\mathrm{yx}} - \xi_1^0\xi_2^0\mathbf{Q}_{\mathrm{yy}} + \mathbf{Q}_{\mathrm{XY}},\\
\mathbf{W}_3 &= {\xi_2^0}^2\mathbf{Q}_{\mathrm{xx}} + \xi_1^0\xi_2^0\mathbf{Q}_{\mathrm{xy}} + \xi_1^0\xi_2^0\mathbf{Q}_{\mathrm{yx}} + {\xi_1^0}^2\mathbf{Q}_{\mathrm{yy}} + \mathbf{Q}_{\mathrm{YY}},\\
\mathbf{W}_4 &= \xi_1^0\xi_2^0\mathbf{Q}_{\mathrm{xx}} + {\xi_1^0}^2\mathbf{Q}_{\mathrm{yx}} - {\xi_2^0}^2\mathbf{Q}_{\mathrm{xy}} - \xi_1^0\xi_2^0\mathbf{Q}_{\mathrm{yy}} + \mathbf{Q}_{\mathrm{YX}}.
\end{aligned} \tag{5.309}$$

If the cofactor matrices are regular, then matrices $\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3$ and $\mathbf{W}_4$ are also regular and invertible. Consequently, a solution for the vectors of Lagrange multipliers can be found by

$$\mathbf{k}_1 = \mathbf{W}_5\left(\xi_2\mathbf{x}_{\mathrm{c}} + \xi_1\mathbf{y}_{\mathrm{c}} + t_y\mathbf{e} - \mathbf{Y}_{\mathrm{c}}\right) - \mathbf{W}_6\left(\xi_1\mathbf{x}_{\mathrm{c}} - \xi_2\mathbf{y}_{\mathrm{c}} + t_x\mathbf{e} - \mathbf{X}_{\mathrm{c}}\right) \tag{5.310}$$

and

$$\mathbf{k}_2 = \mathbf{W}_7\left(\xi_1\mathbf{x}_{\mathrm{c}} - \xi_2\mathbf{y}_{\mathrm{c}} + t_x\mathbf{e} - \mathbf{X}_{\mathrm{c}}\right) - \mathbf{W}_8\left(\xi_2\mathbf{x}_{\mathrm{c}} + \xi_1\mathbf{y}_{\mathrm{c}} + t_y\mathbf{e} - \mathbf{Y}_{\mathrm{c}}\right), \tag{5.311}$$

after introducing the matrices

$$\begin{aligned}
\mathbf{W}_5 &= \left(\mathbf{W}_1 - \mathbf{W}_2\mathbf{W}_3^{-1}\mathbf{W}_4\right)^{-1}\mathbf{W}_2\mathbf{W}_3^{-1},\\
\mathbf{W}_6 &= \left(\mathbf{W}_1 - \mathbf{W}_2\mathbf{W}_3^{-1}\mathbf{W}_4\right)^{-1},\\
\mathbf{W}_7 &= \left(\mathbf{W}_3 - \mathbf{W}_4\mathbf{W}_1^{-1}\mathbf{W}_2\right)^{-1}\mathbf{W}_4\mathbf{W}_1^{-1},\\
\mathbf{W}_8 &= \left(\mathbf{W}_3 - \mathbf{W}_4\mathbf{W}_1^{-1}\mathbf{W}_2\right)^{-1}.
\end{aligned} \tag{5.312}$$

Substituting the derived vectors of Lagrange multipliers $\mathbf{k}_1$ and $\mathbf{k}_2$ into the normal equations (5.297) and (5.298) yields the translation vectors

$$t_x\left[\mathbf{e}^{\mathrm{T}}\mathbf{W}_6\mathbf{e}\right] = t_y\left[\mathbf{e}^{\mathrm{T}}\mathbf{W}_5\mathbf{e}\right] + \left[\mathbf{e}^{\mathrm{T}}\mathbf{W}_5\left(\xi_2\mathrm{x}+\xi_1\mathrm{y}-\mathrm{Y}\right) - \mathbf{e}^{\mathrm{T}}\mathbf{W}_6\left(\xi_1\mathrm{x}-\xi_2\mathrm{y}-\mathrm{X}\right)\right] \tag{5.313}$$

and

$$t_y\left[\mathbf{e}^{\mathrm{T}}\mathbf{W}_8\mathbf{e}\right] = t_x\left[\mathbf{e}^{\mathrm{T}}\mathbf{W}_7\mathbf{e}\right] + \left[\mathbf{e}^{\mathrm{T}}\mathbf{W}_7\left(\xi_1\mathrm{x}-\xi_2\mathrm{y}-\mathrm{X}\right) - \mathbf{e}^{\mathrm{T}}\mathbf{W}_8\left(\xi_2\mathrm{x}+\xi_1\mathrm{y}-\mathrm{Y}\right)\right]. \tag{5.314}$$

Furthermore, estimates for the transformation parameters ξ_1 and ξ_2 can be computed by substituting the residual vectors, the vectors of Lagrange multipliers, as well as the translation parameters into equations (5.295) and (5.296). This results in the reduced normal equations

$$\xi_1 f_1 + \xi_2 f_2 + f_3 = 0 \tag{5.315}$$

and

$$\xi_2 f_5 + \xi_1 f_4 + f_6 = 0, \tag{5.316}$$

with the respective quantities

$$\begin{aligned}
f_1 &= \mathbf{k}_1^{\mathrm{T}}\mathbf{Q}_{\mathrm{xx}}\mathbf{k}_1 + \mathbf{k}_1^{\mathrm{T}}\mathbf{Q}_{\mathrm{xy}}\mathbf{k}_2 + \mathbf{k}_2^{\mathrm{T}}\mathbf{Q}_{\mathrm{yx}}\mathbf{k}_1 + \mathbf{k}_2^{\mathrm{T}}\mathbf{Q}_{\mathrm{yy}}\mathbf{k}_2 + \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_5\mathbf{y}_{\mathrm{c}} - \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_6\mathbf{x}_{\mathrm{c}} + \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_7\mathbf{x}_{\mathrm{c}} - \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_8\mathbf{y}_{\mathrm{c}},\\
f_2 &= \mathbf{k}_1^{\mathrm{T}}\mathbf{Q}_{\mathrm{xx}}\mathbf{k}_2 - \mathbf{k}_1^{\mathrm{T}}\mathbf{Q}_{\mathrm{xy}}\mathbf{k}_1 + \mathbf{k}_2^{\mathrm{T}}\mathbf{Q}_{\mathrm{yx}}\mathbf{k}_2 - \mathbf{k}_2^{\mathrm{T}}\mathbf{Q}_{\mathrm{yy}}\mathbf{k}_1 + \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_5\mathbf{x}_{\mathrm{c}} + \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_6\mathbf{y}_{\mathrm{c}} - \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_7\mathbf{y}_{\mathrm{c}} - \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_8\mathbf{x}_{\mathrm{c}},\\
f_3 &= \left(\mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_7 - \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_6\right)\left(t_x\mathbf{e} - \mathbf{X}_{\mathrm{c}}\right) + \left(\mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_5 - \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_8\right)\left(t_y\mathbf{e} - \mathbf{Y}_{\mathrm{c}}\right),\\
f_4 &= \mathbf{k}_1^{\mathrm{T}}\mathbf{Q}_{\mathrm{yx}}\mathbf{k}_1 + \mathbf{k}_1^{\mathrm{T}}\mathbf{Q}_{\mathrm{yy}}\mathbf{k}_2 - \mathbf{k}_2^{\mathrm{T}}\mathbf{Q}_{\mathrm{xx}}\mathbf{k}_1 - \mathbf{k}_2^{\mathrm{T}}\mathbf{Q}_{\mathrm{xy}}\mathbf{k}_2 + \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_5\mathbf{y}_{\mathrm{c}} - \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_6\mathbf{x}_{\mathrm{c}} - \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_7\mathbf{x}_{\mathrm{c}} + \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_8\mathbf{y}_{\mathrm{c}},\\
f_5 &= \mathbf{k}_1^{\mathrm{T}}\mathbf{Q}_{\mathrm{yx}}\mathbf{k}_2 + \mathbf{k}_1^{\mathrm{T}}\mathbf{Q}_{\mathrm{yy}}\mathbf{k}_2 - \mathbf{k}_2^{\mathrm{T}}\mathbf{Q}_{\mathrm{xx}}\mathbf{k}_2 + \mathbf{k}_2^{\mathrm{T}}\mathbf{Q}_{\mathrm{xy}}\mathbf{k}_1 + \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_5\mathbf{x}_{\mathrm{c}} + \mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_6\mathbf{y}_{\mathrm{c}} + \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_7\mathbf{y}_{\mathrm{c}} + \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_8\mathbf{x}_{\mathrm{c}},\\
f_6 &= \left(-\mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_6 - \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_7\right)\left(t_x\mathbf{e} - \mathbf{X}_{\mathrm{c}}\right) + \left(\mathbf{y}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_5 + \mathbf{x}_{\mathrm{c}}^{\mathrm{T}}\mathbf{W}_8\right)\left(t_y\mathbf{e} - \mathbf{Y}_{\mathrm{c}}\right).
\end{aligned} \tag{5.317}$$

Finally, solving equation (5.316) for ξ_2 and introducing it in (5.315) leads to the least squares estimate for

$$\hat{\xi}_1 = \left(f_1 f_5 - f_2 f_4\right)^{-1}\left(f_2 f_6 - f_3 f_5\right) \tag{5.318}$$

and

$$\hat{\xi}_2 = -\hat{\xi}_1\frac{f_4}{f_5} - \frac{f_6}{f_5}. \tag{5.319}$$

Choosing appropriate approximate values for the unknown transformation parameters, an iterative process could give the solution to this adjustment problem. Equations (5.318) and (5.319) are pseudo-linear with the auxiliary functions $f_1, f_2, \ldots, f_6$ being approximated in each iteration step. Algorithm 8 includes the presented iterative least squares solution for the 2D similarity transformation of coordinates with correlated observations.

Algorithm 8 Least squares 2D similarity transformation of coordinates with correlated observations

```
Choose approximate values for ξ1^0 and ξ2^0.
Set threshold ϵ for the break-off condition of the iteration process.
Set parameters d_ξ1 = |ξ̂1 − ξ1^0| = ∞ and d_ξ2 = |ξ̂2 − ξ2^0| = ∞, for entering the iteration process.
while d_ξ1 = |ξ̂1 − ξ1^0| > ϵ or d_ξ2 = |ξ̂2 − ξ2^0| > ϵ do
    Compute the auxiliary matrices W1, W2, ..., W8.
    Estimate the translation parameters t̂x and t̂y.
    Compute the vectors of Lagrange multipliers k1 and k2.
    Compute the coefficients f1, f2, ..., f6.
    Estimate parameters ξ̂1 and ξ̂2.
    Compute parameter d_ξ1 = |ξ̂1 − ξ1^0| and d_ξ2 = |ξ̂2 − ξ2^0|.
    Update the approximate values with the estimated ones (ξ1^0 = ξ̂1, ξ2^0 = ξ̂2) .
end while
return ξ̂1, ξ̂1, t̂x and t̂y.
```

Solution for singular cofactor matrices

An iterative least squares solution is possible also for the case of singular cofactor matrices, following the same procedure as in the previous application cases. Grouping equations (5.307) and (5.308), together with the normal equations (5.295)-(5.298), results in the system of equations

$$
\begin{aligned}
\mathbf{W}_1\mathbf{k}_1 + \mathbf{W}_2\mathbf{k}_2 &= -\left(\xi_1\mathbf{x}_\mathrm{c} - \xi_2\mathbf{y}_\mathrm{c} + t_x\mathbf{e} - \mathbf{X}_\mathrm{c}\right),\\
\mathbf{W}_3\mathbf{k}_2 + \mathbf{W}_4\mathbf{k}_1 &= -\left(\xi_2\mathbf{x}_\mathrm{c} + \xi_1\mathbf{y}_\mathrm{c} + t_y\mathbf{e} - \mathbf{Y}_\mathrm{c}\right),\\
\mathbf{k}_1^\mathrm{T}\left(\mathbf{x}_\mathrm{c} + \mathbf{v}_\mathrm{x}\right) + \mathbf{k}_2^\mathrm{T}\left(\mathbf{y}_\mathrm{c} + \mathbf{v}_\mathrm{y}\right) &= 0,\\
\mathbf{k}_1^\mathrm{T}\left(\mathbf{y}_\mathrm{c} + \mathbf{v}_\mathrm{y}\right) - \mathbf{k}_2^\mathrm{T}\left(\mathbf{x}_\mathrm{c} + \mathbf{v}_\mathrm{x}\right) &= 0,\\
\mathbf{k}_1^\mathrm{T}\mathbf{e} &= 0,\\
\mathbf{k}_2^\mathrm{T}\mathbf{e} &= 0.
\end{aligned}
\tag{5.320}
$$

Introducing approximate values for the residual vectors $\mathbf{v}_\mathrm{x}^0$, $\mathbf{v}_\mathrm{y}^0$, $\mathbf{v}_\mathrm{X}^0$ and $\mathbf{v}_\mathrm{X}^0$, the developed equation system can be expressed by

$$\begin{bmatrix} \mathbf{W}_1 & \mathbf{W}_2 & \mathbf{x}_\mathrm{c} & -\mathbf{y}_\mathrm{c} & \mathbf{e} & 0 \\ \mathbf{W}_4 & \mathbf{W}_3 & \mathbf{y}_\mathrm{c} & \mathbf{x}_\mathrm{c} & 0 & \mathbf{e} \\ \left(\mathbf{x}_\mathrm{c}+\mathbf{v}_\mathrm{x}^0\right)^\mathrm{T} & \left(\mathbf{y}_\mathrm{c}+\mathbf{v}_\mathrm{y}^0\right)^\mathrm{T} & 0 & 0 & 0 & 0 \\ \left(\mathbf{y}_\mathrm{c}+\mathbf{v}_\mathrm{y}^0\right)^\mathrm{T} & -\left(\mathbf{x}_\mathrm{c}+\mathbf{v}_\mathrm{x}^0\right)^\mathrm{T} & 0 & 0 & 0 & 0 \\ \mathbf{e}^\mathrm{T} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{e}^\mathrm{T} & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{k}_1 \\ \mathbf{k}_2 \\ \xi_1 \\ \xi_2 \\ t_x \\ t_y \end{bmatrix} = \begin{bmatrix} \mathbf{X}_\mathrm{c} \\ \mathbf{Y}_\mathrm{c} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{5.321}$$

which can be equivalently formulated as

$$\mathbf{N} \begin{bmatrix} \mathbf{k} \\ \mathbf{X} \end{bmatrix} = \mathbf{n}. \tag{5.322}$$

The introduced matrices are

$$\mathbf{N} = \begin{bmatrix} \mathbf{W}_1 & \mathbf{W}_2 & \mathbf{x}_\mathrm{c} & -\mathbf{y}_\mathrm{c} & \mathbf{e} & 0 \\ \mathbf{W}_4 & \mathbf{W}_3 & \mathbf{y}_\mathrm{c} & \mathbf{x}_\mathrm{c} & 0 & \mathbf{e} \\ \left(\mathbf{x}_\mathrm{c}+\mathbf{v}_\mathrm{x}^0\right)^\mathrm{T} & \left(\mathbf{y}_\mathrm{c}+\mathbf{v}_\mathrm{y}^0\right)^\mathrm{T} & 0 & 0 & 0 & 0 \\ \left(\mathbf{y}_\mathrm{c}+\mathbf{v}_\mathrm{y}^0\right)^\mathrm{T} & -\left(\mathbf{x}_\mathrm{c}+\mathbf{v}_\mathrm{x}^0\right)^\mathrm{T} & 0 & 0 & 0 & 0 \\ \mathbf{e}^\mathrm{T} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{e}^\mathrm{T} & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{n} = \begin{bmatrix} \mathbf{X}_\mathrm{c} \\ \mathbf{Y}_\mathrm{c} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{5.323}$$

with the vector of unknown transformation parameters

$$\mathbf{X} = \begin{bmatrix} \xi_1 \\ \xi_2 \\ t_x \\ t_y \end{bmatrix}. \tag{5.324}$$

The least squares solution of this adjustment problem is

$$\begin{bmatrix} \hat{\mathbf{k}} \\ \hat{\mathbf{X}} \end{bmatrix} = \mathbf{N}^{-1}\mathbf{n}. \tag{5.325}$$

The estimated parameters for the 2D similarity transformation can be utilized as new approximations and the procedure is repeated until the necessary predefined condition is met.

Solution with a symmetric normal matrix N

Similarly to the cases of subsections 5.2.4.2 and 5.3.4, a symmetric matrix $\mathbf{N}$ can be obtained by adding the terms $\xi_1\mathbf{v}_\mathrm{x} - \xi_2\mathbf{v}_\mathrm{y}$ and $\xi_2\mathbf{v}_\mathrm{x} + \xi_1\mathbf{v}_\mathrm{y}$ to both sides of equations (5.307) and (5.308) respectively. In this way, the equation system (5.321) becomes

$$\begin{bmatrix} \mathbf{W} & \mathbf{A} \\ \mathbf{A}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{k} \\ \mathbf{X} \end{bmatrix} = \begin{bmatrix} \mathbf{w} \\ \mathbf{0} \end{bmatrix}, \tag{5.326}$$

with the relevant matrices being expressed for this problem as

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_1 & \mathbf{W}_2 \\ \mathbf{W}_4 & \mathbf{W}_3 \end{bmatrix}, \tag{5.327}$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}^0 & -\mathbf{y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{y}}^0 & \mathbf{e} & 0 \\ \mathbf{y}_{\mathrm{c}} + \mathbf{v}_{\mathrm{y}}^0 & \mathbf{x}_{\mathrm{c}} + \mathbf{v}_{\mathrm{x}}^0 & 0 & \mathbf{e} \end{bmatrix} \tag{5.328}$$

and

$$\mathbf{w} = -\mathbf{z}_{\mathrm{c}} + a^0 \mathbf{v}_{\mathrm{x}}^0 + b^0 \mathbf{v}_{\mathrm{y}}^0. \tag{5.329}$$

A solution of this adjustment problem can be computed by equation (5.325), after introducing

$$\mathbf{N} = \begin{bmatrix} \mathbf{W} & \mathbf{A} \\ \mathbf{A}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \quad , \quad \mathbf{n} = \begin{bmatrix} \mathbf{w} \\ \mathbf{0} \end{bmatrix}, \tag{5.330}$$

with $\mathbf{N}$ being symmetric.

The rank deficiency of matrix $\mathbf{W}$, respectively of the matrices $\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3$ and $\mathbf{W}_4$, depends on the cofactor matrices of the problem and is important for the inversion of matrix $\mathbf{N}$. Similar to the adjustment cases from subsections 5.2.4.2 and 5.3.4, the presented criterion (5.102) will ensure the existence of a unique solution.

The developed iterative procedure is presented in Algorithm 9 for obtaining a weighted least squares solution for the 2D transformation parameters, when singular cofactor matrices are given.

Algorithm 9 Least squares 2D similarity transformation of coordinates with singular cofactor matrices

```
Choose approximate values for ξ1^0, ξ2^0, v_x^0, v_y^0 and v_X^0, v_Y^0 .
Set threshold ϵ for the break-off condition of the iteration process.
Define parameters d_ξ1 = |ξ̂1 − ξ1^0| = ∞ and d_ξ2 = |ξ̂2 − ξ2^0| = ∞, for entering the iteration process.
while d_ξ1 = |ξ̂1 − ξ1^0| > ϵ or d_ξ2 = |ξ̂2 − ξ2^0| > ϵ do
    Compute matrices W, A and vector w.
    Build matrix N and vector n.
    Estimate the vector of unknown parameters [k̂; X̂].
    Compute the residual vectors v_x, v_y and v_X, v_Y.
    Compute parameter d_ξ1 = |ξ̂1 − ξ1^0| and d_ξ2 = |ξ̂2 − ξ2^0|.
    Update the approximate values with the estimated ones, with ξ1^0 = ξ̂1, ξ2^0 = ξ̂2, v_x^0 = v_x, v_y^0 = v_y,
    v_X^0 = v_X and v_Y^0 = v_Y.
end while
return ξ̂1, ξ̂1, t̂_x and t̂_y.
```

5.5 Discussion of weighted nonlinear least squares solutions

In this chapter, three adjustment cases that belong to a class of nonlinear least squares problems with a direct solution have been examined: the fitting of straight line in 2D, the fitting of a plane in 3D and the 2D similarity transformation of coordinates. Further, four individual weighting scenarios were presumed in each adjustment problem that often occur in practice: constant weights for the coordinates in each direction, individual weights for the coordinates of each point, individual weight for each coordinate and individually weighted and correlated coordinates. A thorough analysis of these weighted least squares problems has shown that in certain cases a direct solution is still possible. Otherwise, an iterative solution could be always developed without performing any kind of linearization of the problem.

A direct solution has been proven to be always possible for the first two weighting cases, regarding the discussed class of least squares problems. The estimated unknown parameters have been obtained by parametrizing appropriately the mathematical model and minimizing a clearly defined Lagrange function. This led to a system of normal equations that can have a nontrivial solution if the determinant is equal to zero or equivalently by solving an eigenvalue problem. Therefore, two novel systematic approaches have been established for the direct solution of the first two weighting cases, based on the same solution strategy of (Malissiovas et al. 2016).

For the last two weighting cases, it has been demonstrated why a direct least squares solution is not possible. Subsequently, iterative algorithms have been presented that are applicable in all cases without making use of linearization of the original problem. The general idea of the established iterative systematic approach is based on the minimization of a Lagrange function and the solution of a reduced system of normal equations, following (Petrović et al. 1983). In addition, a simple extension to this systematic approach has been demonstrated for the solution of these adjustment problems when singular cofactor matrices are given. The developed iterative algorithms produce the WTLS solution and can be compared to those presented by Fang (2011) and Snow (2012).

6 Numerical Investigations

The developed methodologies and implemented algorithms of the previous chapters are tested here for two application examples: this of fitting a straight line to measured points in 2D and for estimating the 2D similarity transformation parameters between two measured groups of homologous points. Different weighting cases are examined for each adjustment problem. Moreover, the presented solutions for the examples in this chapter have been tested and found to be numerically equal to the least squares solution within the GHM.

6.1 Fitting of a straight line in 2D

This section illustrates the least squares solution for fitting a straight line in 2D. The dataset of the measured point coordinates is listed in Table 6.1. It originates from the work of Pearson (1901) and since then it has been utilized by many authors. (Snow, 2012, p. 67) noticed that York (1966) introduced unusual weights for the observed coordinates and solved the problem iteratively. Different algorithms for estimating the least squares solution of the same dataset using York's stochastic model has been presented at least by Neri et al. (1989), Schaffrin and Wieser (2008), Shen et al. (2011) or Amiri-Simkooei and Jazaeri (2012). Snow (2012) introduced also correlations between the measured coordinates and solved the problem for regular and singular cofactor matrices.

TABLE 6.1: Example dataset of measured points in 2D.

Point No.	x-coord. [m]	y-coord. [m]
1	0	5.9
2	0.9	5.4
3	1.8	4.4
4	2.6	4.6
5	3.3	3.5
6	4.4	3.7
7	5.2	2.8
8	6.1	2.8
9	6.5	2.4
10	7.4	1.5

The measured coordinates are in both directions of x and y under the influence of random errors. Least squares solutions for fitting a straight line to the points are presented below for four different weighting cases.

Equally weighted coordinates

For the first weighting case the measured coordinates are uncorrelated and have been obtained with equal precision, i.e. equal weights:

$$p_{x_i} = p_{y_i} = 1. \tag{6.1}$$

For solving the nonlinear problem of fitting a straight line to the ten points of the presented dataset, Pearson (1901) proposed a direct approach using a functional model equivalent to equation (4.32). The least squares solution for the slope of the requested line has been found in that article to be $\hat{\beta} = -0.546$. A solution has been also obtained utilizing the algorithm of Neitzel and Petrovic (2008). The results from the GHM are listed in Table 6.2.

Table 6.2: Solution within the GHM using the algorithm of Neitzel and Petrovic (2008).

Estimated parameter	GHM solution
$\hat{\beta}$ (slope $\hat{a}$)	-0.545561197521
$\hat{\gamma}$ (y-intercept $\hat{b}$)	5.7840437745301

Further, a direct least squares solution was obtained for the unknown line parameters a, b and c, following the developed methodology of section 4.2.1. The Lagrange function of equation (4.11) leads to a homogeneous system of equations with one unknown parameter k, that can be estimated by solving

$$\begin{vmatrix} (56.396 - k) & -30.43 \\ -30.43 & (17.22 - k) \end{vmatrix} = 0. \tag{6.2}$$

This yields a quadratic equation with the solutions for the unknown parameters $k_{min} = 0.618572759437049$ and $k_{max} = 72.9974272240563$. The results for the line parameters can be found in Table 6.3. Parameters $\hat{a}$, $\hat{b}$ and $\hat{c}$ were utilized further to compute parameters $\hat{\beta}$ and $\hat{\gamma}$.

Table 6.3: Direct least squares solution (section 4.2.1).

Estimated parameter	least squares solution
$\hat{a}$	-0.4789242860482
$\hat{b}$	-0.8778562115935
$\hat{c}$	5.0775587555999
Computed parameter	
$\hat{\beta} = -\frac{\hat{a}}{\hat{b}}$	-0.545561197521
$\hat{\gamma} = -\frac{\hat{c}}{\hat{b}}$	5.7840437745301

An estimate for the unknown line parameters was derived using the TLS approach. The determinant of the generalized eigenvalue problem was built following the procedure for the eigenvalue/eigenvector decomposition of matrix **G** (equation 4.43). This results in the characteristic equation of the eigenvalues (quadratic equation) with the solutions $\lambda_{min} = 0.618572759437049$ and $\lambda_{max} = 72.997427240563$. The TLS solution for the unknown line parameters is presented in Table 6.4.

TABLE 6.4: TLS solution (section 4.2.2).

Estimated parameter	TLS solution
$\hat{\beta}$	-0.545561197521
$\hat{\gamma}$	5.7840437745301

The developed direct least squares solution is, as expected, identical to the TLS. Both solutions are numerically consistent with the result of the linearized GHM. The requested straight line is depicted in Figure 6.1, together with the measured points and their estimated residuals.

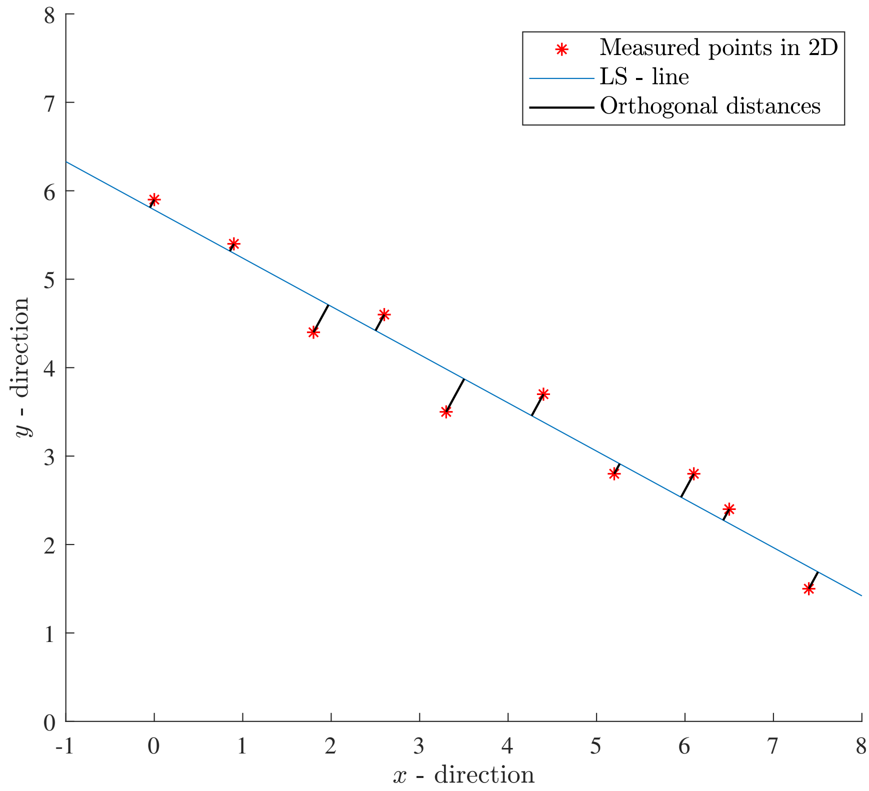

FIGURE 6.1: Fitting a straight line to points in 2D, with observed x and y coordinates of equal precision.

Equally weighted coordinates in each direction

In this weighting case the measured coordinates are uncorrelated and have been obtained with the same precision in each direction. The postulated weights are

$$p_{x_i} = 0.5 \quad \text{and} \quad p_{y_i} = 1.5. \tag{6.3}$$

A direct solution can be derived in this case from the developed methodology of section 5.2.1. The results are presented in Table 6.5.

TABLE 6.5: Direct least squares solution (section 5.2.1)

Estimated parameter	Direct least squares solution
$\hat{a}$	-0.4832580303705
$\hat{b}$	-0.8754779700726
$\hat{c}$	5.0853141652839
Computed parameter	
$\hat{\beta} = -\frac{\hat{a}}{\hat{b}}$	-0.5519933646422
$\hat{\gamma} = -\frac{\hat{c}}{\hat{b}}$	5.8086146529331

For comparison reasons the least squares solution has been computed again with the algorithm of (Neitzel and Petrovic 2008), taking into account the current stochastic model. Both the developed direct solution and the one within the GHM are numerically identical. The estimated straight line is depicted in Figure 6.2.

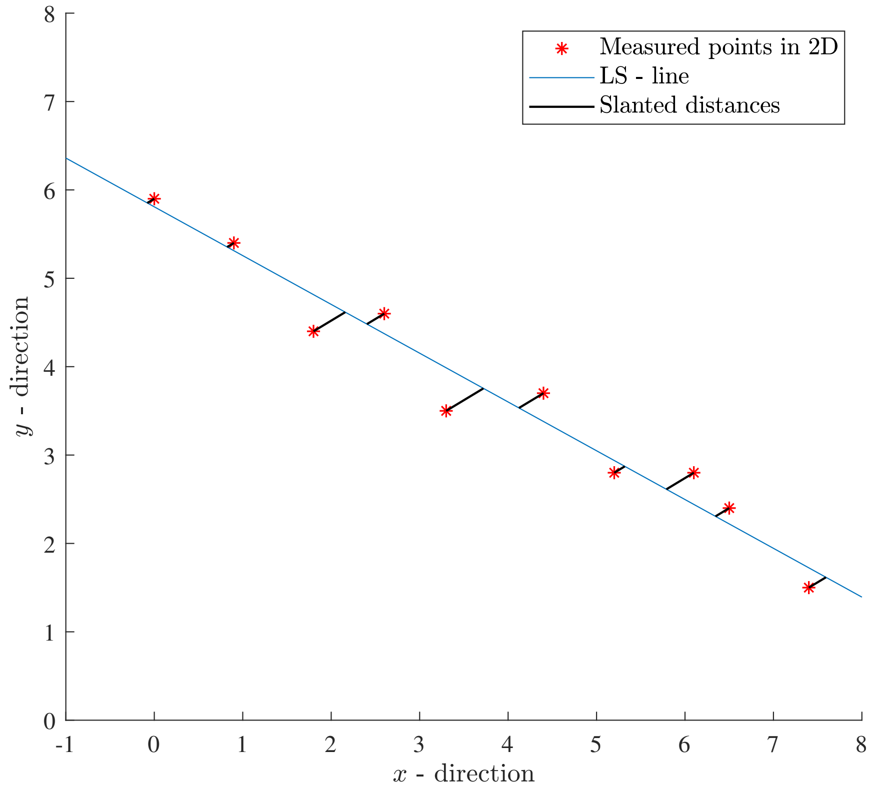

FIGURE 6.2: Fitting a straight line to points in 2D, with observed x and y coordinates and p_x, p_y individual constant weights for each coordinate axis.

Individually weighted points

For this weighting example the measured coordinates are uncorrelated and have been obtained with individual precision for each measured point. The postulated weights are listed in Table 6.6.

TABLE 6.6: Individual weights for each point.

Point No.	$p_x = \frac{1}{{\sigma_x}^2}$	$p_y = \frac{1}{{\sigma_y}^2}$
1	1	1
2	1.2	1.2
3	0.8	0.8
4	1.1	1.1
5	0.9	0.9
6	1.15	1.15
7	1	1
8	0.93	0.93
9	1.25	1.25
10	1.13	1.13

A direct weighted least squares solution has been obtained following the procedure of section 5.2.2. The results for the estimated line parameters can be found in Table 6.7.

TABLE 6.7: Direct least squares solution (section 5.2.2).

Estimated parameter	Direct least squares solution
$\hat{a}$	-0.4824660036697
$\hat{b}$	-0.875914696362
$\hat{c}$	5.1014648040614
Computed parameter	
$\hat{\beta} = -\frac{\hat{a}}{\hat{b}}$	-0.5508139156399
$\hat{\gamma} = -\frac{\hat{c}}{\hat{b}}$	5.8241571071355

The computed straight line is shown in Figure 6.3.

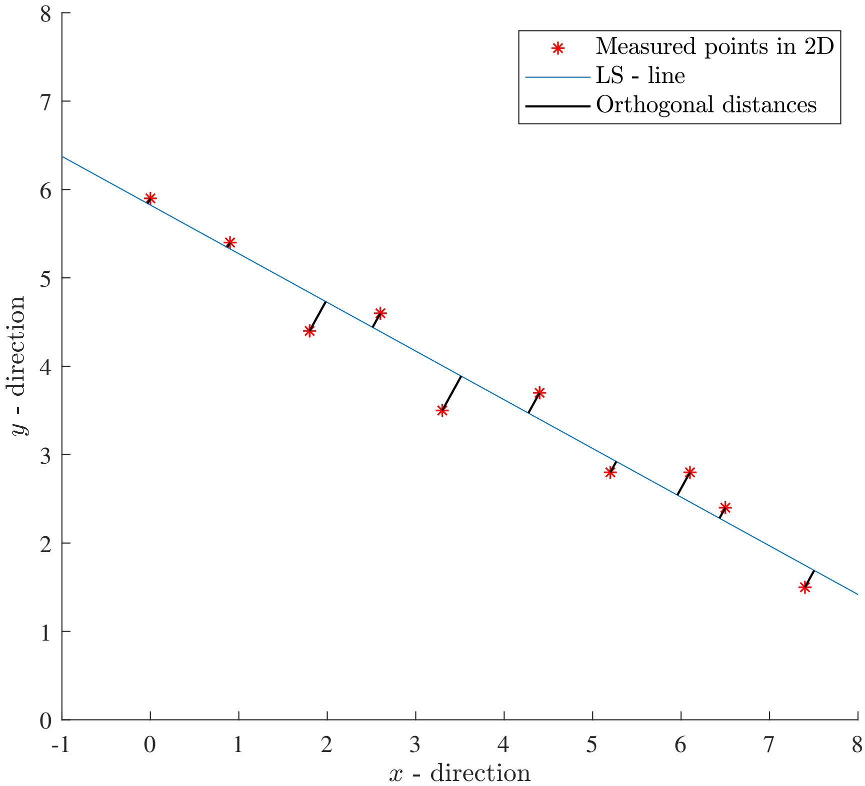

FIGURE 6.3: Fitting a straight line to points in 2D, with individual weight for the coordinates of each point.

Individually weighted 2D coordinates

The stochastic model of York (1966) has been adopted for this adjustment example. The measured coordinates are uncorrelated and have been obtained with individual precision. The postulated weights are listed in Table 6.8.

TABLE 6.8: Individual weights for each coordinate.

Point No.	$p_x = \frac{1}{{\sigma_x}^2}$	$p_y = \frac{1}{{\sigma_y}^2}$
1	1000	1
2	1000	1.8
3	500	4
4	800	8
5	200	20
6	80	20
7	60	70
8	20	70
9	1.8	100
10	1	500

An iterative weighted least squares solution has been derived from the developed approach of section 5.2.3, by employing Algorithm 1. The results are listed in Table 6.9.

TABLE 6.9: Iterative least squares solution using Algorithm 1 (section 5.2.3).

Estimated parameter	Iterative least squares solution
$\hat{a}$	0.4805334074462
$\hat{b}$	1
$\hat{c}$	-5.4799102240329
Computed parameter	
$\hat{\beta} = -\frac{\hat{a}}{\hat{b}}$	-0.4805334074462
$\hat{\gamma} = -\frac{\hat{c}}{\hat{b}}$	5.4799102240329

The presented solution has been found to be numerically identical with the iterative least squares solution of York (1966), Neri et al. (1989), the pseudoquadratic algorithm from (Petrović et al. 1983) and the WTLS algrithm of Schaffrin and Wieser (2008). The estimated line is depicted in Figure 6.4.

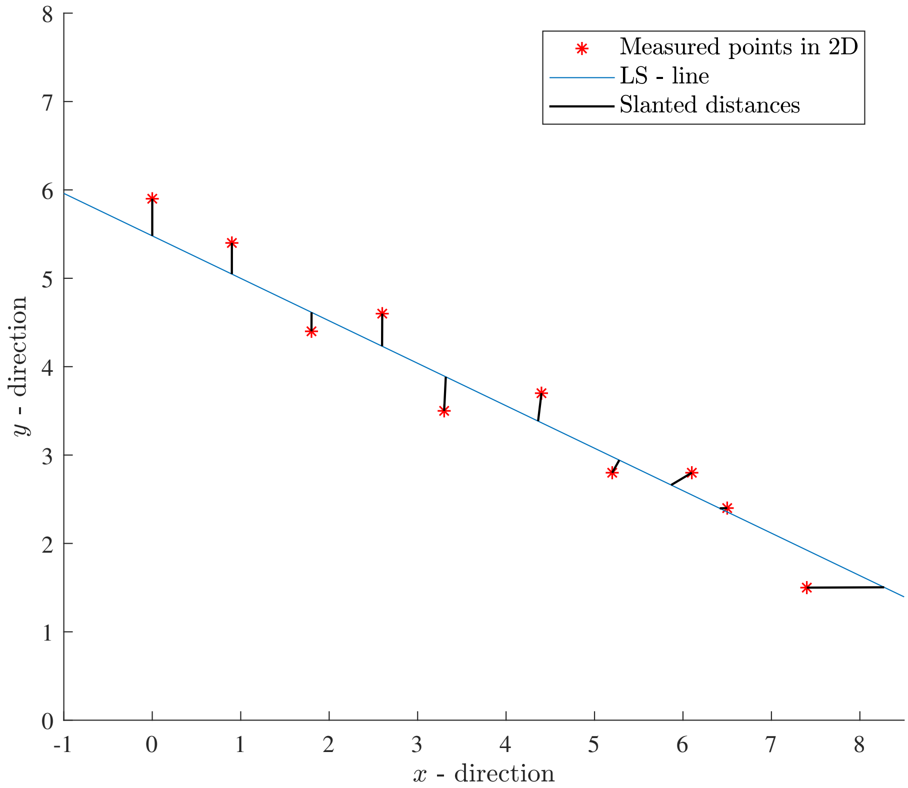

FIGURE 6.4: Fitting a straight line to points in 2D, with individual weight for each measured coordinate.

Individually weighted and correlated 2D coordinates

In addition to the stochastic model of Table 6.8, correlations between the measured coordinates of each point are postulated in this example. This is for instance the case when polar coordinates of points have been originally measured, while their Cartesian coordinates are utilized in the adjustment, together with their stochastic properties from a linear error propagation. For comparison reasons, the necessary correlations between the point coordinates have been taken directly from the numerical investigations of (Snow 2012, pp. 68-70) (i.e. the case of a regular cofactor matrix) and are listed in Table 6.10.

TABLE 6.10: Individual weights for each coordinate and correlations for each point.

Point No.	$p_x = \frac{1}{{\sigma_x}^2}$	$p_y = \frac{1}{{\sigma_y}^2}$	ρ_{xy}
1	1000	1	-0.165956
2	1000	1.8	0.440649
3	500	4	-0.999771
4	800	8	-0.395335
5	200	20	-0.706488
6	80	20	-0.815323
7	60	70	-0.627480
8	20	70	-0.308879
9	1.8	100	-0.206465
10	1	500	0.077633

The variance-covariance matrix for this adjustment problem can be computed by

$$\mathbf{\Sigma}_{\mathrm{LL}} = \begin{bmatrix} \mathbf{\Sigma}_{\mathrm{xx}} & \mathbf{\Sigma}_{\mathrm{xy}} \\ \mathbf{\Sigma}_{\mathrm{yx}} & \mathbf{\Sigma}_{\mathrm{yy}} \end{bmatrix}, \tag{6.4}$$

with

$$\mathbf{\Sigma}_{\mathrm{xx}} = \begin{bmatrix} \sigma_{x_1}^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{x_2}^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{x_3}^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{x_4}^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{x_5}^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{x_6}^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{x_7}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{x_8}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{x_9}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{x_{10}}^2 \end{bmatrix}, \tag{6.5}$$

$$\mathbf{\Sigma}_{\mathrm{yy}} = \begin{bmatrix} \sigma_{y_1}^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{y_2}^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{y_3}^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{y_4}^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{y_5}^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{y_6}^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{y_7}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{y_8}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{y_9}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{y_{10}}^2 \end{bmatrix} \tag{6.6}$$

and

$$\mathbf{\Sigma}_{\mathrm{xy}} = \mathbf{\Sigma}_{\mathrm{yx}} = \begin{bmatrix} \sigma_{x_1y_1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{x_2y_2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{x_3y_3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{x_4y_4} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{x_5y_5} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{x_6y_6} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{x_7y_7} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{x_8y_8} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{x_9y_9} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{x_{10}y_{10}} \end{bmatrix} . \tag{6.7}$$

The individual covariances between the coordinates of each point are

$$\sigma_{xy} = \rho_{xy}\ \sigma_x\ \sigma_y. \tag{6.8}$$

Setting the variance of the universal weight equal to one ($\sigma_0^2 = 1$) the cofactor matrix of the observations is

$$\mathbf{Q}_{\mathrm{LL}} = \begin{bmatrix} \mathbf{Q}_{\mathrm{xx}} & \mathbf{Q}_{\mathrm{xy}} \\ \mathbf{Q}_{\mathrm{yx}} & \mathbf{Q}_{\mathrm{yy}} \end{bmatrix} = \sigma_0^2 \mathbf{\Sigma}_{\mathrm{LL}} = \mathbf{\Sigma}_{\mathrm{LL}}. \tag{6.9}$$

For a regular cofactor matrix it is possible to compute the weight matrix

$$\mathbf{P} = \mathbf{Q}_{\mathrm{LL}}^{-1}. \tag{6.10}$$

A least squares solution for the unknown line parameters can be estimated iteratively, utilizing Algorithm 2 from the developed approach of section 5.2.4.1. The results are presented in Table 6.11.

TABLE 6.11: Iterative least squares solution using Algorithm 2 (section 5.2.4.1).

Estimated parameter	Iterative least squares solution
$\hat{a}$	0.4592286797279
$\hat{b}$	1
$\hat{c}$	-5.357272562041
Computed parameter	
$\hat{\beta} = -\frac{\hat{a}}{\hat{b}}$	-0.4592286797279
$\hat{\gamma} = -\frac{\hat{c}}{\hat{b}}$	5.357272562041

The results from the proposed approach have been compared and found to be numerically identical with the solution from the WTLS algorithm presented in (Snow 2012, p. 72). The requested straight line is depicted in Figure 6.5.

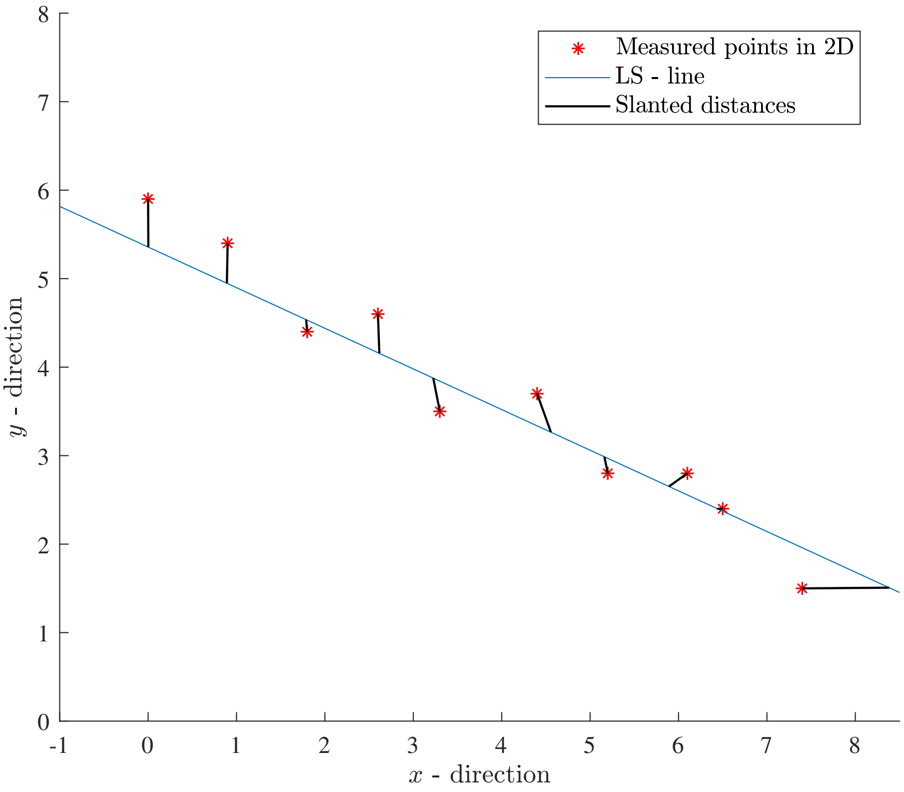

Figure 6.5: Fitting a straight line to points in 2D, with individually weighted and correlated coordinates for each point.

Solution with a singular cofactor matrix

A singular variance-covariance matrix is postulated in this example for the measured point coordinates including correlations between the observations. This is the case when, for example, the 2D Cartesian coordinates of the points have been obtained by a least squares adjustment of a free network and their stochastic properties by a linearized error propagation. The same stochastic model as in (Snow 2012, pp. 71-72) is utilized here, for the case of a singular cofactor matrix which satisfies the NS criterion, in order to compare the results of the proposed approach. The necessary variances and covariances can be found in Appendix A.1 and lead to a singular matrix $\mathbf{W}$ with rank deficiency equal to 2. However, applying the developed criterion from equation (5.102) results in

$$\text{rank}\left([\mathbf{W} \mid \mathbf{A}]\right) = 10 = n, \tag{6.11}$$

which ensures that a unque solution still exists, with

- rank of $\mathbf{W} = 8 < n$, with n = number of condition equations;
- rank of $\mathbf{A} = 2 = m$, with m = number of unknown parameters;
- redundancy : $r_d = n - m = 10 - 2 = 8$;

Here it can be seen that the rank of matrix $\mathbf{W}$ is smaller than the number of condition equations n, caused by the rank deficiency of the introduced cofactor matrices.

A least squares solution for the unknown line parameters has been estimated iteratively, using Algorithm 3 from section 5.2.4.2. The results are shown in Table 6.12.

TABLE 6.12: Iterative least squares solution using Algorithm 3 (section 5.2.4.2).

Estimated parameter	Iterative least squares solution
$\hat{a}$	0.4931726182468
$\hat{b}$	1
$\hat{c}$	-5.54275204298
Computed parameter	
$\hat{\beta} = -\frac{\hat{a}}{\hat{b}}$	-0.4931726182468
$\hat{\gamma} = -\frac{\hat{c}}{\hat{b}}$	5.54275204298

The solution for the line parameters is numerically equal to the WTLS solution from (Snow 2012, p. 72) and the solution within the GHM for the case of a singular cofactor matrix. The computed straight line is depicted in Figure 6.6.

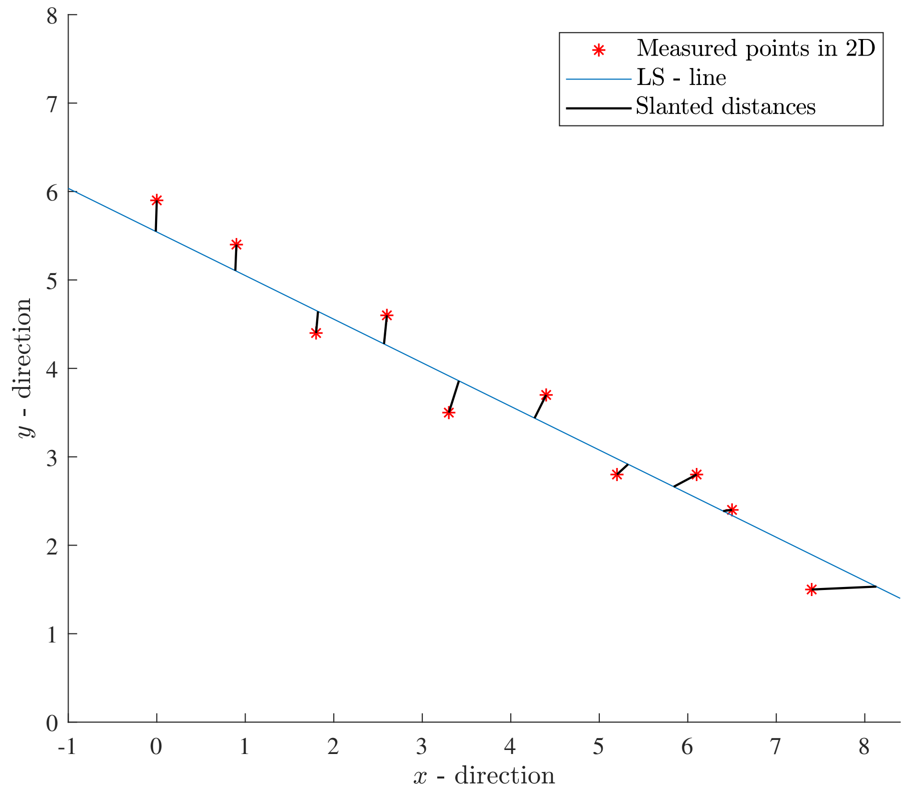

FIGURE 6.6: Fitting a straight line to points in 2D, with a singular cofactor matrix.

6.2 2D similarity transformation of coordinates

The least squares solution of the 2D similarity transformation is presented in this subsection following the developed methodologies and algorithms of chapters 4 and 5. The coordinates of four homologous points have been measured in two coordinate systems, the target XY and the source xy system and are listed in Table 6.13.

TABLE 6.13: Example dataset for the 2D similarity transformation

Point No.	Target S.		Source S.	
i	$X_i[m]$	$Y_i[m]$	$x_i[m]$	$y_i[m]$
1	-117.478	0	17.856	144.794
2	117.472	0	252.637	154.448
3	0.015	-117.41	140.089	32.326
4	-0.014	117.451	130.40	267.027

This dataset originates from (Mikhail et al. 2001, pp. 397-402) and has been utilized in the past at least by Felus and Schaffrin (2005), Neitzel (2010) and Malissiovas et al. (2016).

Equally weighted coordinates

In the first weighting case of this numerical example the coordinates of the homologous points are equally weighted and uncorrelated, i.e. equal weights:

$$p_{X_i} = p_{Y_i} = p_{x_i} = p_{y_i} = 1. \tag{6.12}$$

A GHM has been employed by Neitzel (2010) for deriving the least squares solution for the unknown transformation parameters between the coordinate systems. The results from the GHM are presented in Table 6.14.

TABLE 6.14: Results from Neitzel (2010).

Estimated parameter	GHM solution
Parameter $\hat{\xi}_1$	0.99900748077781
Parameter $\hat{\xi}_2$	-0.04109806319405
Scale factor $\hat{\mu}$	0.99985248784424
Rotational angle $\hat{\phi}$	-2°21′20.72″
Translation parameter $\hat{t}_x$	-141.2628 mm
Translation parameter $\hat{t}_y$	-143.9316 mm

The developed direct least squares solution for the 2D similarity transformation of section 4.5.1 is presented in Table 6.15. For comparison reasons, the estimates for α, β and γ can be substituted in equation (4.110) to derive the parameters ξ_1 and ξ_2. Furthermore, the rotational angle ϕ as well as the scale factor μ can be computed by substituting $\hat{\xi_1}$ and $\hat{\xi_2}$ into equation (4.102). Thus, the rotational angle is

$$\hat{\phi} = \arctan\left(\frac{\hat{\xi_2}}{\hat{\xi_1}}\right) \tag{6.13}$$

and the scale factor

$$\hat{\mu} = \frac{\hat{\xi_1}}{\cos\hat{\phi}}. \tag{6.14}$$

Similarly, the translation parameters t_x and t_y can be derived from equation (4.106).

TABLE 6.15: Direct least squares solution for the 2D similarity transformation

Estimated parameter	Direct least squares solution
$\hat{\alpha}$	-0.4832580303705
$\hat{\beta}$	-0.8754779700726
$\hat{\gamma}$	5.0853141652839
Computed parameter	
$\hat{\xi_1} = -\frac{\hat{\alpha}}{\hat{\gamma}}$	0.99900748077781
$\hat{\xi_2} = -\frac{\hat{\beta}}{\hat{\gamma}}$	-0.04109806319405
$\hat{\mu}$	0.99985248784424
$\hat{\phi}$	$-2°21'20.723943558''$
$\hat{t}_x$	-141.2627900259449
$\hat{t}_y$	-143.9316426333377

In addition, the TLS solution of section 4.5.2 has been computed and found to be identical to the solution from the proposed direct least squares approach. Both coincide numerically with the results from the GHM, as it can be seen from Tables 6.14 and 6.15.

Equally weighted coordinates in each coordinate system

Equal weights between the coordinates of each coordinate system are assumed for the second weighting case, with

$$p_{XY_i} = 1.1 \quad \text{and} \quad p_{xy_i} = 0.9. \tag{6.15}$$

A direct weighted least squares solution has been derived here, following the developed methodology of section 5.4.1. The results are presented in Table 6.16.

TABLE 6.16: Direct least squares solution (section 5.4.1)

Estimated parameter	Direct least squares solution
$\hat{\alpha}$	-0.7064570685103
$\hat{\beta}$	0.0290628626954
$\hat{\gamma}$	0.7071589357166

Computed parameter	
$\hat{\xi}_1 = -\frac{\hat{\alpha}}{\hat{\gamma}}$	0.9990074830836
$\hat{\xi}_2 = -\frac{\hat{\beta}}{\hat{\gamma}}$	-0.0410980632889
$\hat{\mu}$	0.99985249015199
$\hat{\phi}$	-2°21′20.723943556″
$\hat{t}_x$	-141.2627903519875
$\hat{t}_y$	-143.9316429655667

Equally weighted homologous points in both systems

For this weighting example the coordinates of the homologous points have been measured with individual precisions. The weights of this stochastic model are listed in Table 6.17.

TABLE 6.17: Individual weights for homologous points in both systems.

Point No.	Target S.		Source S.	
i	$p_{X\,i} = \dfrac{1}{\sigma_{X\,i}^2}$	$p_{Y\,i} = \dfrac{1}{\sigma_{Y\,i}^2}$	$p_{x\,i} = \dfrac{1}{\sigma_{x\,i}^2}$	$p_{y\,i} = \dfrac{1}{\sigma_{y\,i}^2}$
1	0.9	0.9	0.9	0.9
2	1.05	1.05	1.05	1.05
3	0.85	0.85	0.85	0.85
4	1.3	1.3	1.3	1.3

A direct weighted least squares solution has been estimated for this example following the solution strategy of section 5.4.2. Table 6.18 contains the results.

TABLE 6.18: Direct least squares solution (section 5.4.2)

Estimated parameter	Direct least squares solution
$\hat{\alpha}$	-0.7064686408794
$\hat{\beta}$	0.0290681956735
$\hat{\gamma}$	0.7071471554453
Computed parameter	
$\hat{\xi}_1 = -\frac{\hat{\alpha}}{\hat{\gamma}}$	0.9990404902845
$\hat{\xi}_2 = -\frac{\hat{\beta}}{\hat{\gamma}}$	-0.0411062894755
$\hat{\mu}$	0.99988580761122
$\hat{\phi}$	-2°21′22.139597905″
$\hat{t}_x$	-141.2687384001714
$\hat{t}_y$	-143.9337541051444

Individually weighted coordinates

Individual precision for each measured coordinate has been postulated in this case. The introduced weights of this stochastic model are listed in Table 6.19.

TABLE 6.19: Individual weight for each coordinate.

Point No.	Target S.		Source S.	
i	$p_{X\,i} = \frac{1}{\sigma_{X\,i}^2}$	$p_{Y\,i} = \frac{1}{\sigma_{Y\,i}^2}$	$p_{x\,i} = \frac{1}{\sigma_{x\,i}^2}$	$p_{y\,i} = \frac{1}{\sigma_{y\,i}^2}$
1	0.95	1.2	1	0.8
2	1	0.75	1.15	0.9
3	0.95	1.3	0.7	0.85
4	1	0.85	1.2	1

An iterative weighted least squares solution for the unknown transformation parameters has been computed by using Algorithm 7 from section 5.4.3. The results are presented in Table 6.20.

Table 6.20: Iterative least squares solution (section 5.4.3)

Estimated parameter	Iterative least squares solution
$\hat{\xi}_1$	0.99899700973504
$\hat{\xi}_2$	-0.041113617539
$\hat{t}_x$	-141.2636321979249
$\hat{t}_y$	-143.9334627584676

Computed parameter	
$\hat{\mu}$	0.99984266512622
$\hat{\phi}$	-2°21′24.018841465″

Individually weighted and correlated coordinates in each coordinate system

In this adjustment example correlations are present only between the coordinates of each point and in both coordinate systems. The correlation coefficients have been computed randomly, like in (Snow, 2012) for the example of fitting a straight line in 2D. Here the *randn* function has been used from the programming language GNU Octave (version 4.0.0). The introduced weights, together with the correlation coefficients are listed in Table 6.21 for the coordinates in the target system and in Table 6.22 for the coordinates in the source system.

Table 6.21: Weights and correlations for the coordinates of the points in the target system.

Point No.	Target S.		
i	$p_{X\,i} = \frac{1}{\sigma_{X\,i}^2}$	$p_{Y\,i} = \frac{1}{\sigma_{Y\,i}^2}$	$\rho_{XY\,i}$
1	0.95	1.2	0.2
2	1	0.75	-0.3
3	0.95	1.3	-0.5
4	1	0.85	0.2

Table 6.22: Weights and correlations for the coordinates of the points in the source system.

Point No.	Source S.		
i	$p_{x\,i} = \frac{1}{\sigma_{x\,i}^2}$	$p_{y\,i} = \frac{1}{\sigma_{y\,i}^2}$	$\rho_{xy\,i}$
1	1	0.8	0.2
2	1.15	0.9	-0.4
3	0.7	0.85	-0.3
4	1.2	1	0.2

An iterative weighted least squares solution was derived for this adjustment problem utilizing Algorithm 8 from section 5.4.4. The results are presented in Table 6.23.

TABLE 6.23: Iterative least squares solution (section 5.4.4)

Estimated parameter	Iterative least squares solution
$\hat{\xi}_1$	0.9990276226893
$\hat{\xi}_2$	-0.0411159727
$\hat{t}_x$	-141.26679907
$\hat{t}_y$	-143.93227343
Computed parameter	
$\hat{\mu}$	0.99987334903342
$\hat{\phi}$	-2°21′24.244598358″

Solution with a singular cofactor matrix

A singular cofactor matrix is postulated in the last weighting example of the 2D similarity transformation. For the sake of comparison, the dataset from the numerical example in (Snow 2012, pp. 77-82) and (Neitzel and Schaffrin 2017) is adopted. The observed coordinates are listed in Table 6.24.

TABLE 6.24: Example dataset from Neitzel and Schaffrin (2016).

Point No.	Target S.		Source S.	
i	$X_i[m]$	$Y_i[m]$	$x_i[m]$	$y_i[m]$
1	400.0040	100.0072	453.8001	137.6099
2	500.0019	299.9994	521.2865	350.7972
3	399.9925	399.9933	406.8728	433.9247
4	100.0059	400.0022	110.5545	386.9880
5	99.9956	99.9978	157.4861	90.6802

The coordinates of the homologous points in the two coordinate systems are the outcome of a free network adjustment. An approximate solution for the variances and covariances of these coordinates has been also computed using a linearized error propagation. As it is explained by Neitzel and Schaffrin (2017), the resulting cofactor matrices are fully populated (i.e. the off-diagonal elements are not zero) and rank deficient, but still fulfilling the NS criterion. The introduced singular cofactor matrices for this numerical example can be found in Appendix A.2 and lead to a singular matrix $\mathbf{W}$ with rank deficiency equal to 2. Also for this numerical example the developed criterion from equation (5.102) leads to

$$\text{rank}\left([\mathbf{W} \mid \mathbf{A}]\right) = 10 = n, \tag{6.16}$$

which ensures that a unqe solution for the unknown parameters exists, with

- rank of $\mathbf{W} = 8 < n$, with n = number of condition equations;
- rank of $\mathbf{A} = 4 = m$, with m = number of unknown parameters;

- redundancy : $r_d = n - m = 10 - 4 = 6$;

An iterative least squares solution has been computed for this adjustment case using Algorithm 9 from section 5.4.4. Table 6.25 contains the resulting estimates for the transformation parameters.

TABLE 6.25: Iterative least squares solution (section 5.4.4)

Estimated parameter	Iterative least squares solution
$\hat{\xi}_1$	0.9876550155542
$\hat{\xi}_2$	-0.1564292113176
$\hat{t}_x$	-69.726354301821
$\hat{t}_y$	35.0782153796499
Computed parameter	
$\hat{\mu}$	0.99996626338233
$\hat{\phi}$	$-9°0'0.004803006''$

The developed iterative procedure provides the exact numerical result as the least squares solution within the GHM presented by Neitzel and Schaffrin (2017) and the WTLS solution from (Snow 2012, p. 80).

7 Conclusion and outlook

7.1 Conclusion

The fundamental principles of adjustment calculus and the method of least squares have been briefly discussed in chapter 2, setting the basis for the methodological developments of this thesis. A review of related works in adjustment calculus has shown that the mathematical modelling of the measurement results embodies the most fundamental parts of every adjustment problem (i.e. the functional and stochastic model). Only a correct mathematical model could lead to meaningful estimates for the unknown parameters and the residuals of each problem. For the purposes of this research an unambiguous definition has been given in chapter 2 concerning linear and nonlinear least squares problems. Additionally, the evaluation of the adjustment results has been discussed in terms of precision and reliability. Various approaches were presented that are common in geodetic literature for the computation and correct interpretation of the stochastic parameters of the estimated unknowns.

The main subject of this thesis is the least squares solution of a class of nonlinear adjustment problems. For this reason, two solution strategies were highlighted in chapter 3. The first is related to the traditional approaches that are commonly used in geodesy and two famous adjustment models, namely the GMM and the GHM. Based on the Gauss-Newton approach, a least squares solution can be derived iteratively and by linearizing the nonlinear functional model. The second strategy includes the most modern algorithms that have been developed in the last decades by the mathematical/statistical community. In the TLS literature the discussed problems are often expressed within the EIV model and various algorithmic approaches have been presented for a solution. Depending on the stochastic model, the presented TLS solutions are direct and cover the cases of equally weighted and uncorrelated measurements, while the WTLS solutions are iterative and can deal usually with more general weighting cases. Chapter 3 comes to the conclusion that TLS and WTLS provide the (weighted) least squares solution of the discussed class of problems. Hence, it is in agreement with the views of Petrovic (2003) and Neitzel (2010), that TLS is not a new method but a special case of the least squares method.

In the fourth chapter of this thesis, the solution of four individual nonlinear adjustment problems is discussed: the fitting of a straight line in 2D and 3D, the fitting of a plane in 3D and the 2D similarity transformation of coordinates. A mathematical relationship between direct least squares and TLS solutions has been presented that was based on the publication of Malissiovas et al. (2016). This shows that TLS produces the least squares solution of a problem using SVD, while the exact solution can also be achieved by following the standard procedure for the least squares solution of a problem, i.e. by minimizing the sum of squared residuals. Additionally, a new solution strategy has been established for the direct least squares solution of the investigated class of problems. The developments of this chapter give an overview of these adjustment problems that can be transformed into an eigenvalue problem and thus, it clarifies in which cases a solution from TLS is possible. This chapter demonstrates that TLS is an algorithmic approach for obtaining the least squares solution and not a method.

The findings in chapter 5 provide an overview of possible weighted least squares solutions for the discussed class of nonlinear adjustment problems, i.e. the WTLS solution for the mathematical/statistical community. Different weighting scenarios and correlations between the observations were postulated for each problem. It has been shown that for certain weighting cases a direct solution still exists. For these cases two novel direct approaches have been proposed. Further, general weight matrices have been examined including also cofactor matrices with correlations between the measurements. New algorithms have been developed and presented for the iterative weighted least squares solution of this class of problems without linearizing the original problem in any step of the procedure. In addition, singular cofactor matrices that still fulfill the criterion of Neitzel and Schaffrin (2016) have been taken into account. The presented algorithms can handle the latter stochastic model without the need of any special treatment of the adjustment problem.

The implemented algorithms are based on the established solution strategies for obtaining a (weighted) least squares solution for the investigated class of nonlinear problems. The direct approaches can be employed in engineering tasks for which efficiency is important, i.e. no need for starting values for the unknown parameters or iterations to obtain a minimum solution for the objective function. For instance, straight lines or planes can be fitted directly to measured 3D point clouds from laser scanners, or the similarity transformation parameters between several sets of homologous points can be estimated, if such a stochastic model is postulated that can lead to a direct solution. In case of more general weight matrices or taking correlations between the measurements into account in the stochastic model, a weighted least squares solution is provided by the presented iterative algorithms.

7.2 Outlook

Complex adjustment problems can be further tackled in future research using the knowledge that was acquired so far. Thus, all the developed algorithms can be possibly extended in order to provide elegant solutions for other adjustment cases such as:

- Variance Component Estimation (VCE) of individual groups of observations,
- Tykhonov regularization of ill-posed problems.

Moreover, the discussed (weighted) least squares solutions can be considered as "optimal" only if the random errors that influence the measurements are normally distributed. In the presence of outliers (or blunders), however, other adjustment methods may be preferred that are more robust in the sense that the solution is not falsified by a small amount of outliers. Thus, the developed algorithms and solution strategies can be extended to become more robust against outliers. Future developments can be based on the following objectives:

- Robust estimators have been studied since decades by geodesists, mathematicians and statisticians. These estimators can be divided into two main groups, the M-estimators (based on the maximum likelihood method) with a breakdown point of 5 to 10% and estimators with higher one, like for example the L_1-norm which can reach a breakdown point of 50% of blunders. It is worth to mention the rigorous solution via linear programming (Dantzig 1963, Dantzig and Thapa 2006), for example using the simplex algorithm presented in (Dantzig 1949) or (Barrodale and Roberts 1974). However, the minimization of L_1-norm with linear programming has been, interestingly, neglected by the geodetic literature with some exceptions, for example (Fuchs 1980) or (Fuchs 1982). It is still not known if such algorithms can be developed for a robust solution of adjustment problems within the EIV model.

- An approximate solution by minimizing a L_p-norm could be obtained by an iterative procedure of reweighted least squares, which is almost exclusively used by geodesists. Some first examples can be found in the algorithm of Schlossmacher (1973) and the contributions of Krarup et al. (1980) and Somogyi and Závoti (1993). In fact, reweighted least squares can possibly fail to provide a correct solution in some cases, as has already been pointed out by Neitzel (2004) for the case of L_1. A thorough analysis of the robustness of various L_p-norms has been presented by Marx (2013), who noticed that L_p for $1.2 < p < 1.5$ may be less resistant to outliers than L_1 and proposed $L_{1.05}$ as an alternative solution. New robust algorithms can be implemented for the detection of blunders, based on the combination of a reweighted approach and the direct solutions of weighted nonlinear least squares from this thesis.

- Most of the mentioned procedures fail to provide a resistant solution when leverage points[1] exist in the dataset. Rigorous methods for the identification of erroneous data can be global optimization methods based on systematic or stochastic search, as demonstrated by the study of Marx (2015) for the detection of blunders by means of a Monte Carlo simulation. Combinatorial approaches can also be employed in leverage point's cases. Examples of combinatorial approaches include the maximum subsample (MSS) method developed theoretically by (Neitzel, 2004, p. 109) and employed successfully by Neitzel and Marx (2007), Wujanz et al. (2016) and Wujanz (2016) for the detection of deformations using laser scanning data. The employment of such methods for the implementation of modern and robust algorithms is of great interest.

[1] See e.g. (Everitt and Skrondal 2010) for a definition.

Appendices

A Stochastic models for the numerical investigations

A.1 Singular cofactor matrix for fitting a straight line in 2D

A singular cofactor matrix is presented in this appendix for the application example of fitting a straight line in 2D. It was computed by utilizing the correlation coefficients from (Snow 2012, pp. 94-95), together with the necessary precisions of the measured coordinates from Table 6.8. The variance-covariance matrix for this adjustment problem is

$$\mathbf{Q}_{\mathrm{LL}} = \begin{bmatrix} \mathbf{Q}_{\mathrm{xx}} & \mathbf{Q}_{\mathrm{xy}} \\ \mathbf{Q}_{\mathrm{yx}} & \mathbf{Q}_{\mathrm{yy}} \end{bmatrix}, \tag{A.1}$$

with the sub-matrices:

$$\mathbf{Q}_{\mathrm{xx}} = \left[\begin{matrix}
0.001 & -0.00010729581873004 & -0.000184367806311119 & -0.000162605765335955 & -0.000396422204114424 & \cdots \\
-0.00010729581873004 & 0.001 & -0.000209334092987561 & -0.00018594935307068 & -0.000458502562871592 & \cdots \\
-0.000184367806311119 & -0.000209334092987561 & 0.002 & -0.000326944096945427 & -0.000819451184602405 & \cdots \\
-0.000162605765335955 & -0.00018594935307068 & -0.000326944096945427 & 0.00125 & -0.000762997332323675 & \cdots \\
-0.000396422204114424 & -0.000458502562871592 & -0.000819451184602405 & -0.000762997332323675 & 0.005 & \cdots \\
-0.0002625284166187 & -0.000318111572725898 & -0.00060532385276285 & -0.000608691217742683 & -0.00177606225648547 & \cdots \\
-1.30268774106986e-06 & -6.44923418774174e-05 & -0.000275369953484654 & -0.000452568996672073 & -0.00192208514323818 & \cdots \\
0.00055638919876439 & 0.000525619226802249 & 0.0006397035186675 & 0.000228426867755593 & -0.000754067056197251 & \cdots \\
0.00130709043198385 & 0.00133828286919678 & 0.00195079375824933 & 0.00127603026024317 & 0.00137539801300953 & \cdots \\
0.00375954407416456 & 0.00404574061392281 & 0.00646160268527631 & 0.00507413806293484 & 0.00993112604965441 & \cdots
\end{matrix}\right.$$

$$\left.\begin{matrix}
-0.0002625284166187 & -1.30268774106986e-06 & 0.00055638919876439 & 0.00130709043198385 & 0.00375954407416456 \\
-0.000318111572725898 & -6.44923418774174e-05 & 0.000525619226802249 & 0.00133828286919678 & 0.00404574061392281 \\
-0.00060532385276285 & -0.000275369953484654 & 0.0006397035186675 & 0.00195079375824933 & 0.00646160268527631 \\
-0.000608691217742684 & -0.000452568996672073 & 0.000228426867755593 & 0.00127603026024317 & 0.00507413806293484 \\
-0.00177606225648547 & -0.00192208514323818 & -0.000754067056197251 & 0.00137539801300953 & 0.00993112604965441 \\
0.0125 & -0.00356447023778097 & -0.0041685755760135 & -0.00394653521323417 & -0.000250424830498049 \\
-0.00356447023778097 & 0.0166666666666667 & -0.0159730654839819 & -0.0211377773455841 & -0.0296837502900246 \\
-0.0041685755760135 & -0.0159730654839819 & 0.05 & -0.0415061974365333 & -0.069564033429064 \\
-0.00394653521323417 & -0.0211377773455841 & -0.0415061974365333 & 0.555555555555556 & -0.114601731317248 \\
-0.000250424830498049 & -0.0296837502900246 & -0.069564033429064 & -0.114601731317248 & 1
\end{matrix}\right], \tag{A.2}$$

$$\mathbf{Q}_{\mathrm{yy}} = \left[\begin{matrix}
1 & -0.0799735814606215 & -0.06518386303754 & -0.0514204579136473 & -0.0396422204114424 & \cdots \\
-0.0799735814606215 & 0.555555555555556 & -0.0551643771471813 & -0.0438286828378432 & -0.0341747632812917 & \cdots \\
-0.06518386303754 & -0.0551643771471813 & 0.25 & -0.0365534612806129 & -0.0289719744741855 & \cdots \\
-0.0514204579136473 & -0.0438286828378432 & -0.0365534612806129 & 0.125 & -0.0241280941877523 & \cdots \\
-0.0396422204114424 & -0.0341747632812917 & -0.0289719744741855 & -0.0241280941877523 & 0.05 & \cdots \\
-0.0166037549406539 & -0.0149959233498933 & -0.0135354528317981 & -0.0121738243548537 & -0.0112328039935045 & \cdots \\
-3.8138791846177e-05 & -0.00140733827808684 & -0.00285034981466199 & -0.00418997473652902 & -0.00562728909449856 & \cdots \\
0.00940469397241429 & 0.00662217980206098 & 0.00382295973501637 & 0.0012209929672505 & -0.00127460596179143 & \cdots \\
0.005545515048479 & 0.00423202202024702 & 0.002926190637374 & 0.00171197424195029 & 0.000583531957097723 & \cdots \\
0.00531679821802292 & 0.00426458505408133 & 0.00323080134263815 & 0.00226922352718828 & 0.0014044733149058 & \cdots
\end{matrix}\right.$$

$$
\begin{matrix}
-0.0166037549406539 & -3.8138791846177e-05 & 0.00940469397241429 & 0.005545515048479 & 0.00531679821802292 \\
-0.0149959233498933 & -0.00140733827808684 & 0.00662217980206098 & 0.00423202202024702 & 0.00426458505408133 \\
-0.0135354528317981 & -0.00285034981466199 & 0.00382295973501637 & 0.002926190637374 & 0.00323080134263815 \\
-0.0121738243548537 & -0.00418997473652902 & 0.0012209929672505 & 0.00171197424195029 & 0.00226922352718828 \\
-0.0112328039935045 & -0.00562728909449856 & -0.00127460596179143 & 0.000583531957097723 & 0.0014044733149058 \\
0.05 & -0.00660011638235733 & -0.00445639474180467 & -0.00105896652148659 & -2.23986777699e-05 \\
-0.00660011638235733 & 0.0142857142857143 & -0.00790461742025186 & -0.00262556138650223 & -0.00122902402483588 \\
-0.00445639474180467 & -0.00790461742025186 & 0.0142857142857143 & -0.00297656367844959 & -0.00166289845870223 \\
-0.00105896652148659 & -0.00262556138650223 & -0.00297656367844959 & 0.01 & -0.000687610387903489 \\
-2.23986777699e-05 & -0.00122902402483588 & -0.00166289845870223 & -0.000687610387903489 & 0.002
\end{matrix}
\left.\vphantom{\begin{matrix}1\\1\\1\\1\\1\\1\\1\\1\\1\\1\end{matrix}}\right] , \tag{A.3}
$$

and

$$
\mathbf{Q}_{\mathrm{xy}} = \left[
\begin{matrix}
0.0316227766016838 & -0.0025289867005658 & -0.00206129473887088 & -0.00162605765335955 & -0.00125359708006575 & \cdots \\
-0.00339299170599481 & 0.0235702260395516 & -0.00234042630964225 & -0.0018594935307068 & -0.00144991241169878 & \cdots \\
-0.00583022195151901 & -0.00493405188950133 & 0.0223606797749979 & -0.00326944096945427 & -0.0025913321746667 & \cdots \\
-0.00514204579136473 & -0.00438286828378432 & -0.00365534612806129 & 0.0125 & -0.00241280941877523 & \cdots \\
-0.0125359708006575 & -0.0108070090465971 & -0.00916174276506853 & -0.00762997332323675 & 0.0158113883008419 & \cdots \\
-0.00830187747032693 & -0.00749796167494667 & -0.00676772641589903 & -0.00608691217742684 & -0.00561640199675225 & \cdots \\
-4.11946034176041e-05 & -0.00152009907587077 & -0.00307872967476321 & -0.00452568996672073 & -0.00607816690940393 & \cdots \\
0.0175945713361161 & 0.0123889639864633 & 0.00715210276593168 & 0.00228426867755593 & -0.0023845694060815 & \cdots \\
0.0413338287288236 & 0.0315436297318278 & 0.021810537267639 & 0.0127603026024317 & 0.00434939041038 & \cdots \\
0.11888722238149 & 0.0953590207675548 & 0.072242914239365 & 0.0507413806293484 & 0.0314049780471384 & \cdots
\end{matrix}
\right.
$$

$$
\begin{matrix}
-0.0005250568332374 & -1.20605449440977e-06 & 0.000297402536496859 & 0.000175364583519327 & 0.000168131922284769 \\
-0.000636223145451796 & -5.97083063915161e-05 & 0.000280955294656434 & 0.000179549488118848 & 0.000180931020641263 \\
-0.0012106477055257 & -0.000254943037809526 & 0.000341935913709648 & 0.000261726447211668 & 0.00028897165695746 \\
-0.00121738243548537 & -0.000418997473652902 & 0.00012209929672505 & 0.000171197424195029 & 0.000226922352718828 \\
-0.00355212451297094 & -0.0017795050590842 & -0.000403065795849046 & 0.000184529007192446 & 0.000444133458802924 \\
0.025 & -0.00330005819117866 & -0.00222819737090234 & 0.000529483260743293 & -1.119933888495e-05 \\
-0.00712894047556193 & 0.0154303349962092 & -0.00853796263679499 & -0.00283593042227886 & -0.00132749766951248 \\
-0.008337151152027 & -0.0147881850800537 & 0.0267261241912424 & -0.00556864073733696 & -0.0031109981507291 \\
-0.00789307042646833 & -0.0195697791310586 & -0.0221859957479004 & 0.074535599249993 & -0.00512514523129067 \\
-0.000500849660996098 & -0.0274818126551341 & -0.0371835399333781 & -0.015375435693872 & 0.0447213595499958
\end{matrix}
\left.\vphantom{\begin{matrix}1\\1\\1\\1\\1\\1\\1\\1\\1\\1\end{matrix}}\right] , \tag{A.4}
$$

with $\mathbf{Q}_{\mathrm{xy}} = \mathbf{Q}_{\mathrm{yx}}^{\mathrm{T}}$.

A.2 Singular cofactor matrix for the 2D similarity transformation

A stochastic model is postulated in the last investigated weighting case for the 2D similarity transformation, that involves a singular cofactor matrix. This matrix has been firstly introduced in (Snow 2012, pp. 96-97) and (Neitzel and Schaffrin 2017). However, it is reordered in such a way, so that it fits to the needs and the formulation of the stochastic model in this work (see section 5.4.4). Therefore, two cofactor matrices are given. The first matrix is related to the coordinates in the source system

$$
\mathbf{Q}_{\mathrm{LL}_1} = \begin{bmatrix} \mathbf{Q}_{\mathrm{xx}} & \mathbf{Q}_{\mathrm{xy}} \\ \mathbf{Q}_{\mathrm{yx}} & \mathbf{Q}_{\mathrm{yy}} \end{bmatrix} , \tag{A.5}
$$

with the relevant sub-matrices:

$$\mathbf{Q}_{\mathrm{xx}} = \begin{bmatrix} 36.370281457026799 & -10.717856095227498 & -8.652980417908310 & -3.008722473688215 & -13.990722470202799 \\ -10.717856095227498 & 31.714850184591498 & -9.754412805230141 & -7.170133519153175 & -4.072447764980719 \\ -8.652980417908310 & -9.754412805230141 & 29.490003562291800 & -9.264714951570749 & -1.817895387582575 \\ -3.008722473688215 & -7.170133519153175 & -9.264714951570749 & 26.031496206858098 & -6.587925262445999 \\ -13.990722470202799 & -4.072447764980719 & -1.817895387582575 & -6.587925262445999 & 26.468990885212101 \end{bmatrix} 10^{-6}[\mathrm{m}^2], \quad (\mathrm{A.6})$$

$$\mathbf{Q}_{\mathrm{yy}} = \begin{bmatrix} 29.082186568548597 & -12.471413471703498 & -6.363232761238930 & -2.835610062891940 & -7.411930272714205 \\ -12.471413471703498 & 30.958038019578300 & -17.720699070083000 & -0.556828023714815 & -0.209097454076965 \\ -6.363232761238930 & -17.720699070083000 & 38.734832754164593 & -10.531447467909800 & -4.119453454932779 \\ -2.835610062891940 & -0.556828023714815 & -10.531447467909800 & 32.063142073587599 & -18.139256519071097 \\ -7.411930272714205 & -0.209097454076965 & -4.119453454932779 & -18.139256519071097 & 29.879737700795101 \end{bmatrix} 10^{-6}[\mathrm{m}^2], \quad (\mathrm{A.7})$$

$$\mathbf{Q}_{\mathrm{xy}} = \mathbf{Q}_{\mathrm{yx}}^{\mathrm{T}} = \begin{bmatrix} -5.470847531049374 & -4.151968395749720 & -6.309575380973525 & 5.061541661693445 & 10.870849646079201 \\ 5.848221379655184 & 3.083310192383325 & -3.437514270957704 & -5.891999249139355 & 0.397981948058545 \\ 4.362573210487724 & 3.061227368299260 & 6.203628967652264 & -10.232648328119298 & -3.394781218319955 \\ -0.456888271105760 & 5.116298153814835 & 2.979330360651054 & -1.237264484444125 & -6.401475758915990 \\ -4.283058787987756 & -7.108867318747670 & 0.564130323627930 & 12.300370400009299 & -1.472574616901790 \end{bmatrix} 10^{-6}[\mathrm{m}^2]. \quad (\mathrm{A.8})$$

The second cofactor matrix concerns the coordinates in the target system

$$\mathbf{Q}_{\mathrm{LL}2} = \begin{bmatrix} \mathbf{Q}_{\mathrm{XX}} & \mathbf{Q}_{\mathrm{XY}} \\ \mathbf{Q}_{\mathrm{YX}} & \mathbf{Q}_{\mathrm{YY}} \end{bmatrix}, \quad (\mathrm{A.9})$$

with

$$\mathbf{Q}_{\mathrm{XX}} = \begin{bmatrix} 6.794487571800590 & -2.067492652515560 & -1.752371987000400 & -0.451547525907080 & -2.523075406377545 \\ -2.067492652515560 & 6.429318039548140 & -1.970512300406600 & -1.403742194865920 & -0.987570891760060 \\ -1.752371987000400 & -1.970512300406600 & 6.229399700497529 & -2.051290948290685 & -0.455224464799860 \\ -0.451547525907080 & -1.403742194865920 & -2.051290948290685 & 5.079972517721780 & -1.173391848658100 \\ -2.523075406377545 & -0.987570891760060 & -0.455224464799860 & -1.173391848658100 & 5.139262611595570 \end{bmatrix} 10^{-6}[\mathrm{m}^2], \quad (\mathrm{A.10})$$

$$\mathbf{Q}_{\mathrm{YY}} = \begin{bmatrix} 6.094989272208480 & -2.499242822649745 & -1.204799645049460 & -0.699300070526570 & -1.691646733982720 \\ -2.499242822649745 & 5.912760395723919 & -3.440142089389230 & -0.117912645153740 & 0.144537161468785 \\ -1.204799645049460 & -3.440142089389230 & 7.206462860853270 & -1.847407593285470 & -0.714113533129125 \\ -0.699300070526570 & -0.117912645153740 & -1.847407593285470 & 6.360602565647050 & -3.695982256681265 \\ -1.691646733982720 & 0.144537161468785 & -0.714113533129125 & -3.695982256681265 & 5.957205362324320 \end{bmatrix} 10^{-6}[\mathrm{m}^2], \quad (\mathrm{A.11})$$

$$\mathbf{Q}_{\mathrm{XY}} = \mathbf{Q}_{\mathrm{YX}}^{\mathrm{T}} = \begin{bmatrix} -1.246373797420340 & -0.879050061336250 & -1.163579541268595 & 0.979701730455350 & 2.309301669569825 \\ 1.090128341653400 & 0.554510615436640 & -0.917558590359990 & -0.955341882018735 & 0.228261515288690 \\ 0.938081379301460 & 0.362129743332930 & 1.443374807628205 & -2.018715860691035 & -0.724870069571550 \\ -0.106812374044770 & 1.212589988419975 & 0.582948103761090 & -0.048153307431655 & -1.640572410704635 \\ -0.675023549489755 & -1.250180285853295 & 0.054815220239305 & 2.042509319686065 & -0.172120704582330 \end{bmatrix} 10^{-6}[\mathrm{m}^2] \quad (\mathrm{A.12})$$

Bibliography

Abatzoglou T., Mendel J. and Harada G., 1991. The constrained Total Least-Squares technique and its application to harmonic superresolution. *Trans. Signal Process.*, 39, 1070–1087.

Adcock R., 1878. A problem in least squares. *The Analyst*, 5, 53–54.

Alkhatib H., 2007. *On Monte Carlo methods with applications to the current satellite gravity missions.* PhD dissertation, Landwirtschaftliche Fakultät, Universität Bonn.

Alkhatib H. and Schuh W., 2007. Integration of the Monte Carlo covariance estimation strategy into tailored solution procedures for large-scale least squares problems. *Journal of Geodesy*, 81, 53–66.

Amiri-Simkooei A. and Jazaeri S., 2012. Weighted total least squares formulated by standard least squares theory. *Journal of Geodetic Science*, 2, 113–124.

Barrodale I. and Roberts F., 1974. Algorithm 478: Solution of an overdetermined system of equations in the L_1 norm. *ACM*, 17, 319–320.

Bickel P.J. and Ritov Y., 1987. Efficient estimation in the errors in variables model. *Ann. Stat.*, 15, 2, 513–540.

Bjerhammar A., 1973. *Theory of errors and generalized matrix inverses.* Elsevier, Amsterdam, London, New York.

Björck A., 2015. *Numerical Methods in Matrix Computations*, Texts in Applied Mathematics 59. Springer International Publishing.

Bronshtein I., Semendyayev K., Musiol G. and Muehlig H., 2005. *Handbook of Mathematics.* Springer, Berlin Heidelberg New York, fifth edition.

Cross P., 1994. *Advanced least squares applied to position fixing.* Working paper no. 6. North East London Polytechnic, Department of Land Surveying.

Dantzig G., 1949. Programming of Independent Activities II. *Mathematical Model, Econometrica*, 17, 200–211.

Dantzig G., 1963. *Linear programming and extensions.* Princeton university press.

Dantzig G. and Thapa M., 2006. *Linear programming 2: Theory and extensions.* Springer Science & Business Media.

Dekking F., Kraaikamp C., Lopuhaä H. and Meester L., 2005. *A Modern Introduction to Probability and Statistics: Understanding why and how.* Springer Science & Business Media.

Deming W., 1931. The application of least squares. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 11(68), 146–158.

Deming W., 1934. On the application of least squares.-II. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 7(114), 804–829.

Deming W., 1964. *Statistical adjustment of data.* Dover publications, Inc., New York.

Drixler E., 1993. *Analyse der Form und Lage von Objekten im Raum*, volume C(409). Deutsche Geodätische Kommission bei der Bayerischen Akademie der Wissenschaften.

Everitt B.S. and Skrondal A., 2010. *The Cambridge dictionary of statistics.* Cambridge university press, fourth edition.

Fang X., 2011. *Weighted Total Least Squares Solutions for Applications in Geodesy.* PhD dissertation, Dept. of Geodesy and Geoinformatics, Leibniz University Hannover, Germany.

Felus Y. and Burtch R., 2009. On symmetrical three-dimensional datum conversion. GPS solutions, 13, 1, 65–74.

Felus Y. and Schaffrin B., 2005. Performing similarity transformations using the errors-in-variables-model. *Proceedings of the ASPRS Meeting*, Washington DC, 35, 751–762.

Fuchs H., 1980. *Untersuchungen zur Ausgleichung durch Minimieren der Absolutsumme der Verbesserungen.* PhD dissertation, TU Graz.

Fuchs H., 1982. Contributions to the Adjustment by Minimizing the Sum of Absolute Residuals. *Manuscripta Geodaetica*, 7, 3, 151–207.

Gauss C., 1809. *Theoria motus corporum coelestium in sectionibus conicis solem anbientium.* F. Perthes and I. H. Besser, Hamburg.

Gauss C., 1823. *Theoria combinationis observationum erroribus minimis obnoxiae.* Henricus Dieterich, Gottingae.

Ghilani C.D., 2010. *Adjustment computations: spatial data analysis.* John Wiley and Sons, Hoboken, NJ, USA, fifth edition.

Golub G. and Van Loan C., 1980. An analysis of the total least squares problem. *SIAM Journal on Numerical Analysis*, 17, 6, 883–893.

Golub G. and Van Loan C., 1989. *Matrix Computation.* The Johns Hopkins University Press, Baltimore, Maryland, second edition.

Golub G. and Van Loan C., 1996. *Matrix Computation.* The Johns Hopkins University Press, Baltimore, Maryland, third edition.

Gonin R., 1989. *Nonlinear Lp-Norm Estimation.* CRC Press.

Groen P., 1996. An introduction to total least squares. *Nieuw Archief voor Wiskunde*, 14, 237–253.

Hampel F., 1980. Robuste Schätzungen: Ein anwendungsorientierter Überblick. *Biometrica*, 22, 3–21.

Helmert F., 1872. *Die Ausgleichungsrechnung nach der Methode der kleinsten Quadrate.* B.G Teubner, Leipzig, Berlin.

Helmert F., 1924. *Die Ausgleichungsrechnung nach der Methode der kleinsten Quadrate (Mit Anwendungen auf die Geodäsie, die Physik und die Theorie der Messinstrumente).* B.G Teubner, Leipzig, Berlin, third edition.

Huber P., 1964. Robust estimation of a location parameter. *Annals of mathematical statistics*, 35, 73–101.

Jäger R., Müller T., Saler H. and Schwäble R., 2005. *Klassische und robuste Ausgleichungsverfahren. Ein Leitfaden für Ausbildung und Praxis von Geodäten und Geoinformatikern.* Herbert Wichmann Verlag, Heidelberg.

Jovičić D., Lapaine M. and Petrović S., 1982. Prilagodjavanje pravca skupu točaka prostora (Fitting a straight line to a set of points in space, in Croatian). *Geodetski list*, 36(59), 260–266.

Julier S. and Uhlmann J., 1996. *A general method for approximating nonlinear transformations of probability distributions.* Technical Report, RRG, Department of Engineering Science, University of Oxford.

Julier S. and Uhlmann J., 2000. A new method for the nonlinear transformation of means and covariances in filters and estimator. *IEEE T Automat Contr*, 45, 477–478.

Julier S., Uhlmann J. and Whyte H.D., 1995. A new approach for filtering nonlinear systems. *Proceedings of the 1995 American Control Conference, IEEE, New York*, 3, 1628–1632.

Kampmann G. and Renner B., 2004. Vergleich verschiedener Methoden zur Bestimmung ausgleichender Ebenen und Geraden. *Allgemenine Vermessungs-Nachrichten*, 2, 56–67.

Koch K.R. and Pope A.J., 1969. Least Squares Adjustment with Zero Variances. *zfv - Zeitschrift für Geodäsie, Geoinformation und Landmanagement*, 10, 390–393.

Krakiwsky E.J., 1975. *A synthesis of recent advances in the method of least squares.* Department of Geodesy and Geomatics Engineering, Lecture notes 42, University of New Brunswick.

Krarup T., Juhl J. and Kubik K., 1980. Götterdämmerung over Least Squares Adjustment. *14th congress of the international society of photogrammetry,* Hamburg, B3, 369–378.

Kupferer S., 2004. Verschiedene Ansätze zur Schätzung einer ausgleichenden Raumgeraden. *Allgemenine Vermessungs-Nachrichten*, 5, 162–170.

Lawson C. and Hanson R., 1974. *Solving Least Squares Problems.* SIAM, Philadelphia.

Lenzmann L. and Lenzmann E., 2004. Strenge Auswertung des nichtlinearen Gauss-Helmert-Modells. *AVN*, 111, 68–73.

Linkwitz K., 1960. Über die Systematik verschiedener Formen der Ausgleichungsrechnung. *zfv - Zeitschrift für Vermessungswesen*, 5, 6, 7–10.

Linkwitz K., 1976. Über einige Ausgleichungsprobleme und ihre Lösung mit Hilfe Matrizen-Eigenwerten.

Linnik Y., 1961. *Method of Least Squares and Principles of the Theory of Observations (Translated from the Russian by Regina C. Elandt, Ph.D.).* Pergammon press, Oxford, London, New York, Paris.

Lösler M., Bähr H. and Ulrich T., 2016. Verfahren zur Transformation von Parametern und Unsicherheiten bei nichtlinearen Zusammenhängen. *Photogrammetrie Laserscanning Optische 3D-Messtechnik. Beiträge der Oldenburger 3D-Tage 2016. Wichmann*, 274–285.

Madsen K., Nielsen H. and Tingelff O., 2004. *Methods for non-linear least squares problems.* Informatics and Mathematical Modelling, Technical University of Denmark.

Mahboub V., 2012. On structured weighted total least squares for geodetic transformations. *Journal of Geodesy*, 86(5), 359–367.

Malissiovas G., Neitzel F. and Petrovic S., 2016. Götterdämmerung over total least squares. *Journal of Geodetic Science*, 6(1), 43–60.

Markovsky I. and Van Huffel S., 2007. Overview of total least-squares methods. *Signal Processing*, 87, 2283–2302.

Marx C., 2013. On resistant $\mathrm{L_p}$-norm estimation by means of iteratively reweighted least squares. *Journal of Applied Geodesy*, 7, 43–60.

Marx C., 2015. Outlier detection by means of monte carlo estimation including resistant scale estimation. *Journal of Applied Geodesy*, 9(2), 123–142.

Marx C., 2017. A weighted adjustment of a similarity transformation between two point sets containing errors. *Journal of Geodetic Science*, 7, 105–112.

Meissl P., 1982. *Least squares adjustment: a modern approach. Mitteilungen der geodätischen Institute der Technischen Universität Graz, Folge 43.* Hochschülerschaft an der Technischen Universität Graz, Ges.m.b.H.

Merimman M., 1877. *Elements of the method of Least squares adjustment.* Cambridge: Printed by C.J. Clay, M.A. at the university press.

Mihajlovic D. and Cvijetinovic Z., 2016. Weighted coordinate trasnformation formulated by standard least-squares theory. *Survey Review*, 0, 1–18.

Mikhail E., Bethel J. and McGlone C., 2001. *Introduction to Modern Photogrammetry.* Wiley: New York Chichester.

Mikhail E.M. and Ackermann F., 1976. *Observations and Least Squares.* Thomas Y. Crowell Company, Inc.

Montgomery D.C. and Runger G.C., 2010. *Applied statistics and probability for engineers.* John Wiley & Sons.

Neitzel F., 2004. *Identifizierung konsistenter Datengruppen am Beispiel der Kongruenzuntersuchung geodätischer Netze*, volume C(565). Deutsche Geodätische Kommission bei der Bayerischen Akademie der Wissenschaften.

Neitzel F., 2010. Generalisation of total least squares on example of unweighted and weighted similarity transformation. *Journal of Geodesy*, 84(12), 751–762.

Neitzel F. and Marx C., 2007. Deformationsanalyse und regionale Anpassung eines historischen Geodatenbestandes. *Entwicklerforum Geoinformationstechnik 2007*, 243–255.

Neitzel F. and Petrovic S., 2008. Total Least Squares (TLS) im Kontext der Ausgleichung nach kleinsten Quadraten am Beispiel der ausgleichenden Geraden. *zfv - Zeitschrift für Geodäsie, Geoinformation und Landmanagement*, 133, 141–148.

Neitzel F. and Schaffrin B., 2016. On the Gauss-Helmert model with a singular dispersion matrix where BQ is of smaller rank than B. *Journal of Computational and Applied Mathematics*, 291, 458–467.

Neitzel F. and Schaffrin B., 2017. Adjusting a 2D Helmert transformation within a Gauss-Helmert model with a singular dispersion matrix where BQ is of smaller rank than B. *Acta Geodaetica et Geophysica, Montanistica Hungarica*, 52, 479–496.

Neri F., Saitta G. and Chiofalo S., 1989. An accurate and straightforward approach to line regression analysis of error-affected experimental data. *Journal of physics E: scientific instruments*, 22, 4, 215–217.

Niemeier W., 2008. *Ausgleichungsrechnung.* Walter de Gruyter, New York, second edition.

Pasioti A., 2015. *Investigation of non-linear least squares problems using the example of circle fitting (Master's Thesis).* Technische Universität Berlin, Institut of Geodesy and Geoinformation Science.

Pearson K., 1901. On Lines and Planes of Closest Fit to Systems of Points in Space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2, 11, 559–572.

Perović G., 2005. *Least Squares (MONOGRAPH).* Faculty of Civil Engineering, University of Belgrade, first edition.

Petrovic S., 2003. *Parameterschätzung für unvollständige funktionale Modelle in der Geodäsie*, volume C(563). Habilitation, Deutsche Geodätische Kommission bei der Bayerischen Akademie der Wissenschaften.

Petrović S., Lapaine M., Jovičić D. and Žarinac Frančula B., 1983. Prilagodjavanje pravca (Fitting of a straight line, in Croatian). In: *Proceedings of the 5th international symposium "Computer at the university", Cavtat*, 529–535.

Pope A., 1972. Some pitfalls to be avoided in the iterative adjustment of nonlinear problems. In: *Proceedings of the 38th Annual Meeting of the American Society of Photogrammetry*, Washington, DC, 449–477.

Pope A.J., 1974. Two approaches to Nonlinear Least Squares Adjustments. *The Canadian Surveyor*, 28, 5, 663–669.

Reinking J., 2008. Total Least Squares? *zfv - Zeitschrift für Geodäsie, Geoinformation und Landmanagement*, 6, 384–389.

Schaffrin B., 2006. A note on Constrained Total Least-Squares estimation. *Linear Algebra and its Applications*, 417, 1, 245–258.

Schaffrin B., 2007. Connecting the Dots: The Straight-Line Case Revisited. *zfv - Zeitschrift für Geodäsie, Geoinformation und Landmanagement*, 132, 385–394.

Schaffrin B., Lee I., Felus Y. and Choi Y., 2006. Total least-squares(TLS) for geodetic straight-line and plane adjustment. *Bollettino di geodesia e scienze affini*, 65, 3, 141–168.

Schaffrin B., Neitzel F., Uzun S. and Mahboub V., 2012. Modifying cadzow's algorithm to generate the optimal TLS-solution for the structured EIV-model of a similarity transformation. *Journal of Geodetic Science*, 2, 2, 98–106.

Schaffrin B. and Snow K., 2014. The case of the Homogeneous Errors-In-Variables Model. *Journal of Geodetic Science*, 4, 1, 166–173.

Schaffrin B. and Wieser A., 2008. On weighted total least-squares adjustment for linear regression. *Journal of Geodesy*, 82, 7, 415–421.

Schlossmacher E., 1973. An iterative technique for absolute deviations curve fitting. *Journal of the American Statistical Association*, 68, 344, 857–859.

Shen Y., Li B. and Chen Y., 2011. An iterative solution of weighted total least-squares adjustment. *Journal of Geodesy*, 85, 4, 229–238.

Snow K., 2012. *Topics in Total Least-Squares within the Errors-In-Variables Model: Singular Cofactor Matrices and Prior Information*. PhD Dissertation, the Ohio State University.

Snow K. and Schaffrin B., 2016. Line fitting in euclidian 3d-space. *Studia Geophysica et Geodaetica*, 60, 2, 210–227.

Somogyi J. and Závoti J., 1993. Robust estimation with iteratively reweighted least-squares method. *Acta Geodaetica et Geophysica, Montanistica Hungarica*, 28, 413–420.

Späth H., 2004. Zur numerischen Berechnung der Trägheitsgeraden und der Trägheitsebene. *Allgemenine Vermessungs-Nachrichten*, 7, 273–275.

Stigler S., 1981. Gauss and the invention of least squares. *The Annals of Statistics*, 9, 3, 465–474.

Taylor J., 1982. *An Introduction to Error Analysis*. University science books, Sausalito, California, second edition.

Teunissen P., 1985. *The geometry of geodetic inverse linear mapping and nonlinear adjustment*. Netherlands Geodetic Commission, Publications on Geodesy, New Series, Vol. 8, No. , Delft.

Teunissen P., 1990. Nonlinear least squares. *Manuscripta Geodaetica*, 15, 137–150.

Teunissen P. and Knickmeyer E., 1988. Nonlinearity and least squares. *CISM Journal ACSGC*, 42, 4, 321–330.

Van Huffel S., 2004. Total Least Squares and Errors-in-Variables Modeling: Bridging the Gap between Statistics, Computational Mathematics and Engineering. *COMPSTAT Proceedings in Computational Statistics*, 17, 539–555.

Van Huffel S. and Vandewalle J., 1989. Algebraic connections between the least squares and total least squares problems. *Numerische Mathematik*, 55, 4, 431–449.

Van Huffel S. and Vandewalle J., 1991. *The Total Least Squares Problem, computational aspects and analysis*. SIAM, Philadelphia.

Wells D. and Krakiwsky E.J., 1971. *The method of least squares.* Department of Geodesy and Geomatics Engineering, Lecture notes 18, University of New Brunswick.

Williamson J.H., 1968. Least-squares fitting of a straight line. *Canadian Journal of Physics*, 46, 16, 1845–1847.

Wolf H., 1978. Das geodätische Gauß-Helmert-Modell und seine Eigenschaften. *zfv - Zeitschrift für Vermessungswesen*, 41–43.

Wolf H., 1979. Singuläre Kovarianzen im Gauss-Helmert-Modell. *zfv - Zeitschrift für Vermessungswesen*, 10, 390–393.

Wujanz D., 2016. *Terrestrial laser scanning for geodetic deformation monitoring*, volume C(775). Deutsche Geodätische Kommission bei der Bayerischen Akademie der Wissenschaften.

Wujanz D., Krueger D. and Neitzel F., 2016. Identification of stable areas in unreferenced laser scans for deformation measurement. *The Photogrammetric Record*, 31, 155, 261–280.

York D., 1966. Least-squares fitting of a straight line. *Canadian Journal of Physics*, 44, 5, 1079–1086.

York D., 1968. Least-squares fitting of a straight line with correlated errors. *Earth and planetary science letters*, 5, 320–324.