Lecture Notes

Adjustment Theory

Nico Sneeuw, Friedhelm Krumm, Matthias Roth

Geodätisches Institut
Universität Stuttgart

http://www.gis.uni-stuttgart.de

Rev. 4.46

October 12, 2017

These are lecture notes in progress.
Please contact us ({sneeuw, krumm, roth}@gis.uni-stuttgart.de) for remarks, errors, suggestions, etc.

Contents

1 Introduction **6**
1.1 Adjustment theory –a first look 6
1.2 Historical development 9

2 Least squares adjustment **13**
2.1 Adjustment with observation equations 13
2.2 Adjustment with condition equations 17
2.3 Synthesis 20

3 Generalizations **21**
3.1 Higher dimensions: the A-model (observation equations) 21
3.2 The datum problem 24
3.3 Linearization of non-linear observation equations 26
3.4 Higher dimensions: the B-model (Condition equations) 32
3.5 Linearization of non-linear condition equations 36
3.6 Higher dimensions: the mixed model 37

4 Weighted least squares **38**
4.1 Weighted observation equations 38
4.1.1 Geometry 40
4.1.2 Application to adjustment problems 41
4.1.3 Higher dimensions 41
4.2 Weighted condition equations 41
4.3 Stochastics 44
4.4 Best Linear Unbiased Estimation (BLUE) 45

5 Geomatics examples **47**
5.1 A-Model: Adjustment of observation equations 47
5.1.1 Planar triangle 47
5.1.2 Distance Network 49
5.1.3 Distance and Direction Network (1) 55
5.1.4 Distance and Direction Network (2a) 58
5.1.5 Free Adjustment: Distance and Direction Network (2b) 62
5.1.6 Overconstrained adjustment: Distance, direction and angle network 65
5.1.7 Polynomial fit 72
5.2 B-Model: Adjustment of condition equations 78
5.2.1 Planar triangle 1 78
5.2.2 Planar triangle 2 79

5.3 Mixed model . . . 81
5.3.1 Straight line fit using A-model with pseudo observation equations . . . 81
5.3.2 Straight line fit using extended B-Model . . . 83
5.3.3 2D Similarity Transformation . . . 85
5.3.4 2D Affine Transformation Model I . . . 90
5.3.5 2D Affine Transformation Model II . . . 94
5.3.6 Ellipse fit under various restrictions . . . 97

6 Statistics **106**
6.1 Expectation of sum of squared residuals . . . 106
6.2 Basics . . . 107
6.3 Hypotheses . . . 108
6.4 Distributions . . . 110

7 Statistical Testing **113**
7.1 Global model test: a first approach . . . 113
7.2 Testing procedure . . . 115
7.3 DIA-Testprinciple . . . 121
7.4 Internal reliability . . . 122
7.5 External reliability . . . 125
7.6 Reliability: a synthesis . . . 126

8 Recursive estimation **128**
8.1 Partitioned model . . . 128
8.1.1 Batch / offline / Stapel / standard . . . 128
8.1.2 Recursive / sequential / real-time . . . 128
8.1.3 Recursive formulation . . . 129
8.1.4 Formulation using condition equations . . . 129
8.2 More general . . . 130

A Partitioning **131**
A.1 Inverse Partitioning Method (IPM) . . . 131
A.2 Inverse Partitioning Method: special case 1 . . . 131
A.3 Inverse Partitioning Method: special case 2 . . . 132

B Statistical Tables **134**
B.1 Standard Normal Distribution z . . . 134
B.2 Central χ^2-Distribution . . . 136
B.3 Non-central χ^2-Distribution . . . 139
B.4 Central t-Distribution . . . 140
B.5 Central F-Distribution . . . 142
B.6 Relation between F-Distribution and other distrubutions . . . 150

C Book recommendations and other material **151**
C.1 Scientific books . . . 151
C.2 Popular science books, literature . . . 153

C.3 Other material 154

1 Introduction

Ausgleichungs-rechnung

Adjustment theory deals with the optimal combination of redundant measurements together with the estimation of unknown parameters.

Teunissen, 2000

1.1 Adjustment theory –a first look

To understand the purpose of adjustment theory consider the following simple highschool example that is supposed to demonstrate how to solve for unknown quantities. In case 0 the price of apples and pears is determined after doing groceries twice. After that we will discuss more interesting shopping scenarios.

Case 0)

$$\begin{cases} 3 \text{ apples} + 4 \text{ pears} = 5.00\,€ \\ 5 \text{ apples} + 2 \text{ pears} = 6.00\,€ \end{cases}$$

$$\text{2 equations in 2 unknowns:} \quad \begin{cases} 5 = 3x_1 + 4x_2 \\ 6 = 5x_1 + 2x_2 \end{cases}$$

$$\text{as matrix-vector system:} \quad \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\text{linear algebra:} \quad y = Ax$$

The determinant of matrix A reads $\det A = 3 \cdot 2 - 5 \cdot 4 = -14$. Thus the above linear system can be inverted:

$$x = A^{-1}y \quad \Longleftrightarrow \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{-14} \begin{pmatrix} 2 & -4 \\ -5 & 3 \end{pmatrix} \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \\ 0.5 \end{pmatrix}$$

So each apple costs 1 € and each pear 50 cents. The price can be determined because there are as many unknowns (the price of apples and the price of pears) as there are observations (shopping twice). The square and regular matrix A is invertible.

Remark 1.1 (terminology) *The left-hand vector y contains the observations. The vector x contains the unknown parameters. The two vectors are linked through the* design matrix *A. The linear model $y = Ax$ is known as the model of* observation equations.

The following cases demonstrate that the idea of determining unknowns from observations is not as straightforward as may seem from the above example.

Case 1a)

If one buys twice as much apples and pears the second time, and if one has to pay twice as much as well, no new information is added to the system of linear equations

$$\left.\begin{aligned} 3\text{a} + 4\text{p} &= 5\,€ \\ 6\text{a} + 8\text{p} &= 10\,€ \end{aligned}\right\} \iff \begin{pmatrix} 5 \\ 10 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 6 & 8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

The matrix A has linearly dependent columns (and rows), i.e. it is singular. Correspondingly $\det A = 0$ and the inverse A^{-1} does not exist. The observations (5 € and 10 €) are *consistent*, but the vector x of unknowns (price per apple or pear) cannot be determined. This situation will return later with so-called *datum problems*. Seemingly trivial, case 1a) is of fundamental importance.

Case 1b)

Suppose the same shopping scenario as above, but now one needs to pay 8 € the second time.

$$y = \begin{pmatrix} 5 \\ 8 \end{pmatrix}$$

In this alternative scenario, the matrix is still singular and x cannot be determined. But worse still, the observations y are inconsistent with the linear model. Mathematically, they do not fulfil the compatibility conditions. In data analysis inconsistency is not necessarily a weakness. In fact, it may add information to the linear system. It might indicate observation errors (in y), for instance a miscalculation of the total grocery bill. Or it might indicate an error in the linear model: the prices may have changed in between, which leads to a different A.

Case 2)

We go back to the consistent and invertible case 0. Suppose a third combination of apples and pears gives an inconsistent result.

$$\begin{pmatrix} 5 \\ 6 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 5 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

The third row is inconsistent with $x_1 = 1, x_2 = \frac{1}{2}$ from case 0. But one can equally maintain that the first row is inconsistent with the second and third. In short, we have redundant and inconsistent information: the number of observations ($m = 3$) is larger than the number of unknowns ($n = 2$). Consequently, matrix A is not a square matrix.

Although a standard inversion is not possible anymore, redundancy is a positive characteristic in engineering disciplines. In data analysis redundancy provides information on the quality of the observations, it strengthens the estimation of the unknowns and allows us to perform statistical tests. Thus, redundancy provides a handle to quality control.

But obviously the inconsistencies have to be eliminated. This is done by spreading them out in an optimal way. This is the task of *adjustment*: to combine redundant and inconsistent data in an optimal way. Two main questions will be addressed in the first part of this course:

- How to combine inconsistent data optimally?

- Which criterion defines what optimal is?

Errors

The inconsistencies may be caused by model errors. If the green grocer changed his prices between two rounds of shopping we need to introduce new parameters. In surveying, however, the observation models are usually well-defined, e.g. the sum of angles in a plane triangle equals π. So usually the inconsistencies arise from observation errors. To make the linear system $y = Ax$ consistent again, we need to introduce an error vector e with the same dimension as the observation vector.

$$\underset{m\times 1}{y} = \underset{m\times n}{A}\;\underset{n\times 1}{x} + \underset{m\times 1}{e} \;. \tag{1.1}$$

Errors go under several names: inconsistencies, residuals, improvements, deviations, discrepancies, and so on.

Remark 1.2 (sign convention) *In many textbooks the error vector is put at the same side of the equation as the observations: $y + e = Ax$. Where to put the e-vector is rather a philosophical question. Practically, though, one should be aware of the definitions used, how the sign of e is defined.*

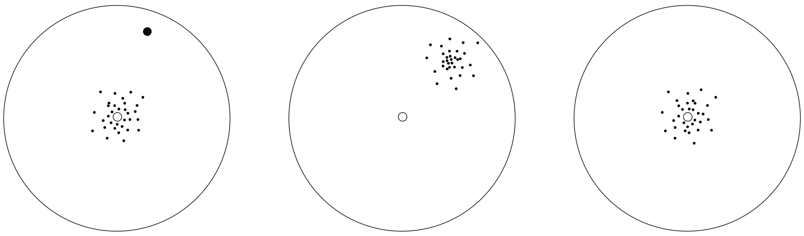
Three different types of errors are usually identified:

grober Fehler
i) *Gross error*, also known as blunder or outlier.

systematischer Fehler
ii) *Systematic error*, or bias.

Zufallsfehler
iii) *Random error*.

These types are visualized in fig. 1.1. In this figure, one can think of the marks left behind by the arrow points in a game of darts, in which one attempts to aim at the bull's eye. Whatever the type,



(a) gross error (b) systematic error (c) random error

Figure 1.1: Different types of errors.

Zufallsvariable
errors are stochastic quantities. Thus, the vector e is a (m-dimensional) *stochastic variable*. The vector of observations is consequently also a stochastic variable. Such quantities will be underlined, if necessary:

$$\underline{y} = Ax + \underline{e}\,.$$

Nevertheless, it will be assumed in the sequel that e is drawn from a distribution of random errors.

1.2 Historical development

The question how to combine redundant and inconsistent data has been treated in many different ways in the past. To compare the different approaches, the following mathematical framework is used:

$$\begin{aligned} &\text{observation model:} & y &= Ax \\ &\text{combination:} & \underset{n\times m}{L}\,\underset{m\times 1}{y} &= \underset{n\times m}{L}\,\underset{m\times n}{A}\,\underset{n\times 1}{x} \\ &\text{invert:} & x &= (LA)^{-1}Ly \\ & & &= By \end{aligned}$$

From a modern viewpoint matrix B is a *left-inverse* of A because $BA = I$. Note that such a left-inverse is not unique, as it depends on the choice of the combination matrix L.

Method of selected points – before 1750

A simple way out of the overdetermined problem is to select only so many observations ("points") as there are unknowns. The remaining unused observations may be used to validate the estimated result. This is the so-called method of selected points. Suppose one uses only the first n observations. Then:

$$\underset{n\times m}{L} = [\ \underset{n\times n}{I} \quad \underset{n\times(m-n)}{0}\ ]$$

The trouble with this approach, obviously, is the arbitrariness of the choice of n observations. There are $\binom{m}{n}$ choices.

From a modern perspective the method of selected points resembles the principle of *cross-validation*. The idea of this principle is to deliberately leave out a limited number of observations during the estimation and to use the estimated parameters to predict values for those observations that were left out. A comparison between actual and predicted observations provides information on the quality of the estimated parameters.

Method of averages – ca. 1750

In 1714 the British government offered the *Longitude Prize* for the precise determination of a ship's longitude. Tobias Mayer's[1] approach was to determine longitude, or rather time, through the motion of the moon. In the course of his investigations he needed to determine the libration of the moon through measurement to lunar surface (craters). This led him to overdetermined systems of observation equations:

$$\underset{27\times 1}{y} = \underset{27\times 3}{A}\,\underset{3\times 1}{x}$$

Mayer called them *equations of conditions*, which is, from today's view point, an unfortunate designation.

[1]Tobias Mayer (1723–1762) made the breakthrough that enabled the lunar distance method to become a practicable way of finding longitude at sea. As a young man, he displayed an interest in cartography and mathematics. In 1750, he was appointed professor in the Georg-August Academy in Göttingen, where he was able to devote more time to his interests in lunar theory and the longitude problem. From 1751 to 1755, he had an extensive correspondence with Leonhard Euler, whose work on differential equations enabled Mayer to calculate lunar distance tables.

Mayer's adjustment strategy:

- distribute the observations into three groups
- sum up the equations within each group
- solve the 3×3-system.

$$\underset{3\times 27}{L} = \begin{pmatrix} 1\ 1 \cdots 1 & 0\ 0\ \ 0\ \ 0 & 0\ 0\ \ 0\ \ 0 \\ 0\ 0\ \ 0\ \ 0 & 1\ 1 \cdots 1 & 0\ 0\ \ 0\ \ 0 \\ 0\ 0\ \ 0\ \ 0 & 0\ 0\ \ 0\ \ 0 & 1\ 1 \cdots 1 \end{pmatrix}$$

Mayer actually believed each aggregate of 9 observations to be "9 times more precise" than a single observation. Today we know that this should be $\sqrt{9} = 3$.

Euler's attempt – 1749

Leonhard Euler[2]

Background:

- Orbital motion of the Saturn under influence of Jupiter
- Stability of the solar system
- Prize (1748) of the Academy of Sciences, Paris

75 observations from the years 1582–1745; 6 unknowns $\Longrightarrow$ Given up!
Euler was mathematician $\longrightarrow$ "Error bounds"

Laplace's attempt – ca. 1787

Laplace[3]

Background: Saturn, too
Reformulated: 4 unknowns
Best data: 24 observations
Approach: like Mayer, but other combinations:

$$\underset{24\times 1}{y} = \underset{24\times 4}{A}\ \underset{4\times 1}{x}$$

$$\underset{4\times 24}{L}\ \underset{24\times 1}{y} = \underset{4\times 24}{L}\ \underset{24\times 4}{A}\ \underset{4\times 1}{x}$$

$$x = (LA)^{-1}Ly$$

[2] Euler (1707–1783) was a Swiss mathematician and physicist. He is considered to be one of the greatest mathematicians who ever lived. Euler was the first to use the term *function* (defined by Leibniz in 1694) to describe an expression involving various arguments; i.e. $y = F(x)$. He is credited with being one of the first to apply calculus to physics.

[3] Pierre-Simon, Marquis de Laplace (1749–1827) was a French mathematician and astronomer who put the final capstone on mathematical astronomy by summarizing and extending the work of his predecessors in his five volume Mécanique Céleste (Celestial Mechanics) (1799–1825). This masterpiece translated the geometrical study of mechanics used by Newton to one based on calculus, known as physical mechanics. He is also the discoverer of Laplace's equation and the Laplace transform, which appear in all branches of mathematical physics – a field he took a leading role in forming. He became count of the Empire in 1806 and was named a marquis in 1817 after the restoration of the Bourbons. Pierre-Simon Laplace was among the most influential scientists in history.

$$\underset{4\times 24}{L} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & 0 & 1 & 1 & 0 & 0 & -1 & 0 & 0 & 1 & 1 & 0 & 0 & -1 & 0 & 0 & 1 & 1 & 0 & -1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & -1 & -1 & 0 & 1 & 1 & 0 & 0 & -1 & -1 & 0 & 1 & 1 & 0 & 0 & -1 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}$$

Method of least absolute deviation – 1760

Roger Boscovich[4]

Ellipticity of the Earth
5 observations (Quito, Cape Town, Rome, Paris, Lapland)
2 unknowns

$$\begin{aligned} M(\varphi) &= \frac{a(1-e^2)}{(1-e^2\sin^2\varphi)^{\frac{3}{2}}} \\ &= a(1-e^2)(1+\frac{3}{2}e^2\sin^2\varphi+\ldots) \\ &\left| \begin{array}{l} M(0) = a(1-e^2) < a \\ M(\frac{\pi}{2}) = a\frac{1-e^2}{(1-e^2)^{\frac{3}{2}}} = \frac{a}{\sqrt{1-e^2}} > a \end{array} \right. \\ &= x_1 + \sin^2\varphi x_2 \end{aligned}$$

First attempt: All $\binom{5}{2} = \frac{5!}{2!(5-2)!} = \frac{5\cdot4\cdot3\cdot2\cdot1}{2\cdot1\cdot3\cdot2\cdot1} = 10$ combinations with 2 observations each.
$\Longrightarrow$ 10 systems of equations (2×2)
$\Longrightarrow$ 10 solutions
Comparison of results.

His result: gross variations of the ellipticity $\Longrightarrow$ reject the ellipsoidal hypothesis.

Second attempt: The mean deviation (or sum of deviations) should be zero:

$$\sum_{i=1}^{5} e_i = 0\,,$$

and the sum of absolute deviations should be minimum:

$$\sum_{i=1}^{5} |e_i| = \min\ .$$

This is an objective adjustment criterion, although its implementation is mathematically difficult. This is the approach of L_1-norm minimization.

Method of least squares – 1805

In 1805 Legendre[5] published his *method of least squares* (in French: *moindres carrés*). The name *least squares* refers to the fact the sum of square residuals is minimized. Legendre developed the method

Methode der kleinsten Quadrate

[4] Rudjer Josip Bošković aka. Roger Boscovich (1711–1787) was a Croatian Jesuit, a mathematician and an innovative physicist, he was active also in astronomy, nature philosophy and poetry as well as technician and geodesist.

[5] Adrien-Marie Legendre (1752–1833) was a French mathematician. He made important contributions to statistics, number theory, abstract algebra and mathematical analysis.

for the determination of orbits of comets and to derive the Earth ellipticity. As will be derived in the next chapter, the matrix L will be the transposed of the design matrix A:

$$
\begin{aligned}
& \mathcal{L} = \sum_{i=1}^{5} e_i^2 = e^{\mathsf{T}} e = (y - Ax)^{\mathsf{T}} (y - Ax) = \min_{\hat{x}} \\
\iff \quad & L = A^{\mathsf{T}} \\
\iff \quad & \underset{n\times 1}{\hat{x}} = (\underbrace{A^{\mathsf{T}}A}_{n\times n})^{-1} \underset{n\times m}{A^{\mathsf{T}}} \underset{m\times 1}{y}
\end{aligned}
$$

After Legendre's publication Gauss states that he already developed and used the method of least squares in 1794. He published his own theory only several years later. A bitter argument over the scientific priority broke out. Nowadays it is acknowledged that Gauss's claim of priority is very likely valid but that he refrained from publication because he found his results still premature.

2 Least squares adjustment

Legendre's method of least squares is actually not a method. Rather, it provides the criterion for the optimal combination of inconsistent data: combine the observations such that the sum of squared residuals is minimal. It was seen already that this criterion defines the combination matrix L:

$$Ly = LAx \quad \Longrightarrow \quad x = (LA)^{-1}Ly\,.$$

But what is so special about $L = A^\mathsf{T}$? In this chapter we will derive the equations of least squares adjustment from several mathematical viewpoints:

- *geometry*: smallest distance (Pythagoras)
- *linear algebra*: orthogonality between the optimal e and the columns of A: $A^\mathsf{T}e = 0$
- *calculus*: minimizing target function → differentiation
- *probability theory*: BLUE (Best Linear Unbiased Estimate)

These viewpoints are elucidated by a simple but fundamental example in which a distance is measured twice.

2.1 Adjustment with observation equations

We will start with the model of the introduction $y = Ax$. This is the *model of observation equations*, in which observations are linearly related to unknowns. vermittelnde Ausgleichung

Suppose that, in order to determine a certain distance, it is measured twice. Let the unknown distance be x and the observations y_1 and y_2:

$$\left.\begin{aligned} y_1 &= x \\ y_2 &= x \end{aligned}\right\} \quad \Longrightarrow \quad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} x \end{pmatrix} \quad \Longrightarrow \quad y = ax \tag{2.1}$$

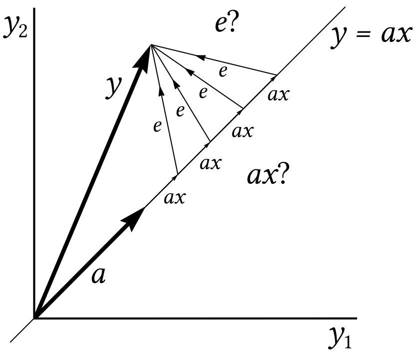
direkte Beobachtungen

If $y_1 = y_2$ the equations are consistent and the parameter x clearly solvable: $x = y_1 = y_2$. If, on the other hand, $y_1 \neq y_2$ the equations are inconsistent and x not solvable directly. Given a limited measurement precision the latter scenario will be more likely. Let's therefore take into account measurement errors e.

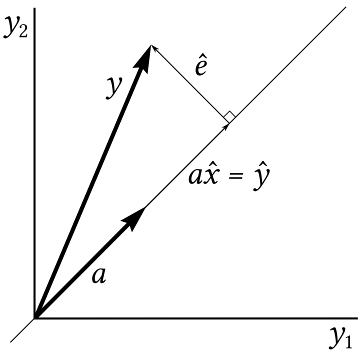
$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} x \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \quad \Longrightarrow \quad y = ax + e \tag{2.2}$$

A geometric view

The column vector a spans up a line $y = ax$ in $\mathbb{R}^2$. This line is the 1D model space or *range space* of A: $\mathcal{R}(A)$. Inconsistency of the observation vector means that y does not lie on this line. Instead, Spaltenraum there is some vector of discrepancies e that connects the observations to the line. Both this vector e and the point on the line, defined by the unknown parameter x, must be found, see the left panel of fig. 2.1. Adjustment of observations is about finding the optimal e and x. An intuitive choice for



(a) Inconsistent data: the observation vector y is not in the model space, i. e. not on the line spanned by a.



(b) Least squares adjustment means orthogonal projection of y onto the line ax. This guarantees the shortest e.

Figure 2.1

"optimality" is to make the vector e as short as possible. The shortest possible e is indicated by a hat: $\hat{e}$. The squared length $\hat{e}^\mathsf{T}\hat{e} = \sum_i \hat{e}_i^2$ is the smallest of all possible $e^\mathsf{T}e = \sum_i e_i^2$, which explains the name *least squares.* If $\hat{e}$ is determined, we will at the same time know the optimal $\hat{x}$.

How do we get the shortest e? The right panel of fig. 2.1 show that the shortest e is perpendicular to a:

$$\hat{e} \perp a$$

Subtracting $\hat{e}$ from the vector of observations y leads to the point $\hat{y} = a\hat{x}$ that is on the line and closest to y. This is the vector of adjusted observations. Being on the line means that $\hat{y}$ is consistent.

If we now substitute $\hat{e} = y - a\hat{x}$, the least squares criterion leads us subsequently to optimal estimates of x, y and e:

orthogonality $\hat{e} \perp a$	$a^\mathsf{T}\hat{e} = 0$	(2.3a)
	$a^\mathsf{T}(y - a\hat{x}) = 0$	(2.3b)
normal equations	$a^\mathsf{T}a\hat{x} = a^\mathsf{T}y$	(2.3c)
LS estimate of x	$\hat{x} = (a^\mathsf{T}a)^{-1}a^\mathsf{T}y$	(2.3d)
LS estimate of y	$\hat{y} = a\hat{x} \;=\; a(a^\mathsf{T}a)^{-1}a^\mathsf{T}y$	(2.3e)
LS estimate of e	$\hat{e} = y - \hat{y} \;=\; [I - a(a^\mathsf{T}a)^{-1}a^\mathsf{T}]y$	(2.3f)
sum square residuals	$\hat{e}^\mathsf{T}\hat{e} = y^\mathsf{T}[I - a(a^\mathsf{T}a)^{-1}a^\mathsf{T}]y$	(2.3g)

Exercise 2.1 *Call the matrix in square brackets P and convince yourself that the sum of squares of the residuals (the squared length of $\hat{e}$) in the last line indeed follows from the line above. Two things should be shown: that P is symmetric, and that $PP = P$.*

The least squares criterion leads us to the above algorithm. Indeed, the combination matrix reads $L = A^\mathsf{T}$.

A calculus view

Let us define the *Lagrangian* or *cost function*:

$$\mathcal{L}_a(x) = \frac{1}{2} e^\mathsf{T} e\,, \tag{2.4}$$

which is half of the sum of square residuals. Its graph would be a parabola. The factor $\frac{1}{2}$ shouldn't worry us. If we find the minimum $\mathcal{L}_a$, then any scaled version of it is also minimized. The task is now to find the $\hat{x}$ that minimizes the Lagrangian. With $e = y - ax$ we get the minimization problem:

$$\begin{aligned}\min_{\hat{x}} \mathcal{L}_a(x) &= \min_{\hat{x}} \frac{1}{2}(y - ax)^\mathsf{T}(y - ax) \\ &= \min_{\hat{x}} \left(\frac{1}{2} y^\mathsf{T} y - x a^\mathsf{T} y + \frac{1}{2} a^\mathsf{T} a x^2 \right)\,.\end{aligned}$$

The term $\frac{1}{2} y^\mathsf{T} y$ is just a constant that doesn't play a role in the minimization. The minimum occurs at the location where the derivative of $\mathcal{L}_a$ is zero (necessary condition):

$$\frac{\mathrm{d}\mathcal{L}_a}{\mathrm{d}x}(\hat{x}) = -a^\mathsf{T} y + a^\mathsf{T} a \hat{x} = 0\,.$$

The solution of this equation, which happens to be the normal equation (2.3c), is the $\hat{x}$ we are looking for:

$$\hat{x} = (a^\mathsf{T} a)^{-1} a^\mathsf{T} y\,.$$

To make sure that the derivative does not give us a maximum, we must check that the second derivative of $\mathcal{L}_a$ is positive at $\hat{x}$ (sufficiency condition):

$$\frac{\mathrm{d}^2\mathcal{L}_a}{\mathrm{d}x^2}(\hat{x}) = a^\mathsf{T} a > 0\,,$$

which is a positive constant for all x indeed.

Projectors

Figure 2.1 shows that the optimal, consistent $\hat{y}$ is obtained by an orthogonal projection of the original y onto the line ax. Mathematically this was translated by (2.3e) as:

$$\begin{aligned} & \hat{y} = a(a^\mathsf{T} a)^{-1} a^\mathsf{T} y && (2.5\text{a}) \\ \Longleftrightarrow \quad & \hat{y} = P_a y && (2.5\text{b}) \\ \text{with} \quad & P_a = a(a^\mathsf{T} a)^{-1} a^\mathsf{T}\,. && (2.5\text{c}) \end{aligned}$$

The matrix P_a is an orthogonal projector. It is an *idempotent* matrix, meaning:

$$P_a P_a = a(a^\mathsf{T} a)^{-1} a^\mathsf{T} a(a^\mathsf{T} a)^{-1} a^\mathsf{T} = P_a\,. \tag{2.6}$$

It projects onto the line ax along a direction orthogonal to a. With this projection in mind, the property $P_a P_a = P_a$ becomes clear: if a vector has been projected already, the second projection has no effect anymore.

Also (2.3f) can be abbreviated:

$$\hat{e} = y - P_a y = (I - P_a)\,y = P_a^\perp y\,,$$

which is also a projection. In order to give $\hat{e}$ the vector y is projected onto a line perpendicular to ax along the direction a. And, of course, $P_a^\perp$ is idempotent as well:

$$P_a^\perp P_a^\perp = (I - P_a)(I - P_a) = I - 2P_a + P_a P_a = I - P_a = P_a^\perp\,.$$

Moreover, the definition (2.5c) makes clear that P_a and $P_a^\perp$ are symmetric. Therefore the square sum of residuals (2.3g) could be simplified to:

$$\hat{e}^\mathsf{T}\hat{e} = y^\mathsf{T} P_a^{\perp\mathsf{T}} P_a^\perp y = y^\mathsf{T} P_a^\perp P_a^\perp y = y^\mathsf{T} P_a^\perp y\,.$$

At a more fundamental level the definition of the orthogonal projector $P_a^\perp = I - P_a$ can be recast into the equation:

$$I = P_a + P_a^\perp\,.$$

zerlegen

Thus, we can *decompose* every vector, say z, into two components: one in component in a subspace defined by P_a, the other mapped onto a subspace by $P_a^\perp$:

$$z = Iz = \left(P_a + P_a^\perp\right) z = P_a z + P_a^\perp z\,.$$

In the case of LS adjustment, the subspaces are defined by the range space $\mathcal{R}(a)$ and its orthogonal complement $\mathcal{R}(a)^\perp$:

$$y = P_a y + P_a^\perp y = \hat{y} + \hat{e}\,,$$

which is visualized in fig. 2.1.

Numerical example

With $a = (1\ 1)^\mathsf{T}$ we will follow the steps from (2.3):

$$\begin{array}{rcll}
(a^\mathsf{T} a)\hat{x} = a^\mathsf{T} y & \longleftrightarrow & 2\hat{x} = y_1 + y_2 & \\[2ex]
\hat{x} = (a^\mathsf{T} a)^{-1} a^\mathsf{T} y & \longleftrightarrow & \hat{x} = \dfrac{1}{2}(y_1 + y_2) & \text{(average)} \\[2ex]
\hat{y} = a(a^\mathsf{T} a)^{-1} a^\mathsf{T} y & \longleftrightarrow & \begin{pmatrix}\hat{y}_1\\ \hat{y}_2\end{pmatrix} = \dfrac{1}{2}\begin{pmatrix} y_1 + y_2\\ y_1 + y_2\end{pmatrix} & \\[2ex]
\hat{e} = y - \hat{y} & \longleftrightarrow & \begin{pmatrix}\hat{e}_1\\ \hat{e}_2\end{pmatrix} = \dfrac{1}{2}\begin{pmatrix} y_1 - y_2\\ -y_1 + y_2\end{pmatrix} & \text{(error distribution)} \\[2ex]
\hat{e}^\mathsf{T}\hat{e} & \longleftrightarrow & \dfrac{1}{2}(y_1 - y_2)^2 & \text{(least squares)}
\end{array}$$

Exercise 2.2 *Verify that the projectors are*

$$P_a = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad P_a^\perp = I - P_a = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

and check the equations $\hat{y} = P_a y$ and $\hat{e} = P_a^\perp y$ with the numerical results above.

2.2 Adjustment with condition equations

In the ideal case, in which the measurements y_1 and y_2 are without error, both observations would be equal: $y_1 = y_2$ or $y_1 - y_2 = 0$. In matrix notation:

$$\begin{pmatrix} 1 & -1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = 0 \quad \Longrightarrow \quad \underset{1\times 2}{b^\mathsf{T}} \underset{2\times 1}{y} = \underset{1\times 1}{0} \; . \tag{2.7}$$

In reality, though, both observations do contain errors, i.e. they are not equal: $y_1 - y_2 \neq 0$ or $b^\mathsf{T} y \neq 0$. Instead of 0 one would obtain a *misclosure* w. If we recast the observation equation into $y - e = ax$, it is clear that it is $(y - e)$ that has to obey the above condition: Widerspruch

$$b^\mathsf{T}(y - e) = 0 \quad \Longrightarrow \quad w := b^\mathsf{T} y = b^\mathsf{T} e \, . \tag{2.8}$$

In this *condition equation* the vector e is unknown. The task of adjustment according to the model of condition equations is to find the smallest possible e that fulfills the condition (2.8). At this stage, the model of condition equations does not involve any kind of parameters x. Bedingungs-gleichung

A geometric view

The condition (2.8) describes a line with normal vector b that goes through the point y. This line is the set of all possible vectors e. We are looking for the shortest e, i.e. the point closest to the origin. Figure 2.2 makes it clear that $\hat{e}$ is perpendicular to the line $b^\mathsf{T} e = w$. So $\hat{e}$ lies on a line through b.

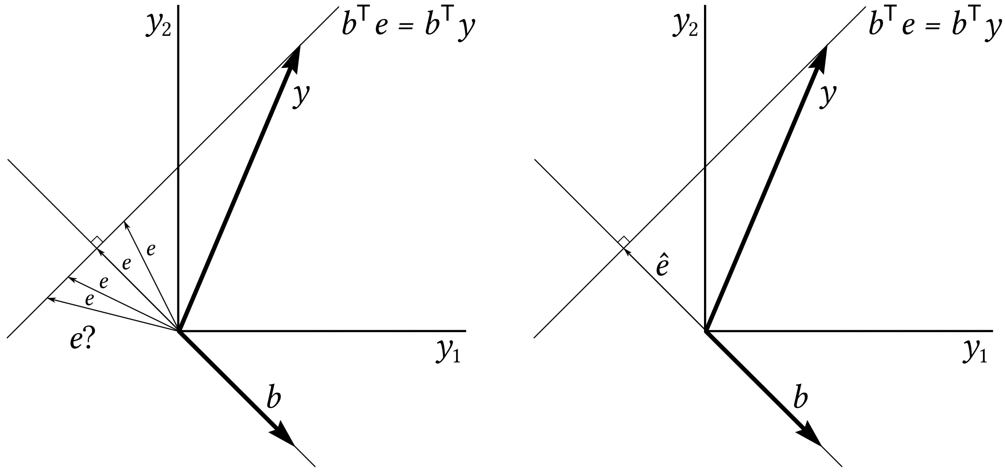
Geometrically, $\hat{e}$ is achieved by projecting y onto a line through b. Knowing the definition of the projectors from the previous section, we here define the following *estimates* by using the projector P_b: Schätzungen

$$\hat{e} = P_b y = b(b^\mathsf{T} b)^{-1} b^\mathsf{T} y = b(b^\mathsf{T} b)^{-1} w \tag{2.9a}$$

$$\begin{aligned} \hat{y} &= y - \hat{e} = y - b(b^\mathsf{T} b)^{-1} b^\mathsf{T} y \\ &= [I - b(b^\mathsf{T} b)^{-1} b^\mathsf{T}] y = P_b^\perp y \end{aligned} \tag{2.9b}$$

$$\hat{e}^\mathsf{T}\hat{e} = y^\mathsf{T} P_b y = y^\mathsf{T} b(b^\mathsf{T} b)^{-1} b^\mathsf{T} y \tag{2.9c}$$

Exercise 2.3 *Confirm that the orthogonal projector P_b is idempotent and verify that the equation for $\hat{e}^\mathsf{T}\hat{e}$ is correct.*



(a) The condition equation describes a line in $\mathbb{R}^2$, perpendicular to b and going through y. We are looking for a point e on this line.

(b) Least squares adjustment with condition equations means orthogonal projection of y onto the line through b. This guarantees the shortest e.

Figure 2.2

Numerical example

With $b^\mathsf{T} = \begin{pmatrix} 1 & -1 \end{pmatrix}$ we get

$$
\begin{aligned}
b^\mathsf{T} b = 2 \quad &\Longrightarrow \quad (b^\mathsf{T} b)^{-1} = \frac{1}{2} \\
P_b = b(b^\mathsf{T} b)^{-1} b^\mathsf{T} &= \frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \begin{pmatrix} 1 & -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \\
\Longrightarrow \quad \hat{e} = P_b y &= \frac{1}{2} \begin{pmatrix} y_1 - y_2 \\ -y_1 + y_2 \end{pmatrix} \\
P_b^\perp = I - P_b &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \\
\Longrightarrow \quad \hat{y} = P_b^\perp y &= \frac{1}{2} \begin{pmatrix} y_1 + y_2 \\ y_1 + y_2 \end{pmatrix}
\end{aligned}
$$

These results for $\hat{y}$ and $\hat{e}$ are the same as those for the adjustment with observation equations. The estimator $\hat{y}$ describes the mean of the two observations, whereas the estimator $\hat{e}$ distributes the inconsistencies equally. Also note that $P_b = P_a^\perp$ and vice versa.

A calculus view

Alternatively we can again determine the optimal e by minimizing the target function $\mathcal{L}_b(e) = e^\mathsf{T} e$, but now under the condition $b^\mathsf{T}(y - e) = 0$:

$$\min_{\hat{e}} \mathcal{L}_b(e) = e^\mathsf{T} e \qquad \text{under} \quad b^\mathsf{T}(y - e) = 0\,, \tag{2.10a}$$

$$\min_{\hat{e},\hat{\lambda}} \mathcal{L}_b(e, \lambda) = \frac{1}{2} e^\mathsf{T} e + \lambda^\mathsf{T}(b^\mathsf{T} y - b^\mathsf{T} e)\,. \tag{2.10b}$$

The main trick here – due to Lagrange – is to not consider the condition as a constraint or limitation of the minimization problem. Instead, the minimization problem is extended. To be precise, the condition is added to the original cost function, multiplied by a factor λ. Such factors are called Lagrangian multipliers. In case of more than one condition, each gets its own multiplier. The target function $\mathcal{L}_b$ is now a function of e and λ.

The minimization problem now exists in finding the $\hat{e}$ and $\hat{\lambda}$ that minimize the extended $\mathcal{L}_b$. Thus we need to derive the partial derivatives of $\mathcal{L}_b$ towards e and λ. Next, we impose the conditions that these partial derivatives are zero when evaluated in $\hat{e}$ and $\hat{\lambda}$.

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial e}(\hat{e}, \hat{\lambda}) = 0 \quad &\Longrightarrow \quad \hat{e} - b\hat{\lambda} = 0 \\ \frac{\partial \mathcal{L}}{\partial \lambda}(\hat{e}, \hat{\lambda}) = 0 \quad &\Longrightarrow \quad b^\mathsf{T} y - b^\mathsf{T} \hat{e} = 0 \end{aligned}$$

In matrix terms, the minimization problem leads to:

$$\begin{pmatrix} I & -b \\ -b^\mathsf{T} & 0 \end{pmatrix} \begin{pmatrix} \hat{e} \\ \hat{\lambda} \end{pmatrix} = \begin{pmatrix} 0 \\ -b^\mathsf{T} y \end{pmatrix}\,. \tag{2.11}$$

Because of the extension of the original minimization problem, this system is square. It might be inverted in a straightforward manner, see also A.1. Instead, we will solve it stepwise. First, rewrite the first line:

$$\hat{e} - b\hat{\lambda} = 0 \quad \Longrightarrow \quad \hat{e} = b\hat{\lambda}\,.$$

This result is then used to eliminate $\hat{e}$ in the second line:

$$b^\mathsf{T} y - b^\mathsf{T} b\hat{\lambda} = 0\,,$$

which is solved by:

$$\hat{\lambda} = (b^\mathsf{T} b)^{-1} b^\mathsf{T} y\,.$$

With this result we go back to the first line:

$$\hat{e} - b(b^\mathsf{T} b)^{-1} b^\mathsf{T} y = 0\,,$$

which is finally solved by:

$$\hat{e} = b(b^\mathsf{T} b)^{-1} b^\mathsf{T} y = P_b y\,.$$

This is the same estimator $\hat{e}$ as (2.9a).

2.3 Synthesis

Both the calculus and geometric approach provide the same LS estimators. This is due to

$$P_a = P_b^\perp \quad \text{and} \quad P_b = P_a^\perp \,,$$

as can be seen in fig. 2.3. The deeper reason is that a is perpendicular to b:

$$b^\mathsf{T} a = \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0 \,, \tag{2.12}$$

which fundamentally connects the model with observation equations to the model with condition equations. Starting with the observation equation, and applying the orthogonality, one ends up with the condition equation:

$$y = ax + e \quad \xrightarrow{b^\mathsf{T}} \quad b^\mathsf{T} y = b^\mathsf{T} ax + b^\mathsf{T} e \quad \xrightarrow{b^\mathsf{T} a = 0} \quad b^\mathsf{T} y = b^\mathsf{T} e \,.$$

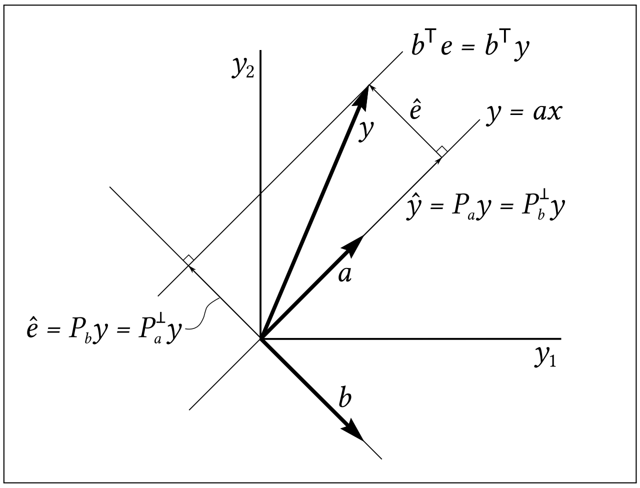


Figure 2.3: Least squares adjustment with observation equations and with condition equations in terms of the projectors P_a and P_b.

3 Generalizations

In this chapter we will apply several generalizations. First we will take the LS adjustment problems to higher dimensions. What we will basically do is replace the vector a by an $(m \times n)$ matrix A and replace the vector b by an $(m \times (m-n))$ matrix B. The basic structure of the projectors and estimators will remain the same.

Moreover, we need to be able to formulate the 2 LS problems with constant terms:

$$y = Ax + a_0 + e \qquad \text{and} \qquad B^\mathsf{T}(y - e) = b_0 \,.$$

Next, we will deal with nonlinear observation equations and nonlinear condition equations. This will involve linearization, the use of approximate values, and iteration.

We will also touch upon the datum problem, which arises if A contains dependent columns. Mathematically we have $\operatorname{rank} A < n$ so that the normal matrix has $\det A^\mathsf{T} A = 0$ and is not invertible.

At the end we will merge both models in order to establish the so-called general model of adjustment theory.

3.1 Higher dimensions: the A-model (observation equations)

The vector of observations y, the vector of inconsistencies e and their respective LS-estimators will be $(m \times 1)$ vectors. The vector x will contain n unknown parameters. Thus the redundancy, that is the number of redundant observations, is:

$$\text{redundancy:} \quad r = m - n \,.$$

Geometry

$y = Ax + e$ is the multidimensional extension of $y = ax + e$ with given (reduced) *vector of observations* y.

Absolutglied-vektor

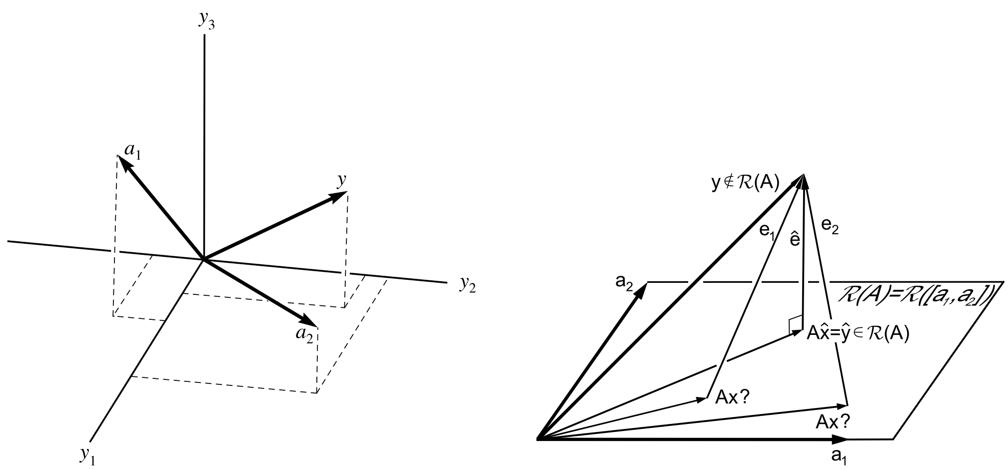
We split A in its n column vectors $\underset{m\times 1}{a_i}$, $i = 1, \ldots, n$

$$\underset{m\times n}{A} = [\ \underset{m\times 1}{a_1}\ ,\ \underset{m\times 1}{a_2}\ ,\ \underset{m\times 1}{a_3}\ , \ldots ,\ \underset{m\times 1}{a_n}\ ]$$

$$\underset{m\times 1}{y} = \sum_{i=1}^{n} \underset{m\times 1}{a_i}\ \underset{1\times 1}{x_i} + \underset{m\times 1}{e}\ ,$$

which span an n-dimensional vector space as a subspace of $\mathbf{E}^m$.

Example: $m = 3$, $n = 2$ ($\underset{m\times 1}{y}$ spans an $\mathbf{E}^3$)



(a) The vectors y, a_1 and a_2 all lie in $\mathbb{R}^3$.

(b) To see that y is inconsistent, the space spanned by a_1 and a_2, is shown as the base plane. It is clear that the observation vector is not in this plane $y \notin \mathcal{R}(A)$, i.e. y cannot be written as a linear combination of a_1 and a_2.

Figure 3.1

$$
\begin{aligned}
\hat{e} &= P_A^{\perp} y = [I - A(A^\mathsf{T}A)^{-1}A^\mathsf{T}]y \\
\hat{y} &= P_A y = A(A^\mathsf{T}A)^{-1}A^\mathsf{T}y = A\hat{x} \\
\hat{x} &= (A^\mathsf{T}A)^{-1}A^\mathsf{T}y
\end{aligned}
$$

genau dann, wenn

$(A^\mathsf{T}A)^{-1}$ exists *iff* $\operatorname{rank} A = n = \operatorname{rank}(A^\mathsf{T}A)$.

Calculus

$$
\begin{aligned}
\mathcal{L}_A(x) &= \frac{1}{2}e^\mathsf{T}e \\
&= \frac{1}{2}(y - Ax)^\mathsf{T}(y - Ax) \\
&= \frac{1}{2}y^\mathsf{T}y - \frac{1}{2}y^\mathsf{T}Ax - \frac{1}{2}x^\mathsf{T}A^\mathsf{T}y + \frac{1}{2}x^\mathsf{T}A^\mathsf{T}Ax \quad \xrightarrow{x} \min
\end{aligned}
$$

$$
\frac{\partial \mathcal{L}}{\partial x}(\hat{x}) = 0 \quad \Longrightarrow \quad \hat{e} = y - \hat{y} = [I - A(A^\mathsf{T}A)^{-1}A^\mathsf{T}]y = P_A^{\perp}y
$$

$P_A^\perp$ idempotent?

$$\begin{aligned}P_A^\perp P_A^\perp &= [I - A(A^\mathsf{T}A)^{-1}A^\mathsf{T}][I - A(A^\mathsf{T}A)^{-1}A^\mathsf{T}]\\ &= I - 2A(A^\mathsf{T}A)^{-1}A^\mathsf{T} + A(A^\mathsf{T}A)^{-1}\underbrace{A^\mathsf{T}A(A^\mathsf{T}A)^{-1}}_{=I}A^\mathsf{T}\\ &= I - A(A^\mathsf{T}A)^{-1}A^\mathsf{T}\\ &= P_A^\perp\end{aligned}$$

$$\hat{y} = P_A y = A(A^\mathsf{T}A)^{-1}A^\mathsf{T}y$$

Example: height network

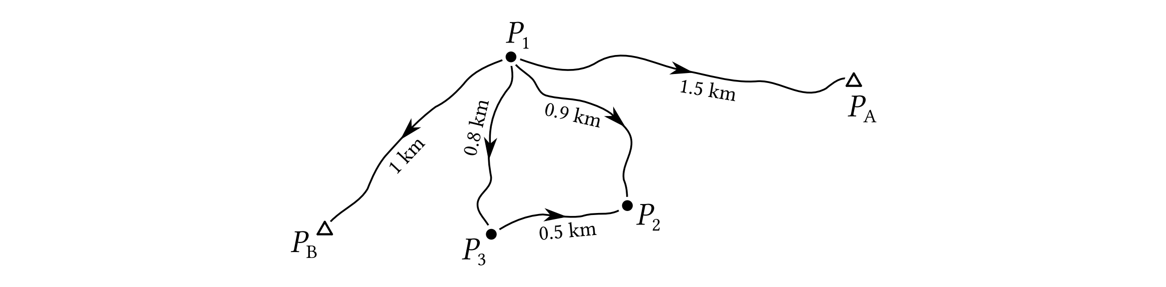


Figure 3.2: Height network with distances between points.

$$h_{1B} = H_B - H_1 + e_{1B}$$

$$h_{13} = H_3 - H_1 + e_{13}$$

$$h_{12} = H_2 - H_1 + e_{12}$$

$$h_{32} = H_2 - H_3 + e_{32}$$

$$h_{1A} = H_A - H_1 + e_{1A}$$

$\Delta h^\mathsf{T} := [h_{1B}, h_{13}, h_{12}, h_{32}, h_{1A}]$ vector of levelled height differences
H_1, H_2, H_3 unknown heights of points P_1, P_2, P_3
H_A, H_B given bench marks

In matrix notation:

$$\begin{pmatrix} h_{1B}\\ h_{13}\\ h_{12}\\ h_{32}\\ h_{1A}\end{pmatrix} = \begin{pmatrix} -1 & 0 & 0\\ -1 & 0 & 1\\ -1 & 1 & 0\\ 0 & 1 & -1\\ -1 & 0 & 0\end{pmatrix}\begin{pmatrix} H_1\\ H_2\\ H_3\end{pmatrix} + \begin{pmatrix} H_B\\ 0\\ 0\\ 0\\ H_A\end{pmatrix} + \begin{pmatrix} e_{1B}\\ e_{13}\\ e_{12}\\ e_{32}\\ e_{1A}\end{pmatrix}$$

$$\begin{pmatrix} h_{1B} - H_B\\ h_{13}\\ h_{12}\\ h_{32}\\ h_{1A} - H_A\end{pmatrix} = \begin{pmatrix} -1 & 0 & 0\\ -1 & 0 & 1\\ -1 & 1 & 0\\ 0 & 1 & -1\\ -1 & 0 & 0\end{pmatrix}\begin{pmatrix} H_1\\ H_2\\ H_3\end{pmatrix} + \begin{pmatrix} e_{1B}\\ e_{13}\\ e_{12}\\ e_{32}\\ e_{1A}\end{pmatrix} \qquad \sim \quad \underset{5\times1}{y} = \underset{5\times3}{A}\,\underset{3\times1}{x} + \underset{5\times1}{e}$$

3.2 The datum problem

So far we have disregarded the fact that the matrix $A^\mathsf{T}A$ might not be invertible because it is rank deficient. From matrix algebra it is known that the rank of the normal equation matrix $N := A^\mathsf{T}A$, $\operatorname{rank} N$, equals the the rank of A, $\operatorname{rank} A$. If it should happen now that – for some reason – matrix A is rank deficient, then the normal equation matrix $N{=}A^\mathsf{T}A$ cannot be inverted. The following statements are equivalent:

- Matrix $\underset{m\times n}{A}$ rank deficient ($\operatorname{rank} A < n$),
- A has linear dependent columns,
- $Ax = 0$ has non-trivial solution $x_{\text{hom}} \neq 0$, i.e. the null space $\mathcal{N}(A)$ of A is not empty,
- $\det(A^\mathsf{T}A) = 0$,
- $A^\mathsf{T}A$ has zero eigenvalues.

Let us investigate this problem of rank deficiency of A and N using levelling observations between points P_1, P_2 and P_3 of the height network shown in fig. 3.2.

$$\left.\begin{aligned} h_{12} &= H_2 - H_1 \\ h_{13} &= H_3 - H_1 \\ h_{32} &= H_2 - H_3 \end{aligned}\right\} \quad \Longrightarrow \quad \begin{pmatrix} h_{12} \\ h_{13} \\ h_{32} \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} H_1 \\ H_2 \\ H_3 \end{pmatrix}$$

$$\Longrightarrow \quad \underset{3\times 1}{y} = \underset{3\times 3}{A}\ \underset{3\times 1}{x}$$

- $m = 3$, $n = 3$, $\operatorname{rank} A = 2 \quad \Longrightarrow \quad d = n - \operatorname{rank} A = 1 \quad \Longrightarrow \quad r = m - (n - d) = 1$,
- $\det A = -1 \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix}, -(-1) \begin{vmatrix} 1 & 0 \\ 1 & -1 \end{vmatrix} = 1 + (-1) = 0$,
- $\Longrightarrow A$ and $N = A^\mathsf{T}A$ are not invertible,
- $d := \dim \mathcal{N}(A) > 0$,
- $Ax = 0$ has a nontrivial solution $\quad \Longrightarrow \quad$ homogeneous solution $x_{\text{hom}} \neq 0$.

$\Longrightarrow x + \lambda x_{\text{hom}}$ is a solution of $y = Ax$ because

$$A\,(x + \lambda x_{\text{hom}}) = Ax + \lambda \underbrace{Ax_{\text{hom}}}_{=0} = Ax = y$$

is fullfilled.

Interpretation:

- Unknown heights can be changed by an arbitrary constant height shift without affecting the observations.
- Observed height differences are not sensitive to the null space $\mathcal{N}(A)$.

Solution approach 1: reduce solution space

- Fix $d = \dim \mathcal{N}(A)$ unknowns and eliminate corresponding columns in A so that the rank of A, $\operatorname{rank} A = n - d$, is full.
- Move fixed unknowns to the observation vector, e.g. fix H_1:

$$\Longrightarrow \quad \begin{pmatrix} h_{12} + H_1 \\ h_{13} + H_1 \\ h_{32} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} H_2 \\ H_3 \end{pmatrix}$$

Solution approach 2: augment solution space

Augment solution space by adding $d = \dim \mathcal{N}(A)$ constraints, e.g.

$$H_1 = 0 \quad \Longrightarrow \quad \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} H_1 \\ H_2 \\ H_3 \end{pmatrix} = 0 \qquad \sim \quad \underset{d\times n}{D^\mathsf{T}} \; \underset{n\times 1}{x} = \underset{d\times 1}{c}$$

In order to remove the rank deficiency of A, matrix D^T must be chosen in such a way that

$$\operatorname{rank}\left([\, \underset{n\times m}{A^\mathsf{T}} \mid \underset{n\times d}{D} \,] \right) = n \,.$$

$AD = 0$, however is not required. As an example, $D^\mathsf{T} = [1, -1, 0]$ is not permitted. The approach of augmenting the solution space is far more flexible as compared to approach 1: no changes of original quantities y, A are necessary. Even curious constraints are allowed as long as datum deficiency is resolved. However, we are faced with the constrained Lagrangian

$$\begin{aligned}
\mathcal{L}_D(x, \lambda) &= \frac{1}{2} e^\mathsf{T} e + \lambda (D^\mathsf{T} x - c) \\
&= \frac{1}{2} y^\mathsf{T} y - y^\mathsf{T} A x + \frac{1}{2} x^\mathsf{T} A^\mathsf{T} A x + \lambda (D^\mathsf{T} x - c) \\
\frac{\partial \mathcal{L}_D}{\partial x} &= -A^\mathsf{T} y + A^\mathsf{T} A x + D\lambda = 0 \\
\frac{\partial \mathcal{L}_D}{\partial \lambda} &= D^\mathsf{T} x - c = 0
\end{aligned}$$

$$\Longrightarrow \quad \underset{(n+d)\times(n+d)}{\begin{pmatrix} A^\mathsf{T} A & D \\ D^\mathsf{T} & 0 \end{pmatrix}} \underset{(n+d)\times 1}{\begin{pmatrix} \hat{x} \\ \hat{\lambda} \end{pmatrix}} = \begin{pmatrix} A^\mathsf{T} y \\ c \end{pmatrix} \quad \Longrightarrow \quad M\hat{z} = v$$

E.g.

$$A = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \quad \Longrightarrow \quad A^\mathsf{T}A = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

$$M = \begin{pmatrix} 2 & 1 & -1 & 1 \\ 1 & 2 & -1 & 0 \\ -1 & -1 & 2 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$\det M = -1 \cdot \det \begin{pmatrix} 1 & 2 & -1 \\ -1 & -1 & 2 \\ 1 & 0 & 0 \end{pmatrix} = -1 \cdot 1 \cdot \det \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = -3$$

$$\Longrightarrow M \text{ regular} \quad \Longrightarrow \quad \hat{z} = M^{-1}v$$

$$\hat{x} = N^{-1} \left\langle A^\mathsf{T}y + Dc - \left\{ D(D^\mathsf{T}N^{-1}D)^{-1} \left[D^\mathsf{T}N^{-1}A^\mathsf{T}y + (D^\mathsf{T}N^{-1}D - I)c \right] \right\} \right\rangle$$

$$N := A^\mathsf{T}A + DD^\mathsf{T}$$

3.3 Linearization of non-linear observation equations

General 1-D-formulation

The functional model

$$y = f(x)\,,$$

expressed by Taylor's theorem, becomes

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n \\ &= f(x_0) + \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x_0} (x - x_0) + \underbrace{\frac{1}{2} \left.\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}\right|_{x_0} (x - x_0)^2 + \ldots}_{\text{negligible if } x - x_0 \text{ small}} \end{aligned}$$

Substracting $f(x_0)$ yields

$$f(x) - f(x_0) = y - y_0 = \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x_0} (x - x_0) + \ldots$$

$$\underbrace{\Delta y = \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_0 (\Delta x)}_{\text{linear model}} + \underbrace{O(\Delta x^2)}_{\substack{\text{terms of higher order} \\ = \text{model errors}}}$$

with $\Delta x := x - x_0$ and $\Delta y := y - y_0$.

General multi-D formulation

$$y_i = f_i(x_j), \quad i = 1,\dots,m;\ j = 1,\dots,n$$
$$x_{j,0} \longrightarrow y_{i,0} = f_i(x_{j,0})$$

$$\begin{aligned}
\Delta y_1 &= \left.\frac{\partial f_1}{\partial x_1}\right|_0 \Delta x_1 + \left.\frac{\partial f_1}{\partial x_2}\right|_0 \Delta x_2 + \dots + \left.\frac{\partial f_1}{\partial x_n}\right|_0 \Delta x_n \\
\Delta y_2 &= \left.\frac{\partial f_2}{\partial x_1}\right|_0 \Delta x_1 + \left.\frac{\partial f_2}{\partial x_2}\right|_0 \Delta x_2 + \dots + \left.\frac{\partial f_2}{\partial x_n}\right|_0 \Delta x_n \\
&\vdots \\
\Delta y_m &= \left.\frac{\partial f_m}{\partial x_1}\right|_0 \Delta x_1 + \left.\frac{\partial f_m}{\partial x_2}\right|_0 \Delta x_2 + \dots + \left.\frac{\partial f_m}{\partial x_n}\right|_0 \Delta x_n .
\end{aligned}$$

Terms of second order and higher have been neglected.

$$\Longrightarrow \begin{pmatrix} \Delta y_1 \\ \Delta y_2 \\ \vdots \\ \Delta y_m \end{pmatrix} = \underbrace{\left.\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}\right|_0}_{\text{Jacobian matrix } A} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_n \end{pmatrix} \qquad \sim \qquad \Delta y = A(x_0)\,\Delta x$$

Planar distance observation:

$$s_{ij} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \quad \xrightarrow{?} \quad y = Ax$$

answer: linearize, Taylor series expansion

Linearization of planar distance observation equation (given Taylor point of expansion is x_i^0, y_i^0, x_j^0, y_j^0 = approximate values of unknown point coordinates); explicit differentiation

$$\underset{\text{“measured”}}{s_{ij}} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} = \sqrt{x_{ij}^2 + y_{ij}^2}$$

$$x_i = x_i^0 + \Delta x_i, \quad y_i = y_i^0 + \Delta y_i,$$
$$x_j = x_j^0 + \Delta x_j, \quad y_j = y_j^0 + \Delta y_j$$

$$\begin{aligned}
s_{ij} &= \sqrt{\left(x_j^0 + \Delta x_j - \left(x_i^0 + \Delta x_i\right)\right)^2 + \left(y_j^0 + \Delta y_j - \left(y_i^0 + \Delta y_i\right)\right)^2} \\
&= \underbrace{\sqrt{\left(x_j^0 - x_i^0\right)^2 + \left(y_j^0 - y_i^0\right)^2}}_{= s_{ij}^0 \ \text{(distance from approximate coordinates)}} + \left.\frac{\partial s_{ij}}{\partial x_i}\right|_0 \Delta x_i + \left.\frac{\partial s_{ij}}{\partial x_j}\right|_0 \Delta x_j + \left.\frac{\partial s_{ij}}{\partial y_i}\right|_0 \Delta y_i + \left.\frac{\partial s_{ij}}{\partial y_j}\right|_0 \Delta y_j
\end{aligned}$$

$$\frac{\partial s_{ij}}{\partial x_i} = \frac{\partial s_{ij}}{\partial x_{ij}}\frac{\partial x_{ij}}{\partial x_i} = \frac{1}{2}\frac{1}{\sqrt{x_{ij}^2 + y_{ij}^2}} 2x_{ij}\,(-1) = -\frac{x_j - x_i}{s_{ij}}$$

$$\frac{\partial s_{ij}}{\partial x_j} = +\frac{x_j - x_i}{s_{ij}}, \qquad \frac{\partial s_{ij}}{\partial y_i} = -\frac{y_j - y_i}{s_{ij}}, \qquad \frac{\partial s_{ij}}{\partial y_j} = +\frac{y_j - y_i}{s_{ij}}$$

$$\Longrightarrow \Delta s_{ij} := \underbrace{s_{ij} - s_{ij}^0}_{\text{“reduced observation”}} = \begin{pmatrix} -\frac{x_j^0 - x_i^0}{s_{ij}^0} & -\frac{y_j^0 - y_i^0}{s_{ij}^0} & \frac{x_j^0 - x_i^0}{s_{ij}^0} & \frac{y_j^0 - y_i^0}{s_{ij}^0} \end{pmatrix} \begin{pmatrix} \Delta x_i \\ \Delta y_i \\ \Delta x_j \\ \Delta y_j \end{pmatrix}$$

$$\Delta y = A(x_0)\,\Delta x$$

Sometimes it is more convenient to use implicit differentiation within the linearization of observation equations.

Depart from $s_{ij}^2 = \left(x_j - x_i\right)^2 + \left(y_j - y_i\right)^2$ instead from s_{ij} and calculate the total differential:

$$\cancel{2} s_{ij}\,\mathrm{d}s_{ij} = \cancel{2}\left(x_j - x_i\right)\left(\mathrm{d}x_j - \mathrm{d}x_i\right) + \cancel{2}\left(y_j - y_i\right)\left(\mathrm{d}y_j - \mathrm{d}y_i\right)$$

Solve for $\mathrm{d}s_{ij}$, introduce approximate value and switch from $\mathrm{d} \longrightarrow \Delta$:

$$\Delta s_{ij} := s_{ij} - s_{ij}^0 = \frac{x_j^0 - x_i^0}{s_{ij}^0}\left(\Delta x_j - \Delta x_i\right) + \frac{y_j^0 - y_i^0}{s_{ij}^0}\left(\Delta y_j - \Delta y_i\right)$$

Grid bearings:

$$T_{ij} = \arctan\frac{x_j - x_i}{y_j - y_i}$$

$\Longrightarrow$ Linearized grid bearing observation equation:

$$\begin{aligned} T_{ij} &= T_{ij}^0 + \frac{1}{1 + \left(\frac{x_j^0 - x_i^0}{y_j^0 - y_i^0}\right)^2}\left(-\frac{1}{y_j^0 - y_i^0}\Delta x_i + \frac{x_j^0 - x_i^0}{(y_j^0 - y_i^0)^2}\Delta y_i + \frac{1}{y_j^0 - y_i^0}\Delta x_j - \frac{x_j^0 - x_i^0}{(y_j^0 - y_i^0)^2}\Delta y_j\right) \\ &= T_{ij}^0 + \frac{(y_j^0 - y_i^0)^2}{(s_{ij}^0)^2}\left(-\frac{1}{y_j^0 - y_i^0}\Delta x_i + \frac{x_j^0 - x_i^0}{(y_j^0 - y_i^0)^2}\Delta y_i + \frac{1}{y_j^0 - y_i^0}\Delta x_j - \frac{x_j^0 - x_i^0}{(y_j^0 - y_i^0)^2}\Delta y_j\right) \\ &= T_{ij}^0 - \frac{y_j^0 - y_i^0}{(s_{ij}^0)^2}\Delta x_i + \frac{x_j^0 - x_i^0}{(s_{ij}^0)^2}\Delta y_i + \frac{y_j^0 - y_i^0}{(s_{ij}^0)^2}\Delta x_j - \frac{x_j^0 - x_i^0}{(s_{ij}^0)^2}\Delta y_j \end{aligned}$$

Directions:

$$r_{ij} = T_{ij} - \omega_i \qquad (\omega_i \text{ additional unknown})$$

$\Longrightarrow$ linearization of bearing observation equation (see also Fig. 3.3)

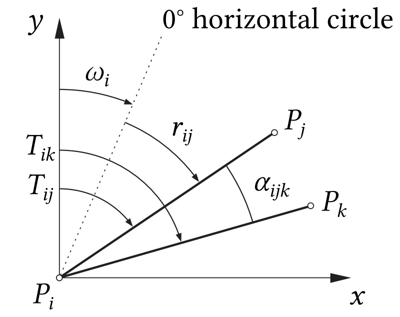


Figure 3.3: Linearization of bearing observation equation, bearing r_{ij}, orientation unknown ω_i.

$$
\begin{aligned}
r_{ij} &= T_{ij} - \omega_i \\
&= \arctan \frac{x_j - x_i}{y_j - y_i} - \omega_i \\
&= r_{ij}^0 - \frac{y_j^0 - y_i^0}{(s_{ij}^0)^2}\Delta x_i + \frac{x_j^0 - x_i^0}{(s_{ij}^0)^2}\Delta y_i + \frac{y_j^0 - y_i^0}{(s_{ij}^0)^2}\Delta x_j - \frac{x_j^0 - x_i^0}{(s_{ij}^0)^2}\Delta y_j - \omega_i
\end{aligned}
$$

Angles:

$$
\begin{aligned}
\alpha_{ijk} &= T_{ik} - T_{ij} \\
&= \arctan \frac{x_k - x_i}{y_k - y_i} - \arctan \frac{x_j - x_i}{y_j - y_i}
\end{aligned}
$$

$\Longrightarrow$ Linearized angle observation equation:

$$
\begin{aligned}
\alpha_{ijk} &= T_{ik}^0 - T_{ij}^0 + \left(-\frac{y_k^0 - y_i^0}{(s_{ik}^0)^2} + \frac{y_j^0 - y_i^0}{(s_{ij}^0)^2} \right) \Delta x_i + \left(\frac{x_k^0 - x_i^0}{(s_{ik}^0)^2} - \frac{x_j^0 - x_i^0}{(s_{ij}^0)^2} \right) \Delta y_i \\
&\quad + \frac{y_k^0 - y_i^0}{(s_{ik}^0)^2}\Delta x_k - \frac{x_k^0 - x_i^0}{(s_{ik}^0)^2}\Delta y_k - \frac{y_j^0 - y_i^0}{(s_{ij}^0)^2}\Delta x_j + \frac{x_j^0 - x_i^0}{(s_{ij}^0)^2}\Delta y_j \\
&= \alpha_{ijk}^0 + \dots
\end{aligned}
$$

3D intersection with additional vertical angles

3D distances:

$$
s_{ij} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2 + (z_j - z_i)^2} \qquad (i = 1, \dots, 4;\ j \equiv P)
$$

… linearization as usual.

Vertical angles:

$$\begin{aligned}\beta_{ij} &= \operatorname{arccot}\frac{\sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}}{z_j - z_i} \qquad \text{other trigonometric relations applicable}\\ &= \operatorname{arccot}\frac{d_{ij}}{z_j - z_i}\\ &= \beta_{ij}^0 - \frac{1}{1 + \left(\frac{d_{ij}}{z_j - z_i}\right)^2} \cdot \ldots\, \Delta x_i + \ldots\, \Delta y_i + \quad \ldots \quad + \ldots\, \Delta z_j\end{aligned}$$

Attention: physical units!

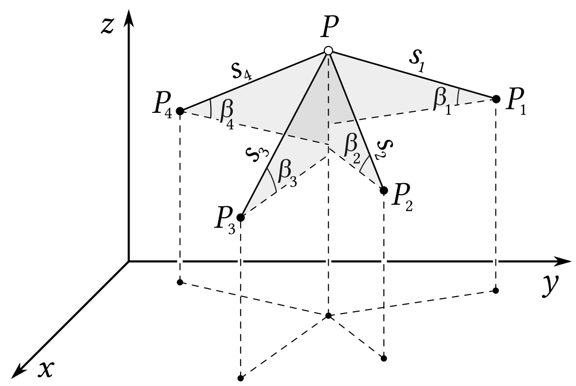


Figure 3.4: 3D intersection and vertical angles.

Iteration (see fig. 3.5)

Linearization (see 3.3) of the functional model $y = f(x)$ yields the linear model:

$$\Delta y = \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x_0} \Delta x + e = A(x_0)\,\Delta x + e\,.$$

The datum problem again

- Matrix $\underset{m\times n}{A}$ is rank deficient ($\operatorname{rank} A < n$),
- A has linear dependent columns,
- $Ax = 0$ has non-trivial solution $x_{\text{hom}} \neq 0$, i.e. the null space $\mathcal{N}(A)$ of A is not empty,
- $\det(A^\mathsf{T} A) = 0$,
- $A^\mathsf{T} A$ has zero eigenvalues.

Example: planar distance network (fig. 3.6)

Rank defect:

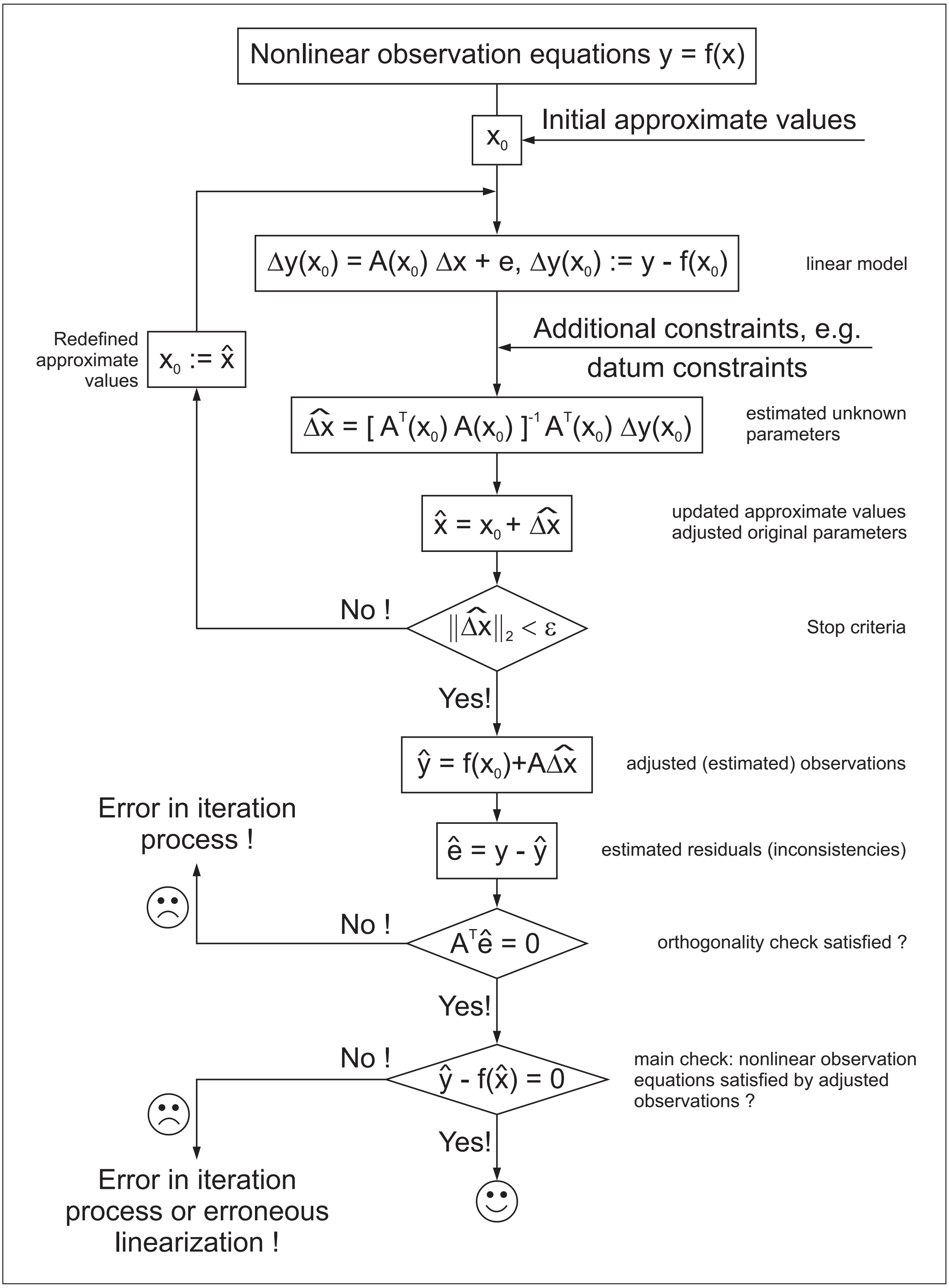


Figure 3.5: Iterative scheme

- Translation $\longrightarrow$ 2 parameters (x-, y-direction),
- Rotation $\longrightarrow$ 1 parameter,

$\Longrightarrow$ total of $d = 3$ parameters,
$\Longrightarrow$ $\operatorname{rank} A = n - d = n - 3$,
9 points $\longrightarrow$ $n - d = 18 - 3 = 15$, $m = 19$, thus $r = 4$.
Conditional adjustment: How many conditions? Answer: r condition equations.

3.4 Higher dimensions: the B-model (Condition equations)

In the *ideal* case we had

$$h_{1\mathrm{B}} - h_{1\mathrm{A}} = (H_\mathrm{B} - H_1) - (H_\mathrm{A} - H_1) = H_\mathrm{B} - H_\mathrm{A}$$
$$h_{13} + h_{32} - h_{12} = (H_3 - H_1) + (H_2 - H_3) - (H_2 - H_1) = 0$$

or

$$\begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} h_{1\mathrm{B}} \\ h_{13} \\ h_{12} \\ h_{32} \\ h_{1\mathrm{A}} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} H_\mathrm{B} \\ 0 \\ 0 \\ 0 \\ H_\mathrm{A} \end{pmatrix} .$$

Due to erroneous observations, a vector e of unknown inconsistencies must be introduced in order to make our linear model consistent.

$$\begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} h_{1\mathrm{B}} - e_{1\mathrm{B}} \\ h_{13} - e_{13} \\ h_{12} - e_{12} \\ h_{32} - e_{32} \\ h_{1\mathrm{A}} - e_{1\mathrm{A}} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} H_\mathrm{B} \\ 0 \\ 0 \\ 0 \\ H_\mathrm{A} \end{pmatrix} .$$

or

$$\underset{2\times 5}{B^\mathsf{T}} \left(\underset{5\times 1}{\Delta h} - \underset{5\times 1}{e} \right) = \underset{2\times 1}{B^\mathsf{T} c} .$$

Connected with this example are the questions

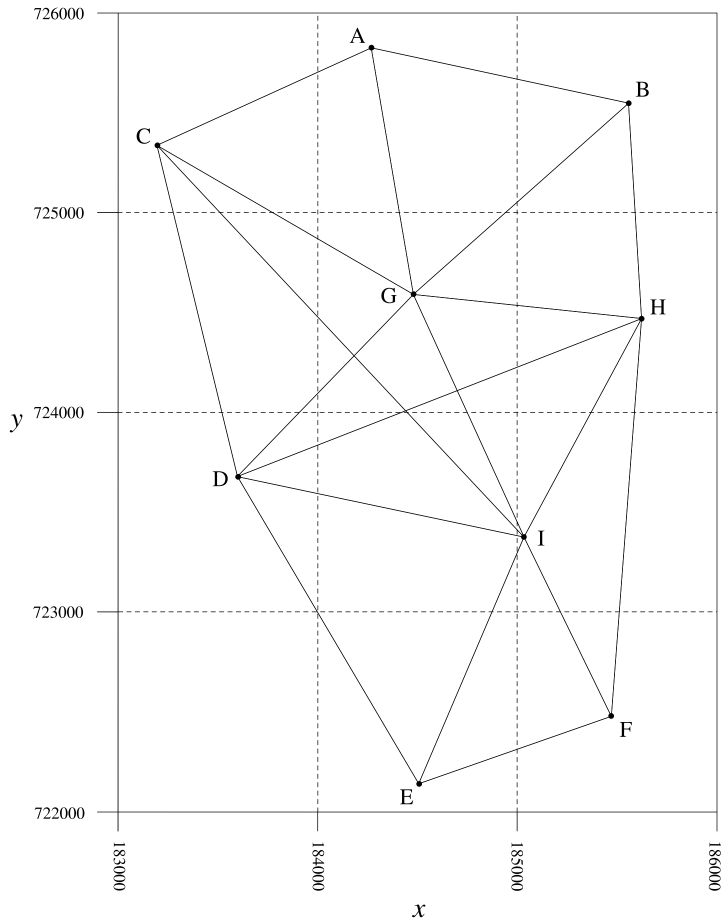
Q 1: How to handle constants like the vector c?

Q 2: How many conditions must be set up?

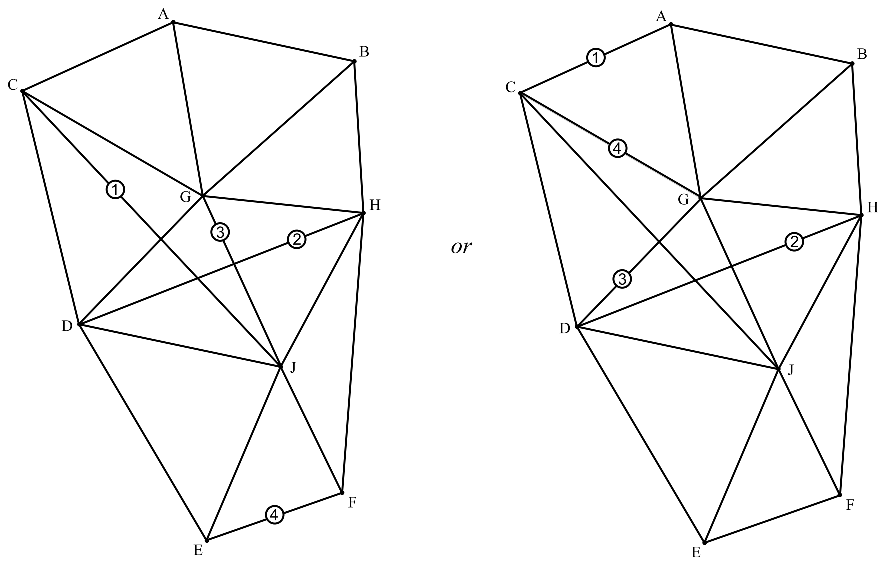
Q 3: Is the solution of the B-model identical to the one of the A-model?

A 1: Starting from

$$B^\mathsf{T}(\Delta h - e) = B^\mathsf{T} c ,$$



(a) distance network



(b) Four lines may be deleted without destabilizing the net.

Figure 3.6

where solely e is unknown, we collect all unknown parts on the left and all known quantities on the right hand side

$$\begin{aligned}
\Longrightarrow B^{\mathsf{T}}\Delta h - B^{\mathsf{T}}e &= B^{\mathsf{T}}c \\
B^{\mathsf{T}}e &= B^{\mathsf{T}}\Delta h - B^{\mathsf{T}}c \\
\underset{r\times m}{B^{\mathsf{T}}}\ \underset{m\times 1}{e} &= \underset{r\times 1}{B^{\mathsf{T}}y} =: w \\
w &: \text{vector of misclosures } w := B^{\mathsf{T}}y \\
y &: \text{reduced vector of observations} \\
r &: \text{number of conditions}
\end{aligned}$$

A 2: The number of conditions equals the redundancy

$$r = m - n$$

Sometimes the number of conditions can hardly be determined without knowledge on the number n of unknowns in the A-model. This will be treated later in more detail together with the so-called datum problem.

A 3:

$$\mathcal{L}_B(e,\lambda) = \frac{1}{2}\underset{1\times m}{e^{\mathsf{T}}}\ \underset{m\times 1}{e} + \underbrace{\underset{1\times r}{\lambda^{\mathsf{T}}}\,(\underset{r\times m}{B^{\mathsf{T}}}\ \underset{m\times 1}{y} - \underset{r\times m}{B^{\mathsf{T}}}\ \underset{m\times 1}{e})}_{1\times 1} \longrightarrow \min_{e,\lambda}$$

$$\begin{aligned}
\frac{\partial\mathcal{L}_B}{\partial e}(\hat{e},\hat{\lambda}) &= \underset{m\times 1}{\hat{e}} - \underset{m\times r}{B}\ \underset{r\times 1}{\hat{\lambda}} = \underset{m\times 1}{0} \\
\frac{\partial\mathcal{L}_B}{\partial \lambda}(\hat{e},\hat{\lambda}) &= -\underset{r\times m}{B^{\mathsf{T}}}\ \underset{m\times 1}{\hat{e}} + \underset{r\times m}{B^{\mathsf{T}}}\ \underset{m\times 1}{y} = \underset{r\times 1}{0} \qquad (w = B^{\mathsf{T}}y) \\
&\Longrightarrow \begin{pmatrix} \underset{m\times m}{I} & \underset{m\times r}{-B} \\ \underset{r\times m}{-B^{\mathsf{T}}} & \underset{r\times r}{0} \end{pmatrix}\begin{pmatrix} \hat{e} \\ \hat{\lambda} \end{pmatrix} = \begin{pmatrix} \underset{m\times 1}{0} \\ \underset{m\times 1}{-w} \end{pmatrix} \\
\hat{e} = B\hat{\lambda} &\Longrightarrow B^{\mathsf{T}}B\hat{\lambda} = w \\
&\Longrightarrow \hat{\lambda} = (B^{\mathsf{T}}B)^{-1}w \qquad \operatorname{rank}(B^{\mathsf{T}}B) = r \\
\Longrightarrow \hat{e} &= B(B^{\mathsf{T}}B)^{-1}w \\
&= B(B^{\mathsf{T}}B)^{-1}B^{\mathsf{T}}y \\
&= P_B y \\
\hat{y} &= y - \hat{e} \\
&= \left[I - B(B^{\mathsf{T}}B)^{-1}B^{\mathsf{T}}\right]y \\
&= P_B^{\perp}y
\end{aligned}$$

For the transition

$$\begin{array}{rcl} \text{parametric model} & \longleftrightarrow & \text{model of condition equations} \\ y = Ax + e & \longleftrightarrow & B^{\mathsf{T}} e = B^{\mathsf{T}} y\,, \end{array}$$

left multiply $y = Ax + e$ by B^{T}:

$$B^{\mathsf{T}} y = B^{\mathsf{T}} Ax + B^{\mathsf{T}} e \quad \Longleftrightarrow \quad B^{\mathsf{T}} A = 0\,.$$

E.g.:

$$\underset{2\times 5}{\begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 1 & 0 \end{pmatrix}} \underset{5\times 3}{\begin{pmatrix} -1 & 0 & 0 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 0 \end{pmatrix}} = \underset{2\times 3}{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}\,.$$

3.5 Linearization of non-linear condition equations

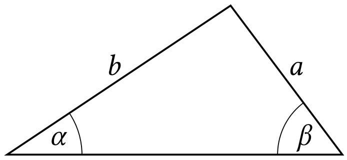


Figure 3.7: Linearization of condition equations

Ideal situation: error free observations

$$\frac{a}{\sin\alpha} = \frac{b}{\sin\beta} \quad \sim \quad a\sin\beta - b\sin\alpha = 0$$

Real situation with "errors" e_a, e_b, e_α, e_β

$$(a - e_a)\sin(\beta - e_\beta) - (b - e_b)\sin(\alpha - e_\alpha) = 0$$

$$\begin{aligned} f(e_a, e_b, e_\alpha, e_\beta) &= f(e_a^0, e_b^0, e_\alpha^0, e_\beta^0) + \left.\frac{\partial f}{\partial e_a}\right|_0 (e_a - e_a^0) + \ldots + \left.\frac{\partial f}{\partial e_\beta}\right|_0 (e_\beta - e_\beta^0) \\ &= f(e_a^0, e_b^0, e_\alpha^0, e_\beta^0) + \left.\frac{\partial f}{\partial e_a}\right|_0 e_a + \ldots + \left.\frac{\partial f}{\partial e_\beta}\right|_0 e_\beta - \left.\frac{\partial f}{\partial e_a}\right|_0 e_a^0 - \ldots - \left.\frac{\partial f}{\partial e_\beta}\right|_0 e_\beta^0 \end{aligned}$$

Model adjustment condition equations

$$w - B^\mathsf{T} e = 0$$

3.6 Higher dimensions: the mixed model (Gauss-Helmert model)

In the A-model, every observation is – in general – a linear or non-linear function of all unknown quantities, i.e.

$$y_i = f_i(x_1, x_2, \dots, x_n) = f_i(x_j) = f_i(x), \quad i = 1, \dots, m;\ j = 1, \dots, n$$

and every observation equation y_i contains just one single inconsistency e_i. In contrast, in the B-model no unknown parameter x exist and we have linear or non-linear relationships between the observations only,

$$f_j(y_i) = f_j(y) = 0, \quad i = 1, \dots, m;\ j = 1, \dots, r\,.$$

However, in many applications, functional relationships exist between both, parameters x and observations y, which can be formulated only as an implicit function

$$f(x_j, y_i) = 0, \quad i = 1, \dots, m;\ j = 1, \dots, n\,.$$

This will lead to a combination of both, A- and B-model, which is known as the *general model of adjustment*, *mixed model* or *Gauss-Helmert model*, in honor of Friedrich Robert Helmert[1].

Example: Best fitting circle with unknown radius r, and unknown centre coordinates u_M, v_M; observations u_i and v_i inconsistent.

$$f(\underbrace{r, u_\text{M}, y_\text{M},}_{\substack{\text{unknown}\\\text{parameters}\\\text{“}x\text{”}}} \underbrace{u_i - e_{u_i}, v_i - e_{v_i}}_{\substack{\text{observations “}y\text{” –}\\\text{inconsistencies “}e\text{”}}}) = (u_i - e_{u_i} - u_\text{M})^2 + (v_i - e_{v_i} - v_\text{M})^2 - r^2 = 0$$

[1] Friedrich Robert Helmert (1843–1917) was a famous German geodesist and mathematician who introduced this model 1872 in his book “Die Ausgleichungsrechnung nach der Methode der kleinsten Quadrate” (Adjustment theory using the method of least squares). He is also known as the father of many mathematical and physical theories of modern geodesy.

4 Weighted least squares

Observations have different weights $\Longleftrightarrow$ different quality.

4.1 Weighted observation equations

Analytical interpretation

Target function:

$$\left.\begin{array}{l} y_1 \longrightarrow w_1 \\ y_2 \longrightarrow w_2 \end{array}\right\} \quad \begin{aligned} \mathcal{L}_a^w &= \frac{1}{2}\left[w_1(y_1 - ax)^2 + w_2(y_2 - ax)^2\right] \\ &= \frac{1}{2}(y - ax)^\mathsf{T}\begin{pmatrix} w_1 & 0 \\ 0 & w_2 \end{pmatrix}(y - ax) \\ &= \frac{1}{2}(y - ax)^\mathsf{T} W (y - ax) \\ &= \frac{1}{2}y^\mathsf{T} W y - y^\mathsf{T} W a x + \frac{1}{2}x^\mathsf{T} a^\mathsf{T} W a x \\ &= \frac{1}{2}e^\mathsf{T} W e \end{aligned}$$

Necessary condition:

$$\begin{aligned} \hat{x}: \quad \min_x \mathcal{L}_a(x) \quad &\Longrightarrow \quad \frac{\partial \mathcal{L}_a}{\partial x}(\hat{x}) \\ &\Longrightarrow \quad \frac{\partial \mathcal{L}}{\partial x} = -y^\mathsf{T} W a + a^\mathsf{T} W a x = -a^\mathsf{T} W y + a^\mathsf{T} W a \hat{x} = 0 \\ &\Longrightarrow \quad a^\mathsf{T} W a \hat{x} = a^\mathsf{T} W y \quad \text{normal equation} \end{aligned}$$

Sufficient condition:

$$\frac{\partial^2 \mathcal{L}}{\partial x^2} = a^\mathsf{T} W a > 0, \quad \text{since } W \text{ is positive definite}$$

Normal equation $\Longrightarrow$

$$\begin{aligned} a^\mathsf{T} W (y - a\hat{x}) = 0 \quad &\Longrightarrow \quad a^\mathsf{T} W \hat{e} = 0 \\ &= \quad \hat{e} \perp W a \end{aligned}$$

normal equations	$a^\mathsf{T} W a \hat{x} = a^\mathsf{T} W y$
WLS estimate of x (weighted least squares)	$\hat{x} = (a^\mathsf{T} W a)^{-1} a^\mathsf{T} W y$
WLS estimate of y	$\hat{y} = a\hat{x} = a(a^\mathsf{T} W a)^{-1} a^\mathsf{T} W y$
WLS estimate of e	$\hat{e} = y - \hat{y} = \left[I - a(a^\mathsf{T} W a)^{-1} a^\mathsf{T} W\right] y$

Example

$$\left.\begin{aligned} a &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ W &= \begin{pmatrix} w_1 & 0 \\ 0 & w_2 \end{pmatrix} \end{aligned}\right\} \quad a^\mathsf{T} W = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} w_1 & 0 \\ 0 & w_2 \end{pmatrix} = \begin{pmatrix} w_1 & w_2 \end{pmatrix}$$

$$a^\mathsf{T} W a = \begin{pmatrix} w_1 & w_2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = w_1 + w_2$$

$\Longrightarrow$

$$\begin{aligned} \hat{x} &= \frac{1}{w_1 + w_2}(w_1 y_1 + w_2 y_2) \qquad \text{(weighted mean)} \\ &= \frac{w_1}{w_1 + w_2} y_1 + \frac{w_2}{w_1 + w_2} y_2 \end{aligned}$$

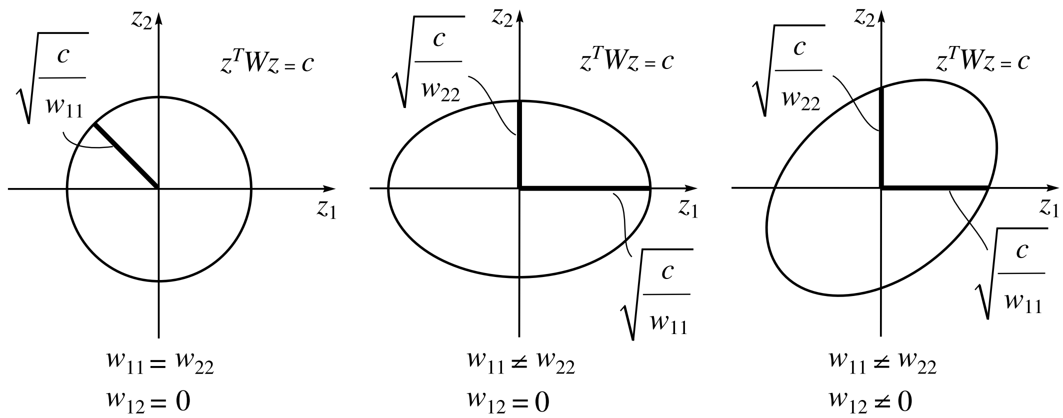
$$\begin{aligned} \begin{pmatrix} \hat{e}_1 \\ \hat{e}_2 \end{pmatrix} &= \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - \begin{pmatrix} \hat{y}_1 \\ \hat{y}_2 \end{pmatrix} \\ &= \frac{1}{w_1 + w_2} \begin{pmatrix} (w_1 + w_2) y_1 - w_1 y_1 - w_2 y_2 \\ (w_1 + w_2) y_2 - w_1 y_1 - w_2 y_2 \end{pmatrix} \\ &= \frac{1}{w_1 + w_2} \begin{pmatrix} w_2 (y_1 - y_2) \\ w_1 (y_2 - y_1) \end{pmatrix} \end{aligned}$$

$$w_1 > w_2 \ : \quad \text{“}y_1 \text{ is more important than } y_2\text{”} \quad \Longrightarrow \quad |\hat{e}_1| < |\hat{e}_2|$$

Projectors

$$\begin{aligned} P_a &= a(a^\mathsf{T} W a)^{-1} a^\mathsf{T} W := P_{a,(Wa)^\perp} \\ P_a P_a &= a(a^\mathsf{T} W a)^{-1} a^\mathsf{T} W a (a^\mathsf{T} W a)^{-1} a^\mathsf{T} W \\ &= a(a^\mathsf{T} W a)^{-1} a^\mathsf{T} W = P_a \end{aligned}$$

P_a idempotent matrix, oblique projector.



(a) Circle

(b) Ellipse aligned with coordinate axes.

(c) General ellipse, not aligned with coordinate axes.

Figure 4.1

4.1.1 Geometry

$$
\begin{aligned}
F(z) = z^{\mathsf{T}} W z &= c \\
w_1 z_1^2 + w_2 z_2^2 &= c \\
\frac{w_1}{c} z_1^2 + \frac{w_2}{c} z_2^2 &= 1 \\
\frac{z_1^2}{\frac{c}{w_1}} + \frac{z_2^2}{\frac{c}{w_2}} &= 1 \\
\frac{z_1^2}{a} + \frac{z_2^2}{b} &= 1 \qquad \text{ellipse equation}
\end{aligned}
$$

A family (c may vary!) of ellipses, the principal exes of which are not aligned with coordinate axes, in general.

Principal axes not aligned with coordinate axes

$$
z^{\mathsf{T}} W z = c \quad \sim \quad z_1^2 w_{11} + 2 w_{12} z_1 z_2 + w_{22} z_2^2 = c
$$

General ellipse

$$
\operatorname{grad} F(z_0) = 2 W z_0 \qquad \text{vector in } z_0\text{, orthogonal to the tangent of the ellipse in } z_0
$$

$$
z - z_0 \perp W z_0 \qquad \text{or} \qquad {z_0}^{\mathsf{T}} W (z - z_0) = 0
$$

4.1.2 Application to adjustment problems

Find a vector starting on line ax, ending in y being parallel to $z - a$ or orthogonal to $a^\mathsf{T} W$: $\hat{e}$

- $\hat{y} = a\hat{x}$ is the projection of y
  - onto a
  - in the direction orthogonal to Wa (along $(Wa)^\perp$)

$$\Longrightarrow \hat{y} = P_{a,(Wa)^\perp} y \quad \text{with} \quad P_{a,(Wa)^\perp} = a(a^\mathsf{T} W a)^{-1} a^\mathsf{T} W$$

- $\hat{e}$ is the projection of y
  - onto $(Wa)^\perp$
  - in direction of a

$$\begin{aligned} \Longrightarrow \hat{e} = P_{(Wa)^\perp, a} y \quad \text{with} \quad P_{(Wa)^\perp, a} &= P^\perp_{a,(Wa)^\perp} \\ &= I - a(a^\mathsf{T} W a)^{-1} a^\mathsf{T} W \\ &= \left[I - a(a^\mathsf{T} W a)^{-1} a^\mathsf{T} W \right] y \end{aligned}$$

- Because of $\hat{e} \not\perp a$ (or $a^\mathsf{T} \hat{e} \neq 0$) projections are oblique projections (or orthogonal projections with respect to the metric W; $\hat{e} \perp Wa$ or $a^\mathsf{T} W \hat{e} = 0$)

4.1.3 Higher dimensions

From one unknown to many unknowns.

$m = 2$

$$\underset{2\times 1}{y} = \underset{2\times 1}{a} \; \underset{1\times 1}{x} + \underset{2\times 1}{e}$$

becomes

$$\underset{m\times 1}{y} = \underset{m\times n}{A} \; \underset{n\times 1}{x} + \underset{m\times 1}{e}$$

Replace a by A!

$$P_{(W_a)^\perp, a} = I - \underset{m\times n}{A} \underbrace{(\underset{}{A^\mathsf{T}} \; \underset{m\times m}{W} \; A)^{-1}}_{n\times n} \underset{n\times m}{A^\mathsf{T}} \; \underset{m\times m}{W}$$

4.2 Weighted condition equations

Geometry

Starting point again: $b^\mathsf{T} a = 0$ ($a \perp b$):

Direction of $(Wa)^\perp$:

$$b^\mathsf{T} a = 0 \Longrightarrow b^\mathsf{T} W^{-1} W a = 0 \Longrightarrow Wa \perp W^{-1} b \Longrightarrow W^{-1} b = (Wa)^\perp$$

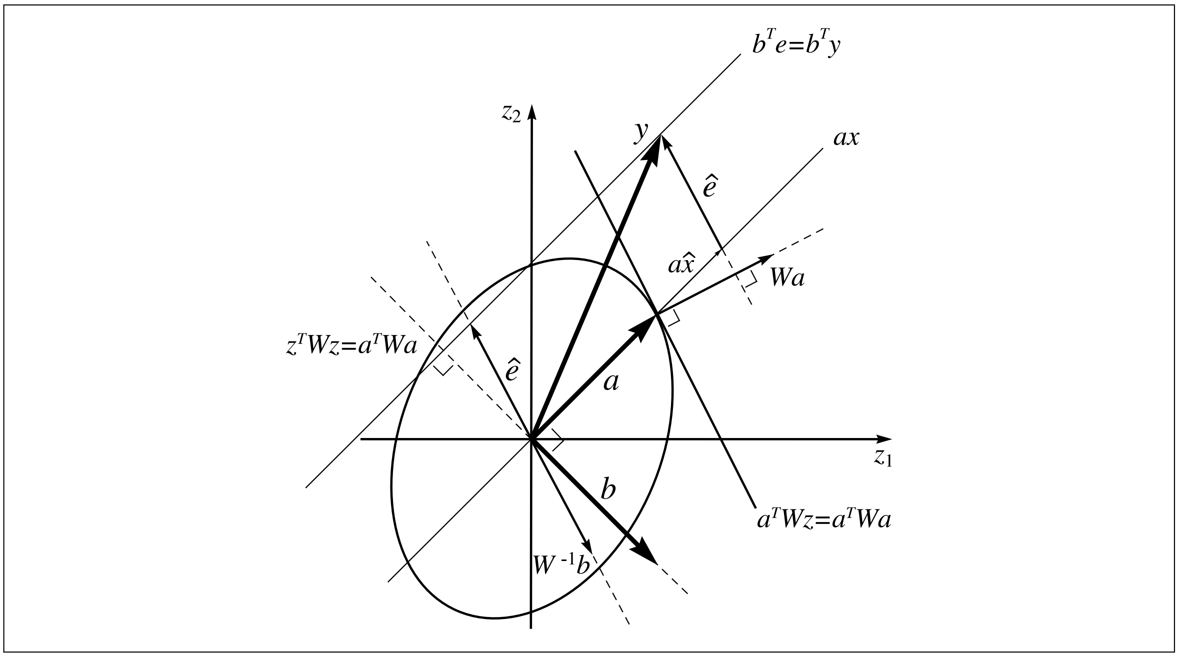


Figure 4.2: weighted condition

Target function to be minimized: $e^\mathsf{T} W e$ under $b^\mathsf{T} e = b^\mathsf{T} y$ or $b^\mathsf{T}(y - e) = 0$.

From all possible e's find that $e = \hat{e}$ which ends on the line $b^\mathsf{T} e = b^\mathsf{T} y$ and generates the smallest $e^\mathsf{T} W e = c$! $\Longrightarrow$ line $b^\mathsf{T} y = b^\mathsf{T} e$ is tangent to $e^\mathsf{T} W e = \hat{e}^\mathsf{T} W \hat{e} = c_{\min}$.

Point of Tangency: normal of the ellipse = normal of the line $b^\mathsf{T} y = b^\mathsf{T} e$ = direction of b $\iff$ $\hat{e}$ is parallel to $W^{-1} b$ $\Longrightarrow$ $\hat{e} = W^{-1} b \alpha$, α an unknown scalar.

Determine α: $\hat{e}$ lies on $b^\mathsf{T} e = b^\mathsf{T} y$

$$\begin{aligned}
&\Longrightarrow \quad b^\mathsf{T} \hat{e} = b^\mathsf{T} W^{-1} b \alpha = b^\mathsf{T} y \\
&\Longrightarrow \quad \alpha = (b^\mathsf{T} W^{-1} b)^{-1} b^\mathsf{T} y \\
&\Longrightarrow \quad \hat{e} = W^{-1} b (b^\mathsf{T} W^{-1} b)^{-1} b^\mathsf{T} y \\
&\Longrightarrow \quad \hat{y} = y - \hat{e} = \left[I - W^{-1} b (b^\mathsf{T} W^{-1} b)^{-1} b^\mathsf{T}\right] y
\end{aligned}$$

Remark: $\hat{e}$ is not the smallest e, orthogonal to $b^\mathsf{T} e = b^\mathsf{T} y$!

Calculus

$$\mathcal{L}_b(e, y) = \frac{1}{2} e^\mathsf{T} W e + \lambda^\mathsf{T}(b^\mathsf{T} y - b^\mathsf{T} e) \qquad \text{etc.}$$

$$\hat{e}: \quad \min_e e^\mathsf{T} W e \qquad \text{under constraint} \quad b^\mathsf{T} e = b^\mathsf{T} y$$

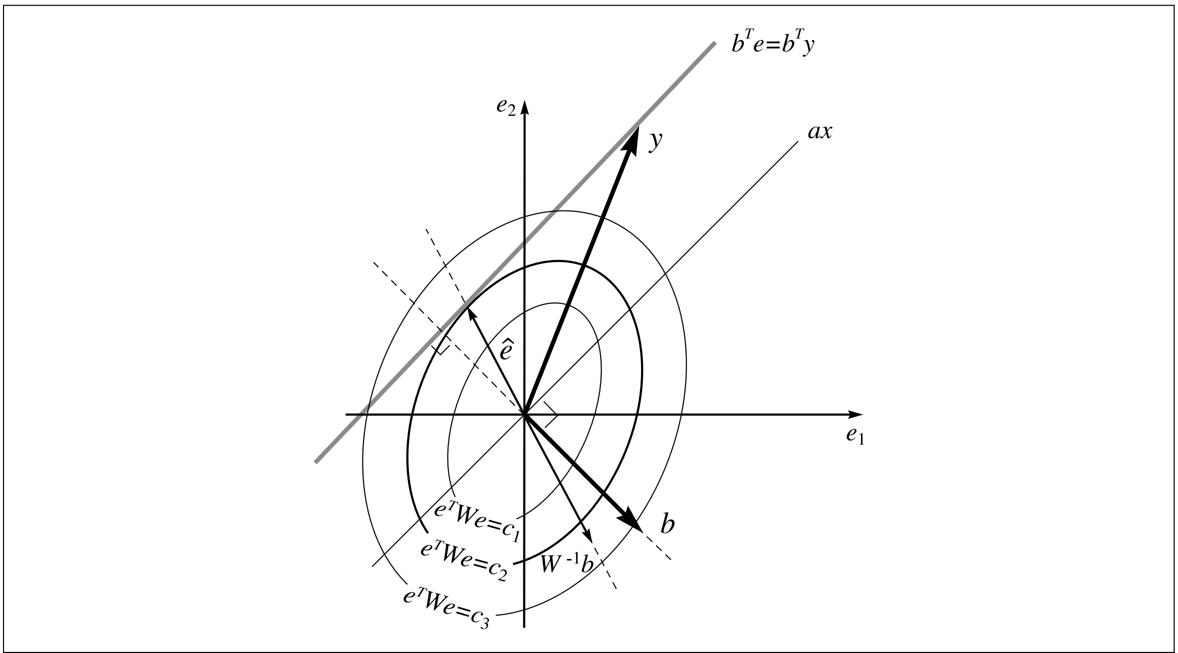


Figure 4.3: possible ellipses

Lagrange:

$$\mathcal{L}_b(e,\lambda) = \frac{1}{2}e^\mathsf{T}We + \lambda^\mathsf{T}(b^\mathsf{T}y - b^\mathsf{T}e)$$

Find e and λ which minimize L_b.

$$\implies \quad \begin{cases} \frac{\partial \mathcal{L}_b}{\partial e}(\hat{e},\hat{\lambda}) = W\hat{e} - b\hat{\lambda} = 0 \\ \frac{\partial \mathcal{L}_b}{\partial \lambda}(\hat{e},\hat{\lambda}) = -b^\mathsf{T}e + b^\mathsf{T}y = 0 \end{cases}$$

$$\iff \quad \begin{pmatrix} W & -b \\ -b^\mathsf{T} & 0 \end{pmatrix} \begin{pmatrix} \hat{e} \\ \hat{\lambda} \end{pmatrix} = \begin{pmatrix} 0 \\ -b^\mathsf{T}y \end{pmatrix}$$

1. row

$$W\hat{e} - b\hat{\lambda} = 0 \implies \hat{e} = W^{-1}b\hat{\lambda}$$

2. row

$$b^\mathsf{T}\hat{e} = b^\mathsf{T}y \implies b^\mathsf{T}W^{-1}b\hat{\lambda} = b^\mathsf{T}y$$

solve for $\hat{\lambda}$

$$\hat{\lambda} = (b^\mathsf{T}W^{-1}b)^{-1}b^\mathsf{T}y$$

substitute in 1. row

$$\hat{e} = W^{-1}b(b^\mathsf{T}W^{-1}b)^{-1}b^\mathsf{T}y$$

$$\hat{y} = y - \hat{e} = \left[I - W^{-1}b(b^\mathsf{T}W^{-1}b)^{-1}b^\mathsf{T}\right]y$$

Higher dimensions

Replace b with B.

$r = m - n$ condition equations, Lagrange multipliers

$$B^{\mathsf{T}} y = B^{\mathsf{T}} e$$

$$\left.\begin{aligned} y &= Ax + e \\ B^{\mathsf{T}} A &= 0 \end{aligned}\right\} \quad \Longrightarrow \quad B^{\mathsf{T}} y = B^{\mathsf{T}} A x + B^{\mathsf{T}} e = B^{\mathsf{T}} e$$

$$\begin{pmatrix} \underset{m \times m}{W} & \underset{m \times r}{B} \\ \underset{r \times m}{B^{\mathsf{T}}} & \underset{r \times r}{0} \end{pmatrix} \begin{pmatrix} \underset{m \times 1}{\hat{e}} \\ \underset{r \times 1}{\hat{\lambda}} \end{pmatrix} = \begin{pmatrix} 0 \\ B^{\mathsf{T}} y \end{pmatrix}$$

$$\begin{aligned} \hat{e} &= W^{-1} B (B^{\mathsf{T}} W^{-1} B)^{-1} B^{\mathsf{T}} y \\ \hat{y} &= \left[I - W^{-1} B (B^{\mathsf{T}} W^{-1} B)^{-1} B^{\mathsf{T}} \right] y \end{aligned}$$

Constant term (RHS)

Ideal case without errors:

$$B^{\mathsf{T}} y = c$$

In reality:

$$B^{\mathsf{T}} (y - e) = c \quad \Longrightarrow \quad B^{\mathsf{T}} e = B^{\mathsf{T}} y - c =: w$$

$$\Longrightarrow \quad \begin{cases} \hat{e} = W^{-1} B (B^{\mathsf{T}} W^{-1} B)^{-1} \underbrace{[B^{\mathsf{T}} y - c]}_{w} \\ \hat{y} = y - \hat{e} = \ldots \end{cases}$$

4.3 Stochastics

Probabilistic formulation

(stochastic quantities are underlined)

$$\begin{array}{lcc} \text{Version 1:} & \underline{y} = Ax + \underline{e}, \quad \mathrm{E}\left\{\underline{e}\right\} = 0 & \mathrm{D}\left\{\underline{e}\right\} = Q_y \\ \text{Version 2:} & \mathrm{E}\left\{\underline{y}\right\} = Ax & \mathrm{D}\left\{\underline{e}\right\} = Q_y \end{array}$$

$$\underbrace{\underbrace{\text{Functional model}}\quad\underbrace{\text{Stochastic model: variance-covariance matrix}}}_{\text{Mathematical model}}$$

Linear Variance-covariance propagation

In general:

$$\underline{z} = L\underline{y}, \qquad Q_z = L Q_y L^{\mathsf{T}}$$

$$
\begin{aligned}
\underline{\hat{x}} &= (A^\mathsf{T} W A)^{-1} A^\mathsf{T} W \underline{y} \\
&= L\underline{y} \\
\longrightarrow \mathrm{E}\left\{\underline{\hat{x}}\right\} &= (A^\mathsf{T} W A)^{-1} A^\mathsf{T} W \,\mathrm{E}\left\{\underline{y}\right\} \\
&= (A^\mathsf{T} W A)^{-1} A^\mathsf{T} W A x \\
&= x \quad \text{(unbiased estimate)} \\
\longrightarrow Q_{\hat{x}} &= L Q_y L^\mathsf{T} \\
&= (A^\mathsf{T} W A)^{-1} A^\mathsf{T} W Q_y W A (A^\mathsf{T} W A)^{-1} \\
\underline{\hat{y}} &= A\underline{\hat{x}} \\
&= P_A \underline{y} \\
\longrightarrow \mathrm{E}\left\{\underline{\hat{y}}\right\} &= A\,\mathrm{E}\left\{\underline{\hat{x}}\right\} = Ax = \mathrm{E}\left\{\underline{y}\right\} \\
\longrightarrow Q_{\hat{y}} &= P_A Q_y {P_A}^\mathsf{T} \\
\underline{\hat{e}} &= \underline{y} - A\underline{\hat{x}} = (I - P_A)\underline{y} \\
\longrightarrow \mathrm{E}\left\{\underline{\hat{e}}\right\} &= \mathrm{E}\left\{\underline{y}\right\} - Ax = 0 \\
\longrightarrow Q_{\hat{e}} &= Q_y - P_A Q_y - Q_y {P_A}^\mathsf{T} + P_A Q_y {P_A}^\mathsf{T}
\end{aligned}
$$

Questions:

- Is $\underline{\hat{x}}$ the best estimator?
- Or: When is $Q_{\hat{x}}$ smallest?

4.4 Best Linear Unbiased Estimation (BLUE)

Best	$Q_{\hat{x}}$ minimal (in LU-Class)
Linear	$\underline{\hat{x}} = L\underline{y}$
Unbiased	$\mathrm{E}\left\{\underline{\hat{x}}\right\} = x$
Estimate	

2D-example (old)

$$
\mathrm{E}\left\{\underline{y}\right\} = ax, \qquad a = \begin{pmatrix} 1 \\ 1 \end{pmatrix}
$$

$$
\mathrm{D}\left\{\underline{y}\right\} = Q_y, \qquad Q_y = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}
$$

L-property:

$$\underline{\hat{x}} = l^\mathsf{T} \underline{y}$$

U-property:

$$\mathrm{E}\left\{\underline{\hat{x}}\right\} = l^\mathsf{T}\,\mathrm{E}\left\{\underline{y}\right\} = l^\mathsf{T} a x = x \quad \Longrightarrow \quad l^\mathsf{T} a = 1$$

B-property:

$$\underline{\hat{x}} = l^\mathsf{T} \underline{y} \quad \Longrightarrow \quad \sigma_{\hat{x}}^2 = l^\mathsf{T} Q_y l$$

Find that l which minimizes $l^\mathsf{T} Q_y l$ and satisfies $l^\mathsf{T} a = 1$!

$$\Longrightarrow \quad \min_l l^\mathsf{T} Q_y l \quad \text{under} \quad l^\mathsf{T} a = 1$$

Solution?

Comparison LS, B-Model

min	$e^\mathsf{T} W e$	$l^\mathsf{T} Q_y l$
under	$b^\mathsf{T} e = b^\mathsf{T} y = w$	$l^\mathsf{T} a = a^\mathsf{T} l = 1$
estimate	$\hat{e} = W^{-1} b (b^\mathsf{T} W^{-1} b)^{-1} w$	$\hat{l} = Q_y^{-1} a (a^\mathsf{T} Q_y^{-1} a)^{-1}$

$$\Longrightarrow \quad \underline{\hat{x}} = \hat{l}^\mathsf{T} \underline{y} = (a^\mathsf{T} Q_y^{-1} a)^{-1} a^\mathsf{T} Q_y^{-1} \underline{y}$$

Higher dimensions

$$a \longrightarrow A, \quad Q_y^{-1} = P_y$$

Gauss coined the variable P from the Latin *pondus*, which means weight.

$$\left.\begin{aligned} \text{BLUE: } \underline{\hat{x}} &= (A^\mathsf{T} P_y A)^{-1} A^\mathsf{T} P_y \underline{y} \\ \text{Det.: } \hat{x} &= (A^\mathsf{T} W A)^{-1} A^\mathsf{T} W y \end{aligned}\right\} \quad \Longrightarrow \quad \text{BLUE, if} \quad W = P_y = Q_y^{-1}$$

Linear Variance-covariance propagation

$$\begin{aligned}
&\underline{\hat{x}} = (A^\mathsf{T} P_y A)^{-1} A^\mathsf{T} P_y \underline{y} \\
&\quad \Longrightarrow \quad Q_{\hat{x}} = (A^\mathsf{T} P_y A)^{-1} A^\mathsf{T} P_y Q_y P_y A (A^\mathsf{T} P_y A)^{-1} = (A^\mathsf{T} P_y A)^{-1} \\
&\underline{\hat{y}} = A \underline{\hat{x}} = P_A \underline{y} \\
&\quad \Longrightarrow \quad Q_{\hat{y}} = A (A^\mathsf{T} P_y A)^{-1} A^\mathsf{T} P_y Q_y = P_A Q_y = P_A Q_y {P_A}^\mathsf{T} = Q_y P_A \\
&\underline{\hat{e}} = (I - P_A) \underline{y} = P_A^\perp \underline{y} = \underline{y} - \underline{\hat{y}} \\
&\quad \Longrightarrow \quad Q_{\hat{e}} = Q_y - P_A Q_y - Q_y {P_A}^\mathsf{T} + P_A Q_y {P_A}^\mathsf{T} = P_A^\perp Q_y = Q_y - Q_{\hat{y}}
\end{aligned}$$

Besides:

$$I = P_A + P_A^\perp \Longrightarrow Q_y = P_A Q_y + P_A^\perp Q_y = Q_{\hat{y}} + Q_{\hat{e}}$$

Note: P_A is a projector, but P_y is a weight matrix.

5 Geomatics examples

Further (simple and more advanced) examples including data files can be found in "Geodetic Network Adjustment Examples" (http://www.gis.uni-stuttgart.de/res/study/addons/adjustment_examples.pdf).

5.1 A-Model: Adjustment of observation equations

5.1.1 Planar triangle

Observations: Angles α, β, γ [°] and distances S_{12}, S_{13}, S_{23} [m]

Auxiliary quantities: Bearings T_{12}, T_{13} [°]

Bearings:

$$T_{ij} = \arctan \frac{x_j - x_i}{y_j - y_i}$$

Angles:

$$\begin{aligned}
\alpha &= T_{12} - T_{13} = \arctan \frac{x_2 - x_1}{y_2 - y_1} - \arctan \frac{x_3 - x_1}{y_3 - y_1} \\
\beta &= T_{23} - T_{21} = \arctan \frac{x_3 - x_2}{y_3 - y_2} - \arctan \frac{x_1 - x_2}{y_1 - y_2} \\
\gamma &= T_{31} - T_{32} = \arctan \frac{x_1 - x_3}{y_1 - y_3} - \arctan \frac{x_2 - x_3}{y_2 - y_3}
\end{aligned}$$

Approximate coordinates:

point	x	y
1	0	0
2	1	0
3	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$

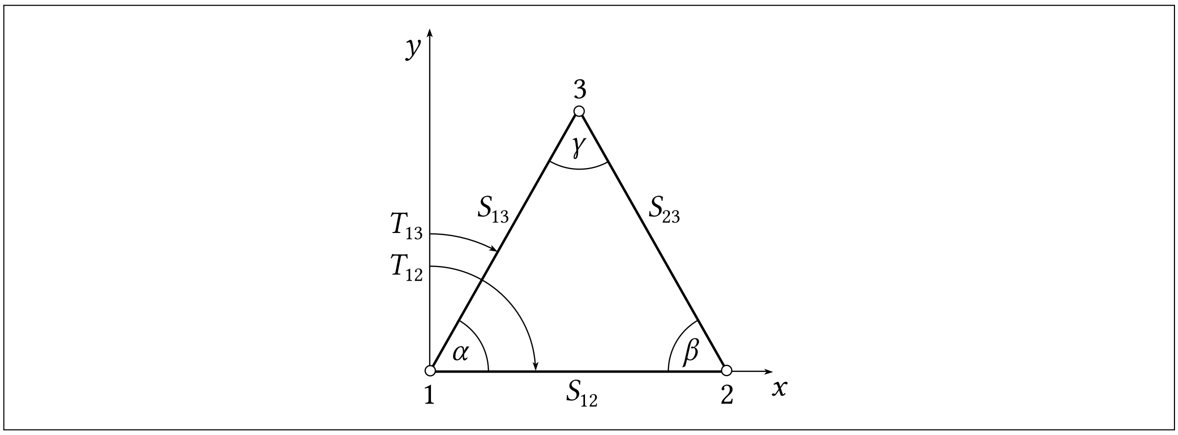


Figure 5.1: Sketch Planar triangle

Approx. coordinates [m]			“Observations” from approx. coordinates		Observations		σ
			S_{12}^0	1 m	S_{12}	1.01 m	±0.01 m
			S_{13}^0	1 m	S_{13}	1.02 m	±0.02 m
	x_0	y_0	S_{23}^0	1 m	S_{23}	0.97 m	±0.01 m
1	0	0	α_0	60°	α	60°	$\pm 1''$
2	1	0	β_0	60°	β	59.7°	$\pm 1'$
3	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	γ_0	60°	γ	60.2°	$\pm 1'$

Observation	y	Designmatrix A						Units	Unknowns
	$\rho := \frac{180°}{\pi}$	$\mathrm{d}x_1$	$\mathrm{d}y_1$	$\mathrm{d}x_2$	$\mathrm{d}y_2$	$\mathrm{d}x_3$	$\mathrm{d}y_3$		[m]
S_{12}	0.01 m	-1	0	1	0	0	0		$\mathrm{d}x_1$
S_{23}	-0.03 m	0	0	$\frac{1}{2}$	$-\frac{\sqrt{3}}{2}$	$-\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	[−]	$\mathrm{d}y_1$
S_{13}	0.02 m	$-\frac{1}{2}$	$-\frac{\sqrt{3}}{2}$	0	0	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$		$\mathrm{d}x_2$
α	0 rad	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	0	-1	$-\frac{\sqrt{3}}{2}$	$\frac{1}{2}$		$\mathrm{d}y_2$
β	$-0.3\frac{\circ}{\rho}$ rad	0	-1	$-\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$[\mathrm{m}^{-1}]$	$\mathrm{d}x_3$
γ	$0.2\frac{\circ}{\rho}$ rad	$-\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	0	-1		$\mathrm{d}y_3$

$\Longrightarrow$ Linearized distance observation equation (Taylor point = point of expansion = set of approximate coordinates).

5.1.2 Distance Network

In this example, measured distances between the points of the network in figure 3.6 are adjusted. The standard deviation of observations is $\sigma_s = \pm 1\,\text{cm}$, the a priori standard deviation $\sigma_0 = \pm 1\,\text{cm}$.

$$\Longrightarrow \quad P_s = \frac{\sigma_0^2}{\sigma_s^2} = 1$$

Table 5.1 contains measured distances (observations y) between respective network points, while table 5.2 contains approximate coordinates of the points. Points A and B are datum points with the minimum number of datum parameters X_A, Y_A, X_B fixed.

leg	length [m]	leg	length [m]
A–B	1309.155	D–H	2179.147
A–C	1188.464	D–I	1461.074
A–G	1267.520	E–F	1031.232
B–G	1447.552	E–I	1353.146
B–H	1077.634	F–H	1991.004
C–D	1715.405	F–I	997.285
C–G	1504.039	G–H	1149.345
C–I	2688.088	G–I	1310.957
D–E	1780.446	H–I	1241.810
D–G	1260.133		

Table 5.1: Observed distances y.

Point ID	X_0 [m]	Y_0 [m]
A	184 270.031	725 830.033
B	185 549.974	725 400.000
C	183 200.000	725 450.000
D	183 800.000	723 550.000
E	184 300.000	722 050.000
F	185 200.000	722 450.000
G	184 500.000	724 400.000
H	185 700.000	724 650.000
I	184 800.000	723 400.000

Table 5.2: Approximate coordinates.

Table 5.3 contains the reduced vector Δy and table 5.4 the estimated parameters at first iteration.

leg	length [m]	leg	length [m]
A–B	−41.098	D–H	−16.303
A–C	52.950	D–I	449.887
A–G	−180.886	E–F	46.346
B–G	−2.429	E–I	−86.472
B–H	312.776	F–H	−265.099
C–D	−277.081	F–I	−33.491
C–G	−167.038	G–H	−76.420
C–I	87.607	G–I	266.926
D–E	199.307	H–I	−298.482
D–G	158.997		

Table 5.3: Reduced observations Δy.

	$\widehat{\Delta\xi}$ [m]		$\widehat{\Delta\xi}$ [m]
X_A	0.000 00	Y_E	70.034 91
Y_A	0.000 00	X_F	220.074 41
X_B	0.000 00	Y_F	12.078 83
Y_B	124.267 84	X_G	−44.702 62
X_C	−29.046 95	Y_G	176.988 00
Y_C	−80.516 22	X_H	−52.354 95
X_D	−241.435 89	Y_H	−206.799 48
Y_D	143.805 90	X_I	190.824 30
X_E	149.155 11	Y_I	−29.001 63

Table 5.4: Estimated parameters.

Tables 5.5 and 5.6 contain the adjusted coordinates and observations ($\hat{y}$) respectively, of the network points after 6 iterations. Table 5.7 shows estimated inconsistencies in measured distances.

$$\Longrightarrow \quad \hat{e}^\mathsf{T} P \hat{e} = 0.035\,\text{cm}^2, \quad (6 \text{ iterations, stop criteria } \|\widehat{\Delta\xi}\| < 10^{-10})$$

point ID	$\hat{X}$ [m]	$\hat{Y}$ [m]
A	184 270.031	725 830.033
B	185 549.974	725 555.019
C	183 185.048	725 344.999
D	183 598.001	723 680.041
E	184 499.996	722 144.987
F	185 469.997	722 495.040
G	184 480.021	724 580.029
H	185 625.005	724 480.000
I	185 030.002	723 390.016

Table 5.5: Adjusted coordinates.

leg	$\hat{y}$ [m]	leg	$\hat{y}$ [mm]
A–B	1309.155	D–H	2179.1462
A–C	1188.464	D– I	1461.0749
A–G	1267.520	E – F	1031.2318
B–G	1447.552	E – I	1353.1463
B–H	1077.634	F –H	1991.0036
C–D	1715.4053	F – I	997.2854
C–G	1504.0395	G–H	1149.3454
C– I	2688.0873	G– I	1310.9572
D–E	1780.4458	H– I	1241.8109
D–G	1260.133		

Table 5.6: Adjusted observations.

leg	$\hat{e}$ [mm]	leg	$\hat{e}$ [mm]
A–B	−0.01	D–H	0.78
A–C	−0.01	D– I	−0.88
A–G	0.00	E – F	0.22
B–G	0.01	E – I	−0.27
B–H	−0.01	F –H	0.43
C–D	−0.27	F – I	−0.39
C–G	−0.46	G–H	−0.35
C– I	0.67	G– I	−0.18
D–E	0.20	H– I	−0.86
D–G	−0.04		

Table 5.7: Estimated inconsistencies.

Figure 5.2 shows the convergence of estimated corrections to approximate coordinates. Fig 5.3 shows the overall convergence in adjustment iteration. Figure 5.4 shows approximate points, adjusted and datum points. Finally in figure 5.5 adjusted and datum points are shown with error ellipses. Table 5.8 shows the *A*-matrix after the first iteration.

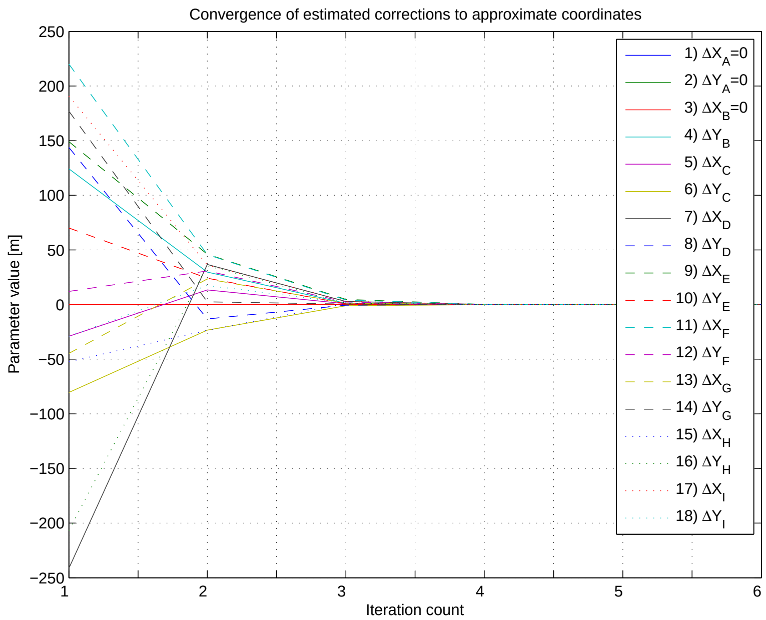


Figure 5.2: Convergence of estimated corrections.

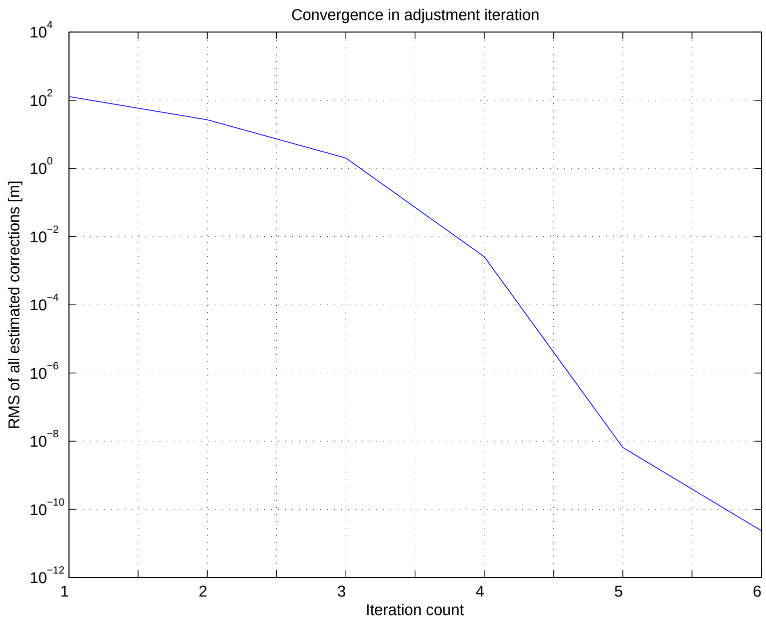


Figure 5.3: Convergence in adjustment iteration.

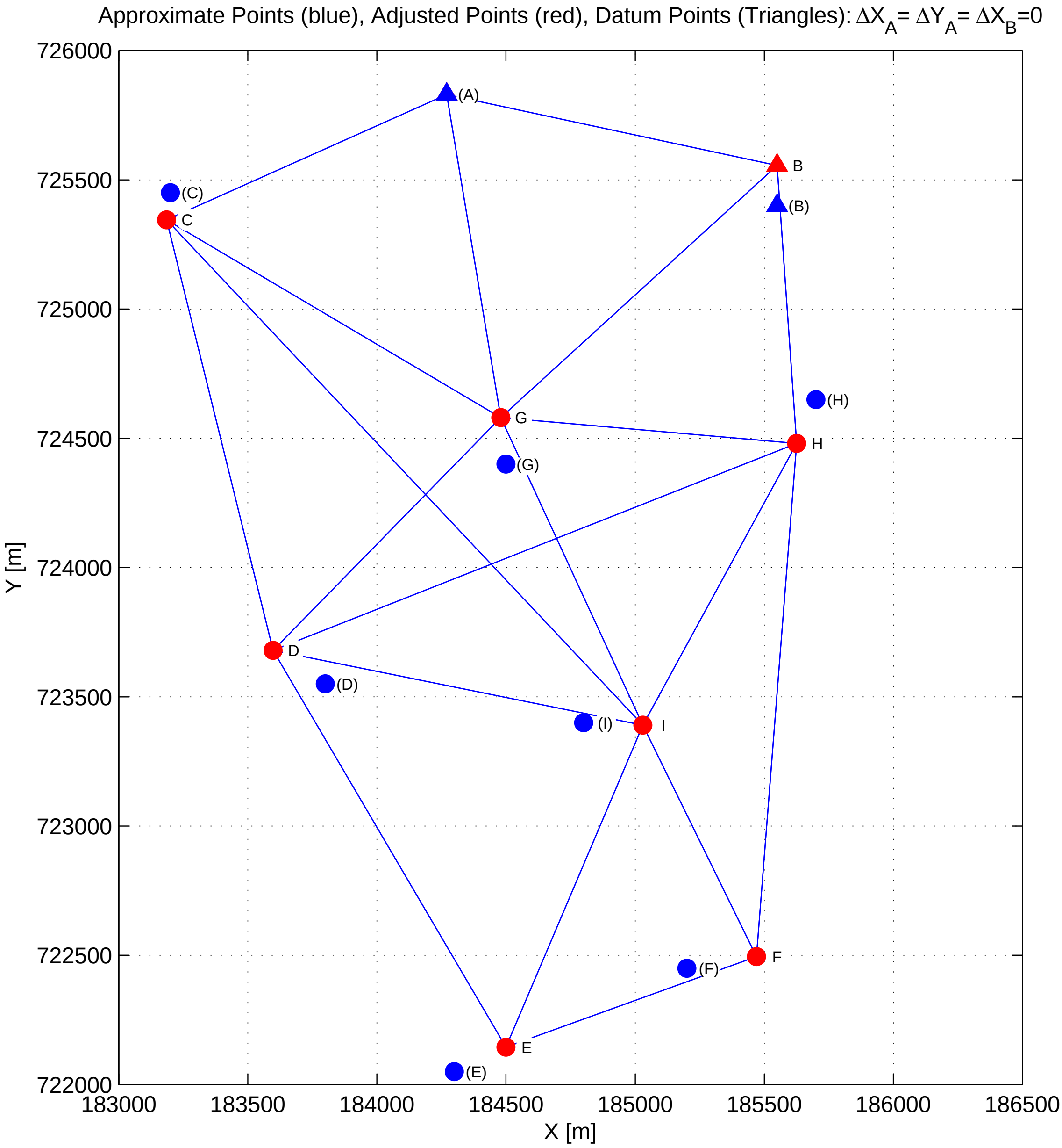


Figure 5.4: Approximate, adjusted and datum points.

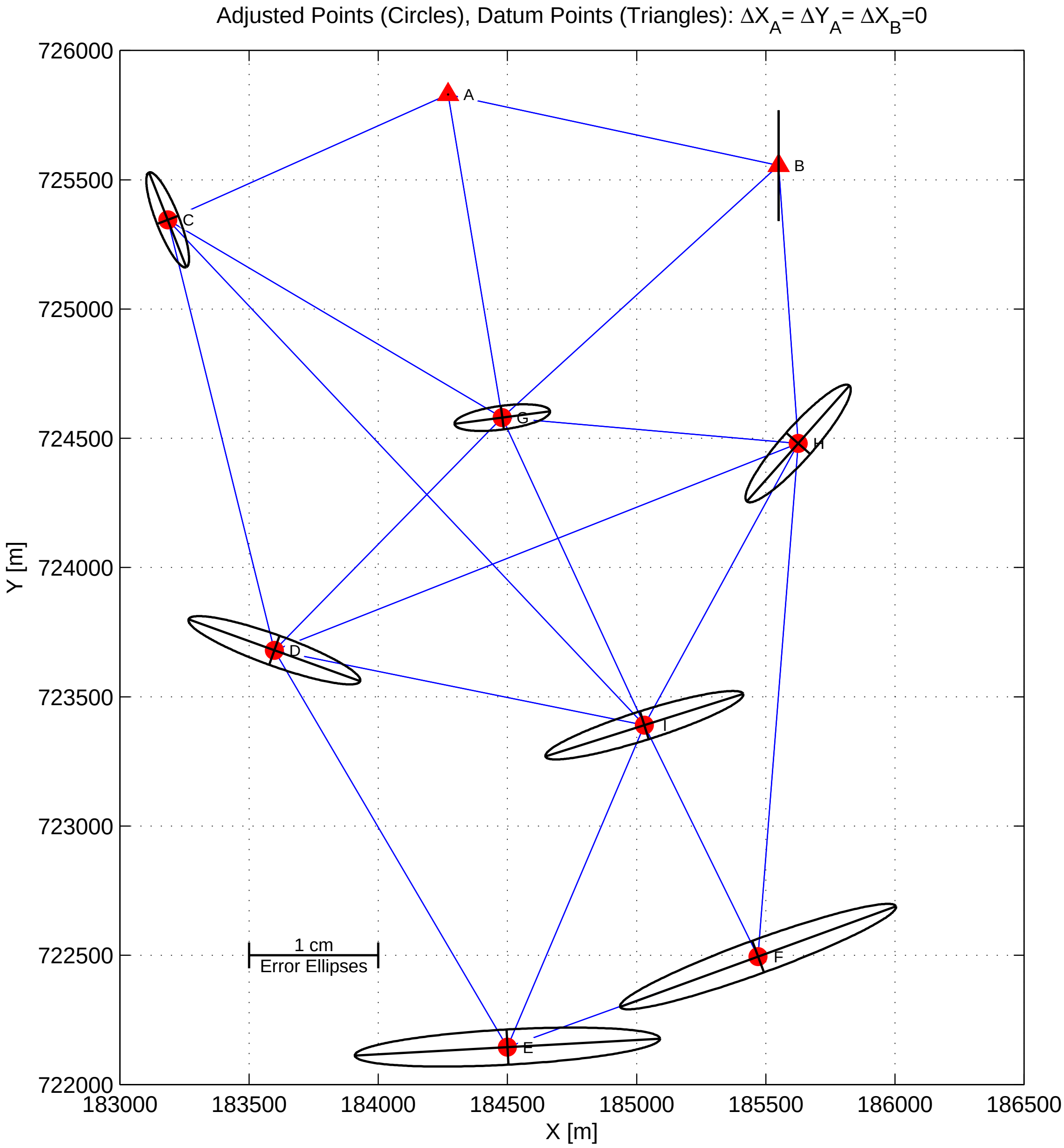


Figure 5.5: Adjusted and datum points with error ellipses.

leg	ΔY_B	ΔX_C	ΔY_C	ΔX_D	ΔY_D	ΔX_E	ΔY_E	ΔX_F	ΔY_F	ΔX_G	ΔY_G	ΔX_H	ΔY_H	ΔX_I	ΔY_I
A–B	−0.318 48	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A–C	0	−0.942 33	−0.334 68	0	0	0	0	0	0	0	0	0	0	0	0
A–G	0	0	0	0	0	0	0	0	0	0.158 77	−0.987 31	0	0	0	0
B–G	0.689 66	0	0	0	0	0	0	0	0	−0.724 13	−0.689 66	0	0	0	0
B–H	0.980 57	0	0	0	0	0	0	0	0	0	0	0.196 15	−0.980 57	0	0
C–D	0	−0.301 13	0.953 58	0.301 13	−0.953 58	0	0	0	0	0	0	0	0	0	0
C–G	0	−0.777 94	0.628 34	0	0	0	0	0	0	0.777 94	−0.628 34	0	0	0	0
C–I	0	−0.615 27	0.788 32	0	0	0	0	0	0	0	0	0	0	0.615 27	−0.788 32
D–E	0	0	0	−0.316 23	0.948 68	0.316 23	−0.948 68	0	0	0	0	0	0	0	0
D–G	0	0	0	−0.635 71	−0.771 93	0	0	0	0	0.635 71	0.771 93	0	0	0	0
D–H	0	0	0	−0.865 43	−0.501 04	0	0	0	0	0	0	0.865 43	0.501 04	0	0
D–I	0	0	0	−0.988 94	0.148 34	0	0	0	0	0	0	0	0	0.988 94	−0.148 34
E–F	0	0	0	0	0	−0.913 81	−0.406 14	0.913 81	0.406 14	0	0	0	0	0	0
E–I	0	0	0	0	0	−0.347 31	−0.937 75	0	0	0	0	0	0	0.347 31	0.937 75
F–H	0	0	0	0	0	0	0	−0.221 62	−0.975 13	0	0	0.221 62	0.975 13	0	0
F–I	0	0	0	0	0	0	0	0.388 06	−0.921 64	0	0	0	0	−0.388 06	0.921 64
G–H	0	0	0	0	0	0	0	0	0	−0.978 98	−0.203 95	0.978 98	0.203 95	0	0
G–I	0	0	0	0	0	0	0	0	0	−0.287 35	0.957 83	0	0	0.287 35	−0.957 83
H–I	0	0	0	0	0	0	0	0	0	0	0	0.584 30	0.811 53	−0.584 30	−0.811 53

Table 5.8: *A*-Matrix (1st iteration).

5.1.3 Distance and Direction Network (1)

A monitoring situation where directions and distances to 4 points A, B, C, and D are measured from point N, see table 5.10. Coordinates (Table 5.9) of points 1 to 4 including N_0 and the orientation $\omega_N^0 = 63.5610$ gon are approximately given (see Jäger, 2005, pg. 241–242).

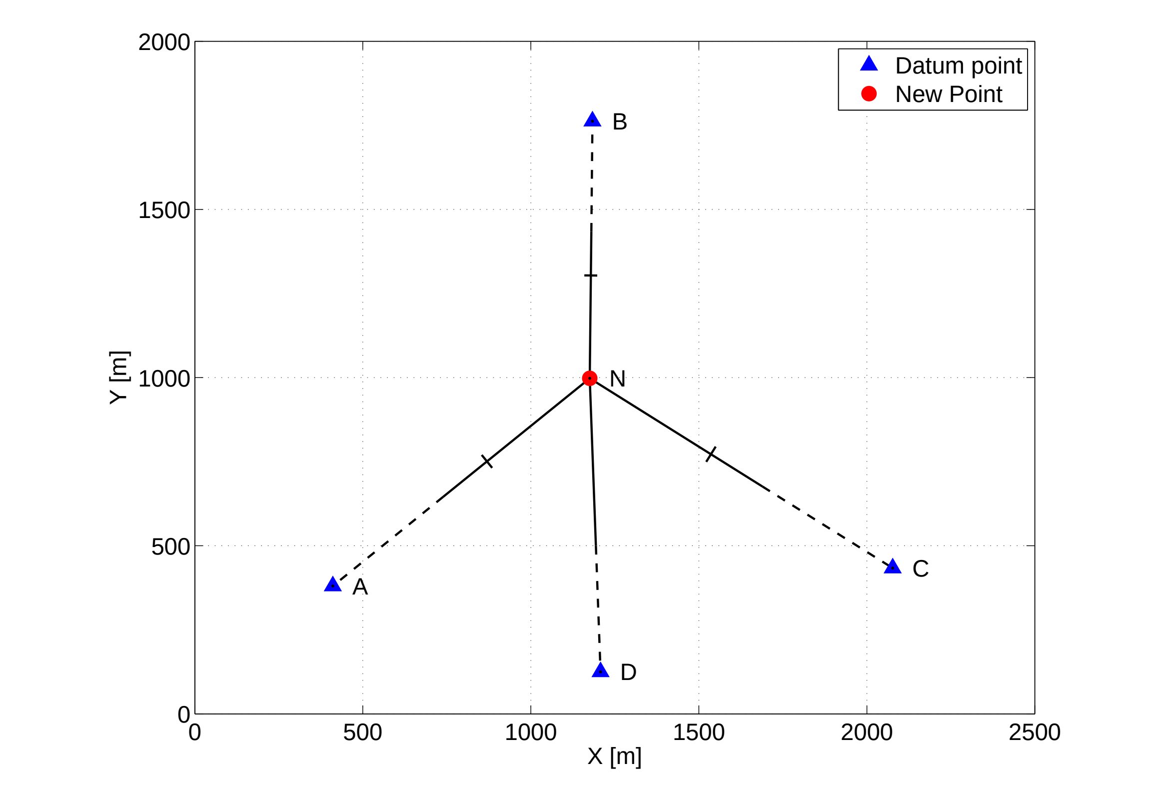


Figure 5.6: Distance and direction network (1).

The standard deviation of observations are assumed to be $\sigma_s = \pm 1$ cm for distances and $\sigma_r = \pm 0.5$ mgon for directions. Thus, for the choice of $\sigma_0 = \pm 1$ m as a priori standard deviation the elements of the weight matrix P are obtained:

$$P_s = \frac{\sigma_0^2}{\sigma_s^2} = 10000, \quad P_r = \frac{\sigma_0^2}{\sigma_r^2} = 1.6211 \cdot 10^{10} \frac{\mathrm{m}^2}{\mathrm{rad}^2}$$

point ID	X [m]	Y [m]
A	410.780	380.130
B	1183.460	1762.670
C	2077.030	433.380
D	1207.570	124.630
N_0	1175.150	997.720

Table 5.9: Coordinates.

measured:		distance [m]	direction [gon]
from	to	$\sigma_s = \pm 1$ cm	$\sigma_r = \pm 0.5$ mgon
N	A	982.690	193.1749
	B	765.000	337.1304
	C	1063.890	72.0344
	D	—	134.0758

Table 5.10: Distance and direction observations.

	leg	ΔX_{N}	ΔY_{N}	$\Delta\omega_0$	phys. unit
distance	N–A	0.777 83	0.628 47	0	$[-]$, $[\frac{\mathrm{m}}{\mathrm{rad}}]$
	N–B	−0.010 86	−0.999 94	0	
	N–C	−0.847 72	0.530 45	0	
direction	N–A	0.000 64	−0.000 79	−1	$[\frac{\mathrm{rad}}{\mathrm{m}}]$, $[-]$
	N–B	−0.001 31	0.000 01	−1	
	N–C	0.000 50	0.000 80	−1	
	N–D	0.001 14	0.000 04	−1	

Table 5.11: Designmatrix A.

leg	Δy [m]
N–A	0.000 37
N–B	0.004 86
N–C	−0.002 46
leg	**Δy [rad]**
N–A	$-7.4935 \cdot 10^{-6}$
N–B	$-2.5420 \cdot 10^{-6}$
N–C	$1.8678 \cdot 10^{-6}$
N–D	$-5.6276 \cdot 10^{-6}$

Table 5.12: Δy.

Table 5.11 shows the Designmatrix A, table 5.12 the reduced observation vector.

Table 5.13 contains the estimated parameters updates after the 1$^{\mathrm{st}}$ iteration and table 5.14 contains the adjusted coordinates for point N, including the adjusted orientation.

	$\widehat{\Delta\xi}$ [m]		$\widehat{\Delta\xi}$ [rad]
$\Delta\hat{X}_{\mathrm{N}_0}$	−0.000 49	$\Delta\hat{\omega}_{\mathrm{N}_0}$	$3.3544 \cdot 10^{-6}$
$\Delta\hat{Y}_{\mathrm{N}_0}$	0.001 60		

Table 5.13: Estimated parameters.

$\hat{X}_{\mathrm{N}}$	1175.150 m	$\hat{\omega}_{\mathrm{N}}$	63.5612 gon
$\hat{Y}_{\mathrm{N}}$	997.722 m		

Table 5.14: Adjusted Coordinates and orientation.

Table 5.15 gives the adjusted observations including the distance s_{ND}, which is approximately given by the adjusted coordinates of point N. Table 5.16 shows the inconsistencies $\hat{e}$ of the observations.

leg	$\hat{s}$ [m]	leg	$\hat{r}$ [rad]
N–A	982.690	N–A	193.1751
N–B	764.994	N–B	337.1304
N–C	1063.894	N–C	72.0341
N–D	873.693	N–D	134.0759

Table 5.15: Adjusted observations.

leg	$\hat{e}$ [m]	leg	$\hat{e}$ [gon]
N–A	−0.0003	N–A	−0.000 16
N–B	0.0065	N–B	$9.6 \cdot 10^{-6}$
N–C	−0.0037	N–C	0.000 27
N–D	–	N–C	−0.000 11

Table 5.16: Inconsistencies.

$$\Longrightarrow \quad \hat{e}^{\mathrm{T}} P \hat{e} = 0.999\,321\,8\,\mathrm{m}^2, \quad \text{(3 iterations, stop criterion } \|\widehat{\Delta\xi}\| < 10^{-10}\text{)}.$$

Finally, tables 5.17 and 5.18 show the standard deviations for coordinates for adjusted point N, as well as for adjusted orientation and observations.

Figure 5.7 shows the situation in detail, including the error ellipse for the adjusted coordinates of point N.

$\hat{\sigma}_{\hat{X}_N}$ [cm]	$\hat{\sigma}_{\hat{Y}_N}$ [cm]	$\hat{\sigma}_{\hat{\omega}_N}$ [mgon]
±0.19	±0.26	±0.13

Table 5.17: Standard deviations of coordinates and orientation.

$\hat{\sigma}_{\hat{r}_{NA}}$	$\hat{\sigma}_{\hat{r}_{NB}}$	$\hat{\sigma}_{\hat{r}_{NC}}$	$\hat{\sigma}_{\hat{r}_{ND}}$	$\hat{\sigma}_{\hat{s}_{NA}}$	$\hat{\sigma}_{\hat{s}_{NB}}$	$\hat{\sigma}_{\hat{s}_{NC}}$	$\hat{\sigma}_{\hat{s}_{ND}}$
[mgon]				[cm]			
±0.19	±0.23	±0.18	±0.17	±0.22	±0.26	±0.21	±0.26

Table 5.18: Standard deviations of directions and distances.

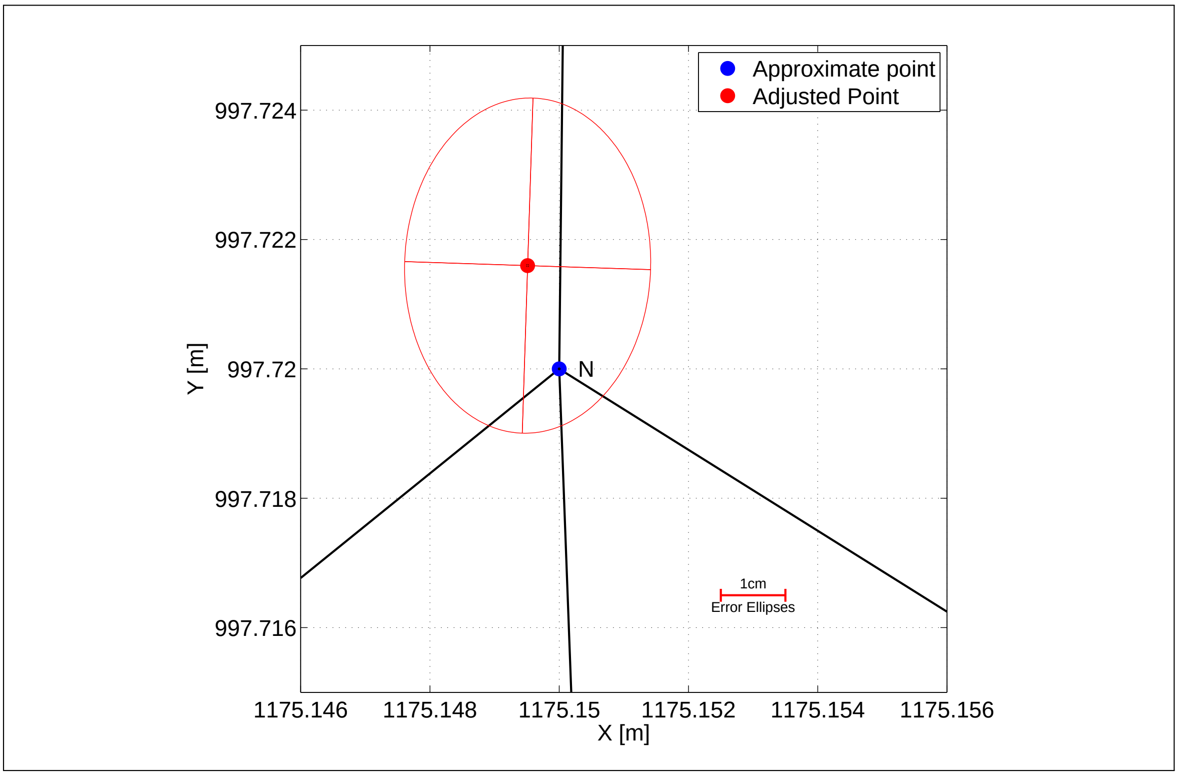


Figure 5.7: Detailed view of point N.

5.1.4 Distance and Direction Network (2a)

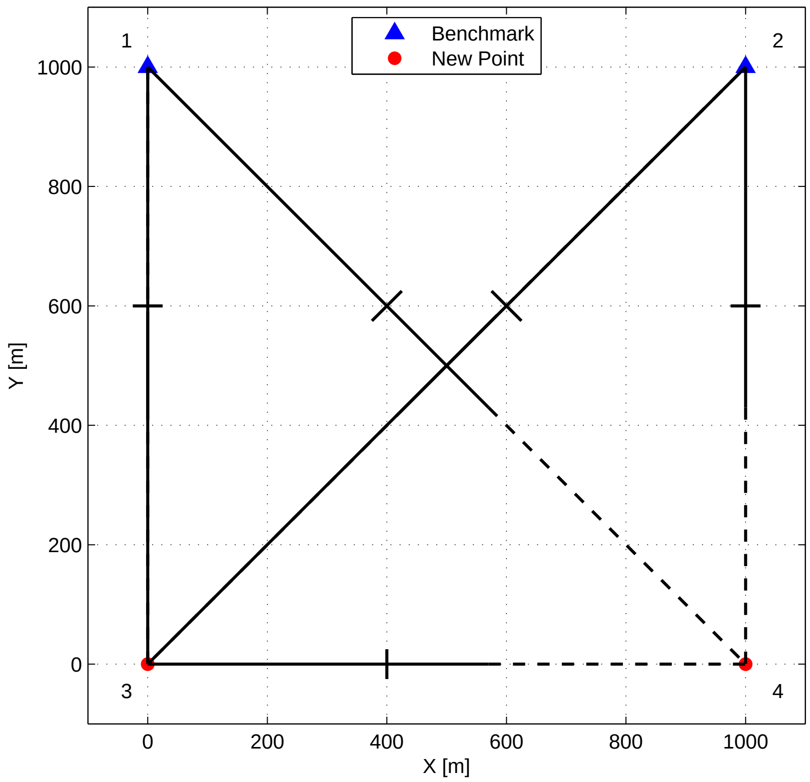


Figure 5.8: Network of Points.

This example (Benning, 2011, pg. 258–261) treats a network (Figure 5.8) of measured distances and directions between two given points and two new points. The standard deviation of observations: $\sigma_s = \pm 1$ cm for distances, $\sigma_r = \pm 1$ mgon for directions and $\sigma_0 = \pm 1$ cm as a priori standard deviation. This give the elements for the weight-matrix P:

$$\Longrightarrow \quad P_s = \frac{\sigma_0^2}{\sigma_s^2} = 1, \qquad P_r = \frac{\sigma_0^2}{\sigma_r^2} = 4.0528 \cdot 10^5 \, \frac{\mathrm{m}^2}{\mathrm{rad}^2}$$

Table 5.19 contains coordinates for benchmarks 1 and 2, table 5.20 approximate coordinates for points 3 and 4.

Point ID	X [m]	Y [m]
1	0.00	1000.00
2	1000.00	1000.00

Table 5.19: Benchmarks.

Point ID	X_0 [m]	Y_0 [m]
3	0.00	0.00
4	1000.00	0.00

Table 5.20: Approximate coordinates.

Table 5.21 contains measured distances (s) between individual points, approximate distances (s_0) and reduced observations (Δs).

leg	s [m]	s_0 [m]	Δs [m]
1–3	1000.02	1000.0000	0.0200
1–4	1414.20	1414.2136	−0.0136
2–3	1414.24	1414.2136	0.0264
2–4	999.98	1000.0000	−0.0200
3–4	1000.00	1000.0000	0.0000

Table 5.21: Distance observations.

Table 5.22 displays direction observations (r), approximate grid bearings (T_0), approximate orientation unknows (ω_0), approximate directions (r_0) and reduced direction observations (Δr_0).

leg	r [gon]	T_0 [gon]	ω_0 [gon]	$r_0 = T_0 - \omega_0$ [gon]	$\Delta r_0 = r - r_0$ [gon]
1–3	50.001	200.000	150.000	50.000	0.001
1–4	0.000	150.000		0.000	0
2–3	49.998	250.000	200.000	50.000	−0.002
2–4	0.000	200.000		0.000	0
3–1	0.000	0.000	0.000	0.000	0
3–2	49.999	50.000		50.000	−0.001
3–4	99.997	100.000		100.000	−0.003

Table 5.22: Direction observations.

Table 5.23 contains the designmatrix A and table 5.24 the reduced observation vector Δy after 1st iteration.

	leg	ΔX_3	ΔY_3	ΔX_4	ΔY_4	$\Delta\omega_1$	$\Delta\omega_2$	$\Delta\omega_3$	phys. unit
distance	1–3	0	−1	0	0	0	0	0	$[-]$, $[\frac{\text{m}}{\text{rad}}]$
	1–4	0	0	0.7071	−0.7071	0	0	0	
	2–3	−0.7071	−0.7071	0	0	0	0	0	
	2–4	0	0	0	−1	0	0	0	
	3–4	−1	0	1	0	0	0	0	
direction	1–3	−0.001	0	0	0	−1	0	0	$[\frac{\text{rad}}{\text{m}}]$, $[-]$
	1–4	0	0	−0.0005	−0.0005	−1	0	0	
	2–3	−0.0005	0.0005	0	0	0	−1	0	
	2–4	0	0	−0.001	0	0	−1	0	
	3–1	−0.001	0	0	0	0	0	−1	
	3–2	−0.0005	0.0005	0	0	0	0	−1	
	3–4	0	0.001	0	−0.001	0	0	−1	

Table 5.23: Designmatrix A.

leg	Δy [m]
1–3	0.02
1–4	−0.0136
2–3	0.0264
2–4	−0.02
3–4	0
leg	**Δy [rad]**
1–3	$1.5708 \cdot 10^{-5}$
1–4	0
2–3	$-3.1416 \cdot 10^{-5}$
2–4	0
3–1	0
3–2	$-1.5708 \cdot 10^{-5}$
3–4	$-4.7124 \cdot 10^{-5}$

Table 5.24: Δy.

Table 5.25 shows the estimated parameter updates after the 1st iteration and table 5.26 contains the adjusted coordinates for point 3 and 4 and adjusted orientations.

Table 5.27 contains the adjusted observations for distances and directions, and 5.28 the estimated inconsistencies $\hat{e}$ of the observations.

$$\Longrightarrow \quad \hat{e}^{\mathsf{T}} P \hat{e} = 1.0463\,\text{cm}^2 \quad \text{(3 iterations with stop criterion } \|\widehat{\Delta\xi}\| < 10^{-12}\text{)}$$

Finally, the tables 5.29, 5.30 and 5.31 contain standard deviations for adjusted coordinates and orientations as well as for adjusted distance and direction observations.

	$\widehat{\Delta\xi}$ [m]		$\widehat{\Delta\xi}$ [gon]
$\Delta\hat{X}_3$	−0.0101	$\Delta\hat{\omega}_1$	−0.0003
$\Delta\hat{Y}_3$	−0.0231	$\Delta\hat{\omega}_2$	0.0011
$\Delta\hat{X}_4$	−0.0096	$\Delta\hat{\omega}_3$	0.0006
$\Delta\hat{Y}_4$	0.0163		

Table 5.25: Estimated parameters.

$\hat{X}_3$	−0.010 m	$\hat{\omega}_1$	149.9997 gon
$\hat{Y}_3$	−0.023 m	$\hat{\omega}_2$	200.0011 gon
$\hat{X}_4$	999.990 m	$\hat{\omega}_3$	0.0006 gon
$\hat{Y}_4$	0.016 m		

Table 5.26: Adjusted coordinates and orientations.

leg	$\hat{s}$ [m]	leg	$\hat{r}$ [gon]
1−3	1000.023	1−3	50.0009
1−4	1414.195	1−4	0.0001
2−3	1414.237	2−3	49.9985
2−4	999.984	2−4	399.9995
3−4	1000.001	3−1	0.0001
		3−2	49.9990
		3−4	99.9969

Table 5.27: Adjusted observations.

leg	$\hat{e}$ [m]	leg	$\hat{e}$ [gon]
1−3	−0.0031	1−3	0.0001
1−4	0.0048	1−4	−0.0001
2−3	0.0029	2−3	−0.0005
2−4	−0.0037	2−4	0.0005
3−4	−0.0005	3−1	−0.0001
		3−2	0.0000
		3−4	0.0001

Table 5.28: Inconsistencies.

$\hat{\sigma}_{\hat{X}_3}$	$\hat{\sigma}_{\hat{Y}_3}$	$\hat{\sigma}_{\hat{X}_4}$	$\hat{\sigma}_{\hat{Y}_4}$	$\hat{\sigma}_{\hat{\omega}_1}$	$\hat{\sigma}_{\hat{\omega}_2}$	$\hat{\sigma}_{\hat{\omega}_3}$
±0.56 cm	±0.41 cm	±0.57 cm	±0.40 cm	±0.44 mgon	±0.44 mgon	±0.41 mgon

Table 5.29: Standard deviations of adjusted coordinates and orientations.

$\hat{\sigma}_{\hat{s}_{13}}$	$\hat{\sigma}_{\hat{s}_{14}}$	$\hat{\sigma}_{\hat{s}_{23}}$	$\hat{\sigma}_{\hat{s}_{24}}$	$\hat{\sigma}_{\hat{s}_{34}}$
±0.41 cm	±0.36 cm	±0.35 cm	±0.40 cm	±0.38 cm

Table 5.30: Standard deviations of distance observations.

$\hat{\sigma}_{\hat{r}_{13}} = \hat{\sigma}_{\hat{r}_{14}}$	$\hat{\sigma}_{\hat{r}_{23}} = \hat{\sigma}_{\hat{r}_{24}}$	$\hat{\sigma}_{\hat{r}_{31}}$	$\hat{\sigma}_{\hat{r}_{32}}$	$\hat{\sigma}_{\hat{r}_{34}}$
±0.34 mgon	±0.35 mgon	±0.30 mgon	±0.28 mgon	±0.32 mgon

Table 5.31: Standard deviations of direction observations.

Figure 5.9 shows network of points including error ellipses, figure 5.10 and figure 5.11 give a detailed view for points 3 and 4.

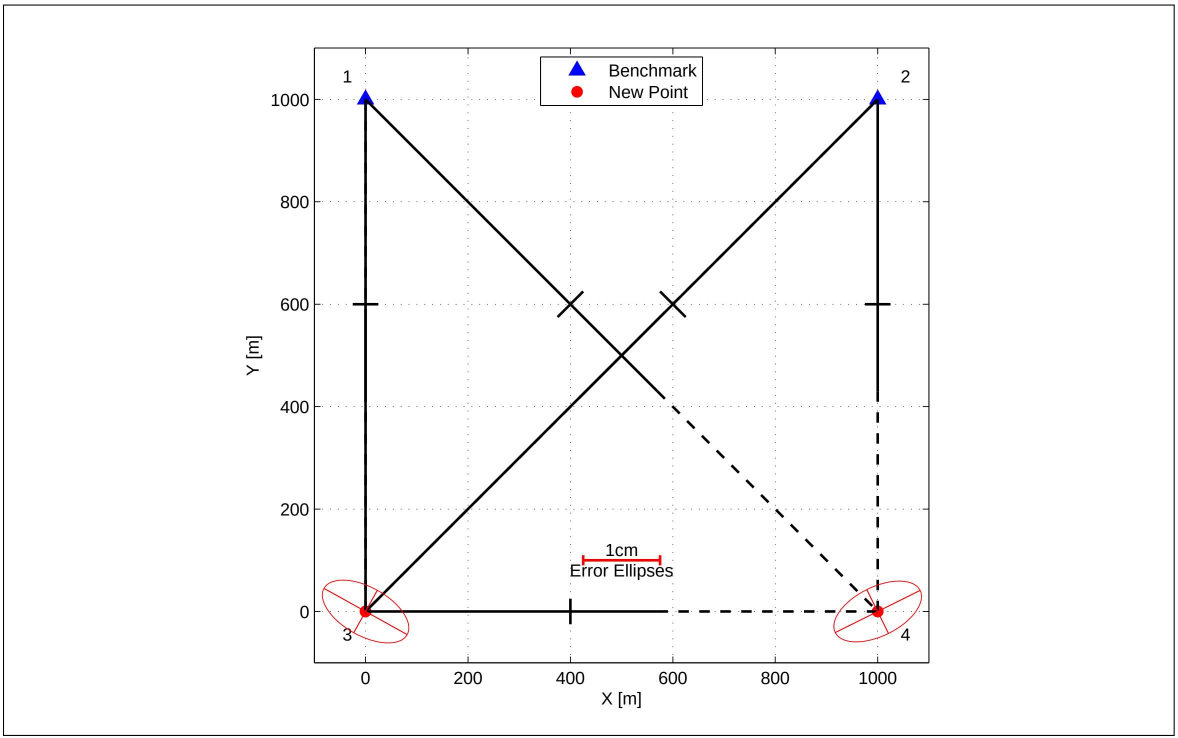


Figure 5.9: Network of points with error ellipses.

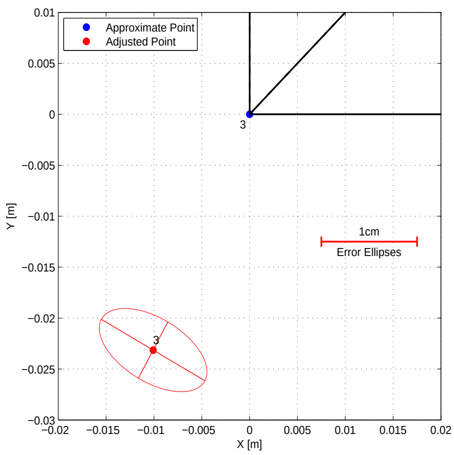


Figure 5.10: Detailed view point 3.

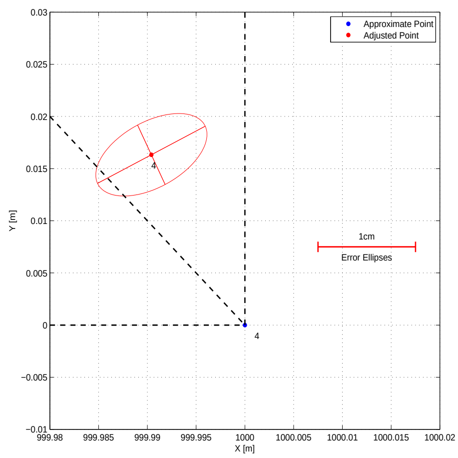


Figure 5.11: Detailed view point 4.

5.1.5 Free Adjustment: Distance and Direction Network (2b)

This example (Benning, 2011, pg. 273–281) processes the same network as before (Figure 5.8). Therefore, observations, weights and reduced observations Δy (Table 5.24) do not change. However, since we deal with a free adjustment here, designmatrix A (Table 5.32) is augmented by four additional columns comprising partial derivatives of those observations which also involve points 1 and 2.

	leg	ΔX_1	ΔY_1	ΔX_2	ΔY_2	ΔX_3	ΔY_3	ΔX_4	ΔY_4	$\Delta\omega_1$	$\Delta\omega_2$	$\Delta\omega_3$	phys. unit
distance	1–3	0	1	0	0	0	−1	0	0	0	0	0	[−], $[\frac{m}{rad}]$
	1–4	−0.707 11	0.707 11	0	0	0	0	0.707 11	−0.707 11	0	0	0	
	2–3	0	0	0.707 11	0.707 11	−0.707 11	−0.707 11	0	0	0	0	0	
	2–4	0	0	0	1	0	0	0	−1	0	0	0	
	3–4	0	0	0	0	−1	0	1	0	0	0	0	
direction	1–3	0.001	0	0	0	−0.001	0	0	0	−1	0	0	$[\frac{rad}{m}]$, [−]
	1–4	0.0005	0.0005	0	0	0	0	−0.0005	−0.0005	−1	0	0	
	2–3	0	0	0.0005	−0.0005	−0.0005	0.0005	0	0	0	−1	0	
	2–4	0	0	0.001	0	0	0	−0.001	0	0	−1	0	
	3–1	0.001	0	0	0	−0.001	0	0	0	0	0	−1	
	3–2	0	0	0.0005	−0.0005	−0.0005	0.0005	0	0	0	0	−1	
	3–4	0	0	0	0	0	0.001	0	−0.001	0	0	−1	

Table 5.32: Designmatrix A for distances and directions.

The free adjustment process makes use of the pseudoinverse N^+ shown in table 5.33 (1[st] iteration).

	ΔX_1	ΔY_1	ΔX_2	ΔY_2	ΔX_3	ΔY_3	ΔX_4	ΔY_4	$\Delta\omega_1$	$\Delta\omega_2$	$\Delta\omega_3$
ΔX_1	0.796 57	0.084 53	−0.762 01	0.144 78	−0.052 62	−0.119 09	0.018 06	−0.110 22	0.000 67	−0.000 63	0.000 12
ΔY_1	0.084 53	0.290 79	−0.209 99	−0.080 26	0.074 64	−0.165 33	0.050 82	−0.045 20	0.000 10	−0.000 22	−0.000 10
ΔX_2	−0.762 01	−0.209 99	0.928 06	−0.036 65	−0.032 15	0.043 95	−0.133 90	0.202 69	−0.000 63	0.000 79	−0.000 12
ΔY_2	0.144 78	−0.080 26	−0.036 65	0.263 51	−0.041 16	−0.027 87	−0.066 96	−0.155 39	0.000 16	−0.000 06	0.000 06
ΔX_3	−0.052 62	0.074 64	−0.032 15	−0.041 16	0.205 97	0.010 13	−0.121 21	−0.043 60	−0.000 08	0.000 00	−0.000 10
ΔY_3	−0.119 09	−0.165 33	0.043 95	−0.027 87	0.010 13	0.240 47	0.065 01	−0.047 27	−0.000 14	0.000 07	0.000 10
ΔX_4	0.018 06	0.050 82	−0.133 90	−0.066 96	−0.121 21	0.065 01	0.237 04	−0.048 87	0.000 04	−0.000 16	0.000 10
ΔY_4	−0.110 22	−0.045 20	0.202 69	−0.155 39	−0.043 60	−0.047 27	−0.048 87	0.247 86	−0.000 12	0.000 21	−0.000 06
$\Delta\omega_1$	0.000 67	0.000 10	−0.000 63	0.000 16	−0.000 08	−0.000 14	0.000 04	−0.000 12	0.000 00	0.000 00	0.000 00
$\Delta\omega_2$	−0.000 63	−0.000 22	0.000 79	−0.000 06	0.000 00	0.000 07	−0.000 16	0.000 21	0.000 00	0.000 00	0.000 00
$\Delta\omega_3$	0.000 12	−0.000 10	−0.000 12	0.000 06	−0.000 10	0.000 10	0.000 10	−0.000 06	0.000 00	0.000 00	0.000 00

Table 5.33: Pseudoinverse N^+.

Table 5.34 shows estimated parameters (1[st] iteration) and table 5.35 contains the adjusted coordinates for point 1–4 and also adjusted orientations.

	$\widehat{\Delta\xi}$ [m]		$\widehat{\Delta\xi}$ [gon]
$\Delta\hat{X}_1$	0.0018	$\Delta\hat{\omega}_1$	−0.0003
$\Delta\hat{Y}_1$	0.0031	$\Delta\hat{\omega}_2$	0.0017
$\Delta\hat{X}_2$	0.0135	$\Delta\hat{\omega}_3$	0.0008
$\Delta\hat{Y}_2$	−0.0014		
$\Delta\hat{X}_3$	−0.0076		
$\Delta\hat{Y}_3$	−0.0184		
$\Delta\hat{X}_4$	−0.0077		
$\Delta\hat{Y}_4$	0.0167		

Table 5.34: Estimated parameters.

	[m]		[gon]
$\hat{X}_1$	0.002	$\hat{\omega}_1$	149.9997
$\hat{Y}_1$	1000.003	$\hat{\omega}_2$	200.0017
$\hat{X}_2$	1000.013	$\hat{\omega}_3$	0.0008
$\hat{Y}_2$	999.999		
$\hat{X}_3$	−0.008		
$\hat{Y}_3$	−0.018		
$\hat{X}_4$	999.992		
$\hat{Y}_4$	0.017		

Table 5.35: Adjusted coordinates and orientations.

Table 5.36 contains the adjusted observations for distances and directions, while table 5.37 shows the estimated inconsistencies $\hat{e}$ of the observations.

leg	$\hat{s}$ [m]	leg	$\hat{r}$ [gon]
1–3	1000.021	1–3	50.0009
1–4	1414.197	1–4	0.0001
2–3	1414.240	2–3	49.9984
2–4	999.982	2–4	399.9996
3–4	1000.000	3–1	399.9998
		3–2	49.9993
		3–4	99.9969

Table 5.36: Adjusted observations.

leg	$\hat{e}$ [m]	leg	$\hat{e}$ [gon]
1–3	−0.0015	1–3	0.0001
1–4	0.0027	1–4	−0.0001
2–3	−0.0004	2–3	−0.0004
2–4	−0.0019	2–4	0.0004
3–4	0.0001	3–1	0.0002
		3–2	−0.0003
		3–4	0.0001

Table 5.37: Inconsistencies.

$$\Longrightarrow \quad \hat{e}^{\mathsf{T}} P \hat{e} = 0.628\,\mathrm{cm}^2 \quad \text{(3 iterations with stop criterion } \|\widehat{\Delta\xi}\| < 10^{-12}\text{)}$$

Finally again, the tables 5.38, 5.39, 5.40 and 5.41 contain standard deviations for adjusted coordinates and orientations as well as for adjusted distance and direction observations.

$\hat{\sigma}_{\hat{X}_1}$	$\hat{\sigma}_{\hat{Y}_1}$	$\hat{\sigma}_{\hat{X}_2}$	$\hat{\sigma}_{\hat{Y}_2}$	$\hat{\sigma}_{\hat{X}_3}$	$\hat{\sigma}_{\hat{Y}_3}$	$\hat{\sigma}_{\hat{X}_4}$	$\hat{\sigma}_{\hat{Y}_4}$
±0.35 cm	±0.21 cm	±0.38 cm	±0.20 cm	±0.18 cm	±0.19 cm	±0.19 cm	±0.20 cm

Table 5.38: Standard deviations of adjusted coordinates.

$\hat{\sigma}_{\hat{\omega}_1}$	$\hat{\sigma}_{\hat{\omega}_2}$	$\hat{\sigma}_{\hat{\omega}_3}$
±0.34 mgon	±0.35 mgon	±0.25 mgon

Table 5.39: Standard deviations of adjusted orientations.

$\hat{\sigma}_{\hat{s}_{13}}$	$\hat{\sigma}_{\hat{s}_{14}}$	$\hat{\sigma}_{\hat{s}_{23}}$	$\hat{\sigma}_{\hat{s}_{24}}$	$\hat{\sigma}_{\hat{s}_{34}}$
±0.37 cm	±0.34 cm	±0.37 cm	±0.36 cm	±0.33 cm

Table 5.40: Standard deviations of adjusted distance observations.

$\hat{\sigma}_{\hat{r}_{13}} = \hat{\sigma}_{\hat{r}_{14}}$	$\hat{\sigma}_{\hat{r}_{23}} = \hat{\sigma}_{\hat{r}_{24}}$	$\hat{\sigma}_{\hat{r}_{31}}$	$\hat{\sigma}_{\hat{r}_{32}}$	$\hat{\sigma}_{\hat{r}_{34}}$
±0.30 mgon	±0.31 mgon	±0.32 mgon	±0.30 mgon	±0.28 mgon

Table 5.41: Standard deviations of adjusted direction observations.

Figure 5.12 shows network of points including error ellipses, figure 5.13, 5.14, 5.15 and 5.16 give a detailed view for points 1, 2, 3 and 4.

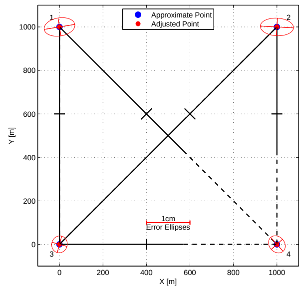


Figure 5.12: Network of points with error ellipses.

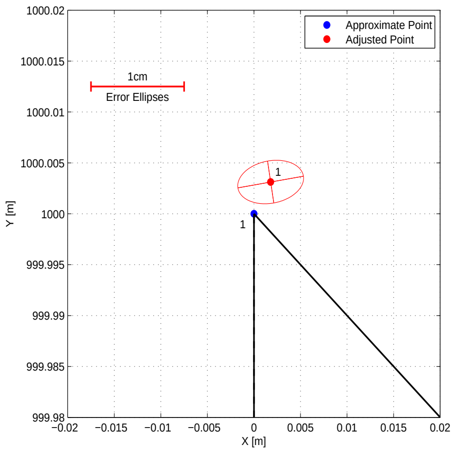


Figure 5.13: Detailed view point 1.

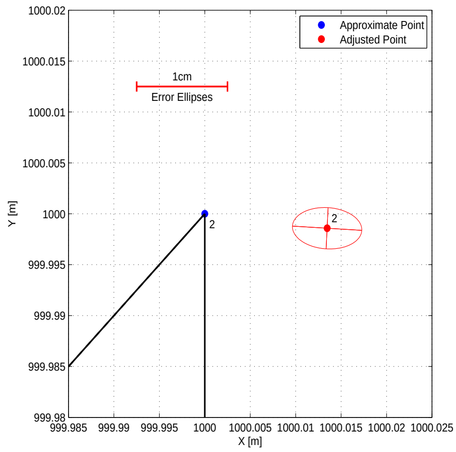


Figure 5.14: Detailed view point 2

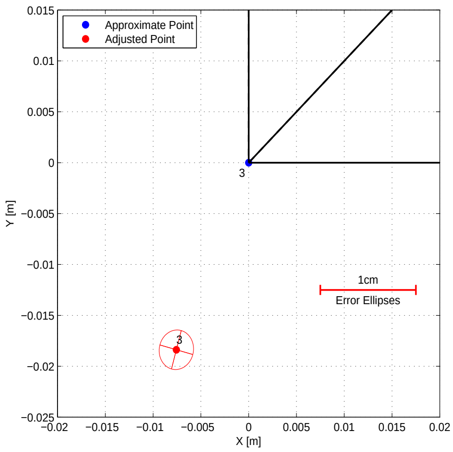


Figure 5.15: Detailed view point 3

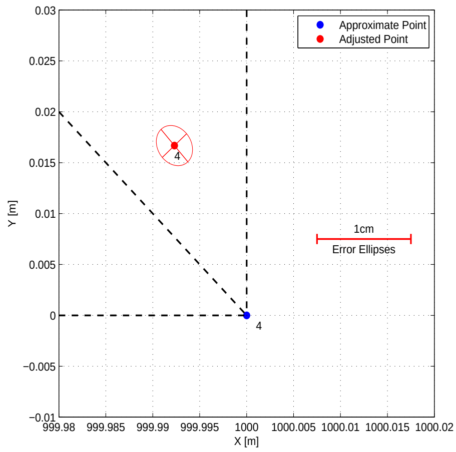


Figure 5.16: Detailed view point 4

5.1.6 Overconstrained adjustment: Distance, direction and angle network

This example, taken from Wolf (1979, pg. 66–78), consists of a 10-point network observed by directions, one distance and one angle, see (5.17). The network is overconstrained because its datum is defined by 6 benchmarks A–F.

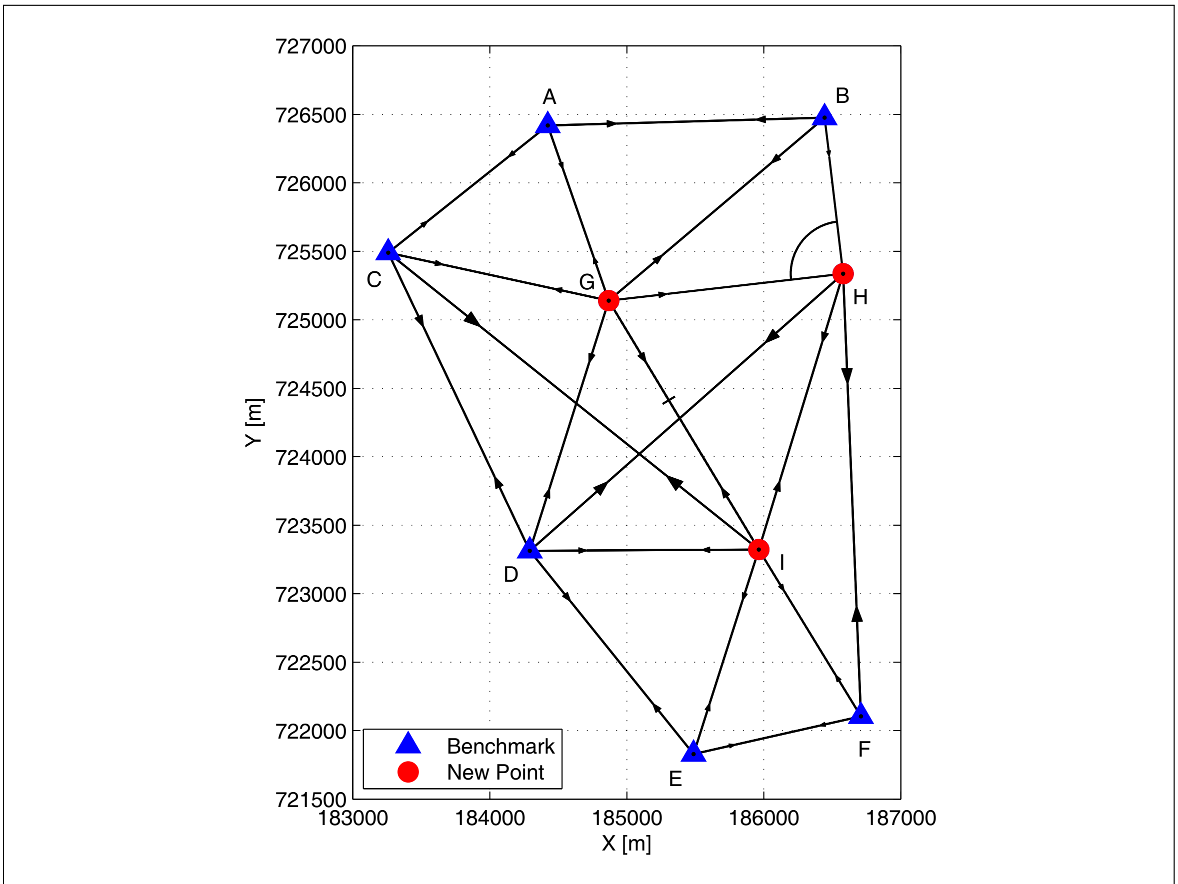


Figure 5.17: Network design

The observations are collected in table 5.43 (2nd column), and the following standard deviations have been assumed:

$$\begin{aligned}\sigma_r &= \pm 2.5\,\text{mgon} \quad &&\text{(directions)},\\ \sigma_\alpha &= \pm 3.5\,\text{mgon} &&\text{(angle)},\\ \sigma_s &= \pm 3\,\text{cm} &&\text{(distance)}.\end{aligned}$$

If the a priori standard deviation is taken as $\sigma_0 = \pm 2.5\,\text{mgon}$ then the weight matrix elements turn out to be

$$P_r = \frac{\sigma_0^2}{\sigma_r^2} = 1, \quad P_\alpha = \frac{\sigma_0^2}{\sigma_\alpha^2} = 0.5102 \quad \text{and} \quad P_s = \frac{\sigma_0^2}{\sigma_s^2} = 1.7135 \cdot 10^{-6}\,\frac{\text{rad}^2}{\text{m}^2}\,.$$

Table 5.42 contains coordinates for benchmarks A to F and approximate coordinates G, H and I.

benchmarks A–F		
point ID	X [m]	Y [m]
A	184 423.28	726 419.33
B	186 444.18	726 476.66
C	183 257.84	725 490.35
D	184 292.00	723 313.00
E	185 487.00	721 829.00
F	186 708.72	722 104.58
approximate coordinates new points G–I		
point ID	X_0 [m]	Y_0 [m]
G	184 868.20	725 139.70
H	186 579.30	725 336.60
I	185 963.07	723 322.02

Table 5.42: Benchmarks and approximate coordinates.

Additionally to observations, table 5.43 includes approximate grid bearings (T_0), angle (α_0) and distance (s_0), approximate orientation unknowns (ω_0), approximate directions (r_0) and reduced observations (Δr_0). Orientation unknowns (ω_0) are mean values calculated from Δr_0.

Table 5.44 contains the designmatrix A after 1$^{\text{st}}$ iteration.

Table 5.45 shows the estimated parameters after 1$^{\text{st}}$ iteration and table 5.46 the adjusted coordinates and orientations.

Table 5.47 contains the estimated inconsistencies $\hat{e}$ of the observations leading to a weighted square sum of residuals

$$\hat{e}^{\mathsf{T}} P \hat{e} = 0.00225\,\text{gon}^2 \quad \text{(4 iterations, stop criteria } \|\widehat{\Delta\xi}\| < 10^{-10}\text{)}$$

Finally, the tables 5.48 and 5.49 contain the standard deviations of coordinates, orientations and observations.

Figures 5.18 and 5.19 show the adjusted network with corresponding absolute and relative error ellipses at/between the new points.

Table 5.50 contains the elements (semi major axis a, semi minor axis b and bearing ϕ of a) for absolute error ellipses for new points and also for relative error ellipses between new points.

Figures 5.20, 5.21 and 5.22 give a detailed view, for the new points G, H and I.

Distance observation [m]						
leg	s	s_0	Δs_0			
G– I	2121.90	2121.96	−0.06	—	—	—
Direction observations [gon]						
leg	r	T_0	$\Delta r_0 = T_0 - r$	ω_0	$r_\omega^0 = T_0 - \omega_0$	$\Delta r_\omega^0 = r - r_\omega^0$
A–B	0.0000	98.1945	98.1945	98.1960	−0.0008	0.0008
A–G	80.5000	178.6975	98.1975		80.5022	−0.0022
A–C	158.9610	257.1571	98.1961		158.9618	−0.0008
B–H	0.0000	192.4898	192.4861	192.4861	0.0073	−0.0073
B–G	62.7260	255.2121	192.4861		62.7296	−0.0036
B–A	105.7120	298.1945	192.4824		105.7120	0.0000
C–A	0.0000	57.1571	57.1571	57.1640	−0.0095	0.0095
C–G	56.4960	113.6491	57.1531		56.4825	0.0135
C– I	85.8450	143.0148	57.1698		85.8482	−0.0032
C–D	114.5950	171.7711	57.1761		114.6045	−0.0095
D–G	0.0000	19.4522	19.4522	19.4476	0.0015	−0.0015
D–H	34.4500	53.8894	19.4394		34.4387	0.0113
D– I	80.2110	99.6564	19.4454		80.2057	0.0053
D–E	137.4020	156.8412	19.4391		137.3905	0.0115
D–C	352.3090	371.7711	19.4621		352.3205	−0.0115
E– I	0.0000	19.6507	19.6507	19.6357	0.0224	−0.0224
E–F	66.2450	85.8763	19.6313		66.2481	−0.0031
E–D	337.2160	356.8412	19.6251		337.2129	0.0031
F–E	0.0000	285.8763	285.8763	285.8642	0.0000	0.0000
F– I	79.1690	365.0151	285.8461		79.1388	0.0302
F–H	111.5820	397.4521	285.8701		111.5758	0.0062
G–B	0.0000	55.2121	55.2121	55.2147	−0.0027	0.0027
G–H	37.4980	92.7064	55.2083		37.4916	0.0064
G– I	110.2580	165.4862	55.2282		110.2715	−0.0135
G–D	164.2320	219.4522	55.2201		164.2374	−0.0054
G–C	258.4410	313.6491	55.2081		258.4344	0.0066
G–A	323.4860	378.6975	55.2115		323.4827	0.0033
H–F	0.0000	197.4521	197.4521	197.4508	0.0013	−0.0013
H– I	21.4450	218.8979	197.4529		21.4471	−0.0021
H–D	56.4420	253.8889	197.4474		56.4386	0.0034
I –H	0.0000	18.8979	18.8979	18.9025	−0.0046	0.0046
I –F	146.1430	165.0151	18.8721		146.1126	0.0304
I –E	200.7330	219.6507	18.9177		200.7482	−0.0152
I –D	280.7560	299.6564	18.9004		280.7539	0.0021
I –C	324.1050	343.0148	18.9098		324.1123	−0.0073
I –G	346.5690	365.4862	18.9172		346.5837	−0.0147
Angle observation [gon]						
leg	α	α_0	$\Delta\alpha_0$			
α_{HGB}	99.7810	99.7834	−0.0024	—	—	—

Table 5.43: Observations.

leg	ΔX_G	ΔY_G	ΔX_H	ΔY_H	ΔX_I	ΔY_I	$\Delta\omega_A$	$\Delta\omega_B$	$\Delta\omega_C$	$\Delta\omega_D$	$\Delta\omega_E$	$\Delta\omega_F$	$\Delta\omega_G$	$\Delta\omega_H$	$\Delta\omega_I$	phys. unit
Distance observation																
G – I	−0.5159	0.8566	0	0	0.5159	−0.8566	0	0	0	0	0	0	0	0	0	[−], [$\frac{m}{rad}$]
Direction observations																
A – B	0	0	0	0	0	0	−1	0	0	0	0	0	0	0	0	[$\frac{rad}{m}$], [−]
A – G	−0.0007	−0.0002	0	0	0	0	−1	0	0	0	0	0	0	0	0	
A – C	0	0	0	0	0	0	−1	0	0	0	0	0	0	0	0	
B – H	0	0	−0.0009	−0.0001	0	0	0	−1	0	0	0	0	0	0	0	
B – G	−0.0003	0.0004	0	0	0	0	0	−1	0	0	0	0	0	0	0	
B – A	0	0	0	0	0	0	0	−1	0	0	0	0	0	0	0	
C – A	0	0	0	0	0	0	0	0	−1	0	0	0	0	0	0	
C – G	−0.0001	−0.0006	0	0	0	0	0	0	−1	0	0	0	0	0	0	
C – I	0	0	0	0	−0.0002	−0.0002	0	0	−1	0	0	0	0	0	0	
C – D	0	0	0	0	0	0	0	0	−1	0	0	0	0	0	0	
D – G	0.0005	−0.0002	0	0	0	0	0	0	0	−1	0	0	0	0	0	
D – H	0	0	0.0002	−0.0002	0	0	0	0	0	−1	0	0	0	0	0	
D – I	0	0	0	0	$3.2 \cdot 10^{-6}$	−0.0006	0	0	0	−1	0	0	0	0	0	
D – E	0	0	0	0	0	0	0	0	0	−1	0	0	0	0	0	
D – C	0	0	0	0	0	0	0	0	0	−1	0	0	0	0	0	
E – I	0	0	0	0	0.0006	−0.0002	0	0	0	0	−1	0	0	0	0	
E – F	0	0	0	0	0	0	0	0	0	0	−1	0	0	0	0	
E – D	0	0	0	0	0	0	0	0	0	0	−1	0	0	0	0	
F – E	0	0	0	0	0	0	0	0	0	0	0	−1	0	0	0	
F – I	0	0	0	0	0.0006	0.0004	0	0	0	0	0	−1	0	0	0	
F – H	0	0	0.0003	$1.2 \cdot 10^{-5}$	0	0	0	0	0	0	0	−1	0	0	0	
G – B	−0.0003	0.0004	0	0	0	0	0	0	0	0	0	0	−1	0	0	
G – H	$-6.6 \cdot 10^{-5}$	0.0006	$6.6 \cdot 10^{-5}$	−0.0006	0	0	0	0	0	0	0	0	−1	0	0	
G – I	0.0004	0.0002	0	0	−0.0004	−0.0002	0	0	0	0	0	0	−1	0	0	
G – D	0.0005	−0.0001	0	0	0	0	0	0	0	0	0	0	−1	0	0	
G – C	−0.0001	−0.0006	0	0	0	0	0	0	0	0	0	0	−1	0	0	
G – A	−0.0007	−0.0002	0	0	0	0	0	0	0	0	0	0	−1	0	0	
H – F	0	0	0.0003	$1.2 \cdot 10^{-5}$	0	0	0	0	0	0	0	0	0	−1	0	
H – I	0	0	0.0005	−0.0001	−0.0005	0.0001	0	0	0	0	0	0	0	−1	0	
H – D	0	0	0.0002	−0.0002	0	0	0	0	0	0	0	0	0	−1	0	
I – H	0	0	0.0005	−0.0001	−0.0005	0.0001	0	0	0	0	0	0	0	0	−1	
I – F	0	0	0	0	0.0006	0.0004	0	0	0	0	0	0	0	0	−1	
I – E	0	0	0	0	0.0006	−0.0002	0	0	0	0	0	0	0	0	−1	
I – D	0	0	0	0	$3.2 \cdot 10^{-6}$	−0.0006	0	0	0	0	0	0	0	0	−1	
I – C	0	0	0	0	−0.0002	−0.0002	0	0	0	0	0	0	0	0	−1	
I – G	0.0004	0.0002	0	0	−0.0004	−0.0002	0	0	0	0	0	0	0	0	−1	
Angle observation																
α_{HGB}	$6.6 \cdot 10^{-5}$	−0.0006	−0.0009	0.0004	0	0	0	0	0	0	0	0	0	0	0	[$\frac{rad}{m}$], [−]

Table 5.44: Designmatrix A (after 1st iteration).

	$\widehat{\Delta\xi}$ [m]
$\Delta\hat{X}_G$	−0.162
$\Delta\hat{Y}_G$	−0.043
$\Delta\hat{X}_H$	0.037
$\Delta\hat{Y}_H$	−0.186
$\Delta\hat{X}_I$	0.145
$\Delta\hat{Y}_I$	0.283
	$\widehat{\Delta\xi}$ [gon]
$\Delta\hat{\omega}_A$	0.0026
$\Delta\hat{\omega}_B$	0.0005
$\Delta\hat{\omega}_C$	−0.0007
$\Delta\hat{\omega}_D$	−0.0024
$\Delta\hat{\omega}_E$	0.0007
$\Delta\hat{\omega}_F$	0.0042
$\Delta\hat{\omega}_G$	0.0002
$\Delta\hat{\omega}_H$	0.0017
$\Delta\hat{\omega}_I$	−0.0024

Table 5.45: Estimated parameters after 1st iteration.

	[m]
$\hat{X}_G$	184 868.038
$\hat{Y}_G$	725 139.657
$\hat{X}_H$	186 579.337
$\hat{Y}_H$	725 336.414
$\hat{X}_I$	185 963.215
$\hat{Y}_I$	723 322.303
	[gon]
$\hat{\omega}_A$	98.1987
$\hat{\omega}_B$	192.4866
$\hat{\omega}_C$	57.1634
$\hat{\omega}_D$	19.4452
$\hat{\omega}_E$	19.6364
$\hat{\omega}_F$	285.8684
$\hat{\omega}_G$	55.2150
$\hat{\omega}_H$	197.4525
$\hat{\omega}_I$	18.9001

Table 5.46: Adjusted coordinates and orientations.

leg	$\hat{s}$ [m]	$\hat{e}$ [m]						
G–I	2121.836	0.0638						
leg	$\hat{r}$ [gon]	$\hat{e}$ [gon]	leg	$\hat{r}$ [gon]	$\hat{e}$ [gon]	leg	$\hat{r}$ [gon]	$\hat{e}$ [gon]
A–B	−0.0042	0.0042	D–I	80.2004	0.0106	G–D	164.2325	−0.0005
A–G	80.5067	−0.0067	D–E	137.3959	0.0061	G–C	258.4371	0.0039
A–C	158.9585	0.0025	D–C	352.3259	−0.0169	G–A	323.4904	−0.0044
B–H	0.0024	−0.0024	E–I	0.0164	−0.0164	H–F	0.0003	−0.0003
B–G	62.7277	−0.0017	E–F	66.2399	0.0051	H–I	21.4464	−0.0014
B–A	105.7079	0.0041	E–D	337.2047	0.0113	H–D	56.4403	0.0017
C–A	−0.0062	0.0062	F–E	0.0079	−0.0079	I–H	−0.0012	0.0012
C–G	56.4887	0.0072	F–I	79.1588	0.0102	I–F	146.1271	0.0159
C–I	85.8457	−0.0007	F–H	111.5843	−0.0023	I–E	200.7527	−0.0197
C–D	114.6078	−0.0128	G–B	−0.0007	0.0007	I–D	280.7455	0.0105
D–G	0.0022	−0.0022	G–H	37.4975	0.0005	I–C	324.1089	−0.0039
D–H	34.4476	0.0024	G–I	110.2583	−0.0003	I–G	346.5731	−0.0041
						leg	$\hat{\alpha}$ [gon]	$\hat{e}$ [gon]
						α_{HGB}	99.7765	0.0045

Table 5.47: Estimated inconsistencies.

	[cm]		[mgon]
$\hat{\sigma}_{\hat{X}_{\mathrm{G}}}$	±11.866	$\hat{\sigma}_{\hat{\omega}_{\mathrm{A}}}$	±6.0023
$\hat{\sigma}_{\hat{Y}_{\mathrm{G}}}$	±13.078	$\hat{\sigma}_{\hat{\omega}_{\mathrm{B}}}$	±6.7376
$\hat{\sigma}_{\hat{X}_{\mathrm{H}}}$	±15.816	$\hat{\sigma}_{\hat{\omega}_{\mathrm{C}}}$	±5.1859
$\hat{\sigma}_{\hat{Y}_{\mathrm{H}}}$	±26.380	$\hat{\sigma}_{\hat{\omega}_{\mathrm{D}}}$	±4.8772
$\hat{\sigma}_{\hat{X}_{\mathrm{I}}}$	±11.470	$\hat{\sigma}_{\hat{\omega}_{\mathrm{E}}}$	±5.9353
$\hat{\sigma}_{\hat{Y}_{\mathrm{I}}}$	±13.537	$\hat{\sigma}_{\hat{\omega}_{\mathrm{F}}}$	±6.1002
		$\hat{\sigma}_{\hat{\omega}_{\mathrm{G}}}$	±4.3863
		$\hat{\sigma}_{\hat{\omega}_{\mathrm{H}}}$	±6.5588
		$\hat{\sigma}_{\hat{\omega}_{\mathrm{I}}}$	±4.3554

Table 5.48: Estimated standard deviations of coordinates and orientations.

	[cm]						
$\hat{\sigma}_{\hat{s}_{GI}}$	±10.2660						
	[mgon]		[mgon]		[mgon]		[mgon]
$\hat{\sigma}_{\hat{r}_{\mathrm{AB}}}$	±6.0023	$\hat{\sigma}_{\hat{r}_{\mathrm{CD}}}$	±5.1859	$\hat{\sigma}_{\hat{r}_{\mathrm{FE}}}$	±6.1002	$\hat{\sigma}_{\hat{r}_{\mathrm{HF}}}$	±6.1800
$\hat{\sigma}_{\hat{r}_{\mathrm{AG}}}$	±6.8033	$\hat{\sigma}_{\hat{r}_{\mathrm{DG}}}$	±5.3020	$\hat{\sigma}_{\hat{r}_{\mathrm{FI}}}$	±6.6864	$\hat{\sigma}_{\hat{r}_{\mathrm{HI}}}$	±6.1533
$\hat{\sigma}_{\hat{r}_{\mathrm{AC}}}$	±6.0023	$\hat{\sigma}_{\hat{r}_{\mathrm{DH}}}$	±5.3830	$\hat{\sigma}_{\hat{r}_{\mathrm{FH}}}$	±6.2494	$\hat{\sigma}_{\hat{r}_{\mathrm{HD}}}$	±6.3495
$\hat{\sigma}_{\hat{r}_{\mathrm{BH}}}$	±8.3169	$\hat{\sigma}_{\hat{r}_{\mathrm{DI}}}$	±5.7282	$\hat{\sigma}_{\hat{r}_{\mathrm{GB}}}$	±6.0317	$\hat{\sigma}_{\hat{r}_{\mathrm{IH}}}$	±6.1128
$\hat{\sigma}_{\hat{r}_{\mathrm{BG}}}$	±6.7802	$\hat{\sigma}_{\hat{r}_{\mathrm{DE}}}$	±4.8772	$\hat{\sigma}_{\hat{r}_{\mathrm{GH}}}$	±8.1933	$\hat{\sigma}_{\hat{r}_{\mathrm{IF}}}$	±6.9589
$\hat{\sigma}_{\hat{r}_{\mathrm{BA}}}$	±6.7376	$\hat{\sigma}_{\hat{r}_{\mathrm{DC}}}$	±4.8772	$\hat{\sigma}_{\hat{r}_{\mathrm{GI}}}$	±6.0859	$\hat{\sigma}_{\hat{r}_{\mathrm{IE}}}$	±5.7188
$\hat{\sigma}_{\hat{r}_{\mathrm{CA}}}$	±5.1859	$\hat{\sigma}_{\hat{r}_{\mathrm{EI}}}$	±6.5637	$\hat{\sigma}_{\hat{r}_{\mathrm{GD}}}$	±6.1072	$\hat{\sigma}_{\hat{r}_{\mathrm{ID}}}$	±5.6981
$\hat{\sigma}_{\hat{r}_{\mathrm{CG}}}$	±6.0882	$\hat{\sigma}_{\hat{r}_{\mathrm{EF}}}$	±5.9353	$\hat{\sigma}_{\hat{r}_{\mathrm{GC}}}$	±6.1493	$\hat{\sigma}_{\hat{r}_{\mathrm{IC}}}$	±4.5646
$\hat{\sigma}_{\hat{r}_{\mathrm{CI}}}$	±5.2157	$\hat{\sigma}_{\hat{r}_{\mathrm{ED}}}$	±5.9353	$\hat{\sigma}_{\hat{r}_{\mathrm{GA}}}$	±6.5839	$\hat{\sigma}_{\hat{r}_{\mathrm{IG}}}$	±5.7435
							[mgon]
						$\hat{\sigma}_{\hat{\alpha}_{\mathrm{HGB}}}$	±9.4045

Table 5.49: Estimated standard deviations of the observations.

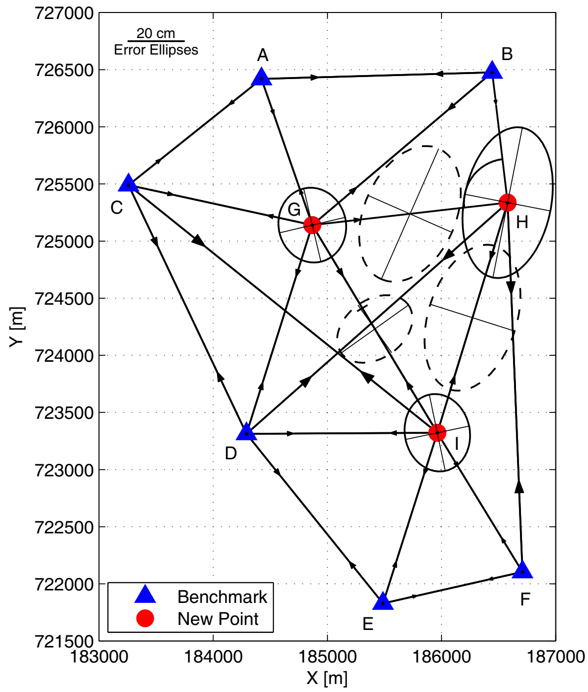


Figure 5.18: Network with error ellipses.

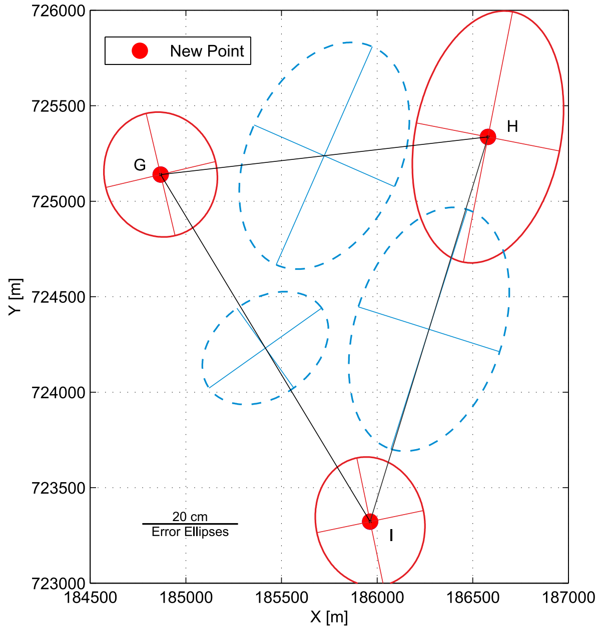


Figure 5.19: Detailed view: absolute and relative error ellipses.

Element's absolute error ellipse			
point	ϕ [gon]	a [cm]	b [cm]
G	185.2077	13.147	11.790
H	12.3417	26.717	15.240
I	186.9145	13.623	11.367
Element's relativ error ellipse			
leg	ϕ [gon]	a [cm]	b [cm]
G–H	26.3811	24.956	16.044
H–I	19.5521	26.328	15.502
G–I	60.6365	14.447	10.237

Table 5.50: Error ellipse elements.

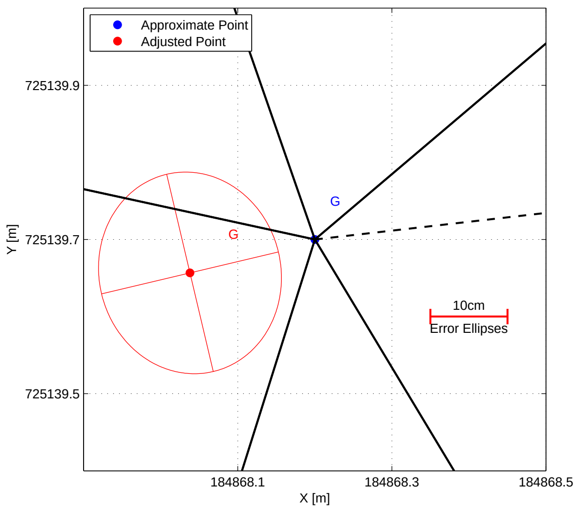


Figure 5.20: Detailed view of point G.

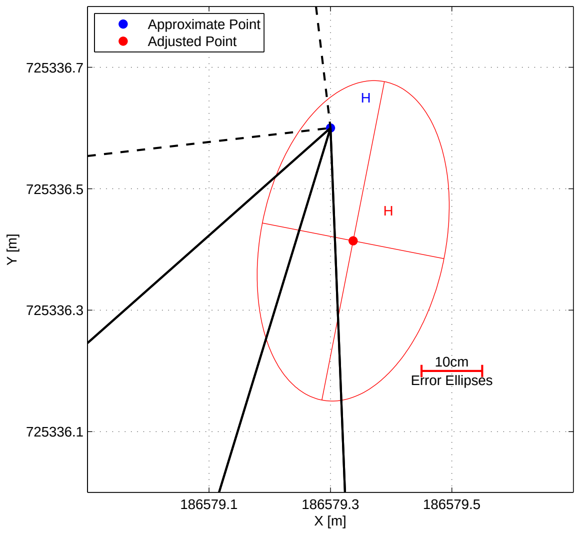


Figure 5.21: Detailed view of point H.

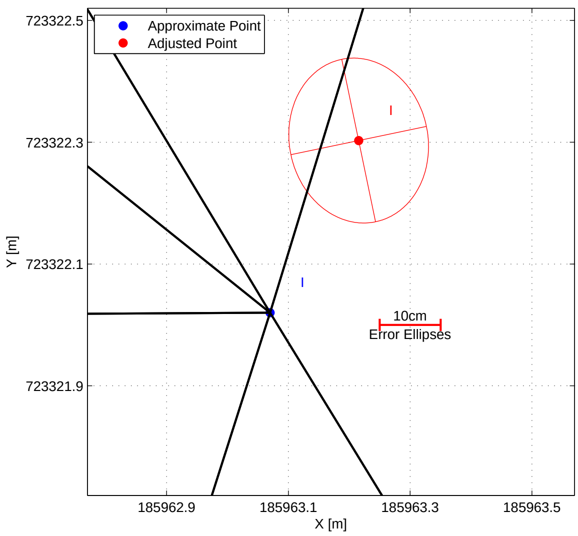


Figure 5.22: Detailed view of point I.

5.1.7 Polynomial fit

Observations: y_i, $i = 1, \ldots, m$.

Given: fixed x-coordinates x_i, $i = 1, \ldots, m$.

Find parameters a_n, $n = 0, \ldots, n_{\max}$ of fitting polynomial

$$f(x) = y = \sum_{n=0}^{n_{\max}} a_n x^n \,.$$

Possible additional restrictions:

(a) tangent in (x_T, y_T) should pass through (x_P, y_P) or

(b) fitting polynomial should pass through (x_Q, y_Q) or

(c) unknown coefficient a_k shall get the numerical value $\tilde{a}_k$.

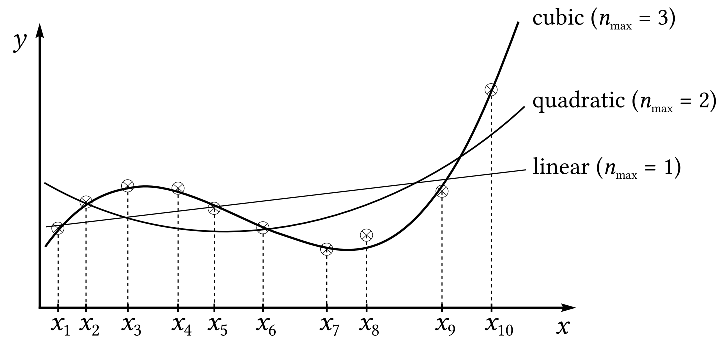


Figure 5.23: Fitting polynomials of different degrees.

Observation equation

$$y_i = \sum_{n=0}^{n_{\max}} a_n x^n + e_i,$$

$$\begin{aligned} y_1 &= a_0 x_1^0 + a_1 x_1^1 + a_2 x_1^2 + \ldots + e_1, \\ &\vdots \\ y_m &= a_0 x_m^0 + a_1 x_m^1 + a_2 x_m^2 + \ldots + e_m \,. \end{aligned}$$

Vandermonde matrix A

$$\underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}}_{y} = \underbrace{\begin{pmatrix} 1 & x_1 & \cdots & x_1^{n_{\max}} \\ 1 & x_2 & \cdots & x_2^{n_{\max}} \\ \vdots & \vdots & & \\ 1 & x_m & \cdots & x_m^{n_{\max}} \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n_{\max}} \end{pmatrix}}_{\xi} + \underbrace{\begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{pmatrix}}_{e} .$$

1. Adjustment principle $e^\mathsf{T} e \longrightarrow \min \xi$.

2. x-coordinates are error free: inconsistences only for y_i.
3. y_i may have matrix Q_y.
4. The smaller $\hat{e}^\mathsf{T}\hat{e}$ for varying $n_{\max}$, the better the fit is. However, the larger $n_{\max}$, the more the polynomial oscillates. Using a sufficiently large value for $n_{\max}$, even $\hat{e}^\mathsf{T}\hat{e} = 0$ can be achieved.

   $\Longrightarrow$ Only low degree polynomials are used.
5. Possible additional restrictions

   a) Tangent in $(x_\mathrm{T}, y_\mathrm{T})$, $x_\mathrm{T} \in x$, shall pass through the point $(x_\mathrm{P}, y_\mathrm{P})$.

   Tangent equation:

$$g(x) = f(x_\mathrm{T}) + f'(x_\mathrm{T})(x - x_\mathrm{T}) \quad \Longrightarrow \quad y_\mathrm{P} = g(x_\mathrm{P}) = f(x_\mathrm{T}) + f'(x_\mathrm{T})(x_\mathrm{P} - x_\mathrm{T}) \,.$$

   Example for $n_{\max} = 2$

$$\begin{aligned} f(x) &= a_0 + a_1 x + a_2 x^2 \qquad \text{parabola} \\ f'(x) &= a_1 + 2a_2 x \end{aligned}$$

   Tangent in $x_\mathrm{T} : g(x) = a_0 + a_1 x_\mathrm{T} + a_2 x_\mathrm{T}^2 + (a_1 + 2a_2 x_\mathrm{T})(x - x_\mathrm{T})$

   Tangent in x_T, passing through $x_\mathrm{P}, y_\mathrm{P}$

$$\begin{aligned} y_\mathrm{P} &= a_0 + a_1 x_\mathrm{T} + a_2 x_\mathrm{T}^2 + (a_1 + 2a_2 x_T)(x_\mathrm{P} - x_\mathrm{T}) \\ &= a_0 + a_1 x_\mathrm{T} + a_2 x_\mathrm{T}^2 + a_1(x_\mathrm{P} - x_\mathrm{T}) + 2a_2(x_\mathrm{P} - x_\mathrm{T})x_\mathrm{T} \\ &= a_0 + x_\mathrm{P} a_1 + x_\mathrm{T}(2x_\mathrm{P} - x_\mathrm{T})a_2 \\ \Longrightarrow B^\mathsf{T}\xi &= y_\mathrm{P}, \quad \text{with} \quad \xi = [a_0, a_1, a_2]^\mathsf{T}, \quad B^\mathsf{T} = \begin{bmatrix} 1 & x_\mathrm{P} & x_\mathrm{T}(2x_\mathrm{P} - x_\mathrm{T}) \end{bmatrix} . \end{aligned}$$

   Include restriction using techniques of Lagrange multipliers or eliminate one unknown coefficient, e. g. a_0, in favor of the other unknown coefficients: $a_0 = y_\mathrm{P} - x_\mathrm{P} a_1 - x_\mathrm{T}(2x_\mathrm{P} - x_\mathrm{T})a_2$.

   General case for a polynomial of degree $n_{\max}$

$$B^\mathsf{T} = \begin{bmatrix} 1 & x_\mathrm{P} & x_\mathrm{T}(2x_\mathrm{P} - x_\mathrm{T}) & \ldots & x_\mathrm{T}^{n_{\max}-1}(n_{\max} x_\mathrm{P} - (n_{\max} - 1)x_\mathrm{T}) \end{bmatrix}$$

   $\Longrightarrow$ Tangent equation with adjusted parameters $\hat{\xi} = \begin{bmatrix} \hat{a}_0, \ldots, \hat{a}_{n_{\max}} \end{bmatrix}^\mathsf{T}$

$$\begin{aligned} y = a_\mathrm{T} x + b_\mathrm{T}, \quad a_\mathrm{T} &:= \frac{\hat{y}_\mathrm{T} - y_\mathrm{P}}{x_\mathrm{T} - x_\mathrm{P}} \quad \text{“tangent slope”}, \\ b_\mathrm{T} &:= \hat{y}_\mathrm{T} - \frac{\hat{y}_\mathrm{T} - y_\mathrm{P}}{x_\mathrm{T} - x_\mathrm{P}} x_\mathrm{T} \quad \text{“axis intercept”}, \\ \hat{y}_\mathrm{T} &= \hat{a}_0 + \hat{a}_1 x_\mathrm{T} + \ldots + \hat{a}_{n_{\max}} x_\mathrm{T}^{n_{\max}} \quad \text{“estimated ordinate”}. \end{aligned}$$

b) adjusted polynomial shall pass through the point (x_Q, y_Q)

$$y_Q = \sum_{n=0}^{n_{\max}} a_n x_Q^n \Longrightarrow B^\mathsf{T}\xi = y_Q, \quad B^\mathsf{T} = \begin{pmatrix} 1 & x_Q & x_Q^2 & \ldots & x_Q^{n_{\max}} \end{pmatrix} .$$

c) The unknown coefficient a_k should have the fixed numerical value $\tilde{a}_k$.

$$B^\mathsf{T}\xi = \tilde{a}_k, \quad B^\mathsf{T} = \begin{bmatrix} 0 & \ldots & \underbrace{1}_{\text{position } k+1} & \ldots \end{bmatrix}$$

or eliminate unknown a_k from ξ by setting it to $\tilde{a}_k$ from the very beginning.

Examples

$$x_i = \begin{bmatrix} -1, & 0, & 1, & 2, & 3, & 4, & 5 \end{bmatrix}^\mathsf{T},$$
$$y_i = \begin{bmatrix} 1.3, & 0.8, & 0.9, & 1.2, & 2.0, & 3.5, & 4.1 \end{bmatrix}^\mathsf{T}.$$

1) No restrictions: see Fig. 5.24.
2) With tangent restriction: see Fig. 5.25.
3) With point restriction: see Fig. 5.26.
4) With coefficient restriction: see Fig. 5.27.

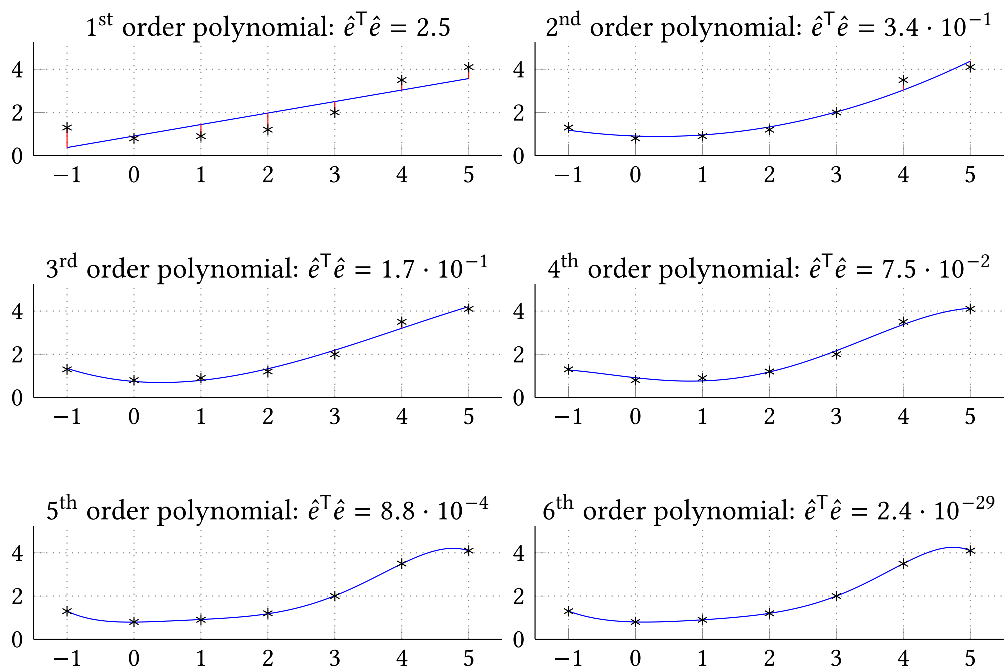


Figure 5.24: Polynomial fit without restrictions.

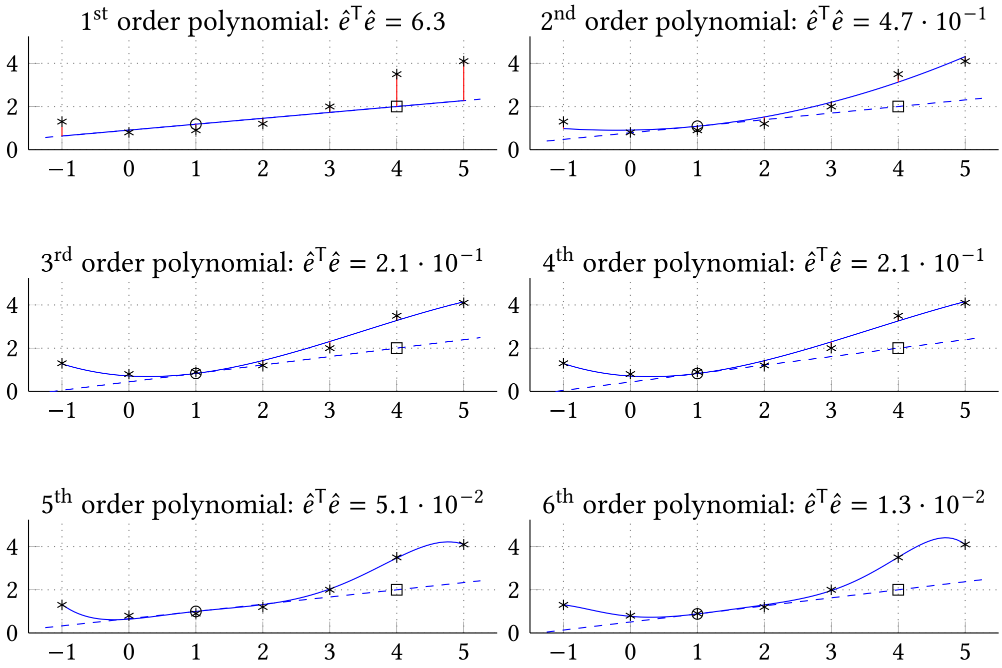


Figure 5.25: Polynomial fit with tangent restriction: tangent in $x_{\mathrm{T}} = 1$, $\hat{y}_{\mathrm{T}}(x_{\mathrm{T}})$ shall pass through the point $x_{\mathrm{P}} = 4$, $y_{\mathrm{P}} = 2$.

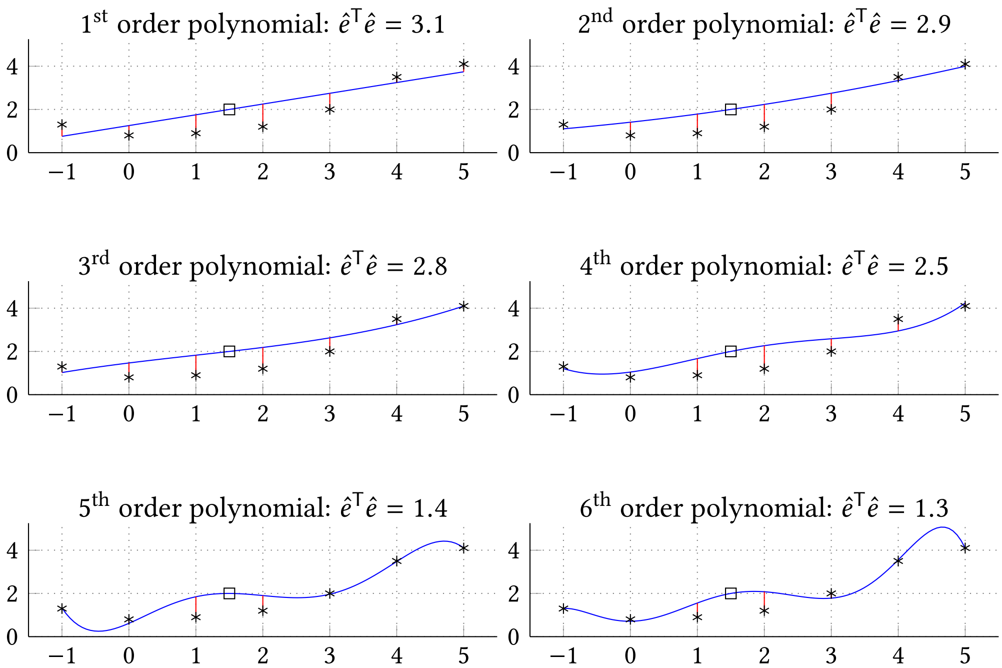


Figure 5.26: Polynomial fit with point restriction: adjusted polynomial shall pass through the point $x_{\mathrm{Q}} = 1.5$, $y_{\mathrm{Q}} = 2$.

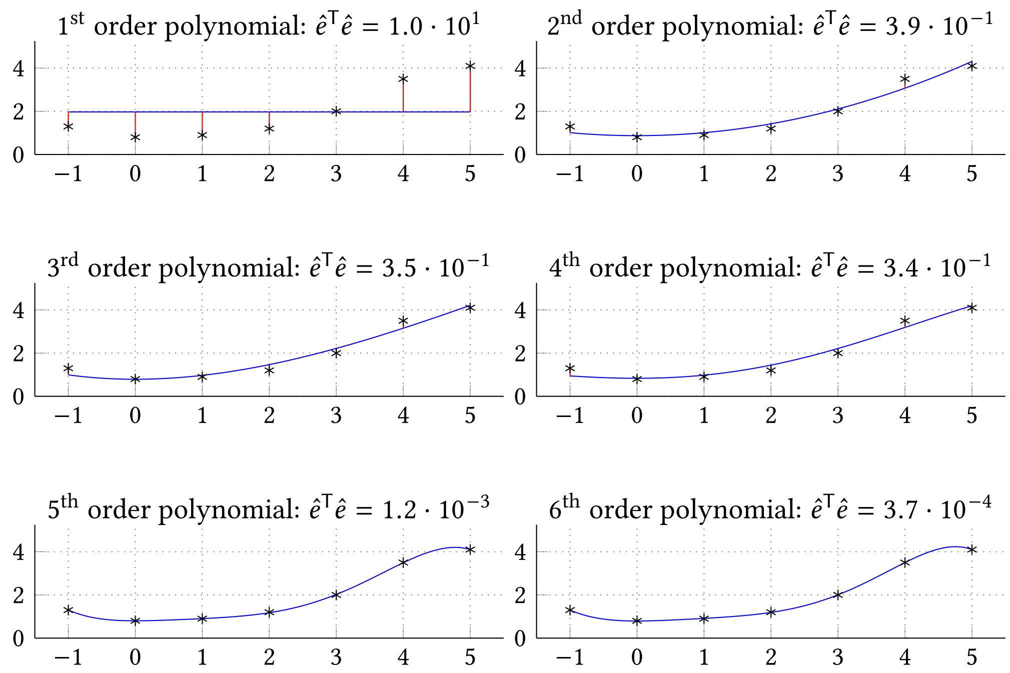


Figure 5.27: Polynomial fit with coefficient restriction: coefficient $\hat{a}_1$ shall vanish, i. e. $\hat{a}_1 = 0$.

More examples: Various straight line fits. For the numerics, the values on page 74 were reused.

1) *Straight line fit* using A-Model, with inconsistencies e_{y_i} in observations y_i ($Q_y^{-1} = I$). Observation equation: $y_i = a_0 + a_1 x_i$.

   Results (see also figure 5.28):

$$\hat{a}_0 = 0.907, \quad \hat{a}_1 = 0.532, \quad \hat{e}^\mathsf{T} P \hat{e} = 2.505,$$

$$\hat{y} = \begin{bmatrix} 0.375, & 0.907, & 1.439, & 1.971, & 2.504, & 3.036, & 3.568 \end{bmatrix}^\mathsf{T},$$
$$\hat{e}_y = \begin{bmatrix} 0.925, & -0.107, & -0.539, & -0.771, & -0.504, & 0.464, & 0.532 \end{bmatrix}^\mathsf{T}.$$

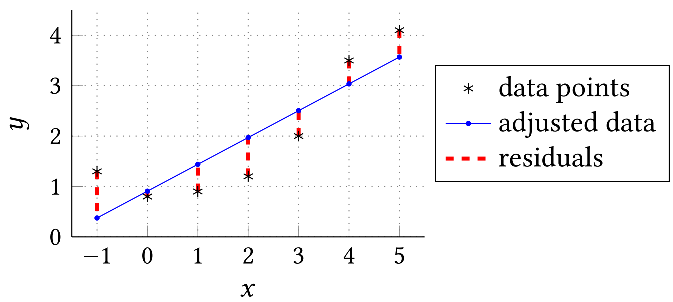


Figure 5.28: A-model with inconsistencies in y_i, uniform weights.

2) *Straight line fit* using A-Model, with inconsistencies e_{x_i} in observations x_i ($Q_x^{-1} = I$). Observation equation: $x_i = a_0 + a_1 y_i$.

   Results (see also figure 5.29):

$$\hat{a}_0 = -0.815, \quad \hat{a}_1 = 1.428, \quad \hat{e}^\mathsf{T} P \hat{e} = 6.723,$$

$$\hat{x} = \begin{bmatrix} 1.041, & 0.327, & 0.470, & 0.898, & 2.041, & 4.183, & 5.040 \end{bmatrix}^\mathsf{T},$$
$$\hat{e}_x = \begin{bmatrix} -2.041, & -0.327, & 0.530, & 1.102, & 0.959, & -0.183, & -0.040 \end{bmatrix}^\mathsf{T}.$$

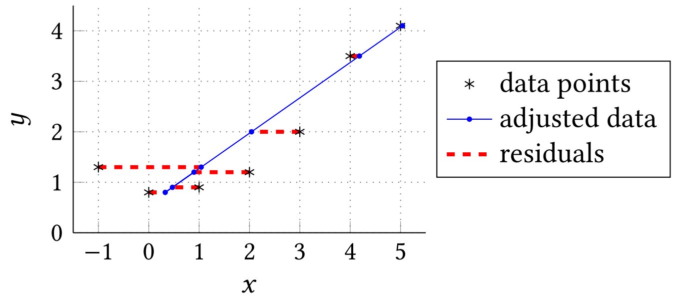


Figure 5.29: A-model with inconsistencies in x_i, uniform weights.

5.2 B-Model: Adjustment of condition equations

5.2.1 Planar triangle 1

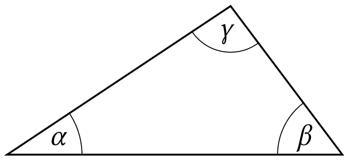


Figure 5.30: Triangle observed by angles

Observations: angles α, β, γ [°]

Unknowns: inconsistencies $e_\alpha, e_\beta, e_\gamma \Longrightarrow$ linear function

$$f(e_\alpha, e_\beta, e_\gamma) = (\alpha - e_\alpha) + (\beta - e_\beta) + (\gamma - e_\gamma) - 180^\circ = 0\,.$$

Model adjustment condition equations

$$B^\mathsf{T}(y - e) - 180^\circ = B^\mathsf{T}y - 180^\circ - B^\mathsf{T}e = w - B^\mathsf{T}e = 0$$

$$\text{with} \quad e = \left(e_\alpha,\, e_\beta,\, e_\gamma\right)^\mathsf{T}, \quad y = \left(\alpha,\, \beta,\, \gamma\right)^\mathsf{T} \quad \text{and} \quad w = B^\mathsf{T}y - 180^\circ \quad (\text{“misclosure”}).$$

5.2.2 Planar triangle 2

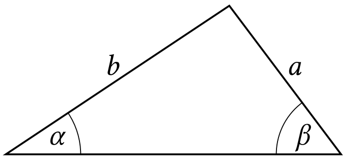


Figure 5.31: Triangle observed by angles and distances.

Observations: angles α, β [°], distances a, b [m].

Unknowns: inconsistencies e_a, e_b, e_α, $e_\beta \Longrightarrow$ non-linear function f

$$f(e_a, e_b, e_\alpha, e_\beta) = (a - e_a)\sin(\beta - e_\beta) - (b - e_b)\sin(\alpha - e_\alpha) = 0,$$

linearized with respect to the "Taylor point" $(e_a^0, e_b^0, e_\alpha^0, e_\beta^0) =: |_0$

$$\begin{aligned} f\left(e_a, e_b, e_\alpha, e_\beta\right) &= f\left(e_a^0, e_b^0, e_\alpha^0, e_\beta^0\right) + \left.\frac{\partial f}{\partial e_a}\right|_0 \left(e_a - e_a^0\right) + \left.\frac{\partial f}{\partial e_b}\right|_0 \left(e_b - e_b^0\right) \\ &\quad + \left.\frac{\partial f}{\partial e_\alpha}\right|_0 \left(e_\alpha - e_\alpha^0\right) + \left.\frac{\partial f}{\partial e_\beta}\right|_0 \left(e_\beta - e_\beta^0\right) \overset{!}{=} 0, \end{aligned}$$

$$\begin{aligned} f\left(e_a^0, e_b^0, e_\alpha^0, e_\beta^0\right) &= (a - e_a^0)\sin(\beta - e_\beta^0) - (b - e_b^0)\sin(\alpha - e_\alpha^0) \\ &= a\sin(\beta - e_\beta^0) - \sin(\beta - e_\beta^0)e_a^0 - b\sin(\alpha - e_\alpha^0) + \sin(\alpha - e_\alpha^0)e_b^0, \end{aligned}$$

$$\begin{aligned} \left.\frac{\partial f}{\partial e_a}\right|_0 \left(e_a - e_a^0\right) &= -\sin(\beta - e_\beta^0)\left(e_a - e_a^0\right) \\ &= -\sin(\beta - e_\beta^0)e_a + \sin(\beta - e_\beta^0)e_a^0, \end{aligned}$$

$$\begin{aligned} \left.\frac{\partial f}{\partial e_b}\right|_0 \left(e_b - e_b^0\right) &= \sin(\alpha - e_\alpha^0)\left(e_b - e_b^0\right) \\ &= \sin(\alpha - e_\alpha^0)e_b - \sin(\alpha - e_\alpha^0)e_b^0, \end{aligned}$$

$$\begin{aligned} \left.\frac{\partial f}{\partial e_\alpha}\right|_0 \left(e_\alpha - e_\alpha^0\right) &= \left(b - e_b^0\right)\cos(\alpha - e_\alpha^0)\left(e_\alpha - e_\alpha^0\right) \\ &= \left(b - e_b^0\right)\cos(\alpha - e_\alpha^0)e_\alpha - \left(b - e_b^0\right)\cos(\alpha - e_\alpha^0)e_\alpha^0, \end{aligned}$$

$$\begin{aligned} \left.\frac{\partial f}{\partial e_\beta}\right|_0 \left(e_\beta - e_\beta^0\right) &= -\left(a - e_a^0\right)\cos(\beta - e_\beta^0)\left(e_\beta - e_\beta^0\right) \\ &= -\left(a - e_a^0\right)\cos(\beta - e_\beta^0)e_\beta + \left(a - e_a^0\right)\cos(\beta - e_\beta^0)e_\beta^0\,. \end{aligned}$$

Model adjustment condition equations

$$w - B^\mathsf{T} e = 0 \quad \text{with} \quad e = \left(e_a,\, e_b,\, e_\alpha,\, e_\beta\right)^\mathsf{T}.$$

Collect the coefficients of all terms with e in $-B^\mathsf{T}$, all remaining terms go into the vector w of misclosures.

$$
\begin{aligned}
\implies \qquad & B^\mathsf{T} = \Big(\sin(\beta - e^0_\beta),\ -\sin(\alpha - e^0_\alpha),\ -(b - e^0_b)\cos(\alpha - e^0_\alpha),\ (a - e^0_a)\cos(\beta - e^0_\beta) \Big)\,, \\
& w = a \sin\left(\beta - e^0_\beta\right) - b \sin\left(\alpha - e^0_\alpha\right) - \left(b - e^0_b\right)\cos(\alpha - e^0_\alpha) e^0_\alpha \\
& \qquad + \left(a - e^0_a\right)\cos(\beta - e^0_\beta) e^0_\beta \qquad \text{(“misclosure”)}
\end{aligned}
$$

Example: observations

$$a = 10, \quad b = 5, \quad \alpha = 60^\circ, \quad \beta = 23.7^\circ$$

with associated weights

$$P_a = 1, \quad P_b = 0.1, \quad P_\alpha = 1^\circ, \quad P_\beta = 0.2^\circ\,.$$

Initial approximate values for unknown inconsistencies

$$e^0_a = e^0_b = 0, \quad e^0_\alpha = e^0_\beta = 0^\circ\,.$$

Results: parameters (after 6 iterations, $\|\widehat{\Delta e}\| < 10^{-12}$)

$$
\begin{gathered}
\hat{e}_a = -5.63 \cdot 10^{-6}, \quad \hat{e}_b = 1.13 \cdot 10^{-4}, \quad \hat{e}_\alpha = 6'26.241'', \quad \hat{e}_\beta = -1^\circ 55' 42.492'', \\
\hat{e}^\mathsf{T} P \hat{e} = 4.017 \cdot 10^{-6}\,.
\end{gathered}
$$

5.3 Mixed model

5.3.1 Straight line fit using A-model with pseudo observation equations

Example: *Straight line fit* using A-Model, with inconsistencies e_{x_i} and e_{y_i} in both observations x_i and y_i ($Q_y^{-1} = Q_x^{-1} = I$, $P = \text{diag}(Q_y^{-1}, Q_x^{-1})$). For the numerics, the values on page 74 have been used.

Unknown parameters a_0, a_1, $\bar{x}_i$, $i = 1, \dots, m$

$$y_i - e_{y_i} = a_0 + a_1(x_i - e_{x_i}) = a_0 + a_1\bar{x}_i \tag{5.1}$$

$$x_i - e_{x_i} = \bar{x}_i \tag{5.2}$$

Approximate values: $a_0 = a_0^0 + \Delta a_0$, $a_1 = a_1^0 + \Delta a_1$, $\bar{x}_i = \bar{x}_i{}^0 + \Delta\bar{x}_i$.

Linearized equations 5.1 and 5.2:

$$\underbrace{y_i - (a_0^0 + a_1^0\bar{x}_i^0)}_{\Delta y_i} - e_{y_i} = \Delta a_0 + a_1^0\Delta\bar{x}_i + \bar{x}_i^0\Delta a_1$$

$$x_i - \bar{x}_i^0 - e_{x_i} = \Delta\bar{x}_i\,.$$

This leads to

$$\Delta y_i - e_{y_i} = \Delta a_0 + a_1^0\Delta\bar{x}_i + \bar{x}_i^0\Delta a_1$$

and

$$\Delta x_i - e_{x_i} = \Delta\bar{x}_i\,.$$

In matrix notation:

$$\underset{2m\times 1}{\begin{pmatrix} \Delta y - e_y \\ \Delta x - e_x \end{pmatrix}} = \underset{2m\times(m+2)}{\begin{pmatrix} A_1 \\ A_2 \end{pmatrix}} \underset{(m+2)\times 1}{\begin{pmatrix} \Delta a_0 \\ \Delta a_1 \\ \Delta\bar{x} \end{pmatrix}} = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \Delta\xi$$

where

$$\underset{m\times(m+2)}{A_1} = \begin{pmatrix} 1 & \bar{x}_1^0 & a_1^0 & \dots & 0 & 0 \\ 1 & \bar{x}_2^0 & 0 & a_1^0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \bar{x}_m^0 & 0 & 0 & \dots & a_1^0 \end{pmatrix}; \qquad \underset{m\times(m+2)}{A_2} = \begin{pmatrix} 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

Results: Initial approximate values for unknown parameters:

$$a_0^0 = 0.800, \quad a_1^0 = 0.550, \quad \bar{x}_i^0 = x_i.$$

Parameters (after 20 iterations, $\|\widehat{\Delta\xi}\| < 10^{-12}$):

$$\hat{a}_0 = 0.829, \quad \hat{a}_1 = 0.571, \quad \hat{e}^\mathsf{T} P\hat{e} = 1.921$$

$$\hat{y} = \begin{bmatrix} 0.514, & 0.822, & 1.277, & 1.782, & 2.409, & 3.209, & 3.787 \end{bmatrix}^\mathsf{T}$$
$$\hat{e}_y = \begin{bmatrix} 0.786, & -0.022, & -0.377, & -0.582, & -0.409, & 0.291, & 0.313 \end{bmatrix}^\mathsf{T}$$
$$\hat{x} = \begin{bmatrix} -0.551, & -0.012, & 0.785, & 1.668, & 2.766, & 4.166, & 5.179 \end{bmatrix}^\mathsf{T}$$
$$\hat{e}_x = \begin{bmatrix} -0.449, & 0.012, & 0.215, & 0.332, & 0.234, & -0.166, & -0.179 \end{bmatrix}^\mathsf{T}$$

See figure 5.32.

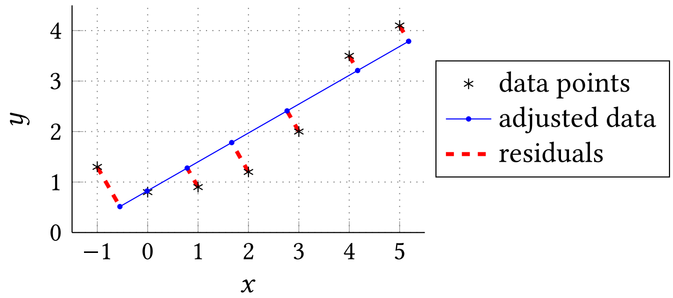


Figure 5.32: A-model with inconsistencies in x_i and y_i, uniform weights.

5.3.2 Straight line fit using extended B-Model

Example: Straight line fit using extended B-Model with $Q_y^{-1} = Q_x^{-1} = I$, $P = \operatorname{diag}(Q_y^{-1}, Q_x^{-1})$.

Non linear condition equation with unknowns a_0, a_1:

$$y_i - e_{y_i} - (a_0 + a_1(x_i - e_{x_i})) = 0$$

Initial approximate values $e^0_{x_i} = 0$, $e^0_{y_i} = 0$, a^0_0, a^0_1 so that

$$e_{x_i} = e^0_{x_i} + \Delta e_{x_i}, \quad e_{y_i} = e^0_{y_i} + \Delta e_{y_i}, \quad a_0 = a^0_0 + \Delta a_0, \quad a_1 = a^0_1 + \Delta a_1 .$$

$$y_i - e^0_{y_i} - \left(a^0_0 + a^0_1(x_i - e^0_{x_i})\right) - \Delta e_{y_i} - \Delta a_0 - \Delta a_1(x_i - e^0_{x_i}) + a^0_1 \Delta e_{x_i} = 0$$

$$y_i - e^0_{y_i} - \left(a^0_0 + a^0_1(x_i - e^0_{x_i})\right) - e_{y_i} + e^0_{y_i} - \begin{bmatrix} 1 & (x_i - e^0_{x_i}) \end{bmatrix} \begin{bmatrix} \Delta a_0 \\ \Delta a_1 \end{bmatrix} + a^0_1 e_{x_i} - a^0_1 e^0_{x_i} = 0$$

$$\underbrace{y_i - (a^0_0 + a^0_1 x_i)}_{w_i} \underbrace{- \begin{bmatrix} 1 & (x_i - e^0_{x_i}) \end{bmatrix}}_{A_i} \underbrace{\begin{bmatrix} \Delta a_0 \\ \Delta a_1 \end{bmatrix}}_{\Delta\xi} + \underbrace{\begin{bmatrix} a^0_1 & -1 \end{bmatrix}}_{B_i^\mathsf{T}} \underbrace{\begin{bmatrix} e_{x_i} \\ e_{y_i} \end{bmatrix}}_{e_i} = 0$$

$$\underset{m\times 2}{A} = - \begin{bmatrix} 1 & x_1 - e^0_{x_1} \\ 1 & x_2 - e^0_{x_2} \\ \vdots & \vdots \\ 1 & x_m - e^0_{x_m} \end{bmatrix}; \qquad \underset{2\times 1}{\Delta\xi} = \begin{bmatrix} \Delta a_0 \\ \Delta a_1 \end{bmatrix}; \qquad \underset{2m\times 1}{e} = \begin{bmatrix} e_{x_1} & \cdots & e_{x_m} & e_{y_1} & \cdots & e_{y_m} \end{bmatrix}^\mathsf{T} = \begin{bmatrix} e_x \\ e_y \end{bmatrix};$$

$$\underset{m\times 2m}{B^\mathsf{T}} = \begin{bmatrix} a^0_1 & & & -1 & & \\ & \ddots & & & \ddots & \\ & & a^0_1 & & & -1 \end{bmatrix} = \begin{bmatrix} a^0_1 I_m, & -I_m \end{bmatrix}; \qquad \underset{m\times 1}{w} = y - (a^0_0 + a^0_1 x) .$$

Lagrangian:

$$\mathcal{L}(\Delta\xi, e, \lambda) = \frac{1}{2} e^\mathsf{T} P e + \lambda^\mathsf{T}(w + A\Delta\xi + B^\mathsf{T} e) \longrightarrow \min_{\Delta\xi, e, \lambda}$$

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial e}(\hat{e}, \hat{\lambda}, \widehat{\Delta\xi}) &= \underset{2m\times 2m}{P}\ \underset{2m\times 1}{\hat{e}} && + \underset{2m\times m}{B}\ \underset{m\times 1}{\hat{\lambda}} = \underset{2m\times 1}{0} \\
\frac{\partial \mathcal{L}}{\partial \Delta\xi}(\hat{e}, \hat{\lambda}, \widehat{\Delta\xi}) &= && \underset{2\times m}{A^\mathsf{T}}\ \underset{m\times 1}{\hat{\lambda}} = \underset{2\times 1}{0} \\
\frac{\partial \mathcal{L}}{\partial \lambda}(\hat{e}, \hat{\lambda}, \widehat{\Delta\xi}) &= \underset{m\times 2m}{B^\mathsf{T}}\ \underset{2m\times 1}{\hat{e}} + \underset{m\times 2}{A}\ \underset{2\times 1}{\widehat{\Delta\xi}} && = -\underset{m\times 1}{w}
\end{aligned}$$

Results (see also figure 5.33):

$$\hat{a}_0 = 0.829, \quad \hat{a}_1 = 0.571, \quad \hat{e}^\mathsf{T} P \hat{e} = 1.921$$

$$\begin{aligned}
\hat{y} &= \begin{bmatrix} 0.514, & 0.822, & 1.277, & 1.782, & 2.409, & 3.209, & 3.787 \end{bmatrix}^\mathsf{T} \\
\hat{e}_y &= \begin{bmatrix} 0.786, & -0.022, & -0.377, & -0.582, & -0.409, & 0.291, & 0.313 \end{bmatrix}^\mathsf{T} \\
\hat{x} &= \begin{bmatrix} -0.551, & -0.012, & 0.785, & 1.668, & 2.766, & 4.166, & 5.179 \end{bmatrix}^\mathsf{T} \\
\hat{e}_x &= \begin{bmatrix} -0.449, & 0.012, & 0.215, & 0.332, & 0.234, & -0.166, & -0.179 \end{bmatrix}^\mathsf{T}
\end{aligned}$$

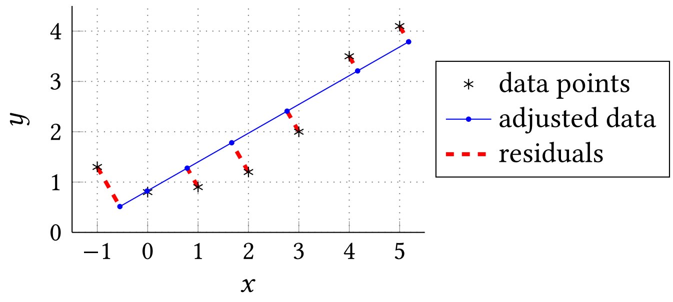


Figure 5.33: Extended B-model with inconsistencies in x_i and y_i, uniform weights.

Example: The following results and figure show the cases for the previous two examples, observations having weights ($P_x = Q_x^{-1} \neq I$, $P_y = Q_y^{-1} \neq I$, $P = \operatorname{diag}(P_x, P_y)$). We introduce the weights

$$\operatorname{diag} P_x = \begin{bmatrix} 3, & 9, & 8, & 4, & 5, & 7, & 10 \end{bmatrix}^\mathsf{T}$$
$$\operatorname{diag} P_y = \begin{bmatrix} 2, & 8, & 7, & 5, & 10, & 8, & 6 \end{bmatrix}^\mathsf{T}$$

Both, the A-model with inconsistencies e_{x_i} and e_{y_i} and the extended B-model, give identical results. Due to $P \neq I$ residuals are not orthogonal to the adjusted line. See figure 5.34.

$$\hat{a}_0 = 0.5512, \quad \hat{a}_1 = 0.6580, \quad \hat{e}^\mathsf{T} P \hat{e} = 7.6931$$

$$\hat{y} = \begin{bmatrix} 0.208, & 0.620, & 1.124, & 1.633, & 2.281, & 3.288, & 3.895 \end{bmatrix}^\mathsf{T}$$
$$\hat{e}_y = \begin{bmatrix} 1.092, & 0.180, & -0.224, & -0.433, & -0.281, & 0.212, & 0.205 \end{bmatrix}^\mathsf{T}$$
$$\hat{x} = \begin{bmatrix} -0.521, & 0.105, & 0.871, & 1.644, & 2.630, & 4.159, & 5.081 \end{bmatrix}^\mathsf{T}$$
$$\hat{e}_x = \begin{bmatrix} -0.479, & -0.105, & 0.129, & 0.356, & 0.370, & -0.159, & -0.081 \end{bmatrix}^\mathsf{T}$$

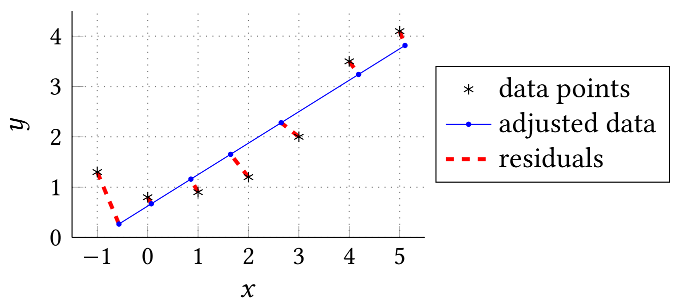


Figure 5.34: Extended B-model with inconsistencies in x_i and y_i, non-uniform weights.

5.3.3 2D Similarity Transformation

The following two tables (see Niemeier, 2008, pg. 374–375) give coordinates with respect to the *source* (u, v)-system and the *target* (x, y)-system. Points 1–4 are identical to both systems (control points). We assume inconsistencies in both *source* and *target* system coordinates and they are uncorrelated having equal unit variances, i. e.

$$P_x = Q_x^{-1} = I, \quad P_y = Q_y^{-1} = I, \quad P_u = Q_u^{-1} = I, \quad P_v = Q_v^{-1} = I, \quad P = \mathrm{diag}(P_x, P_y, P_u, P_v) .$$

Table 5.51: Source coordinates.

Point	u [m]	v [m]
1	14 029.640	12 786.840
2	14 914.630	12 535.560
3	14 771.830	11 404.660
4	13 221.620	11 840.320
13	14 735.090	12 127.380
14	14 253.840	11 923.950
15	13 603.740	11 836.700
16	14 291.760	12 495.310
17	13 931.500	12 307.610

Table 5.52: Target coordinates.

Point	x [m]	y [m]
1	19 405.518	23 159.823
2	20 291.232	22 909.817
3	20 150.035	21 778.202
4	18 598.550	22 211.755

A 2D similarity transformation (*Helmert transformation*) to transform the set of coordinates from the *source system* to the *target system* will be performed.

$$\underbrace{\begin{bmatrix} x_i \\ y_i \end{bmatrix}}_{\text{target}} = \lambda \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \underbrace{\begin{bmatrix} u_i \\ v_i \end{bmatrix}}_{\text{source}} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}$$

Usual adjustment: only target coordinates x_i and y_i have inconsistencies

$$\begin{aligned} x_i - e_{x_i} &= \lambda u_i \cos\alpha + \lambda v_i \sin\alpha + t_x, \\ y_i - e_{y_i} &= -\lambda u_i \sin\alpha + \lambda v_i \cos\alpha + t_y . \end{aligned}$$

Mixed model approach I: **A-model** with inconsistencies in both $[x_i, y_i]$ and $[u_i, v_i]$ coordinates ($i = 1, \ldots, p$, with p number of control points).

$$\begin{aligned} x_i - e_{x_i} &= \lambda \bar{u}_i \cos\alpha + \lambda \bar{v}_i \sin\alpha + t_x, \\ y_i - e_{y_i} &= -\lambda \bar{u}_i \sin\alpha + \lambda \bar{v}_i \cos\alpha + t_y, \\ u_i - e_{u_i} &= \bar{u}_i, \\ v_i - e_{v_i} &= \bar{v}_i . \end{aligned}$$

Approximate values:

$$\lambda = \lambda^0 + \Delta\lambda, \quad \alpha = \alpha^0 + \Delta\alpha, \quad t_x = t_x^0 + \Delta t_x, \quad t_y = t_y^0 + \Delta t_y,$$
$$\bar{u}_i = \bar{u}_i^0 + \Delta\bar{u}_i, \quad \bar{v}_i = \bar{v}_i^0 + \Delta\bar{v}_i.$$

Linearization process:

$$\begin{aligned}
x_i - e_{x_i} &= \underbrace{(\lambda^0 \cos\alpha^0 \bar{u}_i^0 + \lambda^0 \sin\alpha^0 \bar{v}_i^0 + t_x^0)}_{x_i^0} + \Delta t_x + \underbrace{(-\lambda^0 \sin\alpha^0 \bar{u}_i^0 + \lambda^0 \cos\alpha^0 \bar{v}_i^0)}_{a_i} \Delta\alpha \\
&\quad + \underbrace{(\cos\alpha^0 \bar{u}_i^0 + \sin\alpha^0 \bar{v}_i^0)}_{b_i} \Delta\lambda + \lambda^0 \cos\alpha^0 \Delta\bar{u}_i + \lambda^0 \sin\alpha^0 \Delta\bar{v}_i \\
y_i - e_{y_i} &= \underbrace{(-\lambda^0 \sin\alpha^0 \bar{u}_i^0 + \lambda^0 \cos\alpha \bar{v}_i^0 + t_y^0)}_{y_i^0} + \Delta t_y \underbrace{-(\lambda^0 \cos\alpha^0 \bar{u}_i^0 + \lambda^0 \sin\alpha^0 \bar{v}_i^0)}_{c_i} \Delta\alpha \\
&\quad + \underbrace{(-\sin\alpha^0 \bar{u}_i^0 + \cos\alpha^0 \bar{v}_i^0)}_{d_i} \Delta\lambda - \lambda^0 \sin\alpha^0 \Delta\bar{u}_i + \lambda^0 \cos\alpha^0 \Delta\bar{v}_i \\
u_i - e_{u_i} &= \bar{u}_i^0 + \Delta\bar{u}_i^0 \\
v_i - e_{v_i} &= \bar{v}_i^0 + \Delta\bar{v}_i^0
\end{aligned}$$

In matrix form:

$$\underbrace{\begin{bmatrix} x_1 - x_1^0 \\ \vdots \\ x_p - x_p^0 \\ \cdots\cdots \\ y_1 - y_1^0 \\ \vdots \\ y_p - y_p^0 \\ \cdots\cdots \\ u_1 - \bar{u}_1^0 \\ \vdots \\ u_p - \bar{u}_p^0 \\ \cdots\cdots \\ v_1 - \bar{v}_1^0 \\ \vdots \\ v_p - \bar{v}_p^0 \end{bmatrix}}_{\substack{l \\ 4p\times 1}} - \underbrace{\begin{bmatrix} e_{x_1} \\ \vdots \\ e_{x_p} \\ \cdots \\ e_{y_1} \\ \vdots \\ e_{y_p} \\ \cdots \\ e_{u_1} \\ \vdots \\ e_{u_p} \\ \cdots \\ e_{v_1} \\ \vdots \\ e_{v_p} \end{bmatrix}}_{\substack{e \\ 4p\times 1}} = \underbrace{\begin{bmatrix} 1 & 0 & a_1 & b_1 & \lambda^0\cos\alpha^0 & \dots & 0 & \lambda^0\sin\alpha^0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & a_p & b_p & 0 & \dots & \lambda^0\cos\alpha^0 & 0 & \dots & \lambda^0\sin\alpha^0 \\ \hdashline 0 & 1 & c_1 & d_1 & -\lambda^0\sin\alpha^0 & \dots & 0 & \lambda^0\cos\alpha^0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & c_p & d_p & 0 & \dots & -\lambda^0\sin\alpha^0 & 0 & \dots & \lambda^0\cos\alpha^0 \\ \hdashline 0 & 0 & 0 & 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & \dots & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 1 \end{bmatrix}}_{\substack{A \\ 4p\times(2p+4)}} \underbrace{\begin{bmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\alpha \\ \Delta\lambda \\ \Delta\bar{u}_1 \\ \dots \\ \Delta\bar{u}_p \\ \Delta\bar{v}_1 \\ \dots \\ \Delta\bar{v}_p \end{bmatrix}}_{\substack{\Delta\xi \\ (2p+4)\times 1}}$$

Results: by using the initial approximate values for unknown parameters

$$t_x^0 = 5500\,\text{m}, \quad t_y^0 = 10\,200\,\text{m}, \quad \alpha^0 = 1.5'', \quad \lambda^0 = 1, \quad \bar{u}^0 = \left[u_1, \dots, u_4\right]^\mathsf{T}, \quad \bar{v}^0 = \left[v_1, \dots, v_4\right]^\mathsf{T},$$

we obtain the parameters (after 5 Iterations, $\|\widehat{\Delta\xi}\| < 10^{-11}$):

$$\hat{t}_x = 5389.091\,\text{m}, \quad \hat{t}_y = 10\,347.006\,\text{m}, \quad \hat{\alpha} = -5'5.557'', \quad \hat{\lambda} = 1.000\,409\,017,$$
$$\hat{e}^{\mathsf{T}} P \hat{e} = 0.001\,284\,79\,\text{m}^2 .$$

Coordinates of data points in the target system are listed in Tab. 5.53.

Table 5.53: Coordinates of data points in the target system.

Point	x [m]	y [m]
13	20 112.219	22 501.170
14	19 631.075	22 296.944
15	18 980.839	22 208.695
16	19 668.163	22 868.593
17	19 308.035	22 680.283

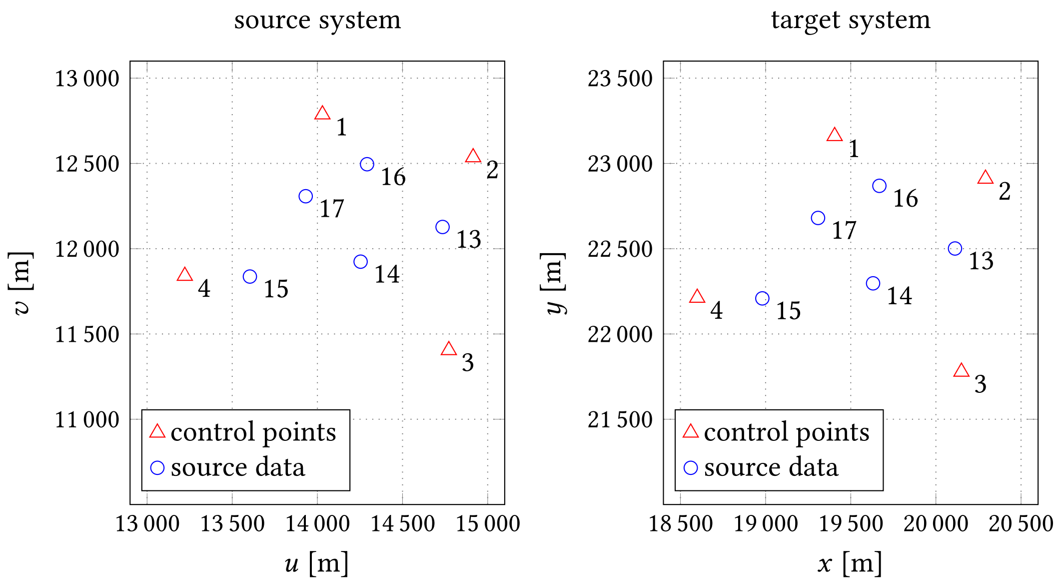


Figure 5.35: 2D similarity transformation: Gauss Markov model; inconsistencies in both source and target system.

Mixed model approach II: **extended B-model** with inconsistencies in both $[x_i, y_i]$ and $[u_i, v_i]$ coordinates ($i = 1, \ldots, p$, with p number of control points).

$$f_{x_i} := x_i - e_{x_i} - (\lambda \cos\alpha(u_i - e_{u_i}) + \lambda \sin\alpha(v_i - e_{v_i}) + t_x) = 0,$$
$$f_{y_i} := y_i - e_{y_i} - \left(-\lambda \sin\alpha(u_i - e_{u_i}) + \lambda \cos\alpha(v_i - e_{v_i}) + t_y\right) = 0 .$$

Linearization using Taylor point $e^0_{x_i}, e^0_{y_i}, e^0_{u_i}, e^0_{v_i}, t^0_x, t^0_y, \alpha^0, \lambda^0$ so that

$$e_{x_i} = e^0_{x_i} + \Delta e_{x_i}, \qquad e_{y_i} = e^0_{y_i} + \Delta e_{y_i}, \qquad e_{u_i} = e^0_{u_i} + \Delta e_{u_i}, \qquad e_{v_i} = e^0_{v_i} + \Delta e_{v_i},$$
$$t_x = t^0_x + \Delta t_x, \qquad t_y = t^0_y + \Delta t_y, \qquad \alpha = \alpha^0 + \Delta\alpha, \qquad \lambda = \lambda^0 + \Delta\lambda\,.$$

$$\begin{bmatrix} f^0_{x_i} \\ f^0_{y_i} \end{bmatrix} + \left.\begin{bmatrix} \frac{\partial f_{x_i}}{\partial e_{x_i}} & \frac{\partial f_{x_i}}{\partial e_{y_i}} & \frac{\partial f_{x_i}}{\partial e_{u_i}} & \frac{\partial f_{x_i}}{\partial e_{v_i}} \\ \frac{\partial f_{y_i}}{\partial e_{x_i}} & \frac{\partial f_{y_i}}{\partial e_{y_i}} & \frac{\partial f_{y_i}}{\partial e_{u_i}} & \frac{\partial f_{y_i}}{\partial e_{v_i}} \end{bmatrix}\right|_0 \begin{bmatrix} \Delta e_{x_i} \\ \Delta e_{y_i} \\ \Delta e_{u_i} \\ \Delta e_{v_i} \end{bmatrix} + \left.\begin{bmatrix} \frac{\partial f_{x_i}}{\partial t_x} & \frac{\partial f_{x_i}}{\partial t_y} & \frac{\partial f_{x_i}}{\partial \alpha} & \frac{\partial f_{x_i}}{\partial \lambda} \\ \frac{\partial f_{y_i}}{\partial t_x} & \frac{\partial f_{y_i}}{\partial t_y} & \frac{\partial f_{y_i}}{\partial \alpha} & \frac{\partial f_{y_i}}{\partial \lambda} \end{bmatrix}\right|_0 \begin{bmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\alpha \\ \Delta\lambda \end{bmatrix} = 0$$

where

$$f^0_{x_i} = x_i - e^0_{x_i} - \left(\lambda^0 \cos\alpha^0 (u_i - e^0_{u_i}) + \lambda^0 \sin\alpha^0 (v_i - e^0_{v_i}) + t^0_x\right)$$
$$f^0_{y_i} = y_i - e^0_{y_i} - \left(-\lambda^0 \sin\alpha^0 (u_i - e^0_{u_i}) + \lambda^0 \cos\alpha^0 (v_i - e^0_{v_i}) + t^0_y\right)\,.$$

First, we replace $\Delta e_{x_i} = e_{x_i} - e^0_{x_i}$, $\Delta e_{y_i} = e_{y_i} - e^0_{y_i}$ etc. and get

$$\begin{bmatrix} f^0_{x_i} \\ f^0_{y_i} \end{bmatrix} + \underbrace{\left.\begin{bmatrix} \frac{\partial f_{x_i}}{\partial e_{x_i}} & \frac{\partial f_{x_i}}{\partial e_{y_i}} & \frac{\partial f_{x_i}}{\partial e_{u_i}} & \frac{\partial f_{x_i}}{\partial e_{v_i}} \\ \frac{\partial f_{y_i}}{\partial e_{x_i}} & \frac{\partial f_{y_i}}{\partial e_{y_i}} & \frac{\partial f_{y_i}}{\partial e_{u_i}} & \frac{\partial f_{y_i}}{\partial e_{v_i}} \end{bmatrix}\right|_0}_{B_i^\mathsf{T}} \begin{bmatrix} e_{x_i} - e^0_{x_i} \\ e_{y_i} - e^0_{y_i} \\ e_{u_i} - e^0_{u_i} \\ e_{v_i} - e^0_{v_i} \end{bmatrix} + \left.\begin{bmatrix} \frac{\partial f_{x_i}}{\partial t_x} & \frac{\partial f_{x_i}}{\partial t_y} & \frac{\partial f_{x_i}}{\partial \alpha} & \frac{\partial f_{x_i}}{\partial \lambda} \\ \frac{\partial f_{y_i}}{\partial t_x} & \frac{\partial f_{y_i}}{\partial t_y} & \frac{\partial f_{y_i}}{\partial \alpha} & \frac{\partial f_{y_i}}{\partial \lambda} \end{bmatrix}\right|_0 \begin{bmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\alpha \\ \Delta\lambda \end{bmatrix} = 0\,.$$

Then, we rearrange the equation to

$$\underbrace{\underbrace{\begin{bmatrix} f^0_{x_i} \\ f^0_{y_i} \end{bmatrix}}_{\underset{2\times 1}{f^0}} - \underset{2\times 4}{B_i^\mathsf{T}} \underbrace{\begin{bmatrix} e^0_{x_i} \\ e^0_{y_i} \\ e^0_{u_i} \\ e^0_{v_i} \end{bmatrix}}_{\underset{4\times 1}{e^0_i}}}_{\underset{2\times 1}{w_i}} + \underset{2\times 4}{B_i^\mathsf{T}} \underbrace{\begin{bmatrix} e_{x_i} \\ e_{y_i} \\ e_{u_i} \\ e_{v_i} \end{bmatrix}}_{\underset{4\times 1}{e_i}} + \underbrace{\left.\begin{bmatrix} \frac{\partial f_{x_i}}{\partial t_x} & \frac{\partial f_{x_i}}{\partial t_y} & \frac{\partial f_{x_i}}{\partial \alpha} & \frac{\partial f_{x_i}}{\partial \lambda} \\ \frac{\partial f_{y_i}}{\partial t_x} & \frac{\partial f_{y_i}}{\partial t_y} & \frac{\partial f_{y_i}}{\partial \alpha} & \frac{\partial f_{y_i}}{\partial \lambda} \end{bmatrix}\right|_0}_{\underset{2\times 4}{A_i}} \underbrace{\begin{bmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\alpha \\ \Delta\lambda \end{bmatrix}}_{\underset{4\times 1}{\Delta\xi}} = 0\,.$$

In matrix notation:

$$\underbrace{\begin{bmatrix} x_1 - x^0_1 \\ \vdots \\ x_p - x^0_p \\ \cdots\cdots \\ y_1 - y^0_1 \\ \vdots \\ y_p - y^0_p \end{bmatrix}}_{\underset{2p\times 1}{w}} + \underbrace{\begin{bmatrix} -I_p & 0_p & \lambda^0\cos\alpha^0 I_p & \lambda^0 \sin\alpha^0 I_p \\ \cdots & \cdots & \cdots & \cdots \\ 0_p & -I_p & -\lambda^0 \sin\alpha^0 I_p & \lambda^0\cos\alpha^0 I_p \end{bmatrix}}_{\underset{2p\times 4p}{B^\mathsf{T}}} \underbrace{\begin{bmatrix} e_{x_1} \\ \vdots \\ e_{x_p} \\ e_{y_1} \\ \vdots \\ e_{y_p} \\ e_{u_1} \\ \vdots \\ e_{u_p} \\ e_{v_1} \\ \vdots \\ e_{v_p} \end{bmatrix}}_{\underset{4p\times 1}{e}} + \underbrace{\begin{bmatrix} -1 & 0 & \lambda^0 a^0_1 & -b^0_1 \\ \vdots & \vdots & \vdots & \vdots \\ -1 & 0 & \lambda^0 a^0_p & -b^0_p \\ \cdots & \cdots & \cdots & \cdots \\ 0 & -1 & \lambda^0 b^0_1 & a^0_1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & -1 & \lambda^0 b^0_p & a^0_p \end{bmatrix}}_{\underset{2p\times 4}{A}} \underbrace{\begin{bmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\alpha \\ \Delta\lambda \end{bmatrix}}_{\underset{4\times 1}{\Delta\xi}} = 0,$$

where I_p is the unit matrix of size $p \times p$ and 0_p the zero matrix of the same size. Additionally, we define $\bar{u}_i^0 = u_i - e_{u_i}^0$ and $\bar{v}_i^0 = v_i - e_{v_i}^0$ to get the abbreviations

$$x_i^0 = \lambda^0(\cos\alpha^0 u_i + \sin\alpha^0 v_i) + t_x^0, \quad y_i^0 = \lambda^0(-\sin\alpha^0 u_i + \cos\alpha^0 v_i) + t_y^0,$$
$$a_i^0 = \sin\alpha^0 \bar{u}_i^0 - \cos\alpha^0 \bar{v}_i^0, \quad b_i^0 = \cos\alpha^0 \bar{u}_i^0 + \sin\alpha^0 \bar{v}_i^0 .$$

Results: by using the following initial approximate values for unknown parameters

$$t_x^0 = 5500\,\mathrm{m}, \quad t_y^0 = 10\,200\,\mathrm{m}, \quad \lambda^0 = 1, \quad \alpha^0 = 1.5'', \quad e_{u_i}^0 = e_{v_i}^0 = 0 \quad \forall i,$$

we get the parameters (after 7 iterations, $\|\widehat{\Delta\xi}\| < 10^{-11}$):

$$\hat{t}_x = 5389.091\,\mathrm{m}, \quad \hat{t}_y = 10\,347.006\,\mathrm{m}, \quad \hat{\alpha} = -5'5.557'', \quad \hat{\lambda} = 1.000\,409\,017,$$
$$\hat{e}^\mathsf{T} P \hat{e} = 0.001\,284\,79\,\mathrm{m}^2 .$$

5.3.4 2D Affine Transformation Model I

Example: 6-parameter affine transformation—model I

The numerical data on page 85 (from Niemeier, 2008, pg. 374–375) are transformed using the 6-parameter affine transformtion

$$\underbrace{\begin{bmatrix} x_i \\ y_i \end{bmatrix}}_{\text{target}} = \underbrace{\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}}_{\text{scale factors}} \underbrace{\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}}_{\text{shear}} \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \underbrace{\begin{bmatrix} u_i \\ v_i \end{bmatrix}}_{\text{source}} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} .$$

Mixed model approach I: **A-model** with inconsistencies in both (x_i, y_i) and (u_i, v_i) coordinates.

$$\begin{aligned} x_i - e_{x_i} &= \lambda_1 \bar{u}_i(\cos\alpha - k\sin\alpha) + \lambda_1 \bar{v}_i(\sin\alpha + k\cos\alpha) + t_x, \\ y_i - e_{y_i} &= -\lambda_2 \bar{u}_i \sin\alpha + \lambda_2 \bar{v}_i \cos\alpha + t_y, \\ u_i - e_{u_i} &= \bar{u}_i, \\ v_i - e_{v_i} &= \bar{v}_i . \end{aligned}$$

Approximate values:

$$\begin{aligned} t_x &= t_x^0 + \Delta t_x, & \alpha &= \alpha^0 + \Delta\alpha, & \bar{u}_i &= \bar{u}_i^0 + \Delta\bar{u}_i, & \lambda_1 &= \lambda_1^0 + \Delta\lambda_1, \\ t_y &= t_y^0 + \Delta t_y, & k &= k^0 + \Delta k, & \bar{v}_i &= \bar{v}_i^0 + \Delta\bar{v}_i & \lambda_2 &= \lambda_2^0 + \Delta\lambda_2, . \end{aligned}$$

Linearization:

$$\begin{aligned} x_i - e_{x_i} = {} & \underbrace{\left(\lambda_1^0 \bar{u}_i^0(\cos\alpha^0 - k^0\sin\alpha^0) + \lambda_1^0 \bar{v}_i^0(\sin\alpha^0 + k^0\cos\alpha^0) + t_x^0\right)}_{x_i^0} + \Delta t_x \\ & + \underbrace{\left(-\lambda_1^0 \bar{u}_i^0(\sin\alpha^0 + k^0\cos\alpha^0) + \lambda_1^0 \bar{v}_i^0(\cos\alpha^0 - k^0\sin\alpha^0)\right)}_{a_i} \Delta\alpha \\ & + \underbrace{\left(\bar{u}_i^0(\cos\alpha^0 - k^0\sin\alpha^0) + \bar{v}_i^0(\sin\alpha^0 + k^0\cos\alpha^0)\right)}_{b_i} \Delta\lambda_1 + \underbrace{\left(-\lambda_1^0 \bar{u}_i^0 \sin\alpha^0 + \lambda_1^0 \bar{v}_i^0 \cos\alpha^0\right)}_{f_i} \Delta k \\ & + \underbrace{\lambda_1^0(\cos\alpha^0 - k^0\sin\alpha^0)}_{g} \Delta\bar{u}_i + \underbrace{\lambda_1^0(\sin\alpha^0 + k^0\cos\alpha^0)}_{h} \Delta\bar{v}_i, \\ y_i - e_{y_i} = {} & \underbrace{\left(-\lambda_2^0 \bar{u}_i^0 \sin\alpha^0 + \lambda_2^0 \bar{v}_i^0 \cos\alpha^0 + t_y^0\right)}_{y_i^0} + \Delta t_y + \underbrace{\left(-\lambda_2^0(\bar{u}_i^0\cos\alpha^0 + \bar{v}_1^0\sin\alpha^0)\right)}_{c_i} \Delta\alpha \\ & + \underbrace{(-\bar{u}_i^0 \sin\alpha^0 + \bar{v}_i^0 \cos\alpha^0)}_{d_i} \Delta\lambda_2 \underbrace{-\lambda_2^0 \sin\alpha^0}_{q} \Delta\bar{u}_i + \underbrace{\lambda_2^0 \cos\alpha^0}_{r} \Delta\bar{v}_i . \end{aligned}$$

In matrix notation:

$$
\underbrace{\begin{bmatrix} x_1 - x_1^0 \\ \vdots \\ x_p - x_p^0 \\ \cdots \\ y_1 - y_1^0 \\ \vdots \\ y_p - y_p^0 \\ \cdots \\ u_1 - \bar{u}_1^0 \\ \vdots \\ u_p - \bar{u}_p^0 \\ \cdots \\ v_1 - \bar{v}_1^0 \\ \vdots \\ v_p - \bar{v}_p^0 \end{bmatrix}}_{\substack{l \\ 4p\times 1}}
-
\underbrace{\begin{bmatrix} e_{x_1} \\ \vdots \\ e_{x_p} \\ \cdots \\ e_{y_1} \\ \vdots \\ e_{y_p} \\ \cdots \\ e_{u_1} \\ \vdots \\ e_{u_p} \\ \cdots \\ e_{v_1} \\ \vdots \\ e_{v_p} \end{bmatrix}}_{\substack{e \\ 4p\times 1}}
=
\underbrace{\begin{bmatrix}
1 & 0 & a_1 & b_1 & 0 & f_1 & g & \dots & 0 & h & \dots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
1 & 0 & a_p & b_p & 0 & f_p & 0 & \dots & g & 0 & \dots & h \\
\cdots & & & & & & & & & & & \cdots \\
0 & 1 & c_1 & 0 & d_1 & 0 & q & \dots & 0 & r & \dots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 1 & c_p & 0 & d_p & 0 & 0 & \dots & q & 0 & \dots & r \\
\cdots & & & & & & & & & & & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & \dots & 0 & 0 & \dots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & 0 & \dots & 0 \\
\cdots & & & & & & & & & & & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 1 & \dots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 1
\end{bmatrix}}_{\substack{A \\ 4p\times(2p+6)}}
\underbrace{\begin{bmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\alpha \\ \Delta\lambda_1 \\ \Delta\lambda_2 \\ \Delta k \\ \Delta\bar{u}_1 \\ \vdots \\ \Delta\bar{u}_p \\ \Delta\bar{v}_1 \\ \vdots \\ \Delta\bar{v}_p \end{bmatrix}}_{\substack{\xi \\ (2p+6)\times 1}} .
$$

Results: by using these initial approximate values for unknown parameters

$$
t_x^0 = 5500\,\mathrm{m}, \quad t_y^0 = 10\,200\,\mathrm{m}, \quad \lambda_1^0 = 1, \quad \lambda_2^0 = 1, \quad \alpha^0 = 1.5'', \quad k^0 = 0,
$$
$$
\bar{u}^0 = \begin{bmatrix} u_1, \dots, u_4 \end{bmatrix}^\mathsf{T}, \quad \bar{v}^0 = \begin{bmatrix} v_1, \dots, v_4 \end{bmatrix}^\mathsf{T}
$$

we obtain the parameters (after 5 iterations, $\|\widehat{\Delta\xi}\| < 10^{-11}$):

$$
\hat{t}_x = 5388.876\,\mathrm{m}, \qquad \hat{\alpha} = -5'7.89'', \qquad \hat{\lambda}_1 = 1.000\,409\,692,
$$
$$
\hat{t}_y = 10\,346.871\,\mathrm{m}, \qquad \hat{k} = 2.8233 \cdot 10^{-5}, \qquad \hat{\lambda}_2 = 1.000\,406\,924, \qquad \hat{e}^\mathsf{T} P \hat{e} = 0.000\,993\,2\,\mathrm{m}^2 .
$$

Coordinates of data points in the target system are listed in Tab. 5.54.

Table 5.54: Coordinates of data points in the target system.

Point	x [m]	y [m]
13	20 112.220	22 501.176
14	19 631.071	22 296.945
15	18 980.833	22 208.689
16	19 668.169	22 868.593
17	19 308.037	22 680.279

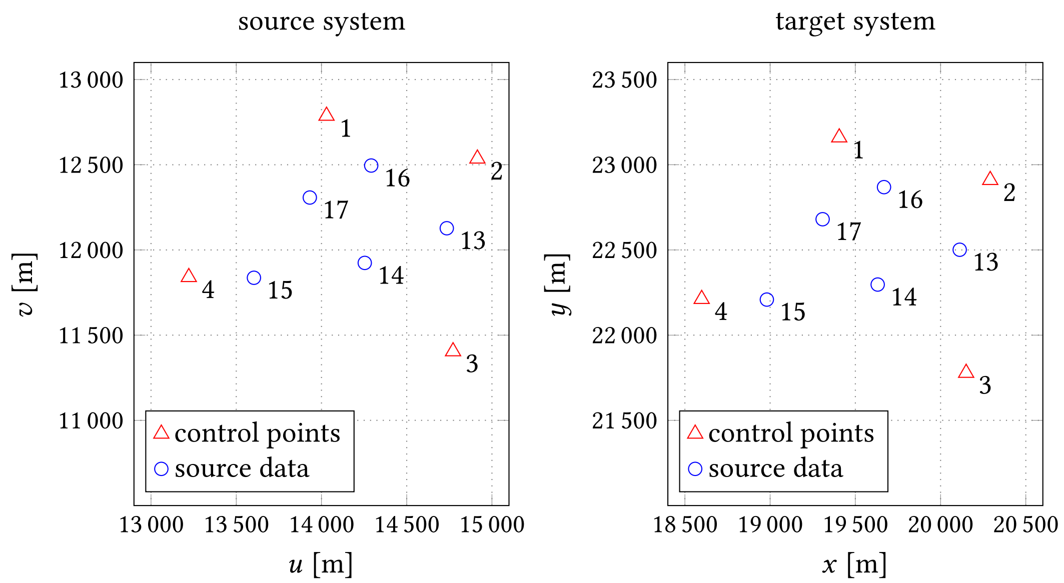


Figure 5.36: 6-parameter affine transformation: Gauss Markov model; inconsistencies in both source and target systems.

Mixed model approach II: **Extended B-model** with inconsistencies in both (x_i, y_i) and (u_i, v_i) coordinates ($i = 1, \ldots, p$ with p number of control points).

$$f_{x_i} := x_i - e_{x_i} - (\lambda_1(u_i - e_{u_i})(\cos\alpha - k\sin\alpha) + \lambda_1(v_i - e_{v_i})(\sin\alpha + k\cos\alpha) + t_x) = 0$$

$$f_{y_i} := y_i - e_{y_i} - \left(-\lambda_2(u_i - e_{u_i})\sin\alpha + \lambda_2(v_i - e_{v_i})\cos\alpha + t_y\right) = 0$$

Linearization using Taylor point $e^0_{x_i}$, $e^0_{y_i}$, $e^0_{u_i}$, $e^0_{v_i}$, t^0_x, t^0_y, α^0, λ^0_1, λ^0_2, k^0 so that

$$\begin{aligned} e_{x_i} &= e^0_{x_i} + \Delta e_{x_i}, & e_{u_i} &= e^0_{u_i} + \Delta e_{u_i}, & t_x &= t^0_x + \Delta t_x, & \alpha &= \alpha^0 + \Delta\alpha, & \lambda_2 &= \lambda^0_2 + \Delta\lambda_2, \\ e_{y_i} &= e^0_{y_i} + \Delta e_{y_i}, & e_{v_i} &= e^0_{v_i} + \Delta e_{v_i}, & t_y &= t^0_y + \Delta t_y, & k &= k^0 + \Delta k, & \lambda_1 &= \lambda^0_1 + \Delta\lambda_1\,. \end{aligned}$$

$$\begin{bmatrix} f^0_{x_i} \\ f^0_{y_i} \end{bmatrix} + \begin{bmatrix} \left.\dfrac{\partial f_{x_i}}{\partial(e_{x_i}, e_{y_i}, e_{u_i}, e_{v_i})}\right|_0 \\ \left.\dfrac{\partial f_{y_i}}{\partial(e_{x_i}, e_{y_i}, e_{u_i}, e_{v_i})}\right|_0 \end{bmatrix} \begin{bmatrix} \Delta e_{x_i} \\ \Delta e_{y_i} \\ \Delta e_{u_i} \\ \Delta e_{v_i} \end{bmatrix} + \begin{bmatrix} \left.\dfrac{\partial f_{x_i}}{\partial(t_x, t_y, \alpha, \lambda_1, \lambda_2, k)}\right|_0 \\ \left.\dfrac{\partial f_{y_i}}{\partial(t_x, t_y, \alpha, \lambda_1, \lambda_2, k)}\right|_0 \end{bmatrix} \begin{bmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\alpha \\ \Delta\lambda_1 \\ \Delta\lambda_2 \\ \Delta k \end{bmatrix} = 0,$$

where

$$f^0_{x_i} = x_i - e^0_{x_i} - \left(\lambda^0_1(u_i - e^0_{u_i})(\cos\alpha^0 - k^0\sin\alpha^0) + \lambda^0_1(v_i - e^0_{v_i})(\sin\alpha^0 + k^0\cos\alpha^0) + t^0_x\right)$$

and $$f^0_{y_i} = y_i - e^0_{y_i} - \left(-\lambda^0_2(u_i - e^0_{u_i})\sin\alpha^0 + \lambda^0_2(v_i - e^0_{v_i})\cos\alpha^0 + t^0_y\right)\,.$$

In matrix notation:

$$
\underbrace{\begin{bmatrix} x_1 - x_1^0 \\ \vdots \\ x_p - x_p^0 \\ \cdots\cdots \\ y_1 - y_1^0 \\ \vdots \\ y_p - y_p^0 \end{bmatrix}}_{\substack{w \\ 2p\times 1}}
+
\underbrace{\begin{bmatrix} -I_p & 0_p & \lambda_1^0 a^0 I_p & \lambda_1^0 b^0 I_p \\ \cdots & \cdots & \cdots & \cdots \\ 0_p & -I_p & -\lambda_2^0 \sin\alpha^0 I_p & \lambda_2^0 \cos\alpha^0 I_p \end{bmatrix}}_{\substack{B^\mathsf{T} \\ 2p\times 4p}}
\underbrace{\begin{bmatrix} e_{x_1} \\ \vdots \\ e_{x_p} \\ e_{y_1} \\ \vdots \\ e_{y_p} \\ e_{u_1} \\ \vdots \\ e_{u_p} \\ e_{v_1} \\ \vdots \\ e_{v_p} \end{bmatrix}}_{\substack{e \\ 4p\times 1}}
+
\underbrace{\begin{bmatrix} -1 & 0 & d_{a,1} & d_{b,1} & 0 & d_{c,1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & 0 & d_{a,p} & d_{b,p} & 0 & d_{c,p} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & -1 & d_{d,1} & 0 & d_{e,1} & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & -1 & d_{d,p} & 0 & d_{e,p} & 0 \end{bmatrix}}_{\substack{A \\ 2p\times 6}}
\underbrace{\begin{bmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\alpha \\ \Delta\lambda_1 \\ \Delta\lambda_2 \\ \Delta k \end{bmatrix}}_{\substack{\Delta\xi \\ 6\times 1}}
= 0,
$$

where I_p is the unit matrix of size $p \times p$ and 0_p the zero matrix of the same size. Additionally, we put the abbreviations $a^0 = \cos\alpha^0 - k^0 \sin\alpha^0$, $b^0 = \sin\alpha^0 + k^0\cos\alpha^0$, $\bar{u}_i^0 = u_i - e_{u_i}^0$ and $\bar{v}_i^0 = v_i - e_{v_i}^0$ to get

$$
\begin{gathered}
x_i^0 = \lambda_1^0 \left(a^0 u_i + b^0 v_i\right) + t_x^0, \quad y_i^0 = \lambda_2^0 \left(-\sin\alpha^0 u_i + \cos\alpha^0 v_i\right) + t_y^0 \\
d_{a,i} = \lambda_1^0 \left(b^0 \bar{u}_i^0 - a^0 \bar{v}_i^0\right), \quad d_{b,i} = -\left(a^0 \bar{u}_i^0 + b^0 \bar{v}_i^0\right), \quad d_{c,i} = \lambda_1^0 \left(\sin\alpha^0 \bar{u}_i^0 - \cos\alpha^0 \bar{v}_i^0\right), \\
d_{d,i} = \lambda_2^0 \left(\cos\alpha^0 \bar{u}_i^0 + \sin\alpha^0 \bar{v}_i^0\right), \quad d_{e,i} = \sin\alpha^0 \bar{u}_i^0 - \cos\alpha^0 \bar{v}_i^0 .
\end{gathered}
$$

Results: with the following initial approximate values for unknown parameters and inconsistencies

$$
t_x^0 = 5500\,\mathrm{m}, \quad t_y^0 = 10\,200\,\mathrm{m}, \quad \lambda_1^0 = 1, \quad \lambda_2^0 = 1, \quad \alpha^0 = 1.5'', \quad k^0 = 0, \quad e_{u_i}^0 = e_{v_i}^0 = 0 \quad \forall i
$$

we get the parameters (after 4 iterations, $\|\widehat{\Delta\xi}\| < 10^{-11}$):

$$
\begin{aligned}
\hat{t}_x &= 5388.876\,\mathrm{m}, & \hat{\alpha} &= -5'7.89'', & \hat{\lambda}_1 &= 1.000\,409\,692, & & \\
\hat{t}_y &= 10\,346.871\,\mathrm{m}, & \hat{k} &= 2.8233 \cdot 10^{-5}, & \hat{\lambda}_2 &= 1.000\,406\,924, & \hat{e}^\mathsf{T} P \hat{e} &= 0.000\,993\,2\,\mathrm{m}^2 .
\end{aligned}
$$

5.3.5 2D Affine Transformation Model II

Example: 6-parameter affine transformation—model II

The same data sets (85) will be analysed using a second model for the 6-parameter affine transformation.

$$\underbrace{\begin{bmatrix} x_i \\ y_i \end{bmatrix}}_{\text{target}} = \underbrace{\begin{bmatrix} \cos\varepsilon & -\sin\delta \\ \sin\varepsilon & \cos\delta \end{bmatrix}}_{\text{rotation angles}} \underbrace{\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}}_{\text{scale factors}} \underbrace{\begin{bmatrix} u \\ v \end{bmatrix}}_{\text{source}} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}$$

Mixed model approach I: **A-model** with inconsistencies in both $[x_i, y_i]$ and $[u_i, v_i]$ coordinates.

$$\begin{aligned} x_i - e_{x_i} &= \lambda_1 \bar{u}_i \cos\varepsilon - \lambda_2 \bar{v}_i \sin\delta + t_x \\ y_i - e_{y_i} &= \lambda_1 \bar{u}_i \sin\varepsilon + \lambda_2 \bar{v}_i \cos\delta + t_y \\ u_i - e_{u_i} &= \bar{u}_i \\ v_i - e_{v_i} &= \bar{v}_i \end{aligned}$$

Approximate values:

$$t_x = t_x^0 + \Delta t_x, \quad t_y = t_y^0 + \Delta t_y, \quad \varepsilon = \varepsilon^0 + \Delta\varepsilon, \quad \delta = \delta^0 + \Delta\delta, \quad \lambda_1 = \lambda_1^0 + \Delta\lambda, \quad \lambda_2 = \lambda_2^0 + \Delta\lambda_2 \,.$$

Linearization process:

$$\begin{aligned} x_i - e_{x_i} &= \underbrace{\left(\lambda_1^0 \bar{u}_i^0 \cos\varepsilon^0 - \lambda_2^0 \bar{v}_i^0 \sin\delta^0 + t_x^0\right)}_{x_i^0} + \Delta t_x - \lambda_1^0 \bar{u}_i^0 \sin\varepsilon^0 \Delta\varepsilon - \lambda_2^0 \bar{v}_i^0 \cos\delta^0 \Delta\delta \\ &\quad + \bar{u}_i^0 \cos\varepsilon^0 \Delta\lambda_1 - \bar{v}_i^0 \sin\delta^0 \Delta\lambda_2 + \lambda_1^0 \cos\varepsilon^0 \Delta\bar{u}_i - \lambda_2^0 \sin\delta^0 \Delta\bar{v}_i \\ y_i - e_{y_i} &= \underbrace{\left(\lambda_1^0 \bar{u}_i^0 \sin\varepsilon^0 + \lambda_2^0 \bar{v}_i^0 \cos\delta^0 + t_y^0\right)}_{y_i^0} + \Delta t_y + \lambda_1^0 \bar{u}_i^0 \cos\varepsilon^0 \Delta\varepsilon - \lambda_2^0 \bar{v}_i^0 \sin\delta^0 \Delta\delta \\ &\quad + \bar{u}_i^0 \sin\varepsilon^0 \Delta\lambda_1 + \bar{v}_i^0 \cos\delta^0 \Delta\lambda_2 + \lambda_1^0 \sin\varepsilon^0 \Delta\bar{u}_i + \lambda_2^0 \cos\delta^0 \Delta\bar{v}_i \\ u_i - e_{u_i} &= \bar{u}_i^0 + \Delta\bar{u}_i \\ v_i - e_{v_i} &= \bar{v}_i^0 + \Delta\bar{v}_i \end{aligned}$$

In matrix notation:

$$
\underbrace{\begin{bmatrix} x_1 - x_1^0 \\ \vdots \\ x_p - x_p^0 \\ \cdots\cdots \\ y_1 - y_1^0 \\ \vdots \\ y_p - y_p^0 \\ \cdots\cdots \\ u_1 - u_1^0 \\ \vdots \\ u_p - u_p^0 \\ \cdots\cdots \\ v_1 - v_1^0 \\ \vdots \\ v_p - v_p^0 \end{bmatrix}}_{\underset{4p\times 1}{l}}
-
\underbrace{\begin{bmatrix} e_{x_1} \\ \vdots \\ e_{x_p} \\ \cdots \\ e_{y_1} \\ \vdots \\ e_{y_p} \\ \cdots \\ e_{u_1} \\ \vdots \\ e_{u_p} \\ \cdots \\ e_{v_1} \\ \vdots \\ e_{v_p} \end{bmatrix}}_{\underset{4p\times 1}{e}}
=
\underbrace{\left[\begin{array}{cccccccccccccc}
1 & 0 & -a_s\bar{u}_1^0 & -b_c\bar{v}_1^0 & c_{c,1} & -d_{s,1} & a_c & \dots & 0 & -b_s & \dots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
1 & 0 & -a_s\bar{u}_p^0 & -b_c\bar{v}_p^0 & c_{c,p} & -d_{s,p} & 0 & \dots & a_c & 0 & \dots & -b_s \\
\hdashline
0 & 1 & a_c\bar{u}_1^0 & -b_s\bar{v}_1^0 & c_{s,1} & d_{c,1} & a_s & \dots & 0 & b_c & \dots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 1 & a_c\bar{u}_p^0 & -b_s\bar{v}_p^0 & c_{s,p} & d_{c,p} & 0 & \dots & a_s & 0 & \dots & b_c \\
\hdashline
0 & 0 & 0 & 0 & 0 & 0 & 1 & \dots & 0 & 0 & \dots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & 0 & \dots & 0 \\
\hdashline
0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 1 & \dots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 1
\end{array}\right]}_{\underset{4p\times(2p+6)}{A}}
\underbrace{\begin{bmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\varepsilon \\ \Delta\delta \\ \Delta\lambda_1 \\ \Delta\lambda_2 \\ \Delta\bar{u}_1 \\ \vdots \\ \Delta\bar{u}_p \\ \Delta\bar{v}_1 \\ \vdots \\ \Delta\bar{v}_p \end{bmatrix}}_{\underset{(2p+6)\times 1}{\Delta\xi}},
$$

where

$$
a_c = \lambda_1^0\cos\varepsilon^0,\qquad b_c = \lambda_2^0\cos\delta^0,\qquad c_{c,i} = \bar{u}_i^0\cos\varepsilon^0,\qquad d_{c,i} = \bar{v}_i^0\cos\delta^0,
$$
$$
a_s = \lambda_1^0\sin\varepsilon^0,\qquad b_s = \lambda_2^0\sin\delta^0,\qquad c_{s,i} = \bar{u}_i^0\sin\varepsilon^0,\qquad d_{s,i} = \bar{v}_i^0\sin\delta^0\,.
$$

Results: Initial approximate values for unknown parameters

$$
t_x^0 = 5500\,\mathrm{m},\quad t_y^0 = 10\,200\,\mathrm{m},\quad \varepsilon^0 = 1.5'',\quad \delta^0 = 3.5'',\quad \lambda_1^0 = 1,\quad \lambda_2^0 = 1,
$$
$$
\bar{u}^0 = \left[u_1,\,\dots,\,u_4\right]^{\mathsf{T}},\quad \bar{v}^0 = \left[v_1,\,\dots,\,v_4\right]^{\mathsf{T}}.
$$

Parameters (after 7 iterations, $\|\widehat{\Delta\xi}\| < 10^{-11}$):

$$
\hat{t}_x = 5388.876\,\mathrm{m},\qquad \hat{\varepsilon} = 5'7.89'',\qquad \hat{\lambda}_1 = 1.000\,409\,734,
$$
$$
\hat{t}_y = 10\,346.871\,\mathrm{m},\qquad \hat{\delta} = 5'2.06'',\qquad \hat{\lambda}_2 = 1.000\,406\,883,\qquad \hat{e}^{\mathsf{T}}P\hat{e} = 0.000\,993\,2\,\mathrm{m}^2\,.
$$

Mixed model approach II: **Extended B-model** with inconsistences in both $[x_i, y_i]$ and $[u_i, v_i]$ coordinates.

$$
f_{x_i} := x_i - e_{x_i} - \left(\lambda_1\,(u_i - e_{u_i})\cos\varepsilon - \lambda_2\,(v_i - e_{v_i})\sin\delta + t_x\right) = 0
$$
$$
f_{y_i} := y_i - e_{y_i} - \Big(\lambda_1\,(u_i - e_{u_i})\sin\varepsilon + \lambda_2\,(v_i - e_{v_i})\cos\delta + t_y\Big) = 0
$$

Initial approximate values:

$$
e_{x_i} = e_{x_i}^0 + \Delta e_{x_i},\quad e_{y_i} = e_{y_i}^0 + \Delta e_{y_i},\quad e_{u_i} = e_{u_i}^0 + \Delta e_{u_i},\quad e_{v_i} = e_{v_i}^0 + \Delta e_{v_i},
$$
$$
t_x = t_x^0 + \Delta t_x,\quad t_y = t_y^0 + \Delta t_y,\quad \varepsilon = \varepsilon^0 + \Delta\varepsilon,\quad \delta = \delta^0 + \Delta\delta,\quad \lambda_1 = \lambda_1^0 + \Delta\lambda_1,\quad \lambda_2 = \lambda_2^0 + \Delta\lambda_2\,.
$$

Linearization:

$$\begin{bmatrix} f^0_{x_i} \\ f^0_{y_i} \end{bmatrix} + \begin{bmatrix} \left.\dfrac{\partial f_{x_i}}{\partial(e_{x_i}, e_{y_i}, e_{u_i}, e_{v_i})}\right|_0 \\ \left.\dfrac{\partial f_{y_i}}{\partial(e_{x_i}, e_{y_i}, e_{u_i}, e_{v_i})}\right|_0 \end{bmatrix} \begin{bmatrix} \Delta e_{x_i} \\ \Delta e_{y_i} \\ \Delta e_{u_i} \\ \Delta e_{v_i} \end{bmatrix} + \begin{bmatrix} \left.\dfrac{\partial f_{x_i}}{\partial(t_x, t_y, \varepsilon, \delta, \lambda_1, \lambda_2)}\right|_0 \\ \left.\dfrac{\partial f_{y_i}}{\partial(t_x, t_y, \varepsilon, \delta, \lambda_1, \lambda_2)}\right|_0 \end{bmatrix} \begin{bmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\varepsilon \\ \Delta\delta \\ \Delta\lambda_1 \\ \Delta\lambda_2 \end{bmatrix} = 0,$$

where

$$\begin{aligned} f^0_{x_i} &= x_i - e^0_{x_i} - \left(\lambda^0_1(u_i - e^0_{u_i})\cos\varepsilon^0 - \lambda^0_2(v_i - e^0_{v_i})\sin\delta^0 + t^0_x\right) \\ \text{and} \quad f^0_{y_i} &= y_i - e^0_{y_i} - \left(\lambda^0_1(u_i - e^0_{u_i})\sin\varepsilon^0 + \lambda^0_2(v_i - e^0_{v_i})\cos\delta^0 + t^0_y\right) \ . \end{aligned}$$

In matrix notation:

$$\underbrace{\begin{bmatrix} x_1 - x^0_1 \\ \vdots \\ x_p - x^0_p \\ \cdots\cdots \\ y_1 - y^0_1 \\ \vdots \\ y_p - y^0_p \end{bmatrix}}_{\substack{w \\ 2p\times 1}} + \underbrace{\begin{bmatrix} -I_p & 0_p & a_c I_p & -b_s I_p \\ \cdots & \cdots & \cdots & \cdots \\ 0_p & -I_p & a_s I_p & b_c I_p \end{bmatrix}}_{\substack{B^\mathsf{T} \\ 2p\times 4p}} \underbrace{\begin{bmatrix} e_{x_1} \\ \vdots \\ e_{x_p} \\ e_{y_1} \\ \vdots \\ e_{y_p} \\ e_{u_1} \\ \vdots \\ e_{u_p} \\ e_{v_1} \\ \vdots \\ e_{v_p} \end{bmatrix}}_{\substack{e \\ 4p\times 1}} + \underbrace{\begin{bmatrix} -1 & 0 & a_s\bar{u}^0_1 & b_c\bar{v}^0_1 & -c_{c,1} & d_{s,1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & 0 & a_s\bar{u}^0_p & b_c\bar{v}^0_p & -c_{c,p} & d_{s,p} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & -1 & -a_c\bar{u}_1 & b_s\bar{v}_1 & -c_{s,1} & -d_{c,1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & -1 & -a_c\bar{u}_p & b_s\bar{v}_p & -c_{s,p} & -d_{c,p} \end{bmatrix}}_{\substack{A \\ 2p\times 6}} \underbrace{\begin{bmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\varepsilon \\ \Delta\delta \\ \Delta\lambda_1 \\ \Delta\lambda_2 \end{bmatrix}}_{\substack{\Delta\xi \\ 6\times 1}} = 0,$$

where I_p is the unit matrix of size $p \times p$ and 0_p the zero matrix of the same size. Additionally, we use the following abbreviations

$$\begin{aligned} a_c &= \lambda^0_1\cos\varepsilon^0, & b_c &= \lambda^0_2\cos\delta^0, & c_{c,i} &= \bar{u}^0_i\cos\varepsilon^0, & d_{c,i} &= \bar{v}^0_i\cos\delta^0, & \bar{u}^0_i &= u_i - e^0_{u_i}, \\ a_s &= \lambda^0_1\sin\varepsilon^0, & b_s &= \lambda^0_2\sin\delta^0, & c_{s,i} &= \bar{u}^0_i\sin\varepsilon^0, & d_{s,i} &= \bar{v}^0_i\sin\delta^0, & \bar{v}^0_i &= v_i - e^0_{u_i} \ . \end{aligned}$$

Results: with the following initial approximate values for unknown parameters

$$t^0_x = 5500\,\text{m}, \quad t^0_y = 10\,200\,\text{m}, \quad \varepsilon^0 = 1.5'', \quad \delta^0 = 3.5'', \quad \lambda^0_1 = 1, \quad \lambda^0_2 = 1,$$
$$e^0_{x_i} = e^0_{y_i} = e^0_{u_i} = e^0_{v_i} = 0 \quad \forall i \ .$$

we get the parameters (after 20 iterations, $\|\widehat{\Delta\xi}\| < 10^{-11}$):

$$\begin{aligned} \hat{t}_x &= 5388.876\,\text{m}, & \hat{\varepsilon} &= 5'7.89'', & \hat{\lambda}_1 &= 1.000\,409\,734, \\ \hat{t}_y &= 10\,346.871\,\text{m}, & \hat{\delta} &= 5'2.06'', & \hat{\lambda}_2 &= 1.000\,406\,882, & \hat{e}^\mathsf{T}P\hat{e} &= 0.000\,993\,2\,\text{m}^2 \ . \end{aligned}$$

5.3.6 Ellipse fit under various restrictions

Example: Best fitting ellipse (here: principal axes aligned with coordinate axes) with unknown semi major axis a, semi minor axis b and centre coordinates $(x_\mathrm{M}, y_\mathrm{M})$; observations x_i and y_i are inconsistent.

$$f(\underbrace{a, b, x_\mathrm{M}, y_\mathrm{M}}_{\substack{\text{unknown}\\ \text{parameters } \xi}}, \underbrace{x_i - e_{x_i}, y_i - e_{y_i}}_{\substack{\text{observations } y\\ -\text{ inconsistencies } e}}) = \frac{(x_i - e_{x_i} - x_\mathrm{M})^2}{a^2} + \frac{\left(y_i - e_{y_i} - y_\mathrm{M}\right)^2}{b^2} - 1 = 0$$

Possible restriction: Best fitting ellipse shall pass through the point $(x_\mathrm{P}, y_\mathrm{P})$

$$g(\underbrace{a, b, x_\mathrm{M}, y_\mathrm{M}}_{\xi}) = \frac{(x_\mathrm{P} - x_\mathrm{M})^2}{a^2} + \frac{(y_\mathrm{P} - y_\mathrm{M})^2}{b^2} - 1 = 0$$

Linearization (with $e^0_{x_i} = e^0_{y_i} = 0$ in the first iteration and given ξ_0)

$$\xi = \xi_0 + \Delta\xi, \qquad e_{x_i} = e^0_{x_i} + \Delta e_{x_i}, \qquad e_{y_i} = e^0_{y_i} + \Delta e_{y_i}$$

$$\begin{aligned}
f(\xi, e) = f(\xi_0, e_0) &+ \left.\frac{\partial f}{\partial \xi}\right|_{x_0, e_0} \Delta\xi + \left.\frac{\partial f}{\partial e}\right|_{\xi_0, e_0} e + O = 0\\
\doteq w &+ A\,\Delta\xi \qquad + B^\mathsf{T} e \qquad = 0\\
g(\xi) = g(\xi_0) &+ \left.\frac{\partial g}{\partial \xi}\right|_{\xi_0} \Delta\xi \qquad + O = 0\\
\doteq w_R &+ R\,\Delta\xi \qquad = 0
\end{aligned}$$

with O terms of higher order.

Linear model and adjustment principle

$$\left.\begin{aligned} A\Delta\xi + B^\mathsf{T} e &= -w\\ R\Delta\xi &= -w_R \end{aligned}\right\} \quad \frac{1}{2} e^\mathsf{T} W e \longrightarrow \min$$

Constrained Lagrangian (m observation equations, n unknown parameters, p inconsistencies, r restrictions)

$$\mathcal{L}_R(\Delta\xi, e, \lambda, \lambda_R) = \frac{1}{2} \underset{1\times p}{e^\mathsf{T}} \underset{p\times p}{W} \underset{p\times 1}{e} + \underset{1\times m}{\lambda^\mathsf{T}} \big(\underset{m\times n}{A} \underset{n\times 1}{\Delta\xi} + \underset{m\times p}{B^\mathsf{T}} \underset{p\times 1}{e} + \underset{m\times 1}{w} \big) + \underset{1\times r}{\lambda_R{}^\mathsf{T}} \big(\underset{r\times n}{R} \underset{n\times 1}{\Delta\xi} + \underset{r\times 1}{w_R} \big) \longrightarrow \min_{\Delta\xi, e, \lambda, \lambda_R}$$

Necessary condition

$$\begin{aligned}
\frac{\partial \mathcal{L}_R}{\partial \Delta\xi}(\widehat{\Delta\xi}, \hat{e}, \hat{\lambda}, \hat{\lambda}_R) &\stackrel{!}{=} 0 &\Longrightarrow&& A^\mathsf{T}\hat{\lambda} + R^\mathsf{T}\hat{\lambda}_R &= 0\\
\frac{\partial \mathcal{L}_R}{\partial e}(\widehat{\Delta\xi}, \hat{e}, \hat{\lambda}, \hat{\lambda}_R) &\stackrel{!}{=} 0 &\Longrightarrow&& W\hat{e} + B\hat{\lambda} &= 0\\
\frac{\partial \mathcal{L}_R}{\partial \lambda}(\widehat{\Delta\xi}, \hat{e}, \hat{\lambda}, \hat{\lambda}_R) &\stackrel{!}{=} 0 &\Longrightarrow&& A\widehat{\Delta\xi} + B^\mathsf{T}\hat{e} &= -w\\
\frac{\partial \mathcal{L}_R}{\partial \lambda_R}(\widehat{\Delta\xi}, \hat{e}, \hat{\lambda}, \hat{\lambda}_R) &\stackrel{!}{=} 0 &\Longrightarrow&& R\widehat{\Delta\xi} &= -w_R
\end{aligned}$$

$$\underset{(p+m+n+r)\times(p+m+n+r)}{\begin{bmatrix} \underset{p\times p}{W} & \underset{p\times m}{B} & \underset{p\times n}{0} & \underset{p\times r}{0}\\ \underset{m\times p}{B^\mathsf{T}} & \underset{m\times m}{0} & \underset{m\times n}{A} & \underset{m\times r}{0}\\ \underset{n\times p}{0} & \underset{n\times m}{A^\mathsf{T}} & \underset{n\times n}{0} & \underset{n\times r}{R^\mathsf{T}}\\ \underset{r\times p}{0} & \underset{r\times m}{0} & \underset{r\times n}{R} & \underset{r\times r}{0} \end{bmatrix}} \underset{(p+m+n+r)\times 1}{\begin{bmatrix} \hat{e}\\ \hat{\lambda}\\ \widehat{\Delta\xi}\\ \hat{\lambda}_R \end{bmatrix}} = \begin{bmatrix} \underset{p\times 1}{0}\\ \underset{m\times 1}{-w}\\ \underset{n\times 1}{0}\\ \underset{r\times 1}{-w_R} \end{bmatrix}$$

1st row multiplied with $-B^\mathsf{T}W^{-1}$ (from left) is added to 2nd row

$$\begin{bmatrix} W & B & 0 & 0\\ 0 & -B^\mathsf{T}W^{-1}B & A & 0\\ 0 & A^\mathsf{T} & 0 & R^\mathsf{T}\\ 0 & 0 & R & 0 \end{bmatrix} \begin{bmatrix} \hat{e}\\ \hat{\lambda}\\ \widehat{\Delta\xi}\\ \hat{\lambda}_R \end{bmatrix} = \begin{bmatrix} 0\\ -w\\ 0\\ -w_R \end{bmatrix}$$

2nd row multiplied with $A^\mathsf{T}\left(B^\mathsf{T}W^{-1}B\right)^{-1}$ (from left) is added to 3rd row

$$\begin{bmatrix} W & B & 0 & 0\\ 0 & -B^\mathsf{T}W^{-1}B & A & 0\\ 0 & 0 & A^\mathsf{T}\left(B^\mathsf{T}W^{-1}B\right)^{-1}A & R^\mathsf{T}\\ 0 & 0 & R & 0 \end{bmatrix} \begin{bmatrix} \hat{e}\\ \hat{\lambda}\\ \widehat{\Delta\xi}\\ \hat{\lambda}_R \end{bmatrix} = \begin{bmatrix} 0\\ -w\\ -A^\mathsf{T}\left(B^\mathsf{T}W^{-1}B\right)^{-1}w\\ -w_R \end{bmatrix}$$

$$\Longrightarrow \quad \begin{bmatrix} A^\mathsf{T}(B^\mathsf{T}W^{-1}B)^{-1}A & R^\mathsf{T}\\ R & 0 \end{bmatrix} \begin{bmatrix} \widehat{\Delta\xi}\\ \hat{\lambda}_R \end{bmatrix} = \begin{bmatrix} -A^\mathsf{T}(B^\mathsf{T}W^{-1}B)^{-1}w\\ -w_R \end{bmatrix}$$

Case 1: $A^\mathsf{T}\left(B^\mathsf{T}W^{-1}B\right)^{-1}A = A^\mathsf{T}M^{-1}A$ is a full-rank matrix. $\Longrightarrow$ Use partitioning formula:

$$\begin{pmatrix} N_{11} & N_{12}\\ N_{21} & N_{22} \end{pmatrix} \begin{pmatrix} Q_{11} & Q_{12}\\ Q_{21} & Q_{22} \end{pmatrix} = \begin{pmatrix} I & 0\\ 0 & I \end{pmatrix} \quad \Longrightarrow \quad \begin{cases} Q_{22} = (N_{22} - N_{21}N_{11}^{-1}N_{12})^{-1}\\ Q_{12} = -N_{11}^{-1}N_{12}Q_{22}\\ Q_{21} = -Q_{22}N_{21}N_{11}^{-1}\\ Q_{11} = N_{11}^{-1} + N_{11}^{-1}N_{12}Q_{22}N_{21}N_{11}^{-1} \end{cases}$$

$$
\begin{aligned}
N_{11} &= A^{\mathsf{T}}(B^{\mathsf{T}}W^{-1}B)^{-1}A = A^{\mathsf{T}}M^{-1}A \\
N_{12} &= R^{\mathsf{T}} \\
N_{21} &= {N_{12}}^{\mathsf{T}} = R \\
N_{22} &= 0 \\
Q_{22} &= [0 - R(A^{\mathsf{T}}M^{-1}A)^{-1}R^{\mathsf{T}}]^{-1} = -[R(A^{\mathsf{T}}M^{-1}A)^{-1}R^{\mathsf{T}}]^{-1} \\
Q_{12} &= (A^{\mathsf{T}}M^{-1}A)^{-1}R^{\mathsf{T}}[R(A^{\mathsf{T}}M^{-1}A)^{-1}R^{\mathsf{T}}]^{-1} = -N_{11}^{-1}N_{12}Q_{22} \\
Q_{21} &= {Q_{12}}^{\mathsf{T}} \\
Q_{11} &= (A^{\mathsf{T}}M^{-1}A)^{-1}\{I - R^{\mathsf{T}}[R(A^{\mathsf{T}}M^{-1}A)^{-1}R^{\mathsf{T}}]^{-1}R(A^{\mathsf{T}}M^{-1}A)^{-1}\} \\
&= N_{11}^{-1} - Q_{12}{N_{12}}^{\mathsf{T}}N_{11}^{-1}
\end{aligned}
$$

$$
\begin{aligned}
\widehat{\Delta\xi} &= -Q_{11}A^{\mathsf{T}}M^{-1}w - Q_{12}w_R \\
&= -(A^{\mathsf{T}}M^{-1}A)^{-1}A^{\mathsf{T}}M^{-1}w \\
&\quad + (A^{\mathsf{T}}M^{-1}A)^{-1}R^{\mathsf{T}}\left(R(A^{\mathsf{T}}M^{-1}A)^{-1}R^{\mathsf{T}}\right)^{-1}R(A^{\mathsf{T}}M^{-1}A)^{-1}A^{\mathsf{T}}M^{-1}w \\
&\quad - (A^{\mathsf{T}}M^{-1}A)^{-1}R^{\mathsf{T}}\left(R(A^{\mathsf{T}}M^{-1}A)^{-1}R^{\mathsf{T}}\right)^{-1}w_R \\
&= -(A^{\mathsf{T}}M^{-1}A)^{-1}A^{\mathsf{T}}M^{-1}w + \delta\widehat{\Delta\xi} \\
&= \widehat{\Delta\xi} \quad \text{(without restrictions } g(\xi) = 0) \quad + \delta\widehat{\Delta\xi}
\end{aligned}
$$

$$
\begin{aligned}
\hat{\lambda}_R &= -Q_{21}A^{\mathsf{T}}M^{-1}w - Q_{22}w_R \\
&= Q_{22}\left(R^{\mathsf{T}}(A^{\mathsf{T}}M^{-1}A)^{-1}A^{\mathsf{T}}M^{-1}w - w_R\right) \\
&= \left(R(A^{\mathsf{T}}M^{-1}A)^{-1}R^{\mathsf{T}}\right)\left(w_R - R^{\mathsf{T}}(A^{\mathsf{T}}M^{-1}A)^{-1}A^{\mathsf{T}}M^{-1}w\right) \\
-w &= -M\hat{\lambda} + A\widehat{\Delta\xi} \\
\implies \quad \hat{\lambda} &= M^{-1}(A\widehat{\Delta\xi} + w) \\
\hat{e} &= W^{-1}B\hat{\lambda} = W^{-1}BM^{-1}(A\widehat{\Delta\xi} + w)
\end{aligned}
$$

Case 2: $A^{\mathsf{T}}(B^{\mathsf{T}}W^{-1}B)^{-1}A = A^{\mathsf{T}}M^{-1}A$ is a rank deficient matrix

$$
\operatorname{rank}(A^{\mathsf{T}}M^{-1}A) = \operatorname{rank} A = n - d
$$

$$
\begin{bmatrix} N & R^{\mathsf{T}} \\ R & 0 \end{bmatrix}^{-1} = \begin{bmatrix} R & S^{\mathsf{T}} \\ S & Q \end{bmatrix}^{-1}
$$

$$NR + R^\mathsf{T} S = I \tag{5.3}$$

$$NS^\mathsf{T} + R^\mathsf{T} Q = 0 \tag{5.4}$$

$$RR = 0 \tag{5.5}$$

$$RS^\mathsf{T} = I \tag{5.6}$$

Since A is rank deficient $AH^\mathsf{T} = 0$ where $H = \text{null}(A) \quad H : d \times n$ therefore

$$\begin{cases} A^\mathsf{T} M^{-1} A H^\mathsf{T} & = 0 \\ NH^\mathsf{T} & = 0 \\ HN & = 0 \end{cases}$$

N is symmetric.

$$\begin{aligned} H \cdot (5.3) \quad &\Longrightarrow \quad H \underbrace{NR}_{0} + HR^\mathsf{T} S = H \quad \Longrightarrow \quad S = (HR^\mathsf{T})^{-1} H \\ H \cdot (5.4) \quad &\Longrightarrow \quad H \underbrace{NS^\mathsf{T}}_{0} + HR^\mathsf{T} Q = 0 \quad \Longrightarrow \quad HR^\mathsf{T} Q = 0 \end{aligned}$$

HR^T full rank $\Longrightarrow Q = 0$.

$$\left.\begin{aligned} (5.3) \quad &\Longrightarrow \quad NR + R^\mathsf{T}(HR^\mathsf{T})^{-1} H = I \\ (5.5) \quad &\Longrightarrow \quad RR = 0 \Longrightarrow R^\mathsf{T} RR = 0 \end{aligned}\right\} \oplus \quad \Longrightarrow \quad (N + R^\mathsf{T} R) R = I - R^\mathsf{T}(HR^\mathsf{T})^{-1} H$$

$$\Longrightarrow \quad R = (N + R^\mathsf{T} R)^{-1} \left(I - R^\mathsf{T}(HR^\mathsf{T})^{-1} H \right)$$

$$\begin{aligned} \widehat{\Delta\xi} &= -RA^\mathsf{T} N^{-1} w + S^\mathsf{T} w_R \\ &= -(N + R^\mathsf{T} R)^{-1} A^\mathsf{T} M^{-1} w \\ &+ \underbrace{(N + R^\mathsf{T} R)^{-1} R^\mathsf{T} (HR^\mathsf{T})^{-1} HA^\mathsf{T} M^{-1} w}_{=0} - S^\mathsf{T} w_R \\ &= -(N + R^\mathsf{T} R)^{-1} A^\mathsf{T} M^{-1} w - H^\mathsf{T} (RH^\mathsf{T})^{-1} w_R \end{aligned}$$

if $w_R = 0$:

$$\widehat{\Delta\xi} = -(N + R^\mathsf{T} R)^{-1} A^\mathsf{T} M^{-1} w$$

$$\begin{aligned} \widehat{\Delta\xi} \longrightarrow \hat{\lambda} &= M^{-1}(A\widehat{\Delta\xi} + w) \\ &= M^{-1}(-(N + R^\mathsf{T} R)^{-1} A^\mathsf{T} M^{-1} w \\ \hat{\lambda} &= -M^{-1} \left((N + R^\mathsf{T} R)^{-1} A^\mathsf{T} M^{-1} - I \right) w \\ \hat{e} &= W^{-1} B \hat{\lambda} \\ &= -W^{-1} B M^{-1} \left((N + R^\mathsf{T} R)^{-1} A^\mathsf{T} M^{-1} - I \right) w \end{aligned}$$

Examples for case 1: $A^\mathsf{T}(B^\mathsf{T}B)^{-1}A$ is a full-rank matrix.

Best fitting ellipse with unknown semi major axes a and b, unknown centre coordinates x_M, y_M, inconsistent observations x_i and y_i, $i = 1, \ldots, m$, no restrictions $g(x)$, ellipse aligned with coordinate axes!

Ellipse equation

$$
\begin{aligned}
f(a, b, x_\mathrm{M}, y_\mathrm{M}, x_i - e_{x_i}, y_i - e_{y_i}) = {} & \left(\frac{x_i - e^0_{x_i} - x^0_\mathrm{M}}{a_0}\right)^2 + \left(\frac{y_i - e^0_{y_i} - y^0_\mathrm{M}}{b_0}\right)^2 - 1 \\
& + \frac{2(x_i - e^0_{x_i} - x^0_\mathrm{M})}{a_0^2} e^0_{x_i} + \frac{2(y_i - e^0_{y_i} - y^0_\mathrm{M})}{b_0^2} e^0_{y_i} \\
& - \frac{2(x_i - e^0_{x_i} - x^0_\mathrm{M})}{a_0^2} \Delta x_\mathrm{M} - \frac{2(y_i - e^0_{y_i} - y^0_\mathrm{M})}{b_0^2} \Delta y_\mathrm{M} \\
& - \frac{2(x_i - e^0_{x_i} - x^0_\mathrm{M})}{a_0^2} e_{x_i} - \frac{2(y_i - e^0_{y_i} - y^0_\mathrm{M})}{b_0^2} e_{y_i} \\
& - \frac{2(x_i - e^0_{x_i} - x^0_\mathrm{M})^2}{a_0^3} \Delta a - \frac{2(y_i - e^0_{y_i} - y^0_\mathrm{M})^2}{b_0^3} \Delta b = 0
\end{aligned}
$$

$$
\begin{aligned}
\Longrightarrow \quad & \begin{bmatrix} \frac{2(x_i - e^0_{x_i} - x^0_\mathrm{M})}{a_0^2} & \frac{2(y_i - e^0_{y_i} - y^0_\mathrm{M})}{b_0^2} \end{bmatrix} \begin{bmatrix} e_{x_i} \\ e_{y_i} \end{bmatrix} \\
& + \begin{bmatrix} -\frac{2(x_i - e^0_{x_i} - x^0_\mathrm{M})}{a_0^2} & -\frac{2(y_i - e^0_{y_i} - y^0_\mathrm{M})}{b_0^2} & -\frac{2(x_i - e^0_{x_i} - x^0_\mathrm{M})^2}{a_0^3} & -\frac{2(y_i - e^0_{y_i} - y^0_\mathrm{M})^2}{b_0^3} \end{bmatrix} \begin{bmatrix} \Delta x_\mathrm{M} \\ \Delta y_\mathrm{M} \\ \Delta a \\ \Delta b \end{bmatrix} \\
& + \frac{(x_i - e^0_{x_i} - x^0_\mathrm{M})^2}{a_0^2} + \frac{(y_i - e^0_{y_i} - y^0_\mathrm{M})^2}{b_0^2} - 1 + \frac{2(x_i - e^0_{x_i} - x^0_\mathrm{M})}{a_0^2} e^0_{x_i} + \frac{2(y_i - e^0_{y_i} - y^0_\mathrm{M})}{b_0^2} e^0_{y_i} = 0
\end{aligned}
$$

$$
\Longrightarrow \qquad p = 2m, \quad W = I_p \quad (p \times p \text{ identity matrix}), \quad n = 4
$$

$$\implies \qquad B^\mathsf{T} e + A\Delta\xi + w = 0$$

$$\underset{m\times p}{B^\mathsf{T}} = -2\begin{bmatrix} \frac{x_1-e^0_{x_1}-x^0_\mathrm{M}}{a_0^2} & \frac{y_1-e^0_{y_1}-y^0_\mathrm{M}}{b_0^2} & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \frac{x_2-e^0_{x_2}-x^0_\mathrm{M}}{a_0^2} & \frac{y_2-e^0_{y_2}-y^0_\mathrm{M}}{b_0^2} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \frac{x_m-e^0_{x_m}-x^0_\mathrm{M}}{a_0^2} & \frac{y_m-e^0_{y_m}-y^0_\mathrm{M}}{b_0^2} \end{bmatrix}$$

$$\underset{m\times 4}{A} = -2\begin{bmatrix} \frac{x_1-e^0_{x_1}-x^0_\mathrm{M}}{a_0^2} & \frac{y_1-e^0_{y_1}-y^0_\mathrm{M}}{b_0^2} & \frac{(x_1-e^0_{x_1}-x^0_\mathrm{M})^2}{a_0^3} & \frac{(y_1-e^0_{y_1}-y^0_\mathrm{M})^2}{b_0^3} \\ \frac{x_2-e^0_{x_2}-x^0_\mathrm{M}}{a_0^2} & \frac{y_2-e^0_{y_2}-y^0_\mathrm{M}}{b_0^2} & \frac{(x_2-e^0_{x_2}-x^0_\mathrm{M})^2}{a_0^3} & \frac{(y_2-e^0_{y_2}-y^0_\mathrm{M})^2}{b_0^3} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{x_m-e^0_{x_m}-x^0_\mathrm{M}}{a_0^2} & \frac{y_m-e^0_{y_m}-y^0_\mathrm{M}}{b_0^2} & \frac{(x_m-e^0_{x_m}-x^0_\mathrm{M})^2}{a_0^3} & \frac{(y_m-e^0_{y_m}-y^0_\mathrm{M})^2}{b_0^3} \end{bmatrix}$$

$$\underset{p\times 1}{e} = \begin{bmatrix} e_{x_1} & e_{y_1} & e_{x_2} & e_{y_2} & \dots & e_{x_m} & e_{y_m} \end{bmatrix}^\mathsf{T}$$

$$\underset{4\times 1}{\Delta\xi} = \begin{bmatrix} \Delta x_\mathrm{M} & \Delta y_\mathrm{M} & \Delta a & \Delta b \end{bmatrix}^\mathsf{T}$$

$$\underset{m\times 1}{w} = \begin{bmatrix} \frac{(x_1-e^0_{x_1}-x^0_\mathrm{M})^2}{a_0^2} + \frac{(y_1-e^0_{y_1}-y^0_\mathrm{M})^2}{b_0^2} + 2\frac{x_1-e^0_{x_1}-x^0_\mathrm{M}}{a_0^2}e^0_{x_1} + 2\frac{y_1-e^0_{y_1}-y^0_\mathrm{M}}{b_0^2}e^0_{y_1} - 1 \\ \frac{(x_2-e^0_{x_2}-x^0_\mathrm{M})^2}{a_0^2} + \frac{(y_2-e^0_{y_2}-y^0_\mathrm{M})^2}{b_0^2} + 2\frac{x_2-e^0_{x_2}-x^0_\mathrm{M}}{a_0^2}e^0_{x_2} + 2\frac{y_2-e^0_{y_2}-y^0_\mathrm{M}}{b_0^2}e^0_{y_2} - 1 \\ \vdots \\ \frac{(x_m-e^0_{x_m}-x^0_\mathrm{M})^2}{a_0^2} + \frac{(y_m-e^0_{y_m}-y^0_\mathrm{M})^2}{b_0^2} + 2\frac{x_m-e^0_{x_m}-x^0_\mathrm{M}}{a_0^2}e^0_{x_m} + 2\frac{y_m-e^0_{y_m}-y^0_\mathrm{M}}{b_0^2}e^0_{y_m} - 1 \end{bmatrix}$$

$$\implies \quad \underset{(3m+4)\times(3m+4)}{\begin{bmatrix} \underset{2m\times 2m}{W} & \underset{2m\times m}{B} & \underset{2m\times n}{0} \\ \underset{m\times 2m}{B^\mathsf{T}} & \underset{m\times m}{0} & \underset{m\times n}{A} \\ \underset{n\times 2m}{0^\mathsf{T}} & \underset{n\times m}{A^\mathsf{T}} & \underset{n\times n}{0} \end{bmatrix}} \underset{(3m+4)\times 1}{\begin{bmatrix} \underset{2m\times 1}{\hat{e}} \\ \underset{m\times 1}{\hat{\lambda}} \\ \underset{n\times 1}{\widehat{\Delta\xi}} \end{bmatrix}} = \underset{(3m+4)\times 1}{\begin{bmatrix} \underset{2m\times 1}{0} \\ \underset{m\times 1}{-w} \\ \underset{n\times 1}{0} \end{bmatrix}}$$

$$\left.\begin{array}{l} W\hat{e} + B\hat{\lambda} = 0 \quad \implies \quad \hat{e} = -W^{-1}B\hat{\lambda} \\ B^\mathsf{T}\hat{e} + A\widehat{\Delta\xi} = -w \end{array}\right\} \implies -B^\mathsf{T}W^{-1}B\hat{\lambda} + A\widehat{\Delta\xi} = -w$$

$$\implies \quad \hat{\lambda} = (B^\mathsf{T}W^{-1}B)^{-1}(A\widehat{\Delta\xi} + w)$$

$$\implies \quad A^\mathsf{T}(B^\mathsf{T}W^{-1}B)^{-1}A\widehat{\Delta\xi} + A^\mathsf{T}(B^\mathsf{T}W^{-1}B)^{-1}w = 0$$

$$\implies \quad \widehat{\Delta\xi} = -\left(A^\mathsf{T}(B^\mathsf{T}W^{-1}B)^{-1}A\right)^{-1} A^\mathsf{T}(B^\mathsf{T}W^{-1}B)^{-1}w$$

$$\begin{aligned} \implies \quad \hat{e} &= -W^{-1}B(B^\mathsf{T}W^{-1}B)^{-1}(A\widehat{\Delta\xi} + w) \\ &= W^{-1}B(B^\mathsf{T}W^{-1}B)^{-1}\left(A\left(A^\mathsf{T}(B^\mathsf{T}W^{-1}B)^{-1}A\right)^{-1}A^\mathsf{T}(B^\mathsf{T}W^{-1}B)^{-1} - I\right)w \end{aligned}$$

Numerics

$$x = \begin{bmatrix} 0, & 50, & 90, & 120, & 130, & -130, & -100, & -50, & 0 \end{bmatrix}^\mathsf{T}$$
$$y = \begin{bmatrix} 120, & 110, & 80, & 0, & -50, & -50, & 60, & 100, & -110 \end{bmatrix}^\mathsf{T}$$

Approximated values:

$$x_M^0 = y_M^0 = 0, \quad a_0 = b_0 = 120\,.$$

Parameters (after 10 iterations: $\|\widehat{\Delta\xi}\| < 10^{-12}$, $\|\widehat{\Delta e}\| < 10^{-12}$):

$$\hat{x}_M = -0.598, \quad \hat{y}_M = -1.942, \quad \hat{a} = 131.087, \quad \hat{b} = 115.131, \quad \hat{e}^\mathsf{T} W \hat{e} = 523.208$$

$$\hat{e}_x = \begin{bmatrix} 0.026, & 1.793, & -0.627, & -10.466, & 9.322, & -8.324, & 6.771, & 1.534, & -0.030 \end{bmatrix}^\mathsf{T}$$
$$\hat{e}_y = \begin{bmatrix} 6.813, & 5.089, & -0.736, & -0.224, & -4.355, & -3.933, & -5.583, & -4.142, & 7.072 \end{bmatrix}^\mathsf{T}$$

See figure 5.37.

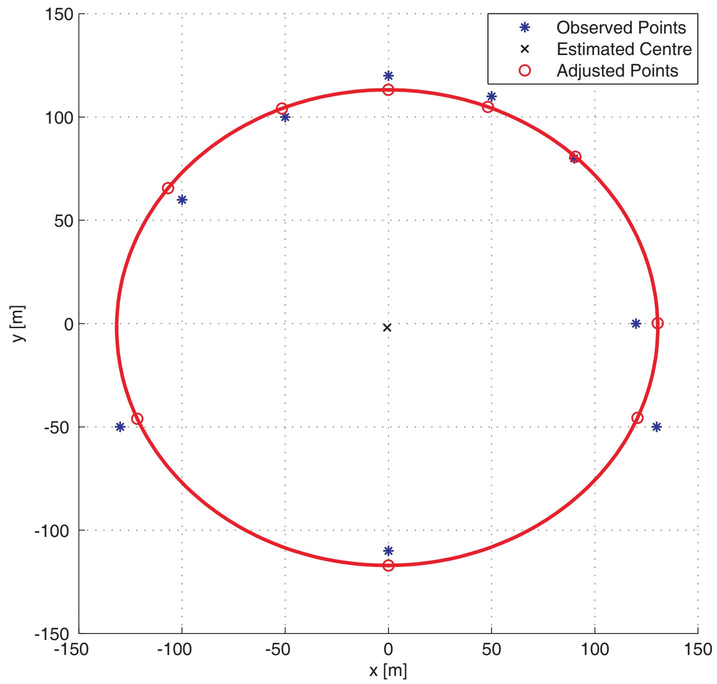


Figure 5.37: Ellipse fit (mixed model), no restriction.

Example 2: as example 1, but with additional (linear) restriction $g(\xi) = 0$ so that $\hat{a} = \hat{b}$ (best fitting circle).

$$g(\xi) = g(x_M, y_M, a, b) = a - b = 0$$
$$\Longrightarrow \quad R = [0, 0, 1, -1], \quad w_R = 0$$

Parameters (after 10 iterations: $||\widehat{\Delta\xi}|| < 10^{-12}$, $||\widehat{\Delta e}|| < 10^{-12}$):

$$\hat{x}_{\mathrm{M}} = 1.119, \quad \hat{y}_{\mathrm{M}} = -3.921, \quad \hat{a} = \hat{b} = 122.939, \quad \hat{e}^{\mathsf{T}} W \hat{e} = 815.668$$

$$\hat{e}_x = \begin{bmatrix} -0.009, & 0.404, & -0.509, & -3.992, & 13.118, & -15.134, & 2.798, & 3.145, & 0.178 \end{bmatrix}^{\mathsf{T}}$$
$$\hat{e}_y = \begin{bmatrix} 0.987, & 0.943, & -0.480, & -0.132, & -4.690, & -5.318, & -1.769, & -6.394, & 16.854 \end{bmatrix}^{\mathsf{T}}$$

See figure 5.38.

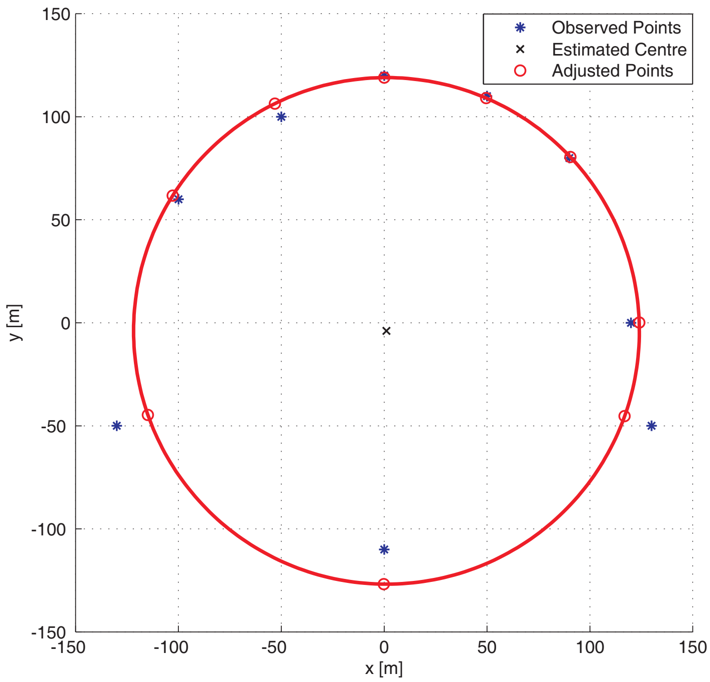


Figure 5.38: Ellipse fit (mixed model), circle restriction.

Example 3: as example 1, but with additional (non-linear) constraint $g(\xi)$ so that best fitting ellipse passes through the point $x_{\mathrm{P}} = 100$, $y_{\mathrm{P}} = -100$.

$$g(\xi) = g(x_m, y_m, a, b) = \frac{(x_{\mathrm{P}} - x_{\mathrm{M}})^2}{a^2} + \frac{(y_{\mathrm{P}} - y_{\mathrm{M}})^2}{b^2} - 1 = 0$$

$$\Longrightarrow \quad R = -2\begin{bmatrix} \frac{x_{\mathrm{P}}-x_{\mathrm{M}}^0}{a_0^2} & \frac{y_{\mathrm{P}}-y_{\mathrm{M}}^0}{b_0^2} & \frac{(x_{\mathrm{P}}-x_{\mathrm{M}}^0)^2}{a_0^3} & \frac{(y_{\mathrm{P}}-y_{\mathrm{M}}^0)^2}{b_0^3} \end{bmatrix}, \quad w_R = \frac{(x_{\mathrm{P}} - x_{\mathrm{M}}^0)^2}{a_0^2} + \frac{(y_{\mathrm{P}} - y_{\mathrm{M}}^0)^2}{b_0^2} - 1$$

Parameters (after 10 iterations: $||\widehat{\Delta\xi}|| < 10^{-12}$, $||\widehat{\Delta e}|| < 10^{-12}$):

$$\hat{x}_{\mathrm{M}} = 5.402, \quad \hat{y}_{\mathrm{M}} = -11.769, \quad \hat{a} = 134.124, \quad \hat{b} = 124.460, \quad \hat{e}' W \hat{e} = 1197.412$$

$$\hat{e}_x = \begin{bmatrix} -0.263, & 1.262, & -2.361, & -18.668, & -2.701, & -6.996, & 2.583, & 0.569, & 1.191 \end{bmatrix}^{\mathsf{T}}$$
$$\hat{e}_y = \begin{bmatrix} 7.401, & 3.985, & -2.988, & -2.287, & 0.966, & -2.275, & -2.051, & -1.336, & 26.079 \end{bmatrix}^{\mathsf{T}}$$

See figure 5.39.

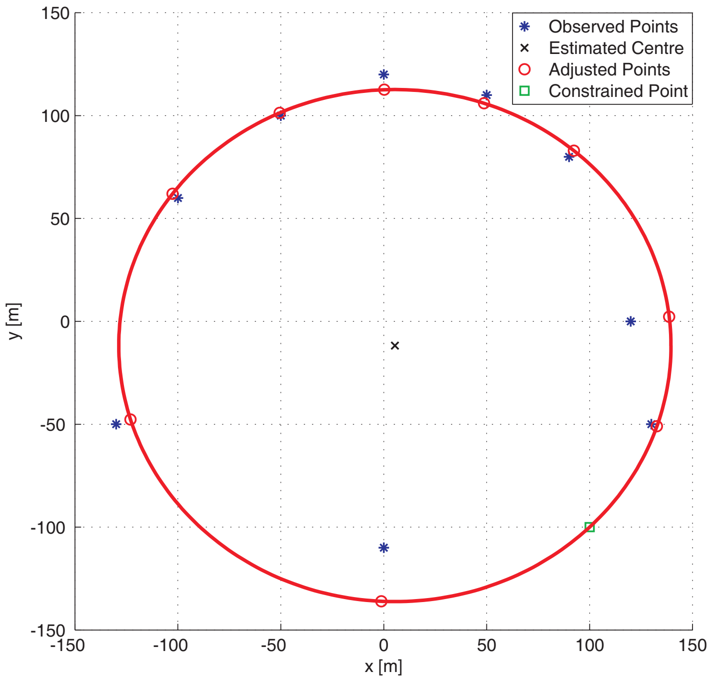


Figure 5.39: Ellipse fit (mixed model), point restriction.

6 Statistics

6.1 Expectation of sum of squared residuals

$$\mathrm{E}\left\{\underline{\hat{e}}^{\mathsf{T}} Q_y^{-1} \underline{\hat{e}}\right\}$$

Note: $\underline{e}^{\mathsf{T}} Q_y^{-1} \underline{e}$ is the quantity to be minimized.

$$\underset{1\times m}{\underline{\hat{e}}^{\mathsf{T}}}\ \underset{m\times m}{Q_y^{-1}}\ \underset{m\times 1}{\underline{\hat{e}}} = \sum_{i=1}^{m}\sum_{j=1}^{m}(P_y)_{ij}\underline{\hat{e}}_i\underline{\hat{e}}_j$$

$$\begin{aligned}
\Longrightarrow \mathrm{E}\left\{\underline{\hat{e}}^{\mathsf{T}} Q_y^{-1} \underline{\hat{e}}\right\} &= \sum_{i=1}^{m}\sum_{j=1}^{m}(P_y)_{ij}\,\mathrm{E}\left\{\underline{\hat{e}}_i\underline{\hat{e}}_j\right\} \\
&= \sum_{i=1}^{m}\sum_{j=1}^{m}(P_y)_{ij}(Q_{\hat{e}})_{ij} \\
&= \sum_{i=1}^{m}\sum_{j=1}^{m}(P_y)_{ij}(Q_{\hat{e}})_{ji} = \sum_{i}\left[P_y Q_{\hat{e}}\right]_{ii} \\
&= \operatorname{trace}(P_y Q_{\hat{e}}) \\
&= \operatorname{trace}(P_y(Q_y - Q_{\hat{y}})) \\
&= \operatorname{trace}(I_m - P_y Q_{\hat{y}}) \\
&= m - \operatorname{trace} P_y Q_{\hat{y}} \\
\operatorname{trace} P_y Q_{\hat{y}} &= \operatorname{trace} Q_{\hat{y}} P_y \\
&= \operatorname{trace} A Q_{\hat{x}} A^{\mathsf{T}} P_y \\
&= \operatorname{trace} A(A^{\mathsf{T}} P_y A)^{-1} A^{\mathsf{T}} P_y \\
&= \operatorname{trace} P_A
\end{aligned}$$

Linear algebra:

$$\operatorname{trace} X = \text{ sum of eigenvalues of } X$$

Q: Eigenvalues of a projector?

$$P_A z = \lambda z \qquad \text{(special) eigenvalue problem}$$

$$\left.\begin{array}{r} P_A P_A z = P_A z = \lambda z \\ P_A P_A z = \lambda P_A z = \lambda^2 z \end{array}\right\} \quad \lambda^2 z = \lambda z \quad \Longrightarrow \quad \lambda(\lambda-1)z = 0 \quad \Longrightarrow \quad \lambda = \begin{cases} 0 \\ 1 \end{cases}$$

$$\Longrightarrow \operatorname{trace} P_A = \text{number of eigenvalues } 1$$

Q: How many eigenvalues $\lambda = 1$?

A:

$$\dim \mathcal{R}(A) = n$$

$$\mathrm{E}\left\{\underline{\hat{e}}^{\mathsf{T}} P_y \underline{\hat{e}}\right\} = m - n \quad (= r \quad \text{redundancy})$$

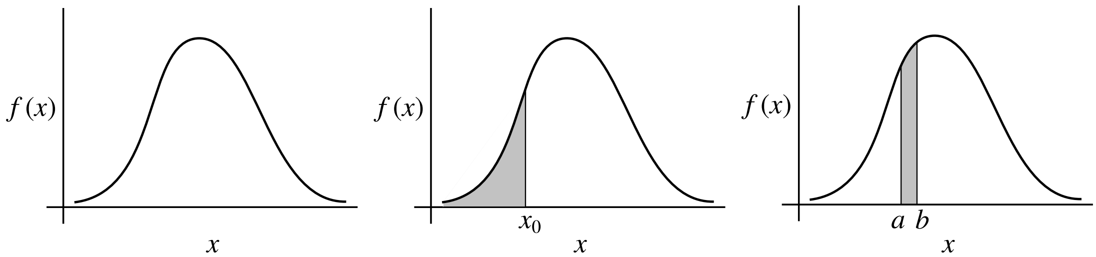
6.2 Basics

Random variable: $\underline{x}$

Realization: x

Probability density function (PDF)

Wahrschein-lichkeitsdichte



(a) Probability density function.
(b) Probability calculations by integrating over the PDF.
(c) Intervall.

Figure 6.1

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$$

Note: not necessarily normal distribution.

$$\mathrm{E}\left\{\underline{x}\right\} =: \mu_x = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x$$

$$\mathrm{D}\left\{\underline{x}\right\} =: \sigma_x^2 = \int_{-\infty}^{\infty} (x - \mu_x)^2 f(x)\,\mathrm{d}x = \mathrm{E}\left\{(x - \mu_x)^2\right\}$$

Probability calculations by integrating over the PDF.

$$P(\underline{x} < x_0) = \int_{-\infty}^{x_0} f(x)\,\mathrm{d}x$$

Verteilungsfunktion

Cumulative distribution or density function (CDF)

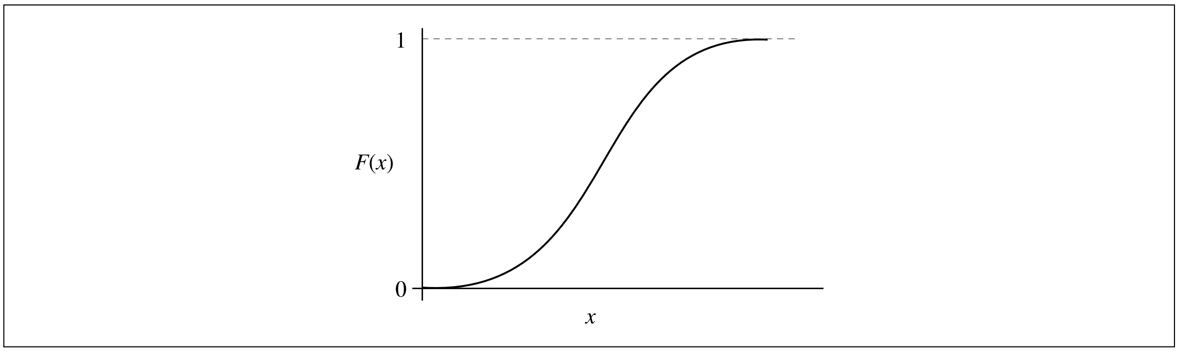


Figure 6.2: Cumulative distribution or density function.

$$F(x) = \int_{-\infty}^{x} f(y)\,\mathrm{d}y = P(\underline{x} < x)$$

e. g.

$$\begin{aligned} P(a \le \underline{x} \le b) &= \int_{a}^{b} f(x)\,\mathrm{d}x = \int_{-\infty}^{b} f(x)\,\mathrm{d}x - \int_{-\infty}^{a} f(x)\,\mathrm{d}x \\ &= F(b) - F(a) \end{aligned}$$

6.3 Hypotheses

Assumption or statement which can be statistically tested.

$$H : \underline{x} \sim f(x)$$

Sicherheitswahrscheinlichkeit

Irrtumswahrscheinlichkeit

Konfidenzbereich (Annahmebereich)

kritischer Bereich (Ablehnungs-, Verwerfungsbereich)

Assumption: $\underline{x}$ is distributed with given $f(x)$.

$$\begin{aligned} P(a \le \underline{x} \le b) = 1 - \alpha &= \textit{confidence level} \\ P(\underline{x} \notin [a; b]) = \alpha &= \textit{significance level} \\ [a; b] &= \textit{confidence region} \\ [-\infty; a] \cup [b; \infty] &= \textit{critical region} \end{aligned}$$

Now: given a realization x of $\underline{x}$. If $a \le x \le b$, there is no reason to reject the hypothesis, otherwise: reject hypothesis. E. g.

$$\hat{\underline{e}} = P_a^{\perp} \underline{y}$$

$$Q_{\hat{e}} = P_a^{\perp} Q_y = Q_y - Q_{\hat{y}}$$

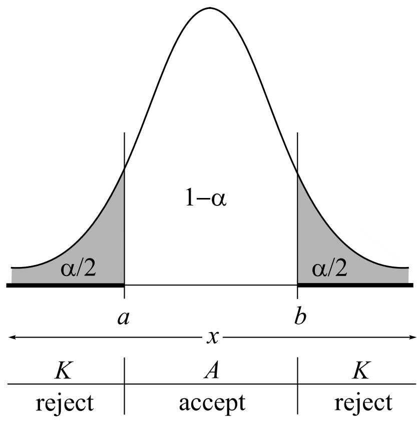


Figure 6.3: Confidence and significance level.

Example Normal distribution

define a, b: determine α

$$
\begin{aligned}
P(\mu - \sigma \leq \underline{x} \leq \mu + \sigma) &= 68.3\% &\implies& \quad \alpha = 0.317 \\
P(\mu - 2\sigma \leq \underline{x} \leq \mu + 2\sigma) &= 95.5\% &\implies& \quad \alpha = 0.045 \\
P(\mu - 3\sigma \leq \underline{x} \leq \mu + 3\sigma) &= 99.7\% &\implies& \quad \alpha = 0.003
\end{aligned}
$$

Matlab: `normpdf`

$$1 - \alpha = F(b) - F(\alpha) = F(\mu + k\sigma) - F(\mu - k\sigma)$$

k = *critical value* kritischer Wert

define α: determine a, b

$$
\begin{aligned}
P(\mu - 1.96\sigma \leq \underline{x} \leq \mu + 1.69\sigma) &= 95\% &\impliedby& \quad \alpha = 0.05 \quad (\approx 2\sigma) \\
P(\mu - 2.58\sigma \leq \underline{x} \leq \mu + 2.58\sigma) &= 99\% &\implies& \quad \alpha = 0.01 \\
P(\mu - 3.29\sigma \leq \underline{x} \leq \mu + 3.29\sigma) &= 99.9\% &\implies& \quad \alpha = 0.001
\end{aligned}
$$

Matlab: `norminv`

Rejection of hypothesis

$\implies$ an alternative hypothesis must hold

$$
\begin{aligned}
H_0 : &\quad \underline{x} \sim f_0(x) \quad \text{null-hypothesis} \\
H_a : &\quad \underline{x} \sim f_a(x) \quad \text{alternative hypothesis}
\end{aligned}
$$

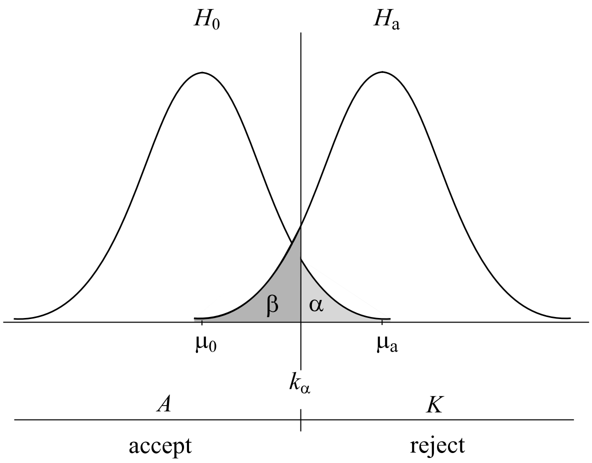


Figure 6.4: Accept or reject hypothesis?

	H_0 true	H_0 false
$x \in K$ $\Longrightarrow$ reject H_0	wrong $\Longrightarrow$ type I error (false alarm) $P(x \in K \vert H_0) = \alpha$	OK
$x \notin K$ $\Longrightarrow$ accept H_0	OK	wrong $\Longrightarrow$ type II error (failed alarm) $P(x \notin K \vert H_a) = \beta$

$$\alpha = \text{ level of significance of test } = \text{ size of test}$$

Testgüte

$$\gamma = 1 - \beta = \textit{power of test}$$

6.4 Distributions

Standard normal distribution (univariate)

$$\underline{x} \sim N(0,1), \qquad f(\underline{x}) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\underline{x}^2},$$

$$\mathrm{E}\left\{\underline{x}\right\} = 0,$$

$$\mathrm{D}\left\{\underline{x}\right\} = \mathrm{E}\left\{\underline{x}^2\right\} = 1 \quad \longleftarrow \quad \underline{x}^2 \sim \chi^2(1,0), \qquad \mathrm{E}\left\{\underline{x}^2\right\} = 1\,.$$

Standard normal (multivariate) → χ^2-distribution

$$\underset{k\text{-vector}}{\underline{x}} \sim N(\underset{k\text{-vector}}{0}, 1), \qquad f(\underline{x}) = \frac{1}{(2\pi)^{\frac{k}{2}}} \exp\left(-\frac{1}{2}\underline{x}^{\mathsf{T}}\underline{x}\right),$$

$$\mathrm{E}\left\{\underset{k\text{-vector}}{\underline{x}}\right\} = \underset{k\text{-vector}}{0},$$

$$\mathrm{D}\left\{\underline{x}\right\} = \mathrm{E}\left\{\underline{x}^2\right\} = 1 \quad \longleftarrow \quad \underline{x}^2 \sim \chi^2(1,0), \qquad \mathrm{E}\left\{\underline{x}\underline{x}^{\mathsf{T}}\right\} = I,$$

$$\underline{x}^{\mathsf{T}}\underline{x} = \underline{x}_1^2 + \underline{x}_2^2 + \ldots + \underline{x}_k^2 \quad \sim \chi^2(k,0),$$

$$\mathrm{E}\left\{\underline{x}^{\mathsf{T}}\underline{x}\right\} = \mathrm{E}\left\{\underline{x}_1^2\right\} + \ldots + \mathrm{E}\left\{\underline{x}_k^2\right\} = k\,.$$

Non-standard normal → central χ^2-distribution

$$\underline{x} \sim N(0, Q_x), \quad Q_x = \begin{pmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_k^2 \end{pmatrix},$$

$$\underline{x}_i \sim N(0, \sigma_i^2), \qquad f(\underline{x}_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2}\frac{\underline{x}_i^2}{\sigma_i^2}\right),$$

$$\underline{y}_i = \frac{\underline{x}_i}{\sigma_i} \sim N(0,1),$$

$$\underline{x}^{\mathsf{T}}Q_x^{-1}\underline{x} = \frac{\underline{x}_1^2}{\sigma_1^2} + \frac{\underline{x}_2^2}{\sigma_2^2} + \ldots + \frac{\underline{x}_k^2}{\sigma_k^2} \sim \chi^2(k,0) \quad \Longrightarrow \quad \mathrm{E}\left\{\underline{x}^{\mathsf{T}}Q_x^{-1}\underline{x}\right\} = k,$$

$$f(\underline{x}) = \frac{1}{(2\pi)^{\frac{k}{2}}(\det Q_x)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\underline{x}^{\mathsf{T}}Q_x^{-1}\underline{x}\right).$$

The same is true when

$$\underline{x} \sim N(0, Q_x) \text{ with } Q_x \text{ full matrix}\,.$$

Non-standard normal → non-central χ^2-distribution

$$\underline{x} \sim N(\mu, I) \quad \Longrightarrow \quad \underline{x}^{\mathsf{T}}\underline{x} \sim \chi^2(k,\lambda),$$

$$\begin{aligned} \mathrm{E}\left\{\underline{x}\right\} &= \mu \\ \mathrm{E}\left\{\underline{x}^{\mathsf{T}}\underline{x}\right\} &= k + \lambda; \qquad \mu^{\mathsf{T}}\mu = \text{ non-centrality parameter} \\ &= \mu_1^2 + \mu_2^2 + \ldots + \mu_k^2 \end{aligned}$$

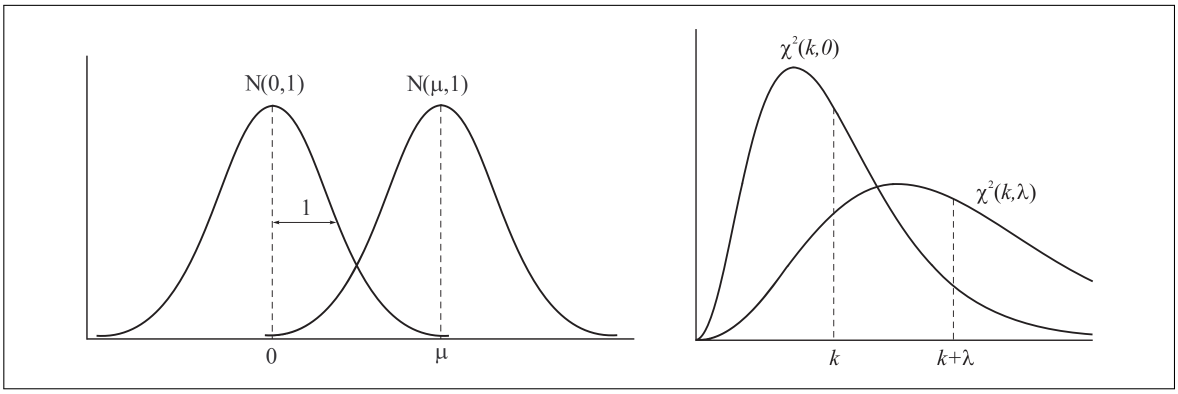


Figure 6.5: Central/Non-central normal and χ^2-distribution

General case

$$
\begin{gathered}
\underline{x} \sim N(\mu, Q_x), \\
f(\underline{x}) = \frac{1}{(2\pi)^{\frac{k}{2}} (\det Q_x)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\underline{x} - \mu) Q_x^{-1} (\underline{x} - \mu)\right), \\
\mathrm{E}\left\{\underline{x}\right\} = \mu, \qquad \mathrm{D}\left\{\underline{x}\right\} = Q_x, \\
\mathrm{E}\left\{\underline{x}^{\mathsf{T}} Q_x^{-1} \underline{x}\right\} = k + \lambda \qquad = \mu^{\mathsf{T}} Q_x^{-1} \mu\,.
\end{gathered}
$$

7 Statistical Testing

7.1 Global model test: a first approach

Statistics of estimated residuals

$$\begin{aligned}
\underline{\hat{e}} &= \underline{y} - \underline{\hat{y}} \\
&= P_A^{\perp}\underline{y} \\
\mathrm{E}\left\{\underline{\hat{e}}\right\} &= 0, \qquad \underline{\hat{e}} \sim N(0, Q_{\hat{e}}) \\
\mathrm{D}\left\{\underline{\hat{e}}\right\} &= Q_{\hat{e}} \\
&= Q_y - Q_{\hat{y}} \\
&= P_A^{\perp} Q_y (P_A^{\perp})^{\mathsf{T}}
\end{aligned}$$

Question: $\underline{\hat{e}}^{\mathsf{T}} Q_{\hat{e}}^{-1} \underline{\hat{e}} \sim \chi^2(m, 0)$ and thus $\mathrm{E}\left\{\underline{\hat{e}}^{\mathsf{T}} Q_{\hat{e}}^{-1} \underline{\hat{e}}\right\} = m$?

No, because $Q_{\hat{e}}$ is singular and therefore not invertible. However, in 6.1:

$$\mathrm{E}\left\{\underline{\hat{e}}^{\mathsf{T}} Q_y^{-1} \underline{\hat{e}}\right\} = \mathrm{trace}(Q_y^{-1} \underbrace{\mathrm{E}\left\{\underline{\hat{e}}\underline{\hat{e}}^{\mathsf{T}}\right\}}_{Q_{\hat{e}}}) = \mathrm{trace}(Q_y^{-1}(Q_y - Q_{\hat{y}})) = m - n$$

Test statistic

As residuals tell us something about the mismatch between data and model, they will be the basis for our testing. In particular the sum of squared estimated residuals will be used as our test statistic $\underline{T}$:

$$\begin{aligned}
\underline{T} &= \underline{\hat{e}}^{\mathsf{T}} Q_y^{-1} \underline{\hat{e}} \sim \chi^2(m - n, 0) \\
\mathrm{E}\left\{\underline{T}\right\} &= m - n
\end{aligned}$$

Thus, we have a test statistic and we know its distribution. This is the starting point for global model testing.

$T > k_\alpha$: reject H_0

In case T—the realization of $\underline{T}$—is larger than a chosen critical value (based on α), the null hypothesis H_0 should be rejected. At this point, we haven't formulated an alternative hypothesis H_a yet. The rejection may be due to:

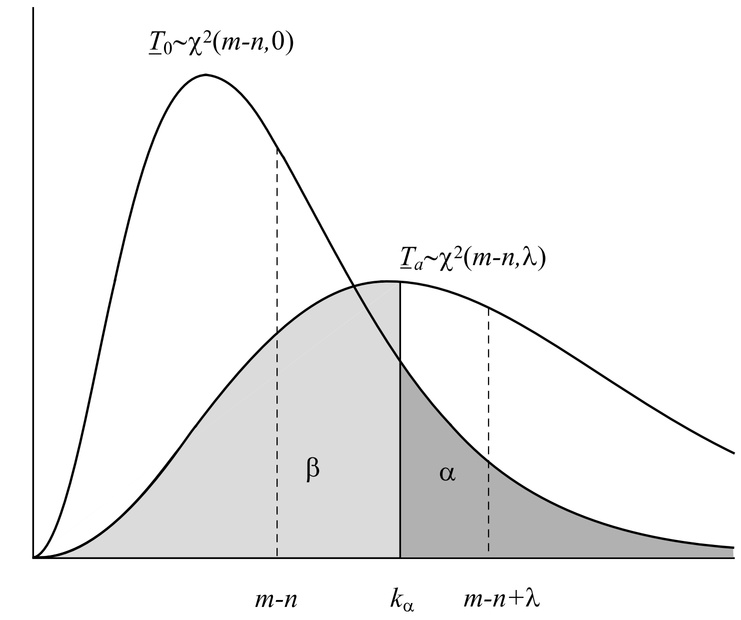


Figure 7.1: Distribution of the test statistic $\underline{T}$ under the null and alternative hypotheses. (Non-centrality parameter λ to be explained later)

- error in the (deterministic) observation model A,
- measurement error: $\mathrm{E}\left\{\underline{e}\right\} \neq 0$,
- wrong assumptions in the stochastic model: $\mathrm{D}\left\{\underline{e}\right\} \nsim Q_y$.

Variance of unit weight

A possible error in the stochastic model would be a wrong scale factor. Let us write $Q_y = \sigma^2 Q$ and see how an unknown variance factor σ^2 propagates through the various estimates:

$$
\begin{aligned}
\underline{\hat{x}} &= (A^{\mathsf{T}} Q_y^{-1} A)^{-1} A^{\mathsf{T}} Q_y^{-1} \underline{y} \\
Q_{\hat{x}} &= (A^{\mathsf{T}} Q_y^{-1} A)^{-1} \\
&\quad \Big\downarrow \begin{array}{l} Q_y = \sigma^2 Q \\ P_y = Q_y^{-1} = \sigma^{-2} Q^{-1} \end{array} \\
\underline{\hat{x}} &= (A^{\mathsf{T}} \sigma^{-2} Q^{-1} A)^{-1} A^{\mathsf{T}} \sigma^{-2} Q^{-1} \underline{y} \\
&= \sigma^2 (A^{\mathsf{T}} Q^{-1} A)^{-1} A^{\mathsf{T}} \sigma^{-2} Q^{-1} \underline{y} \\
&= (A^{\mathsf{T}} Q^{-1} A)^{-1} A^{\mathsf{T}} Q^{-1} \underline{y} \qquad \Longrightarrow \text{independent on } \sigma^2 \\
Q_{\hat{x}} &= \sigma^2 (A^{\mathsf{T}} Q^{-1} A)^{-1} \qquad \Longrightarrow \text{depending on } \sigma^2
\end{aligned}
$$

Thus, the estimate $\hat{x}$ is independent of the variance factor and therefore insensitive to stochastic model errors. However, the covariance matrix $Q_{\hat{x}}$ is scaled by the variance factor. This is also true for functions $\hat{f} = F(\hat{x})$: while $\hat{f}$ is not influenced by σ^2, its covariance-matrix $Q_{\hat{f}}$ is changed accordingly. How about the test statistic $\underline{T}$?

$$\begin{aligned} & \mathrm{E}\left\{\underline{\hat{e}}^\mathsf{T} Q_y^{-1} \underline{\hat{e}}\right\} = \mathrm{E}\left\{\sigma^{-2} \underline{\hat{e}}^\mathsf{T} Q^{-1} \underline{\hat{e}}\right\} = m - n \\ \Longrightarrow\ & \mathrm{E}\left\{\underline{\hat{e}}^\mathsf{T} Q^{-1} \underline{\hat{e}}\right\} = \sigma^2 (m-n) \end{aligned}$$

Alternative test statistic

This leads to a new test statistic:

$$\underline{\hat{\sigma}}^2 = \frac{\underline{\hat{e}}^\mathsf{T} Q^{-1} \underline{\hat{e}}}{m-n} \Longrightarrow \mathrm{E}\left\{\underline{\hat{\sigma}}^2\right\} = \sigma^2 \,,$$

which shows that $\underline{\hat{\sigma}}^2$ is an *unbiased estimate* of σ^2. unverzerrte Schätzung

If we consider Q as the a priori variance-covariance matrix, then $\hat{Q}_y = \underline{\hat{\sigma}}^2$ is the a posteriori one.

Now consider the ratio between a posteriori and a priori variance as an alternative test statistic:

$$\frac{\underline{\hat{\sigma}}^2}{\sigma^2} = \frac{\underline{\hat{e}}^\mathsf{T} \sigma^{-2} Q^{-1} \underline{\hat{e}}}{m-n} = \frac{\underline{\hat{e}}^\mathsf{T} Q_y^{-1} \underline{\hat{e}}}{m-n} \sim \frac{\chi^2(m-n, 0)}{m-n} = F(m-n, \infty, 0)$$

The ratio has a so-called Fisher distribution.

$$\mathrm{E}\left\{\frac{\underline{\hat{\sigma}}^2}{\sigma^2}\right\} = 1$$

7.2 Testing procedure

Null hypothesis and alternative hypothesis

If the null hypothesis is described by $\mathrm{E}\left\{\underline{y}\right\} = Ax$, $\mathrm{D}\left\{\underline{y}\right\} = Q_y$, and if we assume that our stochastic model is correct, then we formulate an alternative hypothesis by augmenting the model. We will add q new parameters ∇ (which is not an operator here). Consequently we will need a design matrix $\underset{m \times q}{C}$ for ∇.

H_0	H_a
$\mathrm{E}\left\{\underline{y}\right\} = Ax; \quad \mathrm{D}\left\{\underline{y}\right\} = Q_y$	$\mathrm{E}\left\{\underline{y}\right\} = Ax + C\nabla; \quad \mathrm{D}\left\{\underline{y}\right\} = Q_y$
	$= \begin{pmatrix} A & C \end{pmatrix} \begin{pmatrix} x \\ \nabla \end{pmatrix}$
$\downarrow$	$\downarrow$
$\hat{\underline{x}}_0$	$\hat{\underline{x}}_\mathrm{a}, \hat{\underline{\nabla}}$
$\downarrow$	$\downarrow$
$\hat{\underline{y}}_0 = A\hat{\underline{x}}_0$	$\hat{\underline{y}}_\mathrm{a} = A\hat{\underline{x}}_\mathrm{a} + C\hat{\underline{\nabla}}$
$\downarrow$	$\downarrow$
$\hat{\underline{e}}_0$	$\hat{\underline{e}}_\mathrm{a}$
$\downarrow$	$\downarrow$
$\hat{\underline{e}}_0^\mathsf{T} Q_y^{-1} \hat{\underline{e}}_0 \sim \chi^2(m-n)$	$\hat{\underline{e}}_\mathrm{a}^\mathsf{T} Q_y^{-1} \hat{\underline{e}}_\mathrm{a} \sim \chi^2(m-n-q)$

H_a more parameters $\Longrightarrow$ sum of squared residuals smaller

$$\hat{\underline{e}}_\mathrm{a}^\mathsf{T} Q_y^{-1} \hat{\underline{e}}_\mathrm{a} < \hat{\underline{e}}_0^\mathsf{T} Q_y^{-1} \hat{\underline{e}}_0$$

$\Longrightarrow$ difference, which is a measure of improvement as test statistic:

$$\underline{T} = \hat{\underline{e}}_0^\mathsf{T} Q_y^{-1} \hat{\underline{e}}_0 - \hat{\underline{e}}_\mathrm{a}^\mathsf{T} Q_y^{-1} \hat{\underline{e}}_\mathrm{a} \qquad \text{First version}$$

How is it distributed?

$$H_0 : \underline{T} \sim \chi^2(q, 0) \quad \text{and} \quad H_\mathrm{a} : \underline{T} \sim \chi^2(q, \lambda)$$

Geometry of H_0 und H_a

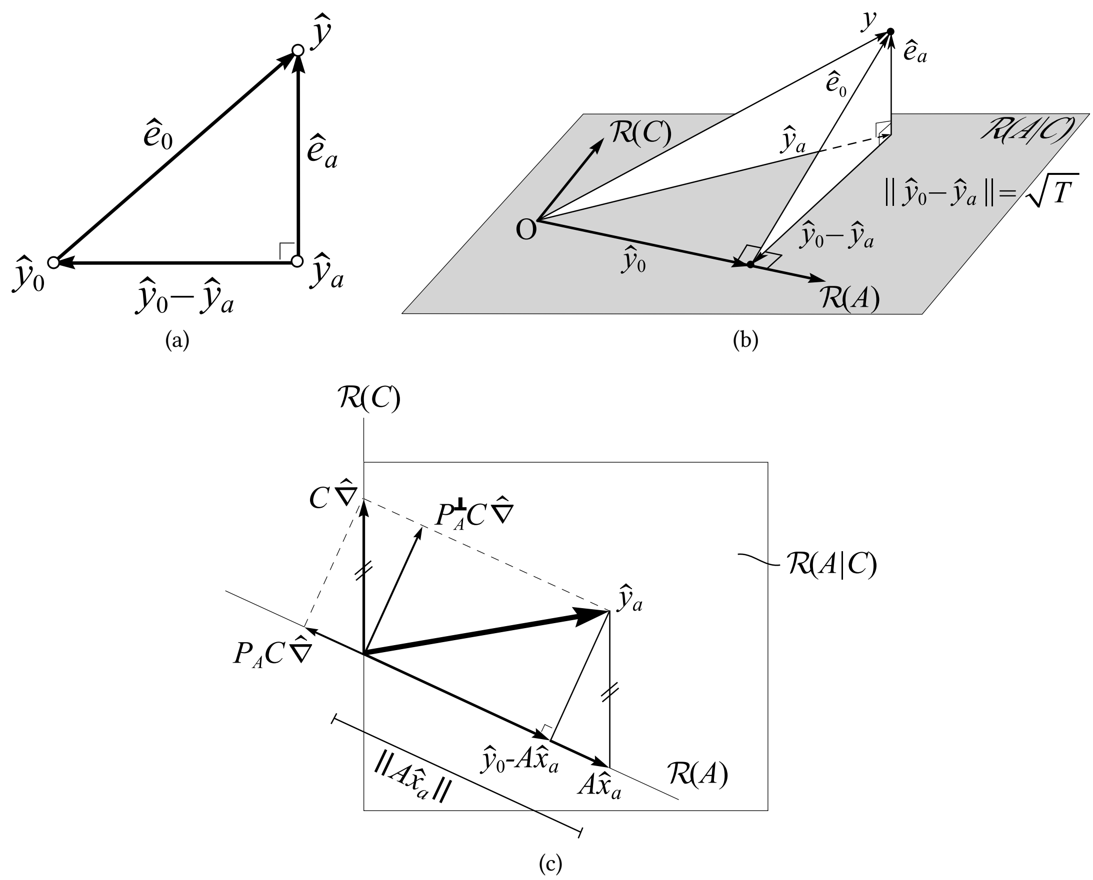


Figure 7.2: Hypotheses H_0 and H_a

$$
\begin{aligned}
\underline{\hat{y}}_\mathrm{a} &= A\underline{\hat{x}}_\mathrm{a} + C\underline{\hat{\nabla}} \\
&= \underbrace{A\underline{\hat{x}}_\mathrm{a} + P_A C\underline{\hat{\nabla}}}_{\underline{\hat{y}}_0} + P_A^{\perp} C\underline{\hat{\nabla}} \\
\Longrightarrow \underline{\hat{y}}_\mathrm{a} - \underline{\hat{y}}_0 &= P_A^{\perp} C\underline{\hat{\nabla}} \\
\Longrightarrow \underline{T} &= P_A^{\perp} C\underline{\hat{\nabla}}^{\mathsf{T}} Q_y^{-1} P_A^{\perp} C\underline{\hat{\nabla}} = \underline{\hat{\nabla}}^{\mathsf{T}} C^{\mathsf{T}} \underbrace{P_A^{\perp\mathsf{T}} Q_y^{-1} P_A^{\perp}}_{=Q_y^{-1}P_A^{\perp} = Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1}} C\underline{\hat{\nabla}} \\
&= \underline{\hat{\nabla}}^{\mathsf{T}} C^{\mathsf{T}} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C\underline{\hat{\nabla}} \qquad \text{Second version} \\
\underline{T} &= \underline{\hat{e}}_0^{\mathsf{T}} Q_y^{-1} \underline{\hat{e}}_0 - \underline{\hat{e}}_\mathrm{a}^{\mathsf{T}} Q_y^{-1} \underline{\hat{e}}_\mathrm{a} \\
&= (\underline{\hat{y}}_0 - \underline{\hat{y}}_\mathrm{a})^{\mathsf{T}} Q_y^{-1} (\underline{\hat{y}}_0 - \underline{\hat{y}}_\mathrm{a}) \qquad \text{Third version} \\
&= \underline{\hat{\nabla}}^{\mathsf{T}} C^{\mathsf{T}} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C\underline{\hat{\nabla}}
\end{aligned}
$$

All versions of $\underline{T}$ require adjustment under H_a

$$(\underline{\hat{e}}_\mathrm{a}, \underline{\hat{y}}_\mathrm{a}, \underline{\hat{\nabla}})$$

Now: Version only with $\underline{\hat{e}}_0$ and C

Normal equations under H_0, H_a

$$
\begin{aligned}
H_0 \quad &: \quad A^\mathsf{T} Q_y^{-1} A \underline{\hat{x}}_0 = A^\mathsf{T} Q_y^{-1} \underline{y} \\
H_\mathrm{a} \quad &: \quad \begin{pmatrix} A^\mathsf{T} \\ C^\mathsf{T} \end{pmatrix} Q_y^{-1} \begin{pmatrix} A & C \end{pmatrix} \begin{pmatrix} \underline{\hat{x}} \\ \underline{\hat{\nabla}} \end{pmatrix} = \begin{pmatrix} A^\mathsf{T} \\ C^\mathsf{T} \end{pmatrix} Q_y^{-1} \underline{y} \\
&\iff \begin{pmatrix} \underset{n \times n}{A^\mathsf{T} Q_y^{-1} A} & \underset{n \times q}{A^\mathsf{T} Q_y^{-1} C} \\ \underset{q \times n}{C^\mathsf{T} Q_y^{-1} A} & \underset{q \times q}{C^\mathsf{T} Q_y^{-1} C} \end{pmatrix} \begin{pmatrix} \underline{\hat{x}}_\mathrm{a} \\ \underline{\hat{\nabla}} \end{pmatrix} = \begin{pmatrix} A^\mathsf{T} Q_y^{-1} \underline{y} \\ C^\mathsf{T} Q_y^{-1} \underline{y} \end{pmatrix}
\end{aligned}
$$

1. row: solve for $\underline{\hat{x}}_\mathrm{a}$

$$
\begin{aligned}
& A^\mathsf{T} Q_y^{-1} A \underline{\hat{x}}_\mathrm{a} + A^\mathsf{T} Q_y^{-1} C \underline{\hat{\nabla}} = A^\mathsf{T} Q_y^{-1} A \underline{\hat{x}}_0 \\
\Longrightarrow\ & \underline{\hat{x}}_\mathrm{a} = \underline{\hat{x}}_0 - (A^\mathsf{T} Q_y^{-1} A)^{-1} A^\mathsf{T} Q_y^{-1} C \underline{\hat{\nabla}} \\
\Longrightarrow\ & A \underline{\hat{x}}_\mathrm{a} = A \underline{\hat{x}}_0 - P_A C \underline{\hat{\nabla}} \\
\Longrightarrow\ & A \underline{\hat{x}}_\mathrm{a} + C \underline{\hat{\nabla}} = A \underline{\hat{x}}_0 + (I - P_A) C \underline{\hat{\nabla}} \\
\Longrightarrow\ & \underline{\hat{y}}_\mathrm{a} = \underline{\hat{y}}_0 + P_A^\perp C \underline{\hat{\nabla}}
\end{aligned}
$$

2. row: substitute $\underline{\hat{x}}_\mathrm{a}$ and solve for $\underline{\hat{\nabla}}$ $\longrightarrow$ laborious derivation!

Result:

$$\underline{\hat{\nabla}} = (C^\mathsf{T} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C)^{-1} C^\mathsf{T} Q_y^{-1} \underline{\hat{e}}_0$$

Substitute $\underline{\hat{\nabla}}$ in second version of $\underline{T}$ $\longrightarrow$ Fourth version

$$
\begin{aligned}
\underline{T} &= \underline{\hat{e}}_0^\mathsf{T} Q_y^{-1} C (\ldots)^{-1} (\ldots) (\ldots)^{-1} C^\mathsf{T} Q_y^{-1} \underline{\hat{e}}_0 \\
&= \underline{\hat{e}}_0^\mathsf{T} Q_y^{-1} C (C^\mathsf{T} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C)^{-1} C^\mathsf{T} Q_y^{-1} \underline{\hat{e}}_0
\end{aligned}
$$

Distribution of $\underline{T}$

Transformation of variables

$$
\begin{aligned}
\underset{q \times 1}{\underline{z}} &= \underset{q \times m}{C^\mathsf{T}} \ \underset{m \times m}{Q_y^{-1}} \ \underset{m \times 1}{\underline{\hat{e}}_0} \\
Q_z &= C^\mathsf{T} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C \\
\underline{\hat{\nabla}} &= Q_z^{-1} \underline{z} \Longrightarrow \underline{z} = Q_z \underline{\hat{\nabla}} \\
\underline{T} &= \underline{z}^\mathsf{T} Q_z^{-1} \underline{z} \sim \chi_q^2
\end{aligned}
$$

H_0	H_a
$\underline{z} \sim N(0, Q_z)$	$\underline{z} \sim N(Q_z\hat{\underline{\nabla}}, Q_z)$
$\underline{T} \sim \chi^2(q, 0)$	$\underline{T} \sim \chi^2(q, \lambda)$
	$\lambda = \nabla^{\mathsf{T}} Q_z Q_z^{-1} Q_z \nabla = \nabla^{\mathsf{T}} C^{\mathsf{T}} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C \nabla$

Summary

Test quantity $\underline{T} = \hat{\underline{e}}_0^{\mathsf{T}} Q_y^{-1} \hat{\underline{e}}_0 - \hat{\underline{e}}_a^{\mathsf{T}} Q_y^{-1} \hat{\underline{e}}_a$ exhibits that H_0 has to be rejected in favor of H_a: Model $\mathrm{E}\{y\} = Ax$ is not suitable. In case H_0 is true $\underline{T}$ is (central) χ^2-distributed with q degrees of freedom, $\underline{T} \sim \chi^2_{q,0}$, otherwise $\underline{T} \sim \chi^2_{q,\lambda}$ with λ being the non-centrality parameter $\lambda = \nabla^{\mathsf{T}} C^{\mathsf{T}} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C \nabla$.

Five alternative versions for $\underline{T}$

$$
\begin{aligned}
&(1) \quad \hat{\underline{e}}_0^{\mathsf{T}} Q_y^{-1} \hat{\underline{e}}_0 - \hat{\underline{e}}_a^{\mathsf{T}} Q_y^{-1} \hat{\underline{e}}_a \\
&(2) \quad (\hat{\underline{y}}_0 - \hat{\underline{y}}_a)^{\mathsf{T}} Q_y^{-1} (\hat{\underline{y}}_0 - \hat{\underline{y}}_a) \\
&(3) \quad \hat{\underline{\nabla}}^{\mathsf{T}} C^{\mathsf{T}} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C \hat{\underline{\nabla}} \\
&(4) \quad \hat{\underline{e}}_0^{\mathsf{T}} Q_y^{-1} C \left(C^{\mathsf{T}} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C \right)^{-1} C^{\mathsf{T}} Q_y^{-1} \hat{\underline{e}}_0 \\
&(5) \quad \underline{z}^{\mathsf{T}} Q_z^{-1} \underline{z}; \quad \underline{z} := C^{\mathsf{T}} Q_y^{-1} \hat{\underline{e}}_0; \quad Q_z = C^{\mathsf{T}} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C
\end{aligned}
$$

Versions (1)–(3) explicitely involve the computation of H_a while cases (4) and (5) require only $\hat{\underline{e}}_0$ and some C.

For the reason that

$$
\begin{aligned}
\underline{z} &\sim N(0, Q_z) \qquad \text{under } H_0 \\
\underline{z} &\sim N(Q_z\nabla, Q_z) \qquad \text{under } H_a
\end{aligned}
$$

test quantity $\underline{T}$ is distributed

$$
\begin{aligned}
\underline{T} &\sim \chi^2(q, 0) \qquad \text{under } H_0 \\
\underline{T} &\sim \chi^2(q, \lambda), \quad \lambda = \nabla^{\mathsf{T}} Q_z \nabla \qquad \text{under } H_a
\end{aligned}
$$

Question: How is the minimal/maximal number of additional parameters ∇?

Answer: Total number of all parameters x and ∇ is $n + q$ which must not exceed number of observations $m \Longrightarrow$

$$
n + q \leq m \Longrightarrow 0 < q \leq m - n
$$

Case (i) $q = m - n$: global model test

$$
\begin{aligned}
\operatorname{rank}(A|C) &\overset{!}{=} n + q = n + (m-n) = m \\
\Longrightarrow &\quad o(A|C) = m \times n + q = m \times m \quad \text{“quadratic”} \\
\Longrightarrow &\quad \text{redundancy} = m - n - q = 0 \\
\Longrightarrow &\quad \underline{\hat{e}}_{\mathrm{a}} = 0 \\
\Longrightarrow &\quad \underline{\hat{y}}_{\mathrm{a}} = \underline{y} \\
\Longrightarrow &\quad \underline{T} = \underline{\hat{e}}_0^{\mathsf{T}} Q_y^{-1} \underline{\hat{e}}_0
\end{aligned}
$$

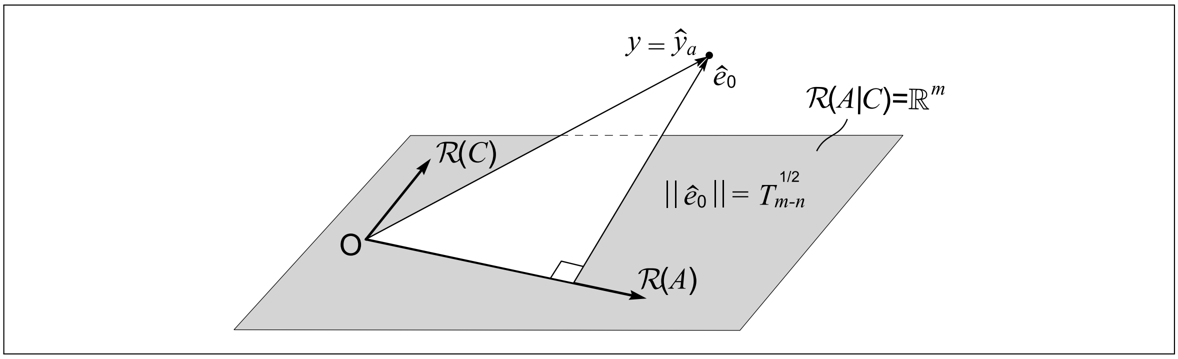


Figure 7.3: Test quantity $\underline{T}$

$$
H_0 : \mathrm{E}\left\{\underline{y}\right\} = Ax \quad \text{versus} \quad H_{\mathrm{a}} : \mathrm{E}\left\{\underline{y}\right\} \in \mathbf{R}^m
$$

$$
\begin{aligned}
&\underline{T} \sim \chi^2_{m-n,0} \\
&\underline{T} \sim \chi^2_{m-n,\lambda}, \quad \lambda = \nabla^{\mathsf{T}} Q_z \nabla
\end{aligned}
$$

For the reason that $\underline{\hat{e}}_{\mathrm{a}} = 0$, it is obviously not necessary to specify any matrix C. The test can always be carried out, that is why it is called overal model test or global test.

Case (ii) $q = 1$**: data snooping**

$\Longrightarrow C$ is an $m \times 1$-vector, ∇ a scalar

$$\begin{aligned}
\underline{T} &= \underline{\hat{e}}_0^\mathsf{T} Q_y^{-1} C \left(C^\mathsf{T} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C \right)^{-1} C^\mathsf{T} Q_y^{-1} \underline{\hat{e}}_0 \\
&= \frac{\left(\underline{\hat{e}}_0^\mathsf{T} Q_y^{-1} C \right)^2}{C^\mathsf{T} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C} \\
&= \frac{\underline{\hat{\nabla}}^2}{\left(C^\mathsf{T} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C \right)^{-1}} \\
&= \frac{\underline{\hat{\nabla}}^2}{\sigma_{\hat{\nabla}}^2} \\
\underline{\hat{\nabla}}^2 &= \frac{C^\mathsf{T} Q_y^{-1} \underline{\hat{e}}}{C^\mathsf{T} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C}
\end{aligned}$$

Important application:

Detection of a gross error (outlier, blunder) in the observations, which leads to a wrong model specification.

$$H_0 : \mathrm{E}\left\{ \underline{y} \right\} = Ax \qquad H_\mathrm{a} : \mathrm{E}\left\{ \underline{y} \right\} = Ax + C\nabla$$

$$C = \Big[0, 0, \ldots, \underbrace{1}_{i\text{-te Pos.}}, 0, \ldots, 0 \Big]^\mathsf{T}$$

Reject H_0 if $\underline{T} = \frac{\hat{\nabla}^2}{\sigma_{\hat{\nabla}}^2} > k_\alpha$ or if $\sqrt{\underline{T}} = \frac{\hat{\nabla}}{\sigma_{\hat{\nabla}}} < -\sqrt{k_\alpha}$ and $\sqrt{\underline{T}} = \frac{\hat{\nabla}}{\sigma_{\hat{\nabla}}} > \sqrt{k_\alpha}$ ($\underline{\hat{\nabla}}$ can be positive or negative)!

Should H_0 be rejected, observation y_i must be checked and corrected, discarded or even be remeasured. The test is performed for every $i = 1, \ldots, m$ if necessary in an iterative manner. The test is called data snooping. For a diagonal matrix Q_y we get

$$\sqrt{\underline{T}} = \frac{\underline{\hat{e}}_i}{\sigma_{\hat{e}_i}}$$

$$H_0 : \sqrt{\underline{T}} \sim N(0, 1) \qquad H_\mathrm{a} : \sqrt{\underline{T}} \sim N(\nabla\sqrt{\underline{T}}, 1)$$

$$\text{with } \nabla\sqrt{\underline{T}} = \sqrt{C^\mathsf{T} Q_y^{-1} Q_{\hat{e}} Q_y^{-1} C}\,\nabla$$

7.3 DIA-Testprinciple

DIA $\Longleftrightarrow$ Detection, Identification, Adaptation

Detection: Check the overall validity of H_0, perform the overall model test, answer the question whether or not we have generally to expect any model error, e. g. an outlier in the data, search for a possible model misspecification.

Identification: Perform data snooping in order to locate a possible gross error. Identify it in the collection of observations. Screen each individual observation for the presence of a blunder.

Adaptation: React to the outcomes of detection and identification step. Perform a corrective action in order to get the null hypothesis accepted. Repair, replace or discard the corrupted observation. Remeasure part of the observations or change the model in order to account for the identified model errors.

Question: How to ensure consistent testing parameters? How can we avoid the situation of a conflict between the overall model test in the detection step and individual test of the identification step?

Answer: Consistency is guaranteed if the probability of detecting an outlier under the alternative hypothesis with $q = 1$ (data snooping) is the same for the global test. Thus, both tests must use the same $\gamma = 1 - \beta$, which is called γ_0 here.

$$\lambda_0 = \lambda\left(\alpha, q = m - n, \gamma = \gamma_0\right) = \lambda\left(\alpha_1, q = 1, \gamma = \gamma_0\right)$$

$q = 1$:

$$\left.\begin{matrix} \gamma_0 = 1 - \beta_0 \\ \alpha_1 \end{matrix}\right\} \Longrightarrow \lambda(\longrightarrow \mu) = \lambda_0$$

$q = m - n$:

$$\left.\begin{matrix} \lambda_0 \\ \gamma_0 = 1 - \beta_0 \end{matrix}\right\} \Longrightarrow \alpha = \alpha_{m-n}$$

e. g.: $\alpha_1 = 1\%$ (usually small), $\beta_0 = 20\% \Longrightarrow \alpha_{m-n} \approx 30\%$

7.4 Internal reliability

innere Zuverlässigkeit

Which model error $C\nabla$ results in the power of test γ_0? Or the other way around: Which model error $C\nabla$ can be just detected with probability γ_0? This question is discussed in the framework of *internal reliability*.

Analysis λ

$$\begin{aligned}
\lambda &= \nabla^\mathsf{T} C^\mathsf{T} Q_y^{-1} Q_{\hat{e}} Q_y^{-1} C\nabla \\
Q_{\hat{e}} &= Q_y - Q_{\hat{y}} \\
&= Q_y - A\left(A^\mathsf{T} Q_y^{-1} A\right)^{-1} A^\mathsf{T} \\
\Longrightarrow \lambda &= \underbrace{\nabla^\mathsf{T} C^\mathsf{T}}_{(C\nabla)^\mathsf{T}} \left[Q_y^{-1} - Q_y^{-1} A\left(A^\mathsf{T} Q_y^{-1} A\right)^{-1} A^\mathsf{T} Q_y^{-1}\right] C\nabla
\end{aligned}$$

Question: For given fixed $\lambda = \lambda_0$, how can $C\nabla$ be manipulated?

Using Q_y:

$$\begin{aligned}\text{“better” observations} &\Longrightarrow Q_y \text{ smaller}\\ &\Longrightarrow Q_y^{-1} \text{ larger}\\ &\Longrightarrow C\nabla \text{ smaller (in order to keep } \lambda = \lambda_0 \text{ constant)}\end{aligned}$$

$\Longrightarrow$ the more precise the observations are, the smaller the model error $C\nabla$ may be. It will be detected with probability γ_0.

Using A:

- more observations $\Longrightarrow$ larger redundancy
  $\Longrightarrow$ for a constant $C\nabla$: λ increases or the other way around for a constant λ, $C\nabla$ gets smaller
- better network design, better configuration, improved distribution of observations, avoid bad geometries in resection problems
  $\Longrightarrow$ $C\nabla$ can be decreased

Minimum Detectable Bias (MDB)

kleinster aufdeckbarer Fehler

$$\underset{m\times 1}{\delta y} := \underset{m\times q}{C}\ \underset{q\times 1}{\nabla} = \mathrm{E}\left\{\underline{y}|H_\mathrm{a}\right\} - \mathrm{E}\left\{\underline{y}|H_0\right\}$$

δy describes the internal reliability; it measures the smallest possible error which can be detected with probability γ.

Question: How can ∇ be determined from $\lambda_0 = (C\nabla)^\mathsf{T} Q_y^{-1} Q_{\hat{e}} Q_y^{-1} C\nabla$?

Case $q = 1$ (datasnooping):

∇ is a scalar, $C = c_i$, $\delta y_i = c_i\nabla$

$$\begin{aligned}\lambda_0 &= c_i{}^\mathsf{T} Q_y^{-1} Q_{\hat{e}} Q_y^{-1} c_i \nabla^2\\ \Longrightarrow |\nabla_i| &= \sqrt{\frac{\lambda_0}{c_i{}^\mathsf{T} Q_y^{-1} Q_{\hat{e}} Q_y^{-1} c_i}}\end{aligned}$$

$|\nabla_i|$ = minimal detectable bias

Assumption: Q_y is diagonal

$$\underset{1\times m}{c_i^{\mathsf{T}} Q_y^{-1}} = \left[0, 0, \ldots, \sigma_{y_i}^{-2}, 0, \ldots\right]$$

$$\begin{aligned}
\Longrightarrow c_i^{\mathsf{T}} Q_y^{-1} Q_{\hat{e}} Q_y^{-1} c_i &= \sigma_{y_i}^{-2} \left[Q_{\hat{e}}\right]_{ii} \sigma_{y_i}^{-2} = \sigma_{y_i}^{-4} \sigma_{\underline{\hat{e}}_i}^{-2} \\
&= \sigma_{y_i}^{-4} \left(\sigma_{y_i}^2 - \sigma_{\hat{y}_i}^2\right) \\
&= \sigma_{y_i}^{-2} \left(1 - \frac{\sigma_{\hat{y}_i}^2}{\sigma_{y_i}^2}\right) \\
\Longrightarrow |\nabla_i| &= \frac{\sqrt{\lambda_0}}{\sqrt{\sigma_{y_i}^{-4} \left(\sigma_{y_i}^2 - \sigma_{\hat{y}_i}^2\right)}} \\
&= \sigma_{y_i} \frac{\sqrt{\lambda_0}}{\sqrt{1 - \frac{\sigma_{\hat{y}_i}^2}{\sigma_{y_i}^2}}} \\
&= \sigma_{y_i} \frac{\sqrt{\lambda_0}}{\sqrt{r_i}}
\end{aligned}$$

a) If no improvement through adjustment

$$\sigma_{\hat{y}_i} = \sigma_{y_i} \Longrightarrow |\nabla_i| = \infty$$

b) If $\sigma_{\hat{y}_i} \ll \sigma_{y_i} : |\nabla_i| = \sigma_{y_i} \sqrt{\lambda_0}$ is detectable

Local redundancy

$$r_i = 1 - \frac{\sigma_{\hat{y}_i}^2}{\sigma_{y_i}^2} = \text{local redundancy number}$$

y_i poorly controlled $\longrightarrow 0 \leq r_i \leq 1 \longleftarrow$ well controlled

$$\sum_{i=1}^{m} r_i = m - n$$

$$\begin{aligned}
r_i &= c_i^{\mathsf{T}} \left(I - Q_{\hat{y}} Q_y^{-1}\right) c_i \\
&= c_i^{\mathsf{T}} \left(I - P_A\right) c_i \\
&= c_i^{\mathsf{T}} P_A^{\perp} c_i \\
\Longrightarrow \sum_i r_i &= \text{trace} P_A^{\perp} = m - n
\end{aligned}$$

NB.: $\mathrm{E}\left\{\underline{\hat{e}}^{\mathsf{T}} Q_y^{-1} \underline{\hat{e}}^{\mathsf{T}}\right\} = m - n$

Mean local redundancy number

$$\bar{r} = \frac{\sum_{i=1}^{m} r_i}{m} = \frac{m-n}{m} \Longrightarrow |\bar{\nabla}_i| = \sigma_{y_i}\sqrt{\frac{\lambda}{\frac{m-n}{m}}}$$

Redundancy

$$\begin{aligned}
\underline{\hat{e}} &= P_A^{\perp}\underline{y} \\
&= \left[I - A\left(A^\mathsf{T} Q_y^{-1} A\right)^{-1} A^\mathsf{T} Q_y^{-1}\right]\underline{y} \\
&= R\underline{y} \qquad R = \text{redundancy matrix} \\
\underline{\hat{e}}_i &= R_{ij}\underline{y}_j \\
&= r_i \underline{y}_i + \dots \\
\Longrightarrow \delta\underline{\hat{e}} &= r_i \delta\underline{y}_i
\end{aligned}$$

$\Longrightarrow$ Local redundancy is a quantity how redundancy is distributed among the single observations or how a model error $\delta y = C\nabla$ is projected onto the residuals.

7.5 External reliability

How does an undetected error corrupt the adjustment results?

$$\begin{aligned}
\delta\underline{y} &:= C\nabla \longrightarrow \delta\underline{\hat{x}}? \\
\underline{\hat{x}} &= (A^\mathsf{T} Q_y^{-1} A)^{-1} A^\mathsf{T} Q_y^{-1}\underline{y} \\
(\underline{\hat{x}} + \delta\underline{\hat{x}}) &= (A^\mathsf{T} Q_y^{-1} A)^{-1} A^\mathsf{T} Q_y^{-1}(\underline{y} + \delta\underline{y}) \\
\delta\underline{\hat{x}} &= (A^\mathsf{T} Q_y^{-1} A)^{-1} A^\mathsf{T} Q_y^{-1}\delta\underline{y}
\end{aligned}$$

Problems:

- $\delta\underline{\hat{x}}$ is a vector-valued quantity
- $\delta\underline{\hat{x}}$ depends on possibly inhomogenous quantities with different physical units.

Remedy: Normalize $\delta\underline{\hat{x}}$ using $Q_{\hat{x}}^{-1} \Longrightarrow$ squared bias-to-noise-ratio

$$\begin{aligned}
\underline{\lambda}_{\hat{x}} &= \delta\underline{\hat{x}}^\mathsf{T} Q_{\hat{x}}^{-1} \delta\underline{\hat{x}} \\
&= \delta\underline{\hat{x}}^\mathsf{T} A^\mathsf{T} Q_y^{-1} A \delta\underline{\hat{x}} \\
&= (P_A \delta y)^\mathsf{T} Q_y^{-1} (P_A \delta\underline{y}) \\
&= \|P_A \delta\underline{y}\|^2_{Q_y^{-1}}
\end{aligned}$$

$\underline{\lambda}_{\hat{x}}$: large $\Longrightarrow$ large influence of a model error $\delta \underline{y}$
$\underline{\lambda}_{\hat{x}}$: small $\Longrightarrow$ insignificant influence of a model error $\delta \underline{y}$

7.6 Reliability: a synthesis

$$\begin{aligned}
\delta \underline{y} = I\delta \underline{y} &= (P_A + P_A^{\perp})\delta \underline{y} = P_A \delta \underline{y} + P_A^{\perp}\delta \underline{y} \\
&\Downarrow \\
\|\delta \underline{y}\|^2_{Q_y^{-1}} &= \|P_A \delta \underline{y}\|^2_{Q_y^{-1}} + \|P_A^{\perp}\delta \underline{y}\|^2_{Q_y^{-1}} \\
\text{or } \delta \underline{y}^{\mathsf{T}} Q_y^{-1} \delta \underline{y} &= \delta \underline{\hat{x}}^{\mathsf{T}} A^{\mathsf{T}} Q_y^{-1} A \delta \underline{\hat{x}} + \delta \underline{y}^{\mathsf{T}} (P_A^{\perp})^{\mathsf{T}} Q_y^{-1} P_A^{\perp} \delta \underline{y} \\
\text{or } \underline{\lambda}_y &= \underline{\lambda}_{\hat{x}} + \underline{\lambda}_0 \\
\Longrightarrow \underline{\lambda}_{\hat{x}} &= \underline{\lambda}_y - \underline{\lambda}_0
\end{aligned}$$

Question: Why is $\|P_A^{\perp}\delta \underline{y}\|^2_{Q_y^{-1}} = \underline{\lambda}_0$?

Answer:

$$\begin{aligned}
\underline{\lambda}_0 = \delta \underline{y}^{\mathsf{T}} Q_y^{-1} Q_{\hat{e}} Q_y^{-1} \delta \underline{y} &= \delta \underline{y}^{\mathsf{T}} Q_y^{-1} P_A^{\perp} \delta \underline{y} \\
&= \delta \underline{y}^{\mathsf{T}} (P_A^{\perp})^{\mathsf{T}} Q_y^{-1} P_A^{\perp} \delta \underline{y} \\
&= (P_A^{\perp} \delta \underline{y})^{\mathsf{T}} Q_y^{-1} P_A^{\perp} \delta \underline{y} \\
&= \|P_A^{\perp}\delta \underline{y}\|^2_{Q_y^{-1}}
\end{aligned}$$

special case

$$q = 1, c_i, Q_y = \text{diagonal}$$

$$\begin{aligned}
\Longrightarrow \underline{\lambda}_{\hat{x}} &= \underline{\lambda}_{y_i} - \underline{\lambda}_0 \\
&= \frac{1}{r_i}\underline{\lambda}_0 - \underline{\lambda}_0 \\
&= \frac{1 - r_i}{r_i}\underline{\lambda}_0 \\
&= \frac{\sigma^2_{\hat{y}_i}\sigma^2_{y_i}}{\sigma^2_{y_i}(\sigma^2_{y_i} - \sigma^2_{\hat{y}_i})}\underline{\lambda}_0 \\
&= \frac{\sigma^2_{\hat{y}_i}}{\sigma^2_{y_i} - \sigma^2_{\hat{y}_i}}\underline{\lambda}_0 \\
&= \frac{1}{\frac{\sigma^2_{y_i}}{\sigma^2_{\hat{y}_i}} - 1}\underline{\lambda}_0
\end{aligned}$$

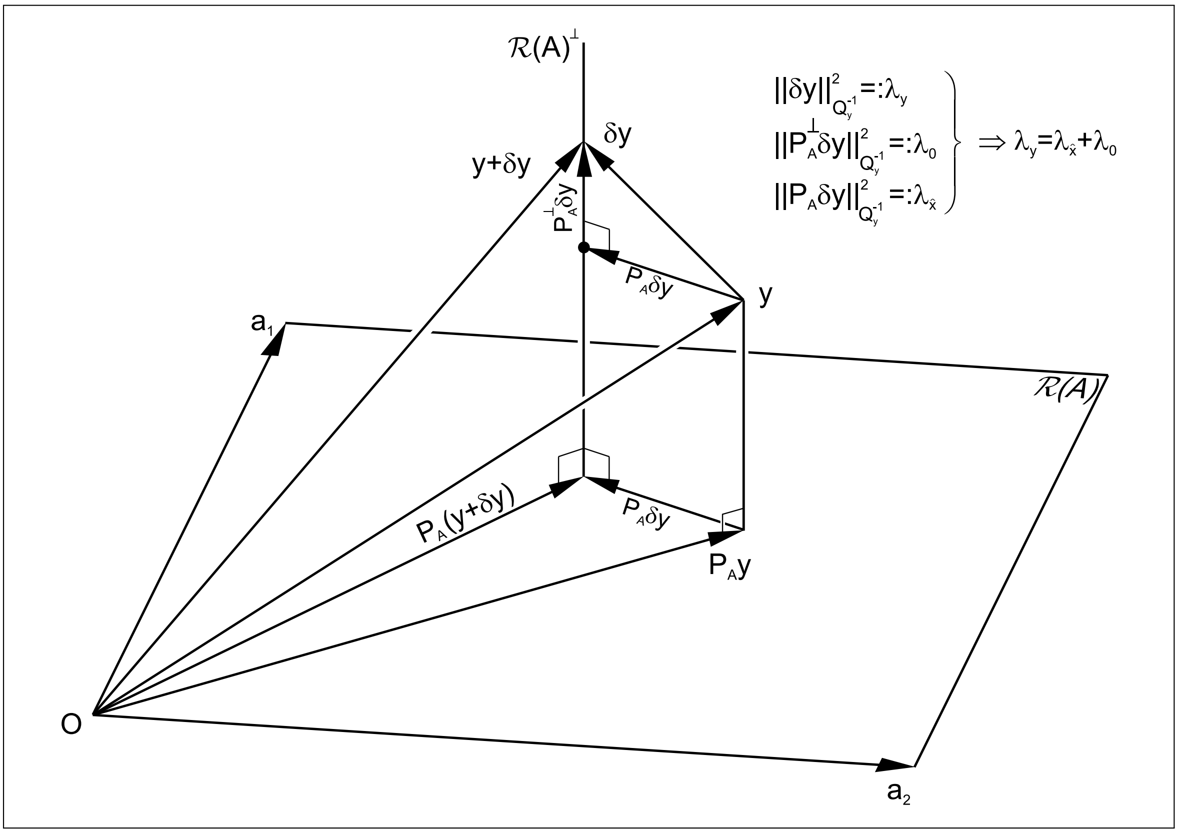


Figure 7.4: Decomposition of λ_y

8 Recursive estimation

8.1 Partitioned model

$$\mathrm{E}\left\{\underline{y}\right\} = \mathrm{E}\left\{\underset{(m_1+m_2)\times 1}{\begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \end{pmatrix}}\right\} = \underset{(m_1+m_2)\times n}{\begin{pmatrix} A_1 \\ A_2 \end{pmatrix}} x; \qquad \mathrm{D}\left\{\begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \end{pmatrix}\right\} = \underset{(m_1+m_2)\times(m_1+m_2)}{\begin{pmatrix} Q_1 & 0 \\ 0 & Q_2 \end{pmatrix}}$$

8.1.1 Batch / offline / Stapel / standard

$$\begin{aligned}
\underline{\hat{x}}_{(1)} &= \left(A_1{}^\mathsf{T} Q_1^{-1} A_1\right) A_1{}^\mathsf{T} Q_1^{-1} \underline{y}_1, \qquad Q_{\hat{x}_{(1)}} = \left(A_1{}^\mathsf{T} Q_1^{-1} A_1\right)^{-1} \\
\underline{\hat{x}}_{(2)} &= \left(A_1{}^\mathsf{T} Q_1^{-1} A_1 + A_2{}^\mathsf{T} Q_2^{-1} A_2\right)^{-1} \left(A_1{}^\mathsf{T} Q_1^{-1} \underline{y}_1 + A_2{}^\mathsf{T} Q_2^{-1} \underline{y}_2\right), \\
Q_{\hat{x}_{(2)}} &= \left(A_1{}^\mathsf{T} Q_1^{-1} A_1 + A_2^\mathsf{T} Q_2^{-1} A_2\right)^{-1}
\end{aligned}$$

8.1.2 Recursive / sequential / real-time

$$\mathrm{E}\left\{\underline{y}_1\right\} = A_1 x \qquad \mathrm{D}\left\{\underline{y}_1\right\} = Q_1$$

$$\Longrightarrow \begin{cases} \underline{\hat{x}}_{(1)} = \left(A_1{}^\mathsf{T} Q_1^{-1} A_1\right)^{-1} A_1{}^\mathsf{T} Q_1^{-1} \underline{y}_1 \\ Q_{\hat{x}_{(1)}} = \left(A_1{}^\mathsf{T} Q_1^{-1} A_1\right)^{-1} \end{cases}$$

$$\begin{cases} \underline{\hat{x}}_{(2)} = \left(A_1{}^\mathsf{T} Q_1^{-1} A_1 + A_2{}^\mathsf{T} Q_2^{-1} A_2\right)^{-1} \left(A_1{}^\mathsf{T} Q_1^{-1} \underline{y}_1 + A_2{}^\mathsf{T} Q_2^{-1} \underline{y}_2\right) \\ \qquad\; = \left(Q_{\hat{x}_{(1)}}^{-1} + A_2{}^\mathsf{T} Q_2^{-1} A_2\right)^{-1} \left(Q_{\hat{x}_{(1)}}^{-1} \underline{\hat{x}}_{(1)} + A_2{}^\mathsf{T} Q_2^{-1} \underline{y}_2\right) \\ Q_{\hat{x}_{(2)}} = \left(Q_{\hat{x}_{(1)}}^{-1} + A_2{}^\mathsf{T} Q_2^{-1} A_2\right)^{-1} \end{cases}$$

Aufdatierungs-
gleichungen

$$\Longrightarrow \begin{cases} \text{measurement update} \\ \text{covariance update} \end{cases}$$

This is also the solution of the problem

$$\mathrm{E}\left\{\begin{pmatrix} \underline{\hat{x}}_{(1)} \\ \underline{y}_2 \end{pmatrix}\right\} = \begin{pmatrix} I \\ A_2 \end{pmatrix} x; \qquad \mathrm{D}\left\{\begin{pmatrix} \underline{\hat{x}}_{(1)} \\ \underline{y}_2 \end{pmatrix}\right\} = \begin{pmatrix} Q_{\hat{x}_{(1)}} & 0 \\ 0 & Q_2 \end{pmatrix}$$

8.1.3 Recursive formulation

Intuitively, it would be nice to have something like $\underline{\hat{x}}_{(2)} = \underline{\hat{x}}_{(1)} + \ldots$

$\Longrightarrow$ Solve

$$Q_{\hat{x}_{(2)}}^{-1} = Q_{\hat{x}_{(1)}}^{-1} + A_2{}^{\mathsf{T}} Q_2^{-1} A_2$$

for $Q_{\hat{x}_{(1)}}^{-1}$

$$\Longrightarrow Q_{\hat{x}_{(1)}}^{-1} = Q_{\hat{x}_{(2)}}^{-1} - A_2{}^{\mathsf{T}} Q_2^{-1} A_2$$

Substitute the result in $\underline{\hat{x}}_{(2)} = \left(Q_{\hat{x}_{(1)}}^{-1} + A_2{}^{\mathsf{T}} Q_2^{-1} A_2\right)^{-1} \left(Q_{\hat{x}_{(1)}}^{-1} \underline{\hat{x}}_{(1)} + A_2{}^{\mathsf{T}} Q_2^{-1} \underline{y}_2\right)$

$$\begin{aligned}
\Longrightarrow \underline{\hat{x}}_{(2)} &= Q_{\hat{x}_{(2)}} \left(Q_{\hat{x}_{(2)}}^{-1} \underline{\hat{x}}_{(1)} - A_2{}^{\mathsf{T}} Q_2^{-1} A_2 \underline{\hat{x}}_{(1)} + A_2{}^{\mathsf{T}} Q_2^{-1} \underline{y}_2\right) \\
&= \underline{\hat{x}}_{(1)} + Q_{\hat{x}_{(2)}} A_2{}^{\mathsf{T}} Q_2^{-1} \left(\underline{y}_2 - A_2 \underline{\hat{x}}_{(1)}\right) \\
&= \underline{\hat{x}}_{(1)} + K \underline{v}_2 \\
\underline{v}_2 &= \underline{y}_2 - A_2 \underline{\hat{x}}_{(1)} \\
K &= Q_{\hat{x}_{(2)}} A_2{}^{\mathsf{T}} Q_2^{-1}
\end{aligned}$$

$A_2 \underline{\hat{x}}_{(1)} \ldots$ predicted observation
$\underline{v}_2 \ldots$ predicted residual (attention!)
$K \ldots$ *gain matrix* Verstärkungs-matrix

disadvantage: too many matrix inversions

$$\underset{n \times n}{Q_{\hat{x}_{(1)}}^{-1}}, \qquad \underset{n \times n}{\left(Q_{\hat{x}_{(1)}}^{-1} + A_2{}^{\mathsf{T}} Q_2^{-1} A_2\right)^{-1}}, \qquad \underset{m_2 \times m_2}{Q_2^{-1}}$$

8.1.4 Formulation using condition equations

$$B^{\mathsf{T}} A = 0$$

$$\Longrightarrow \begin{pmatrix} -A_2 & I \end{pmatrix} \begin{pmatrix} I \\ A_2 \end{pmatrix} = 0$$

$$\begin{pmatrix} -A_2 & I \end{pmatrix} \mathrm{E}\left\{\begin{pmatrix} \underline{\hat{x}}_{(1)} \\ \underline{y}_2 \end{pmatrix}\right\} = 0; \qquad \mathrm{D}\left\{\begin{pmatrix} \underline{\hat{x}}_{(1)} \\ \underline{y}_2 \end{pmatrix}\right\} = \begin{pmatrix} Q_{\hat{x}_{(1)}} & 0 \\ 0 & Q_2 \end{pmatrix}$$

$$\left.\begin{aligned} B^{\mathsf{T}} \mathrm{E}\{\underline{y}\} &= 0 \\ \mathrm{D}\{\underline{y}\} &= Q_y \end{aligned}\right\} \Longrightarrow \underline{\hat{y}} = \left[I - Q_y B \left(B^{\mathsf{T}} Q_y B\right)^{-1} B^{\mathsf{T}}\right] \underline{y}$$

$$\Longrightarrow \begin{pmatrix} \underline{\hat{x}}_{(2)} \\ \underline{\hat{y}}_2 \end{pmatrix} = \left[\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} - \begin{pmatrix} -Q_{\hat{x}_{(1)}} A_2{}^{\mathsf{T}} \\ Q_2 \end{pmatrix} \left(Q_2 + A_2 Q_{\hat{x}_{(1)}} A_2{}^{\mathsf{T}}\right)^{-1} \begin{pmatrix} -A_2 & I \end{pmatrix}\right] \begin{pmatrix} \underline{\hat{x}}_{(1)} \\ \underline{y}_2 \end{pmatrix}$$

$\Longrightarrow$ Measurement update

$$
\begin{aligned}
\underline{\hat{x}}_{(2)} &= \underline{\hat{x}}_{(1)} + Q_{\hat{x}_{(1)}} A_2{}^{\mathsf{T}} \left(Q_2 + A_2 Q_{\hat{x}_{(1)}} A_2{}^{\mathsf{T}} \right)^{-1} \left(\underline{y}_2 - A_2 \underline{\hat{x}}_{(1)} \right) \\
&= \underline{\hat{x}}_{(1)} + K \underline{v}_2 \\
\underset{m_2 \times m_2}{K} &= Q_{\hat{x}_{(1)}} A_2{}^{\mathsf{T}} \left(Q_2 + A_2 Q_{\hat{x}_{(1)}} A_2{}^{\mathsf{T}} \right)^{-1} \\
&= Q_{\hat{x}_{(1)}} A_2{}^{\mathsf{T}} Q_{v_2}^{-1}
\end{aligned}
$$

$\Longrightarrow$ Covariance update

$$
\begin{aligned}
Q_{\hat{x}_{(2)}} &= Q_{\hat{x}_{(1)}} - Q_{\hat{x}_{(1)}} A_2{}^{\mathsf{T}} \left(Q_2 + A_2 Q_{\hat{x}_{(1)}} A_2{}^{\mathsf{T}} \right)^{-1} A_2 Q_{\hat{x}_{(1)}} \\
&= Q_{\hat{x}_{(1)}} - K A_2 Q_{\hat{x}_{(1)}} \\
&= \left(I - K A_2 \right) Q_{\hat{x}_{(1)}}
\end{aligned}
$$

Remark: Variance decreases as more observations are included.

8.2 More general

$$
\mathrm{E}\left\{ \begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \vdots \\ \underline{y}_k \end{pmatrix} \right\} = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_k \end{pmatrix} x; \qquad \mathrm{D}\left\{ \begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \vdots \\ \underline{y}_k \end{pmatrix} \right\} = \begin{pmatrix} Q_1 & & & 0 \\ & Q_2 & & \\ & & \ddots & \\ 0 & & & Q_k \end{pmatrix}
$$

Batch:

$$
\underline{\hat{x}} = \left(\sum_{i=1}^{k} A_i{}^{\mathsf{T}} Q_i^{-1} A_i \right)^{-1} \left(\sum_{i=1}^{k} A_i{}^{\mathsf{T}} Q_i^{-1} \underline{y}_i \right)
$$

Recursive:

$$
\begin{aligned}
\underline{\hat{x}}_{(k)} &= \underline{\hat{x}}_{(k-1)} + K_k \underline{v}_k \\
\underline{v}_k &= \underline{y}_k - A_k \underline{\hat{x}}_{(k-1)} \\
K_k &= Q_{\hat{x}_{(k-1)}} A_k{}^{\mathsf{T}} \left(Q_k + A_k Q_{\hat{x}_{(k-1)}} A_k{}^{\mathsf{T}} \right)^{-1} \\
&= Q_{\hat{x}_{(k-1)}} A_k{}^{\mathsf{T}} Q_{v_k}^{-1} \\
Q_{\hat{x}_{(k)}} &= \left(I - K_k A_k \right) Q_{\hat{x}_{(k-1)}}
\end{aligned}
$$

A Partitioning

A.1 Inverse Partitioning Method (IPM)

$$\begin{bmatrix} \underset{n\times n}{W} & \underset{n\times k}{X} \\ \underset{k\times n}{Y} & \underset{k\times k}{Z} \end{bmatrix} \underbrace{\begin{bmatrix} \underset{n\times n}{A} & \underset{n\times k}{B} \\ \underset{k\times n}{C} & \underset{k\times k}{D} \end{bmatrix}}_{\text{Inverse}} = \begin{bmatrix} I_n & 0 \\ 0 & I_k \end{bmatrix} \qquad \text{A, B, C, D are unknown}$$

(1) $WA + XC = I_n, \qquad \operatorname{rank} W = n$

(2) $WB + XD = 0$

(3) $YA + ZC = 0$

(4) $YB + ZD = I_k$

(5) $W^{-1} \cdot (1): \quad A + W^{-1}XC = W^{-1} \Longrightarrow A = W^{-1} - W^{-1}XC$

(6) Insert (5) into (3) : $\quad YW^{-1} - YW^{-1}XC + ZC = 0 \Longrightarrow C = -(Z - YW^{-1}X)^{-1}YW^{-1}$ (provided $G = Z - YW^{-1}X$ is non-singular)

(7) $D = G^{-1} = (Z - YW^{-1}X)^{-1}$

(8) $B = -W^{-1}XG^{-1} = -W^{-1}X(Z - YW^{-1}X)^{-1}$

A.2 Inverse Partitioning Method: special case 1

$$\begin{bmatrix} I & -b \\ -b^{\mathsf{T}} & 0 \end{bmatrix} \underbrace{\begin{bmatrix} A & B \\ C & D \end{bmatrix}}_{\text{Inverse}} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \qquad \text{A, B, C, D are unknown}$$

(1) $A - bC = I$

(2) $B - bD = 0$

(3) $-b^{\mathsf{T}}A = 0$

(4) $-b^{\mathsf{T}}B = I$

(5) $\underbrace{-b^{\mathsf{T}}A}_{=0} - b^{\mathsf{T}}bC = -b^{\mathsf{T}} \Longrightarrow C = -(b^{\mathsf{T}}b)^{-1}b^{\mathsf{T}}$

(6) $-b^{\mathsf{T}}B + b^{\mathsf{T}}bD = 0 \Longrightarrow I + b^{\mathsf{T}}bD = 0 \Longrightarrow D = -(b^{\mathsf{T}}b)^{-1}$

(7) $A + b(b^{\mathsf{T}}b)^{-1}b^{\mathsf{T}} = I \Longrightarrow A = I - b(b^{\mathsf{T}}b)^{-1}b^{\mathsf{T}}$

(8) $B + b(b^{\mathsf{T}}b)^{-1} = 0 \Longrightarrow B = -b(b^{\mathsf{T}}b)^{-1}$

$$\begin{bmatrix} I & -b \\ -b^{\mathsf{T}} & 0 \end{bmatrix}^{-1} = \begin{bmatrix} I - b(b^{\mathsf{T}}b)^{-1}b^{\mathsf{T}} & -b(b^{\mathsf{T}}b)^{-1} \\ -(b^{\mathsf{T}}b)^{-1}b^{\mathsf{T}} & -(b^{\mathsf{T}}b)^{-1} \end{bmatrix}$$

$$\hat{e} = b(b^{\mathsf{T}}b)^{-1}b^{\mathsf{T}}y$$

$$\hat{\lambda} = (b^{\mathsf{T}}b)^{-1}b^{\mathsf{T}}y$$

A.3 Inverse Partitioning Method: special case 2

(A rank deficient, constraint $D^{\mathsf{T}}x = c$).

The normal matrix of the linear system is symmetric, therefore

$$\begin{pmatrix} A^{\mathsf{T}}A & D \\ D^{\mathsf{T}} & 0 \end{pmatrix} \underbrace{\begin{pmatrix} R & S^{\mathsf{T}} \\ S & Q \end{pmatrix}}_{\text{Inverse}} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

then

$$(A^{\mathsf{T}}A)R + DS = I \tag{A.1}$$

$$(A^{\mathsf{T}}A)S^{\mathsf{T}} + DQ = 0 \tag{A.2}$$

$$D^{\mathsf{T}}R = 0 \tag{A.3}$$

$$D^{\mathsf{T}}S^{\mathsf{T}} = I \tag{A.4}$$

and with $H = \text{null}(A), AH^{\mathsf{T}} = 0$

$$H \cdot (\text{A.1}) \Longrightarrow \underbrace{H(A^{\mathsf{T}}}_{0}A)R + HDS = H \Longrightarrow S = (HD)^{-1}H$$

$$H \cdot (\text{A.2}) \Longrightarrow \underbrace{H(A^{\mathsf{T}}}_{0}A)S^{\mathsf{T}} + HDQ = 0 \Longrightarrow HDQ = 0$$

since HD is a $d \times d$ full-rank matrix

$$HDQ = 0 \Longrightarrow Q = 0$$

$$\begin{aligned} (\text{A.1}) + D \cdot (\text{A.3}) &\Longrightarrow (A^{\mathsf{T}}A)R + D(HD)^{-1}H + DD^{\mathsf{T}}R = I \\ &\Longrightarrow (A^{\mathsf{T}}A + DD^{\mathsf{T}})R = I - D(HD)^{-1}H \\ &\Longrightarrow R = (A^{\mathsf{T}}A + DD^{\mathsf{T}})^{-1}(I - D(HD)^{-1}H) \end{aligned} \tag{A.5}$$

Inserting R and S into the normal equations

$$\begin{cases} \hat{x} = RA^{\mathsf{T}}y = (A^{\mathsf{T}}A + DD^{\mathsf{T}})^{-1}A^{\mathsf{T}}y - (A^{\mathsf{T}}A + DD^{\mathsf{T}})^{-1}D(HD)^{-1}\underbrace{HA^{\mathsf{T}}}_{0}y \\ \hat{\lambda} = SA^{\mathsf{T}}y = (HD)^{-1}\underbrace{HA^{\mathsf{T}}}_{0}y = 0 \end{cases}$$

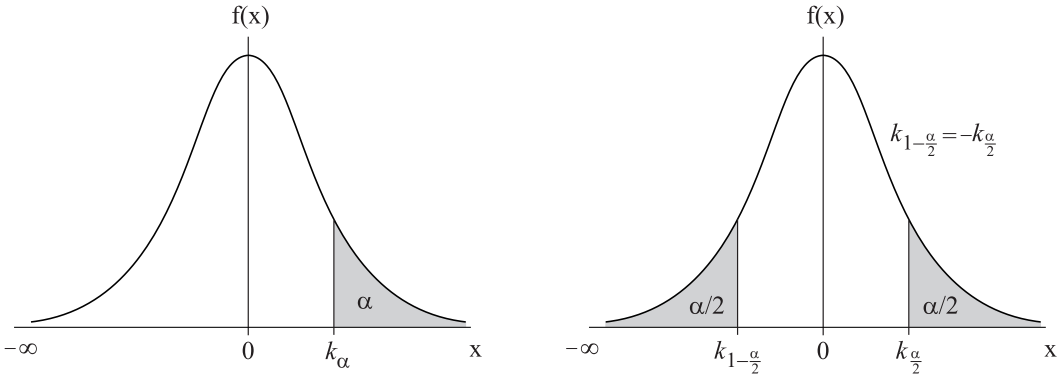
$$\begin{aligned} &\Longrightarrow \hat{\lambda} = 0 \\ &\Longrightarrow \hat{x} = (A^{\mathsf{T}}A + DD^{\mathsf{T}})^{-1}A^{\mathsf{T}}y + H(H^{\mathsf{T}}D)^{-1}c \\ \hat{e} \;&=\; y - A\hat{x} = (I - A(A^{\mathsf{T}}A + DD^{\mathsf{T}})^{-1}A^{\mathsf{T}})y \end{aligned}$$

B Statistical Tables

B.1 Standard Normal Distribution z

Computation of one-sided level of significance $\alpha = 1 - \int_{-\infty}^{k_\alpha} f(x)\,\mathrm{d}x$ and

two-sided level of significance $\alpha = 2 \int_{-\infty}^{k_{1-\alpha/2}} f(x)\,\mathrm{d}x$.



k_α	0	1	2	3	4	5	6	7	8	9
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681

Computation of one-sided level of significance $\alpha = 1 - \int_{-\infty}^{k_\alpha} f(x)\,\mathrm{d}x$ and

two-sided level of significance $\alpha = 2 \int_{-\infty}^{k_{1-\alpha/2}} f(x)\,\mathrm{d}x$ (continued).

k_α	0	1	2	3	4	5	6	7	8	9
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002

Calculation in MATLAB:

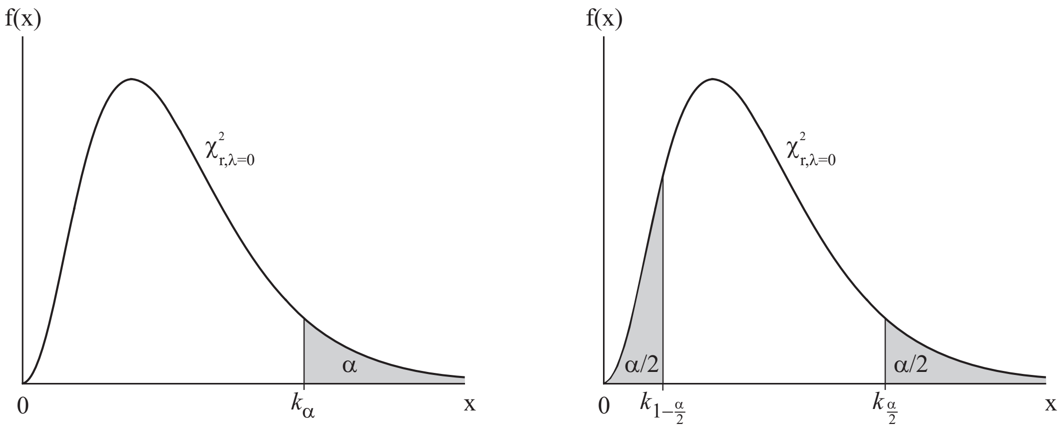
$$\alpha = 1 - \texttt{normcdf}(k_\alpha) \qquad\qquad k_\alpha = \texttt{norminv}(1-\alpha)$$

Example ($k_\alpha = 0.87$, one-sided):

$$\alpha = 0.1922 = 1 - \texttt{normcdf}(0.87) \qquad\qquad k_\alpha = 0.87 = \texttt{norminv}(1-0.1922)$$

B.2 Central χ^2-Distribution

Computation of critical value $k_\alpha = \chi^2_{1-\alpha;r,\lambda=0}$ (r degrees of freedom).



$r\backslash\alpha$	0.995	0.990	0.975	0.950	0.900	0.500
1	0.000	0.000	0.001	0.004	0.016	0.455
2	0.010	0.020	0.051	0.103	0.211	1.386
3	0.072	0.115	0.216	0.352	0.584	2.366
4	0.207	0.297	0.484	0.711	1.064	3.357
5	0.412	0.554	0.831	1.145	1.610	4.351
6	0.676	0.872	1.237	1.635	2.204	5.348
7	0.989	1.239	1.690	2.167	2.833	6.346
8	1.344	1.646	2.180	2.733	3.490	7.344
9	1.735	2.088	2.700	3.325	4.168	8.343
10	2.156	2.558	3.247	3.940	4.865	9.342
11	2.603	3.053	3.816	4.575	5.578	10.34
12	3.074	3.571	4.404	5.226	6.304	11.34
13	3.565	4.107	5.009	5.892	7.042	12.34
14	4.075	4.660	5.629	6.571	7.790	13.34
15	4.601	5.229	6.262	7.261	8.547	14.34
16	5.142	5.812	6.908	7.962	9.312	15.34
17	5.697	6.408	7.564	8.672	10.09	16.34
18	6.265	7.015	8.231	9.390	10.86	17.34
19	6.844	7.633	8.907	10.12	11.65	18.34
20	7.434	8.260	9.591	10.85	12.44	19.34
21	8.034	8.897	10.28	11.59	13.24	20.34
22	8.643	9.542	10.98	12.34	14.04	21.34
23	9.260	10.20	11.69	13.09	14.85	22.34
24	9.886	10.86	12.40	13.85	15.66	23.34
25	10.52	11.52	13.12	14.61	16.47	24.34
26	11.16	12.20	13.84	15.38	17.29	25.34
27	11.81	12.88	14.57	16.15	18.11	26.34
28	12.46	13.56	15.31	16.93	18.94	27.34
29	13.12	14.26	16.05	17.71	19.77	28.34
30	13.79	14.95	16.79	18.49	20.60	29.34

Computation of critical value $k_\alpha = \chi^2_{1-\alpha;r,\lambda=0}$ (continued).

$r\backslash\alpha$	0.995	0.990	0.975	0.950	0.900	0.500
35	17.19	18.51	20.57	22.47	24.80	34.34
40	20.71	22.16	24.43	26.51	29.05	39.34
45	24.31	25.90	28.37	30.61	33.35	44.34
50	27.99	29.71	32.36	34.76	37.69	49.33
60	35.53	37.48	40.48	43.19	46.46	59.33
70	43.28	45.44	48.76	51.74	55.33	69.33
80	51.17	53.54	57.15	60.39	64.28	79.33
90	59.20	61.75	65.65	69.13	73.29	89.33
100	67.33	70.06	74.22	77.93	82.36	99.33

$r\backslash\alpha$	0.100	0.050	0.025	0.010	0.005	0.001
1	2.706	3.841	5.024	6.635	7.879	10.83
2	4.605	5.991	7.378	9.210	10.60	13.82
3	6.251	7.815	9.348	11.34	12.84	16.27
4	7.779	9.488	11.14	13.28	14.86	18.47
5	9.236	11.07	12.83	15.09	16.75	20.52
6	10.64	12.59	14.45	16.81	18.55	22.46
7	12.02	14.07	16.01	18.48	20.28	24.32
8	13.36	15.51	17.53	20.09	21.95	26.12
9	14.68	16.92	19.02	21.67	23.59	27.88
10	15.99	18.31	20.48	23.21	25.19	29.59
11	17.28	19.68	21.92	24.72	26.76	31.26
12	18.55	21.03	23.34	26.22	28.30	32.91
13	19.81	22.36	24.74	27.69	29.82	34.53
14	21.06	23.68	26.12	29.14	31.32	36.12
15	22.31	25.00	27.49	30.58	32.80	37.70
16	23.54	26.30	28.85	32.00	34.27	39.25
17	24.77	27.59	30.19	33.41	35.72	40.79
18	25.99	28.87	31.53	34.81	37.16	42.31
19	27.20	30.14	32.85	36.19	38.58	43.82
20	28.41	31.41	34.17	37.57	40.00	45.31
21	29.62	32.67	35.48	38.93	41.40	46.80
22	30.81	33.92	36.78	40.29	42.80	48.27
23	32.01	35.17	38.08	41.64	44.18	49.73
24	33.20	36.42	39.36	42.98	45.56	51.18
25	34.38	37.65	40.65	44.31	46.93	52.62
26	35.56	38.89	41.92	45.64	48.29	54.05
27	36.74	40.11	43.19	46.96	49.64	55.48
28	37.92	41.34	44.46	48.28	50.99	56.89
29	39.09	42.56	45.72	49.59	52.34	58.30
30	40.26	43.77	46.98	50.89	53.67	59.70

Computation of critical value $k_\alpha = \chi^2_{1-\alpha;r,\lambda=0}$ (continued).

$r \backslash \alpha$	0.100	0.050	0.025	0.010	0.005	0.001
35	46.06	49.80	53.20	57.34	60.27	66.62
40	51.81	55.76	59.34	63.69	66.77	73.40
45	57.51	61.66	65.41	69.96	73.17	80.08
50	63.17	67.50	71.42	76.15	79.49	86.66
60	74.40	79.08	83.30	88.38	91.95	99.61
70	85.53	90.53	95.02	100.4	104.2	112.3
80	96.58	101.9	106.6	112.3	116.3	124.8
90	107.6	113.1	118.1	124.1	128.3	137.2
100	118.5	124.3	129.6	135.8	140.2	149.4

Calculation in Matlab:

$$k_\alpha = \texttt{chi2inv}(1-\alpha, r) \qquad \alpha = 1 - \texttt{chi2cdf}(k_\alpha, r)$$

Example ($\alpha = 0.95$, one-sided):

$$k_\alpha = 11.59 = \texttt{chi2inv}(1-0.95, 21) \qquad \alpha = 0.95 = 1 - \texttt{chi2cdf}(11.59, 21)$$

Example ($\alpha = 0.01$, two-sided):

$$k_{1-\alpha/2} = 41.4 = \texttt{chi2inv}(1-{}^{0.01}/_{2}, 21) \qquad k_{\alpha/2} = 8.034 = \texttt{chi2inv}({}^{0.01}/_{2}, 21)$$

B.3 Non-central χ^2-Distribution

Computation of power of test $\gamma = 1 - \beta$
(Non-centrality parameter λ, $r = 1$ degrees of freedom).

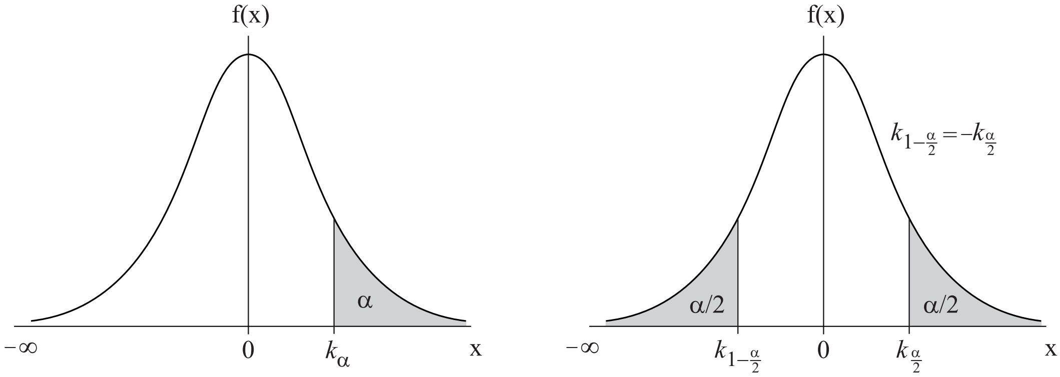
$\lambda\backslash\alpha$	0.100	0.010	0.001	$\lambda\backslash\alpha$	0.100	0.010	0.001
1.000	0.264	0.058	0.011	11.000	0.953	0.771	0.510
1.250	0.302	0.073	0.015	11.250	0.956	0.782	0.525
1.500	0.339	0.088	0.019	11.500	0.960	0.793	0.540
1.750	0.375	0.105	0.025	11.750	0.963	0.803	0.555
2.000	0.410	0.123	0.030	12.000	0.966	0.813	0.569
2.250	0.443	0.141	0.037	12.250	0.968	0.822	0.583
2.500	0.475	0.160	0.044	12.500	0.971	0.831	0.597
2.750	0.506	0.179	0.051	12.750	0.973	0.840	0.610
3.000	0.535	0.199	0.060	13.000	0.975	0.848	0.624
3.250	0.563	0.220	0.068	13.250	0.977	0.856	0.637
3.500	0.590	0.240	0.078	13.500	0.979	0.864	0.649
3.750	0.615	0.261	0.088	13.750	0.980	0.871	0.662
4.000	0.639	0.282	0.098	14.000	0.982	0.878	0.674
4.250	0.662	0.304	0.110	14.250	0.983	0.885	0.686
4.500	0.683	0.325	0.121	14.500	0.985	0.891	0.698
4.750	0.704	0.346	0.133	14.750	0.986	0.897	0.709
5.000	0.723	0.367	0.146	15.000	0.987	0.903	0.720
5.250	0.741	0.388	0.159	15.250	0.988	0.908	0.731
5.500	0.758	0.409	0.172	15.500	0.989	0.913	0.741
5.750	0.774	0.429	0.186	15.750	0.990	0.918	0.751
6.000	0.790	0.450	0.200	16.000	0.991	0.923	0.761
6.250	0.804	0.470	0.215	16.250	0.991	0.927	0.771
6.500	0.817	0.490	0.229	16.500	0.992	0.931	0.780
6.750	0.830	0.509	0.244	16.750	0.993	0.935	0.789
7.000	0.842	0.528	0.260	17.000	0.993	0.939	0.797
7.250	0.853	0.546	0.275	17.250	0.994	0.943	0.806
7.500	0.863	0.565	0.291	17.500	0.994	0.946	0.814
7.750	0.873	0.582	0.306	17.750	0.995	0.949	0.822
8.000	0.882	0.600	0.322	18.000	0.995	0.952	0.829
8.250	0.890	0.617	0.338	18.250	0.996	0.955	0.837
8.500	0.898	0.633	0.354	18.500	0.996	0.958	0.844
8.750	0.905	0.649	0.370	18.750	0.996	0.960	0.851
9.000	0.912	0.664	0.386	19.000	0.997	0.963	0.857
9.250	0.919	0.679	0.402	19.250	0.997	0.965	0.864
9.500	0.925	0.694	0.417	19.500	0.997	0.967	0.870
9.750	0.930	0.708	0.433	19.750	0.997	0.969	0.876
10.000	0.935	0.721	0.449	20.000	0.998	0.971	0.881
10.250	0.940	0.734	0.465	20.250	0.998	0.973	0.887
10.500	0.945	0.747	0.480	20.500	0.998	0.975	0.892
10.750	0.949	0.759	0.495	20.750	0.998	0.976	0.897

Calculation in Matlab:

$$k_\alpha = \texttt{chi2inv}(1-\alpha, r) \qquad\qquad \gamma = 1 - \texttt{ncx2cdf}(k_\alpha, r, \lambda)$$

B.4 Central t-Distribution

Computation of critical value $k_\alpha = t_{1-\alpha;r}$ (r degrees of freedom).



$r\backslash\alpha$	0.100	0.050	0.025	0.010	0.005	0.001
1	3.078	6.314	12.71	31.82	63.66	318.3
2	1.886	2.920	4.303	6.965	9.925	22.33
3	1.638	2.353	3.182	4.541	5.841	10.21
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.893
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.785
8	1.397	1.860	2.306	2.896	3.355	4.501
9	1.383	1.833	2.262	2.821	3.250	4.297
10	1.372	1.812	2.228	2.764	3.169	4.144
11	1.363	1.796	2.201	2.718	3.106	4.025
12	1.356	1.782	2.179	2.681	3.055	3.930
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.624	2.977	3.787
15	1.341	1.753	2.131	2.602	2.947	3.733
16	1.337	1.746	2.120	2.583	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.646
18	1.330	1.734	2.101	2.552	2.878	3.610
19	1.328	1.729	2.093	2.539	2.861	3.579
20	1.325	1.725	2.086	2.528	2.845	3.552

Computation of critical value $k_\alpha = t_{1-\alpha;r}$ (continued).

$r\backslash\alpha$	0.100	0.050	0.025	0.010	0.005	0.001
21	1.323	1.721	2.080	2.518	2.831	3.527
22	1.321	1.717	2.074	2.508	2.819	3.505
23	1.319	1.714	2.069	2.500	2.807	3.485
24	1.318	1.711	2.064	2.492	2.797	3.467
25	1.316	1.708	2.060	2.485	2.787	3.450
26	1.315	1.706	2.056	2.479	2.779	3.435
27	1.314	1.703	2.052	2.473	2.771	3.421
28	1.313	1.701	2.048	2.467	2.763	3.408
29	1.311	1.699	2.045	2.462	2.756	3.396
30	1.310	1.697	2.042	2.457	2.750	3.385
35	1.306	1.690	2.030	2.438	2.724	3.340
40	1.303	1.684	2.021	2.423	2.704	3.307
45	1.301	1.679	2.014	2.412	2.690	3.281
50	1.299	1.676	2.009	2.403	2.678	3.261
60	1.296	1.671	2.000	2.390	2.660	3.232
70	1.294	1.667	1.994	2.381	2.648	3.211
80	1.292	1.664	1.990	2.374	2.639	3.195
90	1.291	1.662	1.987	2.368	2.632	3.183
100	1.290	1.660	1.984	2.364	2.626	3.174
200	1.286	1.653	1.972	2.345	2.601	3.131
500	1.283	1.648	1.965	2.334	2.586	3.107
∞	1.282	1.645	1.960	2.327	2.576	3.091

Calculation in Matlab:

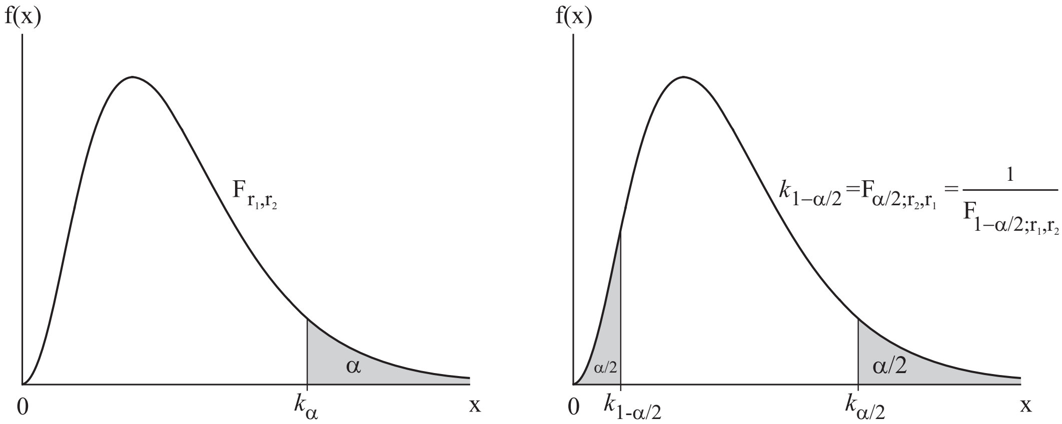
$$k_\alpha = \texttt{tinv}(1-\alpha, r) \qquad\qquad \alpha = 1 - \texttt{tcdf}(k_\alpha, r)$$

Example ($\alpha = 0.005$, one-sided):

$$k_\alpha = 2.898 = \texttt{tinv}(1-0.005, 17) \qquad\qquad \alpha = 0.005 = 1 - \texttt{tcdf}(2.898, 17)$$

B.5 Central F-Distribution

Computation of critical value $k_\alpha = F_{1-\alpha;r_1,r_2,\lambda=0}$ (r_1, r_2 degrees of freedom).



$\alpha = 0.10, \quad 1-\alpha = 0.90$

$r_2\backslash r_1$	1	2	3	4	5	6	7	8	9	10	12
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	60.71
2	8.526	9.000	9.162	9.243	9.293	9.326	9.349	9.367	9.381	9.392	9.408
3	5.538	5.462	5.391	5.343	5.309	5.285	5.266	5.252	5.240	5.230	5.216
4	4.545	4.325	4.191	4.107	4.051	4.010	3.979	3.955	3.936	3.920	3.896
5	4.060	3.780	3.619	3.520	3.453	3.405	3.368	3.339	3.316	3.297	3.268
6	3.776	3.463	3.289	3.181	3.108	3.055	3.014	2.983	2.958	2.937	2.905
7	3.589	3.257	3.074	2.961	2.883	2.827	2.785	2.752	2.725	2.703	2.668
8	3.458	3.113	2.924	2.806	2.726	2.668	2.624	2.589	2.561	2.538	2.502
9	3.360	3.006	2.813	2.693	2.611	2.551	2.505	2.469	2.440	2.416	2.379
10	3.285	2.924	2.728	2.605	2.522	2.461	2.414	2.377	2.347	2.323	2.284
11	3.225	2.860	2.660	2.536	2.451	2.389	2.342	2.304	2.274	2.248	2.209
12	3.177	2.807	2.606	2.480	2.394	2.331	2.283	2.245	2.214	2.188	2.147
13	3.136	2.763	2.560	2.434	2.347	2.283	2.234	2.195	2.164	2.138	2.097
14	3.102	2.726	2.522	2.395	2.307	2.243	2.193	2.154	2.122	2.095	2.054
15	3.073	2.695	2.490	2.361	2.273	2.208	2.158	2.119	2.086	2.059	2.017
16	3.048	2.668	2.462	2.333	2.244	2.178	2.128	2.088	2.055	2.028	1.985
17	3.026	2.645	2.437	2.308	2.218	2.152	2.102	2.061	2.028	2.001	1.958
18	3.007	2.624	2.416	2.286	2.196	2.130	2.079	2.038	2.005	1.977	1.933
19	2.990	2.606	2.397	2.266	2.176	2.109	2.058	2.017	1.984	1.956	1.912
20	2.975	2.589	2.380	2.249	2.158	2.091	2.040	1.999	1.965	1.937	1.892
22	2.949	2.561	2.351	2.219	2.128	2.060	2.008	1.967	1.933	1.904	1.859
24	2.927	2.538	2.327	2.195	2.103	2.035	1.983	1.941	1.906	1.877	1.832
26	2.909	2.519	2.307	2.174	2.082	2.014	1.961	1.919	1.884	1.855	1.809
28	2.894	2.503	2.291	2.157	2.064	1.996	1.943	1.900	1.865	1.836	1.790
30	2.881	2.489	2.276	2.142	2.049	1.980	1.927	1.884	1.849	1.819	1.773
40	2.835	2.440	2.226	2.091	1.997	1.927	1.873	1.829	1.793	1.763	1.715
50	2.809	2.412	2.197	2.061	1.966	1.895	1.840	1.796	1.760	1.729	1.680
60	2.791	2.393	2.177	2.041	1.946	1.875	1.819	1.775	1.738	1.707	1.657
80	2.769	2.370	2.154	2.016	1.921	1.849	1.793	1.748	1.711	1.680	1.629
100	2.756	2.356	2.139	2.002	1.906	1.834	1.778	1.732	1.695	1.663	1.612
200	2.731	2.329	2.111	1.973	1.876	1.804	1.747	1.701	1.663	1.631	1.579
500	2.716	2.313	2.095	1.956	1.859	1.786	1.729	1.683	1.644	1.612	1.559
∞	2.706	2.303	2.084	1.945	1.847	1.774	1.717	1.670	1.632	1.599	1.546

Computation of critical value $k_\alpha = F_{1-\alpha;r_1,r_2,\lambda=0}$ (continued).

$\alpha = 0.10, \quad 1-\alpha = 0.90$

$r_2 \backslash r_1$	14	16	18	20	30	40	50	100	200	500	∞
1	61.07	61.35	61.57	61.74	62.26	62.53	62.69	63.01	63.17	63.26	63.33
2	9.420	9.429	9.436	9.441	9.458	9.466	9.471	9.481	9.486	9.489	9.491
3	5.205	5.196	5.190	5.184	5.168	5.160	5.155	5.144	5.139	5.136	5.134
4	3.878	3.864	3.853	3.844	3.817	3.804	3.795	3.778	3.769	3.764	3.761
5	3.247	3.230	3.217	3.207	3.174	3.157	3.147	3.126	3.116	3.109	3.105
6	2.881	2.863	2.848	2.836	2.800	2.781	2.770	2.746	2.734	2.727	2.722
7	2.643	2.623	2.607	2.595	2.555	2.535	2.523	2.497	2.484	2.476	2.471
8	2.475	2.455	2.438	2.425	2.383	2.361	2.348	2.321	2.307	2.298	2.293
9	2.351	2.329	2.312	2.298	2.255	2.232	2.218	2.189	2.174	2.165	2.159
10	2.255	2.233	2.215	2.201	2.155	2.132	2.117	2.087	2.071	2.062	2.055
11	2.179	2.156	2.138	2.123	2.076	2.052	2.036	2.005	1.989	1.979	1.972
12	2.117	2.094	2.075	2.060	2.011	1.986	1.970	1.938	1.921	1.911	1.904
13	2.066	2.042	2.023	2.007	1.958	1.931	1.915	1.882	1.864	1.853	1.846
14	2.022	1.998	1.978	1.962	1.912	1.885	1.869	1.834	1.816	1.805	1.797
15	1.985	1.961	1.941	1.924	1.873	1.845	1.828	1.793	1.774	1.763	1.755
16	1.953	1.928	1.908	1.891	1.839	1.811	1.793	1.757	1.738	1.726	1.718
17	1.925	1.900	1.879	1.862	1.809	1.781	1.763	1.726	1.706	1.694	1.686
18	1.900	1.875	1.854	1.837	1.783	1.754	1.736	1.698	1.678	1.665	1.657
19	1.878	1.852	1.831	1.814	1.759	1.730	1.711	1.673	1.652	1.639	1.631
20	1.859	1.833	1.811	1.794	1.738	1.708	1.690	1.650	1.629	1.616	1.607
22	1.825	1.798	1.777	1.759	1.702	1.671	1.652	1.611	1.590	1.576	1.567
24	1.797	1.770	1.748	1.730	1.672	1.641	1.621	1.579	1.556	1.542	1.533
26	1.774	1.747	1.724	1.706	1.647	1.615	1.594	1.551	1.528	1.514	1.504
28	1.754	1.726	1.704	1.685	1.625	1.592	1.572	1.528	1.504	1.489	1.478
30	1.737	1.709	1.686	1.667	1.606	1.573	1.552	1.507	1.482	1.467	1.456
40	1.678	1.649	1.625	1.605	1.541	1.506	1.483	1.434	1.406	1.389	1.377
50	1.643	1.613	1.588	1.568	1.502	1.465	1.441	1.388	1.359	1.340	1.327
60	1.619	1.589	1.564	1.543	1.476	1.437	1.413	1.358	1.326	1.306	1.292
80	1.590	1.559	1.534	1.513	1.443	1.403	1.377	1.318	1.284	1.261	1.245
100	1.573	1.542	1.516	1.494	1.423	1.382	1.355	1.293	1.257	1.232	1.214
200	1.539	1.507	1.480	1.458	1.383	1.339	1.310	1.242	1.199	1.168	1.144
500	1.518	1.485	1.458	1.435	1.358	1.313	1.282	1.209	1.160	1.122	1.087
∞	1.505	1.471	1.444	1.421	1.342	1.295	1.263	1.185	1.130	1.082	1.008

Computation of critical value $k_\alpha = F_{1-\alpha;r_1,r_2,\lambda=0}$.

$\alpha = 0.05, \quad 1-\alpha = 0.95$

$r_2 \backslash r_1$	1	2	3	4	5	6	7	8	9	10	12
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41
3	10.13	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812	8.786	8.745
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999	5.964	5.912
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735	4.678
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099	4.060	4.000
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637	3.575
8	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438	3.388	3.347	3.284
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137	3.073
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978	2.913
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896	2.854	2.788
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753	2.687
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671	2.604
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646	2.602	2.534
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544	2.475
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.538	2.494	2.425
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548	2.494	2.450	2.381
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456	2.412	2.342
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423	2.378	2.308
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348	2.278
22	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397	2.342	2.297	2.226
24	4.260	3.403	3.009	2.776	2.621	2.508	2.423	2.355	2.300	2.255	2.183
26	4.225	3.369	2.975	2.743	2.587	2.474	2.388	2.321	2.265	2.220	2.148
28	4.196	3.340	2.947	2.714	2.558	2.445	2.359	2.291	2.236	2.190	2.118
30	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266	2.211	2.165	2.092
40	4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180	2.124	2.077	2.003
50	4.034	3.183	2.790	2.557	2.400	2.286	2.199	2.130	2.073	2.026	1.952
60	4.001	3.150	2.758	2.525	2.368	2.254	2.167	2.097	2.040	1.993	1.917
80	3.960	3.111	2.719	2.486	2.329	2.214	2.126	2.056	1.999	1.951	1.875
100	3.936	3.087	2.696	2.463	2.305	2.191	2.103	2.032	1.975	1.927	1.850
200	3.888	3.041	2.650	2.417	2.259	2.144	2.056	1.985	1.927	1.878	1.801
500	3.860	3.014	2.623	2.390	2.232	2.117	2.028	1.957	1.899	1.850	1.772
∞	3.842	2.996	2.605	2.372	2.214	2.099	2.010	1.939	1.880	1.831	1.752

Computation of critical value $k_\alpha = F_{1-\alpha;r_1,r_2,\lambda=0}$ (continued).

$\alpha = 0.05, \quad 1 - \alpha = 0.95$

$r_2 \backslash r_1$	14	16	18	20	30	40	50	100	200	500	∞
1	245.4	246.5	247.3	248.0	250.1	251.1	251.8	253.0	253.7	254.1	254.3
2	19.42	19.43	19.44	19.45	19.46	19.47	19.48	19.49	19.49	19.49	19.50
3	8.715	8.692	8.675	8.660	8.617	8.594	8.581	8.554	8.540	8.532	8.526
4	5.873	5.844	5.821	5.803	5.746	5.717	5.699	5.664	5.646	5.635	5.628
5	4.636	4.604	4.579	4.558	4.496	4.464	4.444	4.405	4.385	4.373	4.365
6	3.956	3.922	3.896	3.874	3.808	3.774	3.754	3.712	3.690	3.678	3.669
7	3.529	3.494	3.467	3.445	3.376	3.340	3.319	3.275	3.252	3.239	3.230
8	3.237	3.202	3.173	3.150	3.079	3.043	3.020	2.975	2.951	2.937	2.928
9	3.025	2.989	2.960	2.936	2.864	2.826	2.803	2.756	2.731	2.717	2.707
10	2.865	2.828	2.798	2.774	2.700	2.661	2.637	2.588	2.563	2.548	2.538
11	2.739	2.701	2.671	2.646	2.570	2.531	2.507	2.457	2.431	2.415	2.405
12	2.637	2.599	2.568	2.544	2.466	2.426	2.401	2.350	2.323	2.307	2.296
13	2.554	2.515	2.484	2.459	2.380	2.339	2.314	2.261	2.234	2.218	2.206
14	2.484	2.445	2.413	2.388	2.308	2.266	2.241	2.187	2.159	2.142	2.131
15	2.424	2.385	2.353	2.328	2.247	2.204	2.178	2.123	2.095	2.078	2.066
16	2.373	2.333	2.302	2.276	2.194	2.151	2.124	2.068	2.039	2.022	2.010
17	2.329	2.289	2.257	2.230	2.148	2.104	2.077	2.020	1.991	1.973	1.960
18	2.290	2.250	2.217	2.191	2.107	2.063	2.035	1.978	1.948	1.929	1.917
19	2.256	2.215	2.182	2.155	2.071	2.026	1.999	1.940	1.910	1.891	1.878
20	2.225	2.184	2.151	2.124	2.039	1.994	1.966	1.907	1.875	1.856	1.843
22	2.173	2.131	2.098	2.071	1.984	1.938	1.909	1.849	1.817	1.797	1.783
24	2.130	2.088	2.054	2.027	1.939	1.892	1.863	1.800	1.768	1.747	1.733
26	2.094	2.052	2.018	1.990	1.901	1.853	1.823	1.760	1.726	1.705	1.691
28	2.064	2.021	1.987	1.959	1.869	1.820	1.790	1.725	1.691	1.669	1.654
30	2.037	1.995	1.960	1.932	1.841	1.792	1.761	1.695	1.660	1.637	1.622
40	1.948	1.904	1.868	1.839	1.744	1.693	1.660	1.589	1.551	1.526	1.509
50	1.895	1.850	1.814	1.784	1.687	1.634	1.599	1.525	1.484	1.457	1.438
60	1.860	1.815	1.778	1.748	1.649	1.594	1.559	1.481	1.438	1.409	1.389
80	1.817	1.772	1.734	1.703	1.602	1.545	1.508	1.426	1.379	1.347	1.325
100	1.792	1.746	1.708	1.676	1.573	1.515	1.477	1.392	1.342	1.308	1.283
200	1.742	1.694	1.656	1.623	1.516	1.455	1.415	1.321	1.263	1.221	1.189
500	1.712	1.664	1.625	1.592	1.482	1.419	1.376	1.275	1.210	1.159	1.113
∞	1.692	1.644	1.604	1.571	1.459	1.394	1.350	1.244	1.170	1.107	1.010

Computation of critical value $k_\alpha = F_{1-\alpha;r_1,r_2,\lambda=0}$.

$\alpha = 0.025, \quad 1-\alpha = 0.975$

$r_2 \backslash r_1$	1	2	3	4	5	6	7	8	9	10	12
1	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3	968.6	976.7
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34
4	12.22	10.65	9.979	9.605	9.364	9.197	9.074	8.980	8.905	8.844	8.751
5	10.01	8.434	7.764	7.388	7.146	6.978	6.853	6.757	6.681	6.619	6.525
6	8.813	7.260	6.599	6.227	5.988	5.820	5.695	5.600	5.523	5.461	5.366
7	8.073	6.542	5.890	5.523	5.285	5.119	4.995	4.899	4.823	4.761	4.666
8	7.571	6.059	5.416	5.053	4.817	4.652	4.529	4.433	4.357	4.295	4.200
9	7.209	5.715	5.078	4.718	4.484	4.320	4.197	4.102	4.026	3.964	3.868
10	6.937	5.456	4.826	4.468	4.236	4.072	3.950	3.855	3.779	3.717	3.621
11	6.724	5.256	4.630	4.275	4.044	3.881	3.759	3.664	3.588	3.526	3.430
12	6.554	5.096	4.474	4.121	3.891	3.728	3.607	3.512	3.436	3.374	3.277
13	6.414	4.965	4.347	3.996	3.767	3.604	3.483	3.388	3.312	3.250	3.153
14	6.298	4.857	4.242	3.892	3.663	3.501	3.380	3.285	3.209	3.147	3.050
15	6.200	4.765	4.153	3.804	3.576	3.415	3.293	3.199	3.123	3.060	2.963
16	6.115	4.687	4.077	3.729	3.502	3.341	3.219	3.125	3.049	2.986	2.889
17	6.042	4.619	4.011	3.665	3.438	3.277	3.156	3.061	2.985	2.922	2.825
18	5.978	4.560	3.954	3.608	3.382	3.221	3.100	3.005	2.929	2.866	2.769
19	5.922	4.508	3.903	3.559	3.333	3.172	3.051	2.956	2.880	2.817	2.720
20	5.871	4.461	3.859	3.515	3.289	3.128	3.007	2.913	2.837	2.774	2.676
22	5.786	4.383	3.783	3.440	3.215	3.055	2.934	2.839	2.763	2.700	2.602
24	5.717	4.319	3.721	3.379	3.155	2.995	2.874	2.779	2.703	2.640	2.541
26	5.659	4.265	3.670	3.329	3.105	2.945	2.824	2.729	2.653	2.590	2.491
28	5.610	4.221	3.626	3.286	3.063	2.903	2.782	2.687	2.611	2.547	2.448
30	5.568	4.182	3.589	3.250	3.026	2.867	2.746	2.651	2.575	2.511	2.412
40	5.424	4.051	3.463	3.126	2.904	2.744	2.624	2.529	2.452	2.388	2.288
50	5.340	3.975	3.390	3.054	2.833	2.674	2.553	2.458	2.381	2.317	2.216
60	5.286	3.925	3.343	3.008	2.786	2.627	2.507	2.412	2.334	2.270	2.169
80	5.218	3.864	3.284	2.950	2.730	2.571	2.450	2.355	2.277	2.213	2.111
100	5.179	3.828	3.250	2.917	2.696	2.537	2.417	2.321	2.244	2.179	2.077
200	5.100	3.758	3.182	2.850	2.630	2.472	2.351	2.256	2.178	2.113	2.010
500	5.054	3.716	3.142	2.811	2.592	2.434	2.313	2.217	2.139	2.074	1.971
∞	5.024	3.689	3.116	2.786	2.567	2.408	2.288	2.192	2.114	2.048	1.945

Computation of critical value $k_\alpha = F_{1-\alpha;r_1,r_2,\lambda=0}$ (continued).

$\alpha = 0.025, \quad 1-\alpha = 0.975$

$r_2 \backslash r_1$	14	16	18	20	30	40	50	100	200	500	∞
1	982.5	986.9	990.3	993.1	1001.	1006.	1008.	1013.	1016.	1017.	1018.
2	39.43	39.44	39.44	39.45	39.46	39.47	39.48	39.49	39.49	39.50	39.50
3	14.28	14.23	14.20	14.17	14.08	14.04	14.01	13.96	13.93	13.91	13.90
4	8.684	8.633	8.592	8.560	8.461	8.411	8.381	8.319	8.289	8.270	8.257
5	6.456	6.403	6.362	6.329	6.227	6.175	6.144	6.080	6.048	6.028	6.015
6	5.297	5.244	5.202	5.168	5.065	5.012	4.980	4.915	4.882	4.862	4.849
7	4.596	4.543	4.501	4.467	4.362	4.309	4.276	4.210	4.176	4.156	4.142
8	4.130	4.076	4.034	3.999	3.894	3.840	3.807	3.739	3.705	3.684	3.670
9	3.798	3.744	3.701	3.667	3.560	3.505	3.472	3.403	3.368	3.347	3.333
10	3.550	3.496	3.453	3.419	3.311	3.255	3.221	3.152	3.116	3.094	3.080
11	3.359	3.304	3.261	3.226	3.118	3.061	3.027	2.956	2.920	2.898	2.883
12	3.206	3.152	3.108	3.073	2.963	2.906	2.871	2.800	2.763	2.740	2.725
13	3.082	3.027	2.983	2.948	2.837	2.780	2.744	2.671	2.634	2.611	2.596
14	2.979	2.923	2.879	2.844	2.732	2.674	2.638	2.565	2.526	2.503	2.487
15	2.891	2.836	2.792	2.756	2.644	2.585	2.549	2.474	2.435	2.411	2.395
16	2.817	2.761	2.717	2.681	2.568	2.509	2.472	2.396	2.357	2.333	2.316
17	2.753	2.697	2.652	2.616	2.502	2.442	2.405	2.329	2.289	2.264	2.248
18	2.696	2.640	2.596	2.559	2.445	2.384	2.347	2.269	2.229	2.204	2.187
19	2.647	2.591	2.546	2.509	2.394	2.333	2.295	2.217	2.176	2.150	2.133
20	2.603	2.547	2.501	2.464	2.349	2.287	2.249	2.170	2.128	2.103	2.085
22	2.528	2.472	2.426	2.389	2.272	2.210	2.171	2.090	2.047	2.021	2.003
24	2.468	2.411	2.365	2.327	2.209	2.146	2.107	2.024	1.981	1.954	1.935
26	2.417	2.360	2.314	2.276	2.157	2.093	2.053	1.969	1.925	1.897	1.878
28	2.374	2.317	2.270	2.232	2.112	2.048	2.007	1.922	1.877	1.848	1.829
30	2.338	2.280	2.233	2.195	2.074	2.009	1.968	1.882	1.835	1.806	1.787
40	2.213	2.154	2.107	2.068	1.943	1.875	1.832	1.741	1.691	1.659	1.637
50	2.140	2.081	2.033	1.993	1.866	1.796	1.752	1.656	1.603	1.569	1.545
60	2.093	2.033	1.985	1.944	1.815	1.744	1.699	1.599	1.543	1.507	1.482
80	2.035	1.974	1.925	1.884	1.752	1.679	1.632	1.527	1.467	1.428	1.400
100	2.000	1.939	1.890	1.849	1.715	1.640	1.592	1.483	1.420	1.378	1.347
200	1.932	1.870	1.820	1.778	1.640	1.562	1.511	1.393	1.320	1.269	1.229
500	1.892	1.830	1.779	1.736	1.596	1.515	1.462	1.336	1.254	1.192	1.137
∞	1.866	1.803	1.752	1.709	1.566	1.484	1.429	1.296	1.206	1.128	1.012

Computation of critical value $k_\alpha = F_{1-\alpha;r_1,r_2,\lambda=0}$.

$\alpha = 0.01, \quad 1 - \alpha = 0.99$

$r_2 \backslash r_1$	1	2	3	4	5	6	7	8	9	10	12
1	4052.	4999.	5403.	5625.	5764.	5859.	5928.	5981.	6022.	6056.	6106.
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.888
6	13.75	10.92	9.780	9.148	8.746	8.466	8.260	8.102	7.976	7.874	7.718
7	12.25	9.547	8.451	7.847	7.460	7.191	6.993	6.840	6.719	6.620	6.469
8	11.26	8.649	7.591	7.006	6.632	6.371	6.178	6.029	5.911	5.814	5.667
9	10.56	8.022	6.992	6.422	6.057	5.802	5.613	5.467	5.351	5.257	5.111
10	10.04	7.559	6.552	5.994	5.636	5.386	5.200	5.057	4.942	4.849	4.706
11	9.646	7.206	6.217	5.668	5.316	5.069	4.886	4.744	4.632	4.539	4.397
12	9.330	6.927	5.953	5.412	5.064	4.821	4.640	4.499	4.388	4.296	4.155
13	9.074	6.701	5.739	5.205	4.862	4.620	4.441	4.302	4.191	4.100	3.960
14	8.862	6.515	5.564	5.035	4.695	4.456	4.278	4.140	4.030	3.939	3.800
15	8.683	6.359	5.417	4.893	4.556	4.318	4.142	4.004	3.895	3.805	3.666
16	8.531	6.226	5.292	4.773	4.437	4.202	4.026	3.890	3.780	3.691	3.553
17	8.400	6.112	5.185	4.669	4.336	4.102	3.927	3.791	3.682	3.593	3.455
18	8.285	6.013	5.092	4.579	4.248	4.015	3.841	3.705	3.597	3.508	3.371
19	8.185	5.926	5.010	4.500	4.171	3.939	3.765	3.631	3.523	3.434	3.297
20	8.096	5.849	4.938	4.431	4.103	3.871	3.699	3.564	3.457	3.368	3.231
22	7.945	5.719	4.817	4.313	3.988	3.758	3.587	3.453	3.346	3.258	3.121
24	7.823	5.614	4.718	4.218	3.895	3.667	3.496	3.363	3.256	3.168	3.032
26	7.721	5.526	4.637	4.140	3.818	3.591	3.421	3.288	3.182	3.094	2.958
28	7.636	5.453	4.568	4.074	3.754	3.528	3.358	3.226	3.120	3.032	2.896
30	7.562	5.390	4.510	4.018	3.699	3.473	3.304	3.173	3.067	2.979	2.843
40	7.314	5.179	4.313	3.828	3.514	3.291	3.124	2.993	2.888	2.801	2.665
50	7.171	5.057	4.199	3.720	3.408	3.186	3.020	2.890	2.785	2.698	2.562
60	7.077	4.977	4.126	3.649	3.339	3.119	2.953	2.823	2.718	2.632	2.496
80	6.963	4.881	4.036	3.563	3.255	3.036	2.871	2.742	2.637	2.551	2.415
100	6.895	4.824	3.984	3.513	3.206	2.988	2.823	2.694	2.590	2.503	2.368
200	6.763	4.713	3.881	3.414	3.110	2.893	2.730	2.601	2.497	2.411	2.275
500	6.686	4.648	3.821	3.357	3.054	2.838	2.675	2.547	2.443	2.356	2.220
∞	6.635	4.605	3.782	3.319	3.017	2.802	2.640	2.511	2.408	2.321	2.185

Computation of critical value $k_\alpha = F_{1-\alpha;r_1,r_2,\lambda=0}$ (continued).

$\alpha = 0.01, \quad 1-\alpha = 0.99$

$r_2 \backslash r_1$	14	16	18	20	30	40	50	100	200	500	∞
1	6143.	6170.	6192.	6209.	6261.	6287.	6303.	6334.	6350.	6360.	6366.
2	99.43	99.44	99.44	99.45	99.47	99.47	99.48	99.49	99.49	99.50	99.50
3	26.92	26.83	26.75	26.69	26.50	26.41	26.35	26.24	26.18	26.15	26.13
4	14.25	14.15	14.08	14.02	13.84	13.75	13.69	13.58	13.52	13.49	13.46
5	9.770	9.680	9.610	9.553	9.379	9.291	9.238	9.130	9.075	9.042	9.021
6	7.605	7.519	7.451	7.396	7.229	7.143	7.091	6.987	6.934	6.902	6.880
7	6.359	6.275	6.209	6.155	5.992	5.908	5.858	5.755	5.702	5.671	5.650
8	5.559	5.477	5.412	5.359	5.198	5.116	5.065	4.963	4.911	4.880	4.859
9	5.005	4.924	4.860	4.808	4.649	4.567	4.517	4.415	4.363	4.332	4.311
10	4.601	4.520	4.457	4.405	4.247	4.165	4.115	4.014	3.962	3.930	3.909
11	4.293	4.213	4.150	4.099	3.941	3.860	3.810	3.708	3.656	3.624	3.603
12	4.052	3.972	3.909	3.858	3.701	3.619	3.569	3.467	3.414	3.382	3.361
13	3.857	3.778	3.716	3.665	3.507	3.425	3.375	3.272	3.219	3.187	3.166
14	3.698	3.619	3.556	3.505	3.348	3.266	3.215	3.112	3.059	3.026	3.004
15	3.564	3.485	3.423	3.372	3.214	3.132	3.081	2.977	2.923	2.891	2.869
16	3.451	3.372	3.310	3.259	3.101	3.018	2.967	2.863	2.808	2.775	2.753
17	3.353	3.275	3.212	3.162	3.003	2.920	2.869	2.764	2.709	2.676	2.653
18	3.269	3.190	3.128	3.077	2.919	2.835	2.784	2.678	2.623	2.589	2.566
19	3.195	3.116	3.054	3.003	2.844	2.761	2.709	2.602	2.547	2.512	2.489
20	3.130	3.051	2.989	2.938	2.778	2.695	2.643	2.535	2.479	2.445	2.421
22	3.019	2.941	2.879	2.827	2.667	2.583	2.531	2.422	2.365	2.329	2.306
24	2.930	2.852	2.789	2.738	2.577	2.492	2.440	2.329	2.271	2.235	2.211
26	2.857	2.778	2.715	2.664	2.503	2.417	2.364	2.252	2.193	2.156	2.132
28	2.795	2.716	2.653	2.602	2.440	2.354	2.300	2.187	2.127	2.090	2.064
30	2.742	2.663	2.600	2.549	2.386	2.299	2.245	2.131	2.070	2.032	2.006
40	2.563	2.484	2.421	2.369	2.203	2.114	2.058	1.938	1.874	1.833	1.805
50	2.461	2.382	2.318	2.265	2.098	2.007	1.949	1.825	1.757	1.713	1.683
60	2.394	2.315	2.251	2.198	2.028	1.936	1.877	1.749	1.678	1.633	1.601
80	2.313	2.233	2.169	2.115	1.944	1.849	1.788	1.655	1.579	1.530	1.494
100	2.265	2.185	2.120	2.067	1.893	1.797	1.735	1.598	1.518	1.466	1.427
200	2.172	2.091	2.026	1.971	1.794	1.694	1.629	1.481	1.391	1.328	1.279
500	2.117	2.036	1.970	1.915	1.735	1.633	1.566	1.408	1.308	1.232	1.165
∞	2.082	2.000	1.934	1.878	1.697	1.592	1.523	1.358	1.248	1.153	1.015

Calculation in Matlab:

$$k_\alpha = \texttt{finv}(1-\alpha, r_1, r_2)$$

Example ($\alpha = 0.05$, one-sided):

$$k_\alpha = 2.774 = \texttt{finv}(1-0.05, 20, 10)$$

$$k_{1-\alpha} = 0.360 = \frac{1}{\texttt{finv}(1-0.05, 20, 10)} = \texttt{finv}(0.05, 10, 20)$$

B.6 Relation between F-Distribution and other distrubutions

χ^2-Distribution	$\chi^2_{1-\alpha;r} = rF_{1-\alpha;r,\infty}$
Standard Normal Distribution	$z_{1-\alpha/2} = \sqrt{F_{1-\alpha;1,\infty}}$
t-Distribution	$t_{1-\alpha/2;r} = \sqrt{F_{1-\alpha;1,r}}$
τ-Distribution	$\tau_{1-\alpha;q,r-q,\lambda} = \sqrt{\frac{rF_{1-\alpha;q,r-q,\lambda}}{r-q+qF_{1-\alpha;q,r-q,\lambda}}}$

C Book recommendations and other material

C.1 Scientific books

- *Anderson, James M. and Edward M. Mikhail*
  Surveying. Theory and Practice
  7th edition
  McGraw-Hill, 1998
  ISBN 0-07-015914-9
- *Benning, Wilhelm*
  Statistik in Geodäsie, Geoinformation und Bauwesen
  4., überarbeitete und erweiterte Auflage
  Wichmann, 2011
  ISBN 978-3-87907-512-6
- *Beucher, Ottmar*
  Wahrscheinlichkeitsrechnung und Statistik mit MATLAB
  Springer, 2007
  ISBN 978-3-540-72155-0
- *Caspary, Wilhelm and Klaus Wichmann*
  Auswertung von Messdaten. Statistische Methoden für Geo- und Ingenieurwissenschaften
  Oldenbourg, 2007
  ISBN 978-3-486-58351-9
- *Chatterjee, Samprit and Ali S. Hadi*
  Regresson Analysis by Example
  Fourth Edition
  John Wiley & Sons, Inc., 2006
  ISBN 13 978-0-471-74696-6
- *Ghilani Charles D.*
  Adjustment Computations. Spatial Data Analysis
  Fifth Edition
  John Wiley & Sons, Inc., 2010
  ISBN 978-0-470-46491-5
- *Grafarend, Erik W.*
  Linear and Nonlinear Models – Fixed Effects, Random Effects, and Mixed Models

de Gruyter, 2006
ISBN 978-3-11-016216-5

- *Jäger, Reiner et al.*
  Klassische und robuste Ausgleichungsverfahren. Ein Leitfaden für Ausbildung und Praxis von Geodäten und Geoinformatikern
  Wichmann, 2005
  ISBN 3-87907-370-8

- *Koch, Karl-Rudolf*
  Parameter Estimation and Hypothesis Testing in Linear Models
  2^{nd} updated and enlarged edition
  Springer, 1999
  ISBN 978-3-540-65257-1

- *Koch, Karl-Rudolf*
  Parameterschätzung und Hypothesentests in linearen Modellen
  Dritte, bearbeitete Auflage
  Dümmlers, 1997
  ISBN 3-427-78923-3

- *Lay, David C.*
  Linear Algebra and its Applications
  3^{rd} edition
  Addison-Wesley Publishing Company, 2003
  ISBN 0-201-70970-8

- *Magnus, Jan R. and Heinz Neudecker*
  Matrix Differential Calculus with Applications in Statistics and Ecnonometrics
  John Wiley & Sons Ltd., 1988
  ISBN 0-471-91516-5

- *Mikhail, Edward M. and Fritz Ackermann*
  Observations and Least Squares
  IEP-A Dun-Donnelley Publisher, 1976
  ISBN 0-7002-2481-5

- *Niemeier, Wolfgang*
  Ausgleichungsrechnung, statistische Auswertemethoden
  2., überarbeitete und erweiterte Auflage
  de Gruyter, 2008
  ISBN 978-3-11-019055-7

- *Strang, Gilbert*
  Linear Algebra and its Applications
  4^{th} edition
  Harcourt Brace & Company, 2005
  ISBN 0030105676

- *Strang, Gilbert*
  Introduction to Linear Algebra
  4^{th} edition
  Wellesley-Cambridge Press, 2009
  ISBN 078-0-9802327-1-4
- *Strang, Gilbert and Kai Borre*
  Linear Algebra, Geodesy, and GPS
  Wellesley-Cambridge Press, 1997
  ISBN 0-9614088-6-3
- *Teunissen, P. J. G.*
  Adjustment theory – an introduction
  Delft University Press, 2003
  ISBN 13 978-90-407-1974-5
- *Teunissen, P. J. G.*
  Testing theory – an introduction
  Delft University Press, 2000–2006
  ISBN 13 978-90-407-1975-2
- *Teunissen, P. J. G.*
  Dynamic data processing – recursive least-squares
  Delft University Press, 2001
  ISBN 13 978-90-407-1976-9
- *Wolf, H.*
  Ausgleichungsrechnung II, Aufgaben und Beispiele zur praktischen Anwendung
  Ferd. Dümmlers Verlag, 1979
  ISBN 3-427-78361-8

C.2 Popular science books, literature

- *Sobel, Dava*
  Longitude: The True Story of a Lone Genius Who Solved the Greatest Scientific Problem of His Time
  Fourth Estate, 1996
  ISBN 1-85702-502-4
  Deutsche Übersetzung:
  Längengrad. Die wahre Geschichte eines einsamen Genies, welches das größte wissenschaftliche Problem seiner Zeit löste
  Berliner Taschenbuch Verlag, 2003
  ISBN 3-8333-0271-2
- *Kehlmann, Daniel*
  Die Vermessung der Welt

Rowohlt, Reinbek, 2005
ISBN 3-498-03528-2

C.3 Other material

- *Krumm, Friedhelm*
  Geodetic Network Adjustment Examples
  `http://www.gis.uni-stuttgart.de/res/study/addons/adjustment_examples.pdf`
- *Strang, Gilbert*
  Linear Algebra, MIT Course 18.06
  `http://ocw.mit.edu/courses/mathematics/18-06-linear-algebra-spring-2010/`