Lecture Notes

Adjustment Theory

Nico Sneeuw, Friedhelm Krumm
Geodätisches Institut
Universität Stuttgart
http://www.uni-stuttgart.de/gi

Rev. 4.30

April 13, 2012

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

Contents

1. Introduction

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

Ausgleichungs-rechnung

(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:

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

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

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

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.

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

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:

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

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

iii) *Random error*. Zufallsfehler

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.

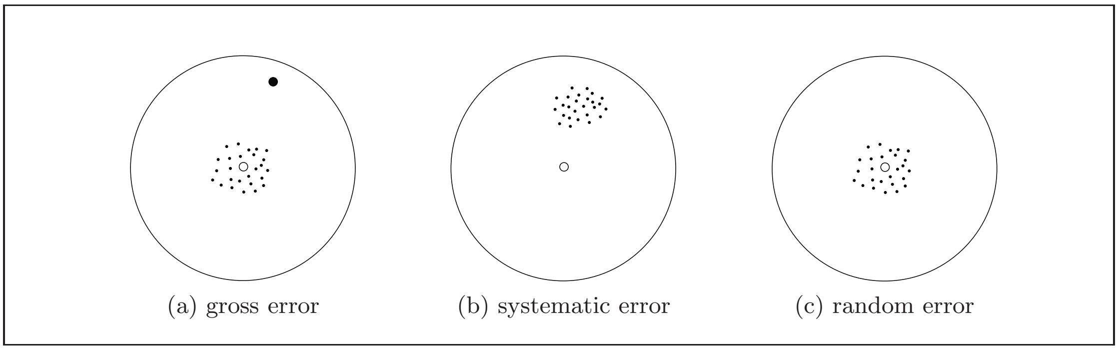


Figure 1.1.: Different types of errors.

Whatever the type, errors are stochastic quantities. Thus, the vector e is a (m-dimensional)
Zufallsvariable *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:

$$\underset{27\times1}{y} = \underset{27\times3}{A} \; \underset{3\times1}{x}$$

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

Mayer's adjustment strategy:

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

[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.

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

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 ⇒ Given up!
Euler was mathematician → "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:

[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.

$$\underset{24\times1}{y} = \underset{24\times4}{A}\ \underset{4\times1}{x}$$

$$\underset{4\times24}{L}\ \underset{24\times1}{y} = \underset{4\times24}{L}\ \underset{24\times4}{A}\ \underset{4\times1}{x}$$

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

$$\underset{4\times24}{L} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & 0 & 1 & 1 & 0 & 0 & -1 & 0 & 0 & 1 & 1 & 0 & 0 & -1 & 0 & 0 & 1 & 1 & 0 & -1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & -1 & -1 & 0 & 1 & 1 & 0 & 0 & -1 & -1 & 0 & 1 & 1 & 0 & 0 & -1 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}$$

Method of least absolute deviation – 1760

Roger Boscovich[4]

Ellipticity of the Earth
5 Observations (Quito, Kapstadt, Rom, Paris, Lappland)
2 unknowns

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

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.
$\Rightarrow$ 10 systems of equations (2×2)
$\Rightarrow$ 10 solutions
Comparison of results.

His result: gross variations of the ellipticity $\Rightarrow$ Reject the ellipsoidal hypothesis.

[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.

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

Methode der kleinsten Quadrate

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 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:

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

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.

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

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:

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

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 $\rightarrow$ 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:

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

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.

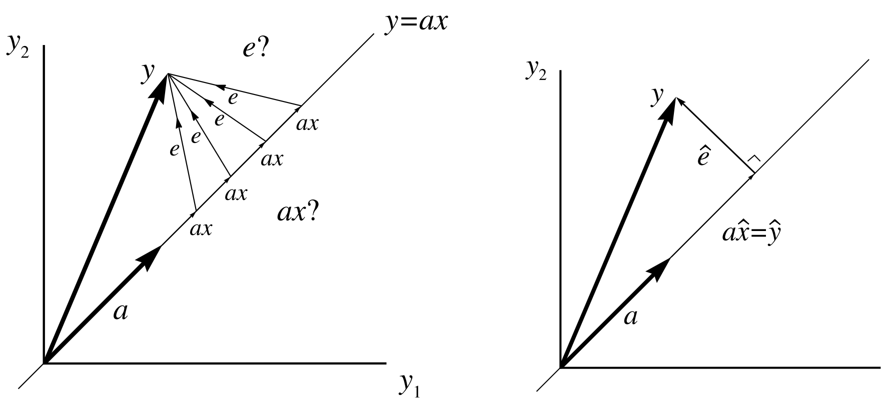
Given a limited measurement precision the latter scenario will be more likely. Let's therefore take into account measurement errors e.

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

A geometric view

Spaltenraum

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, 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.



(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.

Adjustment of observations is about finding the optimal e and x. An intuitive choice for "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:

$$\text{orthogonality } \hat{e} \perp a \qquad a^\mathsf{T}\hat{e} = 0 \tag{2.3a}$$

$$a^\mathsf{T}(y - a\hat{x}) = 0 \tag{2.3b}$$

$$\text{normal equations} \qquad a^\mathsf{T}a\hat{x} = a^\mathsf{T}y \tag{2.3c}$$

$$\text{LS estimate of } x \qquad \hat{x} = (a^\mathsf{T}a)^{-1}a^\mathsf{T}y \tag{2.3d}$$

$$\text{LS estimate of } y \qquad \hat{y} = a\hat{x} \;=\; a(a^\mathsf{T}a)^{-1}a^\mathsf{T}y \tag{2.3e}$$

$$\text{LS estimate of } e \qquad \hat{e} = y - \hat{y} \;=\; [I - a(a^\mathsf{T}a)^{-1}a^\mathsf{T}]y \tag{2.3f}$$

$$\text{sum square residuals} \qquad \hat{e}^\mathsf{T}\hat{e} = y^\mathsf{T}[I - a(a^\mathsf{T}a)^{-1}a^\mathsf{T}]y \tag{2.3g}$$

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:

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

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're 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:

$$\hat{y} = a(a^\mathsf{T}a)^{-1}a^\mathsf{T}y \tag{2.5a}$$

$$\Longleftrightarrow \qquad \hat{y} = P_a y \tag{2.5b}$$

$$\text{with} \qquad P_a = a(a^\mathsf{T}a)^{-1}a^\mathsf{T}\,. \tag{2.5c}$$

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

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

It projects onto the line ax along a direction orthogonal to a. With this projection in mind, the property $P_aP_a = P_a$ 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:

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

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:

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

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 $P_a^\perp$: zerlegen

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

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):

$$(a^\mathsf{T}a)\hat{x} = a^\mathsf{T}y \quad\longleftrightarrow\quad 2\hat{x} = y_1 + y_2$$

$$\hat{x} = (a^\mathsf{T}a)^{-1}a^\mathsf{T}y \quad\longleftrightarrow\quad \hat{x} = \frac{1}{2}(y_1 + y_2) \qquad \text{(average)}$$

$$\hat{y} = a(a^\mathsf{T}a)^{-1}a^\mathsf{T}y \quad\longleftrightarrow\quad \begin{pmatrix}\hat{y}_1\\ \hat{y}_2\end{pmatrix} = \frac{1}{2}\begin{pmatrix} y_1 + y_2\\ y_1 + y_2\end{pmatrix}$$

$$\hat{e} = y - \hat{y} \quad\longleftrightarrow\quad \begin{pmatrix}\hat{e}_1\\ \hat{e}_2\end{pmatrix} = \frac{1}{2}\begin{pmatrix} y_1 - y_2\\ -y_1 + y_2\end{pmatrix} \qquad \text{(error distribution)}$$

$$\hat{e}^\mathsf{T}\hat{e} \quad\longleftrightarrow\quad \frac{1}{2}(y_1 - y_2)^2 \qquad \text{(least squares)}$$

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:

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

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

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

Bedingungsgleichung

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.

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.

Schätzer

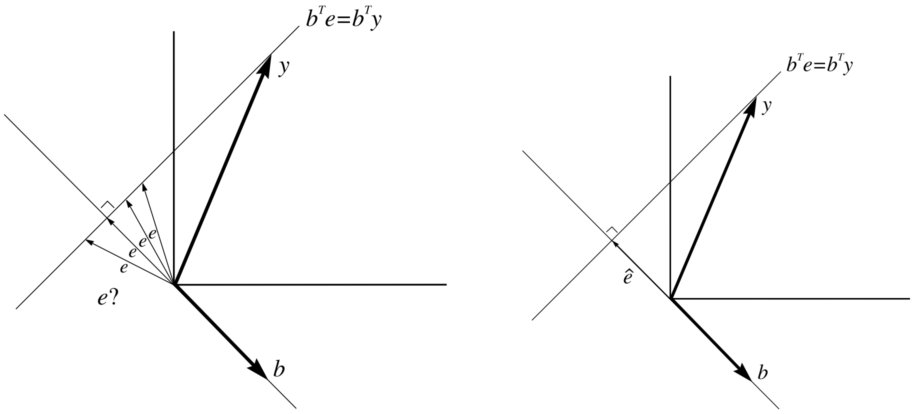
Geometrically, $\hat{e}$ is achieved by projecting y onto a line through b. Knowing the definition of the projectors from the previous section, we here define the following *estimates* by using the projector P_b:

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

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

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

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 $\hat{\lambda}$.

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

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.2. Instead, we will solve it stepwise. First, rewrite the first line:

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

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:

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

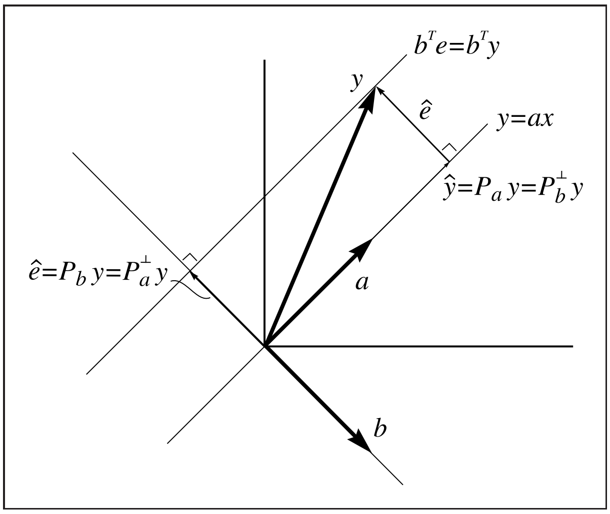


Figure 2.3: Least squares adjustment with observation equations and with condition equations in terms of the projectors P_a and P_b.

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:

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

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 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:

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

Geometry

Absolutgliedvektor

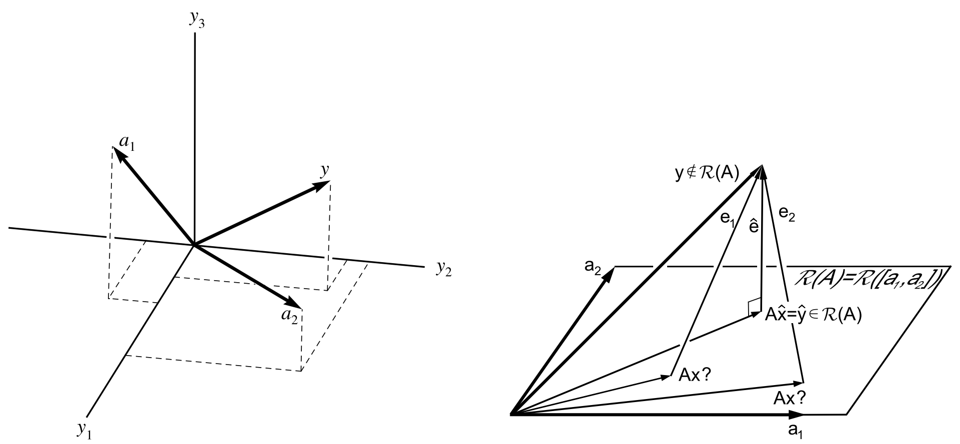
$y = Ax + e$ is the multidimensional extension of $y = ax + e$ with given (reduced) *vector of observations* y.
We split A in its n column vectors $\underset{m\times 1}{a_i}$, $i = 1, \ldots, n$

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

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

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.

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

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

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

$(A^{\mathsf{T}}A)^{-1}$ exists *iff* rank $A = n = \text{rank}(A^{\mathsf{T}}A)$ genau dann, wenn

Calculus

$$
\begin{aligned}
\mathcal{L}_A(x) &= \frac{1}{2}e^\mathsf{T}e \\
&= \frac{1}{2}(y-Ax)^\mathsf{T}(y-Ax) \\
&= \frac{1}{2}y^\mathsf{T}y - \frac{1}{2}y^\mathsf{T}Ax - \frac{1}{2}x^\mathsf{T}A^\mathsf{T}y + \frac{1}{2}x^\mathsf{T}A^\mathsf{T}Ax \\
&\xrightarrow{x} \min \\
\frac{\partial\mathcal{L}}{\partial x}(\hat{x}) = 0 &\Longrightarrow \hat{e} = y - \hat{y} = [I - A(A^\mathsf{T}A)^{-1}A^\mathsf{T}]y = P_A^\perp y
\end{aligned}
$$

$P_A^\perp$ 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

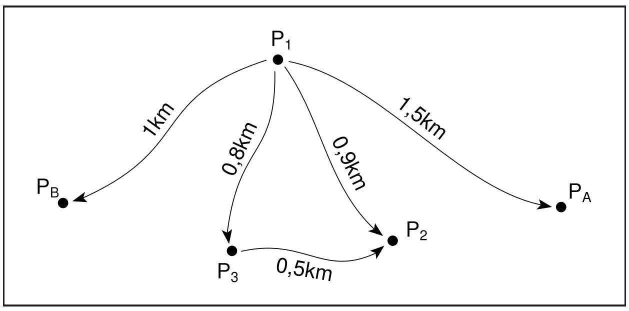


Figure 3.2.

$$
\begin{aligned}
h_{1\mathrm{B}} &= H_{\mathrm{B}} - H_1 + e_{1\mathrm{B}} \\
h_{13} &= H_3 - H_1 + e_{13} \\
h_{12} &= H_2 - H_1 + e_{12} \\
h_{32} &= H_2 - H_3 + e_{32} \\
h_{1\mathrm{A}} &= H_{\mathrm{A}} - H_1 + e_{1\mathrm{A}}
\end{aligned}
$$

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

In matrix notation:

$$
\begin{pmatrix} h_{1\mathrm{B}} \\ h_{13} \\ h_{12} \\ h_{32} \\ h_{1\mathrm{A}} \end{pmatrix}
=
\begin{pmatrix} -1 & 0 & 0 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 0 \end{pmatrix}
\begin{pmatrix} H_1 \\ H_2 \\ H_3 \end{pmatrix}
+
\begin{pmatrix} H_{\mathrm{B}} \\ 0 \\ 0 \\ 0 \\ H_{\mathrm{A}} \end{pmatrix}
+
\begin{pmatrix} e_{1\mathrm{B}} \\ e_{13} \\ e_{12} \\ e_{32} \\ e_{1\mathrm{A}} \end{pmatrix}
\qquad (\text{or } y = Ax + e)
$$

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

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 $x_{\mathrm{hom}} \neq 0$, 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.

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

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

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

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

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

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 $\mathrm{d} = \dim \mathcal{N}(A)$ 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:

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

Solution approach 2: augment solution space

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

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

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

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

$AD = 0$, 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

$$\begin{aligned}
\mathcal{L}_D(x, \lambda) &= \frac{1}{2} e^\mathsf{T} e + \lambda(D^\mathsf{T} x - c) \\
&= \frac{1}{2} y^\mathsf{T} y - y^\mathsf{T} A x + \frac{1}{2} x^\mathsf{T} A^\mathsf{T} A x + \lambda(D^\mathsf{T} x - c) \\
\frac{\partial \mathcal{L}_D}{\partial x} &= -A^\mathsf{T} y + A^\mathsf{T} A x + D\lambda = 0 \\
\frac{\partial \mathcal{L}_D}{\partial \lambda} &= D^\mathsf{T} x - c = 0 \\
&\Longrightarrow \underset{(n+d)\times(n+d)}{\begin{pmatrix} A^\mathsf{T} A & D \\ D^\mathsf{T} & 0 \end{pmatrix}} \underset{(n+d)\times 1}{\begin{pmatrix} \hat{x} \\ \hat{\lambda} \end{pmatrix}} = \begin{pmatrix} A^\mathsf{T} y \\ c \end{pmatrix} \Longrightarrow M\hat{z} = v
\end{aligned}$$

E. g.

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

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

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

$$\Longrightarrow M \text{ regular} \Rightarrow \hat{z} = M^{-1} v$$
$$\hat{x} = N^{-1} \langle A^\mathsf{T} y + Dc - \{D(D^\mathsf{T} N^{-1} D)^{-1} [D^\mathsf{T} N^{-1} A^\mathsf{T} y + (D^\mathsf{T} N^{-1} D - I)c]\} \rangle$$
$$N := A^\mathsf{T} A + DD^\mathsf{T}$$

3.3. Linearization of non-linear observation equations

General 1-D-formulation

$$\begin{aligned} y &= f(x), \quad x = x_0 + \Delta x \\ &= f(x_0) + \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x_0} (x - x_0) + \underbrace{\frac{1}{2} \left.\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}\right|_{x_0} (x - x_0)^2 + \ldots}_{\text{negligible if } x - x_0 \text{ small}} \end{aligned}$$

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

$$\underbrace{\Delta y = \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_0 (x - x_0)}_{\text{linear model}} + \underbrace{\mathcal{O}(\Delta x^2)}_{\substack{\text{terms of higher order} \\ = \text{ model errors}}}, \quad \Delta x := x - x_0$$

General multi-D formulation

$$\begin{aligned} y_i &= f_i(x_j), \quad i = 1, ..., m; j = 1, ..., n \\ x_{j,0} &\longrightarrow y_{i,0} = f_i(x_{j,0}) \end{aligned}$$

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

Terms of second order and higher have been neglected.

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

Planar distance observation:

$$s_{\mathrm{ij}} = \sqrt{(x_{\mathrm{j}} - x_{\mathrm{i}})^2 + (y_{\mathrm{j}} - y_{\mathrm{i}})^2} \xrightarrow{?} y = Ax$$

answer: linearize, Taylor series expansion Linearization of planar distance observation equation (given Taylor point of expansion is $x_{\mathrm{i}}^0, y_{\mathrm{i}}^0, x_{\mathrm{j}}^0, y_{\mathrm{j}}^0$ = approximate values of unknown point coordinates); explicit differentiation

$$\text{“measured”}\ s_{\mathrm{ij}} = \sqrt{(x_{\mathrm{j}} - x_{\mathrm{i}})^2 + (y_{\mathrm{j}} - y_{\mathrm{i}})^2} = \sqrt{x_{\mathrm{ij}}^2 + y_{\mathrm{ij}}^2}$$

$$x_{\mathrm{i}} = x_{\mathrm{i}}^0 + \Delta x_{\mathrm{i}}, \quad y_{\mathrm{i}} = y_{\mathrm{i}}^0 + \Delta y_{\mathrm{i}},$$

$$x_{\mathrm{j}} = x_{\mathrm{j}}^0 + \Delta x_{\mathrm{j}}, \quad y_{\mathrm{j}} = y_{\mathrm{j}}^0 + \Delta y_{\mathrm{j}}$$

$$s_{\mathrm{ij}} = \sqrt{\Big(x_{\mathrm{j}}^0 + \Delta x_{\mathrm{j}} - \big(x_{\mathrm{i}}^0 + \Delta x_{\mathrm{i}}\big)\Big)^2 + \Big(y_{\mathrm{j}}^0 + \Delta y_{\mathrm{j}} - \big(y_{\mathrm{i}}^0 + \Delta y_{\mathrm{i}}\big)\Big)^2}$$

$$= \underbrace{\sqrt{\Big(x_{\mathrm{j}}^0 - x_{\mathrm{i}}^0\Big)^2 + \Big(y_{\mathrm{j}}^0 - y_{\mathrm{i}}^0\Big)^2}}_{s_{\mathrm{ij}}^0 \quad \text{(distance from approximate coordinates)}} +$$

$$+ \left.\frac{\partial s_{\mathrm{ij}}}{\partial x_{\mathrm{i}}}\right|_0 \Delta x_{\mathrm{i}} + \left.\frac{\partial s_{\mathrm{ij}}}{\partial x_{\mathrm{j}}}\right|_0 \Delta x_{\mathrm{j}} + \left.\frac{\partial s_{\mathrm{ij}}}{\partial y_{\mathrm{i}}}\right|_0 \Delta y_{\mathrm{i}} + \left.\frac{\partial s_{\mathrm{ij}}}{\partial y_{\mathrm{j}}}\right|_0 \Delta y_{\mathrm{j}}$$

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

$$\frac{\partial s_{\mathrm{ij}}}{\partial x_{\mathrm{j}}} = +\frac{x_{\mathrm{j}} - x_{\mathrm{i}}}{s_{\mathrm{ij}}}, \qquad \frac{\partial s_{\mathrm{ij}}}{\partial y_{\mathrm{i}}} = -\frac{y_{\mathrm{j}} - y_{\mathrm{i}}}{s_{\mathrm{ij}}}, \qquad \frac{\partial s_{\mathrm{ij}}}{\partial y_{\mathrm{j}}} = +\frac{y_{\mathrm{j}} - y_{\mathrm{A}}}{s_{\mathrm{ij}}}$$

$$\Longrightarrow \Delta s_{\mathrm{ij}} := \underbrace{s_{\mathrm{ij}} - s^0_{\mathrm{ij}}}_{\text{“reduced observation”}} = \begin{pmatrix} -\frac{x^0_{\mathrm{j}}-x^0_{\mathrm{i}}}{s^0_{\mathrm{ij}}} & -\frac{y^0_{\mathrm{j}}-y^0_{\mathrm{i}}}{s^0_{\mathrm{ij}}} & \frac{x^0_{\mathrm{j}}-x^0_{\mathrm{i}}}{s^0_{\mathrm{ij}}} & \frac{y^0_{\mathrm{j}}-y^0_{\mathrm{i}}}{s^0_{\mathrm{ij}}} \end{pmatrix} \begin{pmatrix} \Delta x_{\mathrm{i}} \\ \Delta y_{\mathrm{i}} \\ \Delta x_{\mathrm{j}} \\ \Delta y_{\mathrm{j}} \end{pmatrix}$$

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

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

Depart from $s^2_{\mathrm{ij}} = (x_{\mathrm{j}} - x_{\mathrm{i}})^2 + (y_{\mathrm{j}} - y_{\mathrm{i}})^2$ instead from s_{ij} and calculate the total differential:

$$\not{2} s_{\mathrm{ij}}\,\mathrm{d}s_{\mathrm{ij}} = \not{2}\,(x_{\mathrm{j}} - x_{\mathrm{i}})\,(\,\mathrm{d}x_{\mathrm{j}} - \,\mathrm{d}x_{\mathrm{i}}) + \not{2}\,(y_{\mathrm{j}} - y_{\mathrm{i}})\,(\,\mathrm{d}y_{\mathrm{j}} - \,\mathrm{d}y_{\mathrm{i}})$$

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

$$\Delta s_{\mathrm{ij}} := s_{\mathrm{ij}} - s^0_{\mathrm{ij}} = \frac{x^0_{\mathrm{j}} - x^0_{\mathrm{i}}}{s^0_{\mathrm{ij}}}\,(\Delta x_{\mathrm{j}} - \Delta x_{\mathrm{i}}) + \frac{y^0_{\mathrm{j}} - y^0_{\mathrm{i}}}{s^0_{\mathrm{ij}}}\,(\Delta y_{\mathrm{j}} - \Delta y_{\mathrm{i}})$$

Grid bearings:

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

$\Longrightarrow$ Linearized grid bearing observation equation:

$$\begin{aligned} T_{ij} &= T^0_{ij} + \frac{1}{1 + \left(\frac{x^0_j - x^0_i}{y^0_j - y^0_i}\right)^2} \left(-\frac{1}{y^0_j - y^0_i}\Delta x_i + \frac{x^0_j - x^0_i}{(y^0_j - y^0_i)^2}\Delta y_i + \frac{1}{y^0_j - y^0_i}\Delta x_j - \frac{x^0_j - x^0_i}{(y^0_j - y^0_i)^2}\Delta y_j \right) \\ &= T^0_{ij} + \frac{(y^0_j - y^0_i)^2}{(S^0_{ij})^2} \left(-\frac{1}{y^0_j - y^0_i}\Delta x_i + \frac{x^0_j - x^0_i}{(y^0_j - y^0_i)^2}\Delta y_i + \frac{1}{y^0_j - y^0_i}\Delta x_j - \frac{x^0_j - x^0_i}{(y^0_j - y^0_i)^2}\Delta y_j \right) \\ &= T^0_{ij} - \frac{y^0_j - y^0_i}{(S^0_{ij})^2}\Delta x_i + \frac{x^0_j - x^0_i}{(S^0_{ij})^2}\Delta y_i + \frac{y^0_j - y^0_i}{(S^0_{ij})^2}\Delta x_j + \frac{x^0_j - x^0_i}{(S^0_{ij})^2}\Delta y_j \end{aligned}$$

Directions:

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

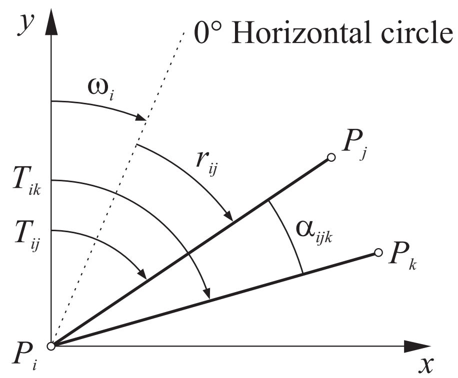


Figure 3.3.: Linearization of bearing observation equation, r_{ij}: bearing, ω_i: orientation unknown

$\Longrightarrow$ Linearization of bearing observation equation

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

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:

$$\begin{aligned}\alpha &= T_{ik}^0 - T_{ij}^0 + \left(-\frac{y_k^0 - y_i^0}{(S_{ik}^0)^2} + \frac{y_j^0 - y_i^0}{(S_{ij}^0)^2} \right) \Delta x_i + \left(\frac{x_k^0 - x_i^0}{(S_{ik}^0)^2} + \frac{x_j^0 - x_i^0}{(S_{ij}^0)^2} \right) \Delta y_i \\ &\quad + \frac{y_k^0 - y_i^0}{(S_{ik}^0)^2} \Delta x_k - \frac{x_k^0 - x_i^0}{(S_{ik}^0)^2} \Delta y_k - \frac{y_j^0 - y_i^0}{(S_{ij}^0)^2} \Delta x_j + \frac{x_j^0 - x_i^0}{(S_{ij}^0)^2} \Delta y_j \\ &= \alpha_0 + \ldots\end{aligned}$$

3D intersection with additional vertical angles

3D distances

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

...Linearization as usual

Vertical angles

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

Attention: physical units!

Iteration (fig. 3.5)

Functional model:

$$y = f(x)$$

$\Longrightarrow$ TAYLOR:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n$$

Linearization:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \mathcal{O}, \qquad \text{where } \mathcal{O} \text{ are small terms/model errors of degree} > 1$$

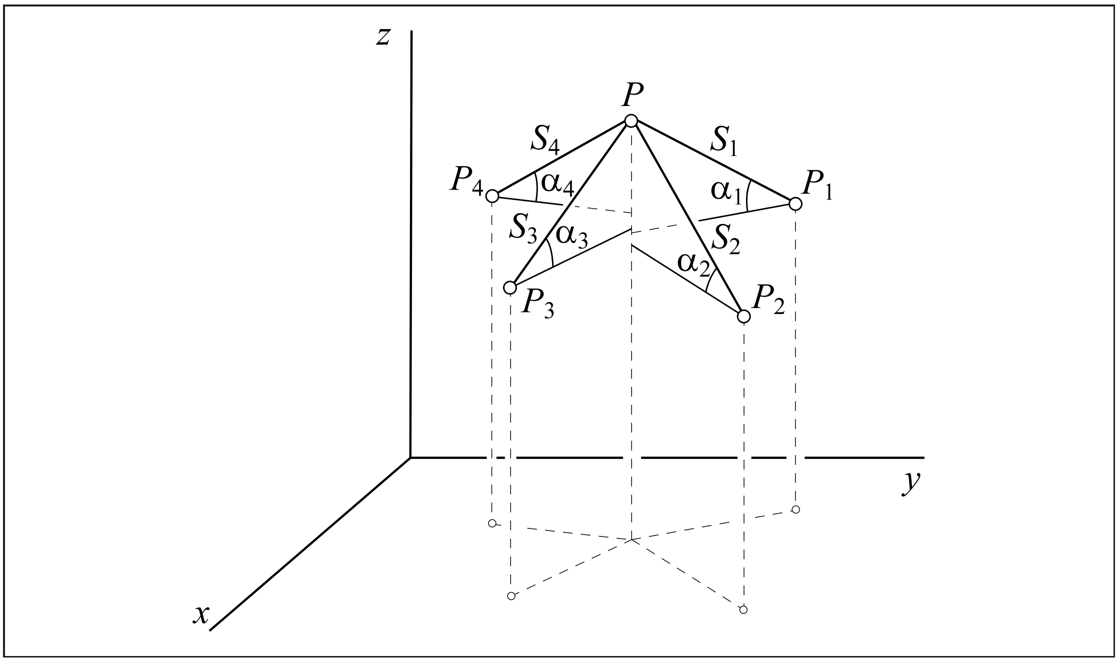


Figure 3.4.: 3D intersection and vertical angles

$$\Longrightarrow f(x) - f(x_0) = f'(x_0)(x - x_0) + \mathcal{O}$$
$$\Delta y = f'(x_0)\Delta x + \mathcal{O}$$
$$\Delta y = \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x_0} \Delta x + \mathcal{O}$$

This results in 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 $\underset{m\times n}{A}$ 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
- $\det(A^{\mathsf{T}}A) = 0$

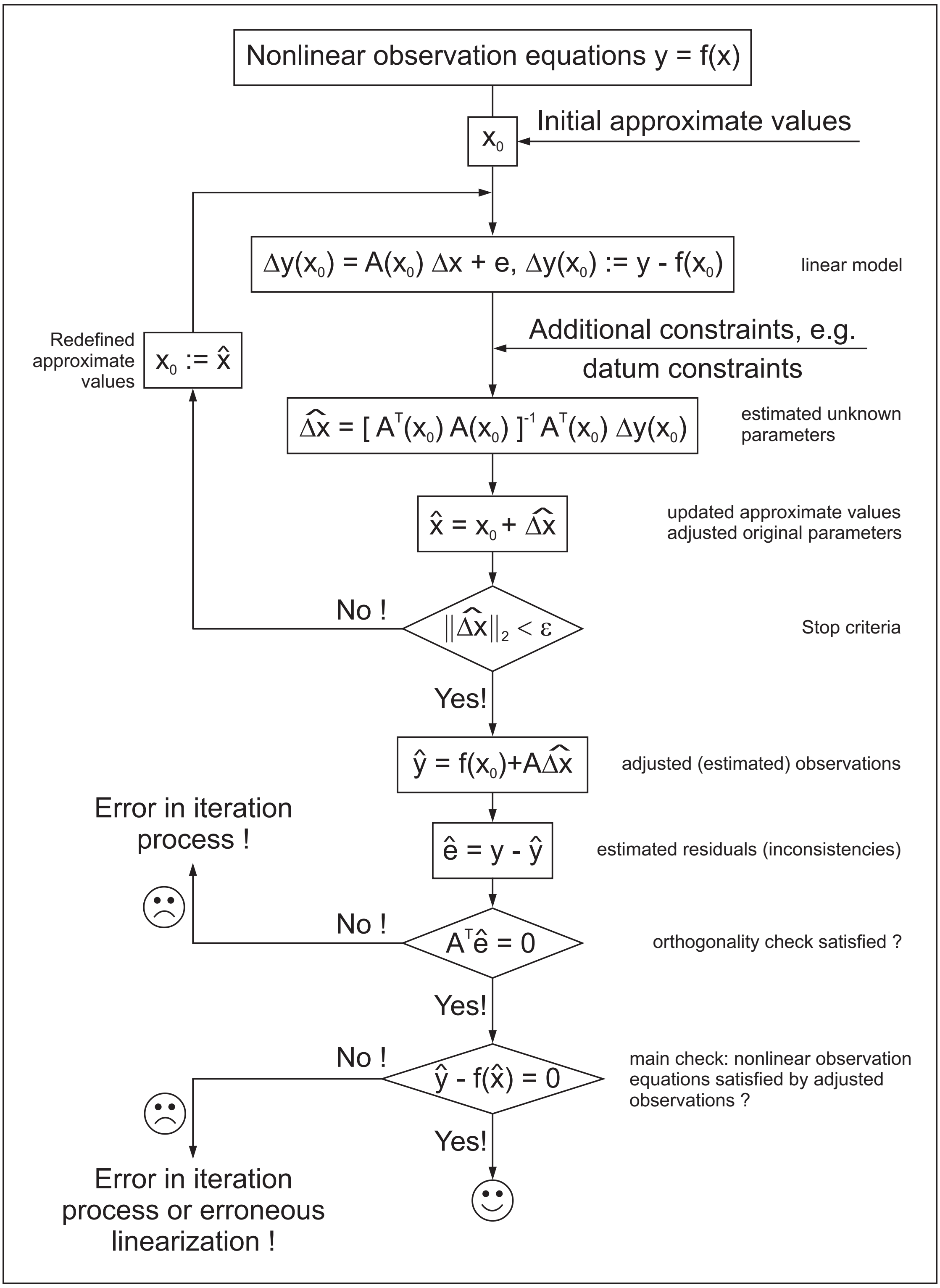


Figure 3.5.: Iterative scheme

- $A^{\mathsf{T}}A$ has zero eigenvalues

Example: planar distance network (fig. 3.6)

Rank defect:

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

- Rotation $\longrightarrow$ 1 parameter

$\Longrightarrow$ total of d = 3 parameters
$\Longrightarrow$ 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

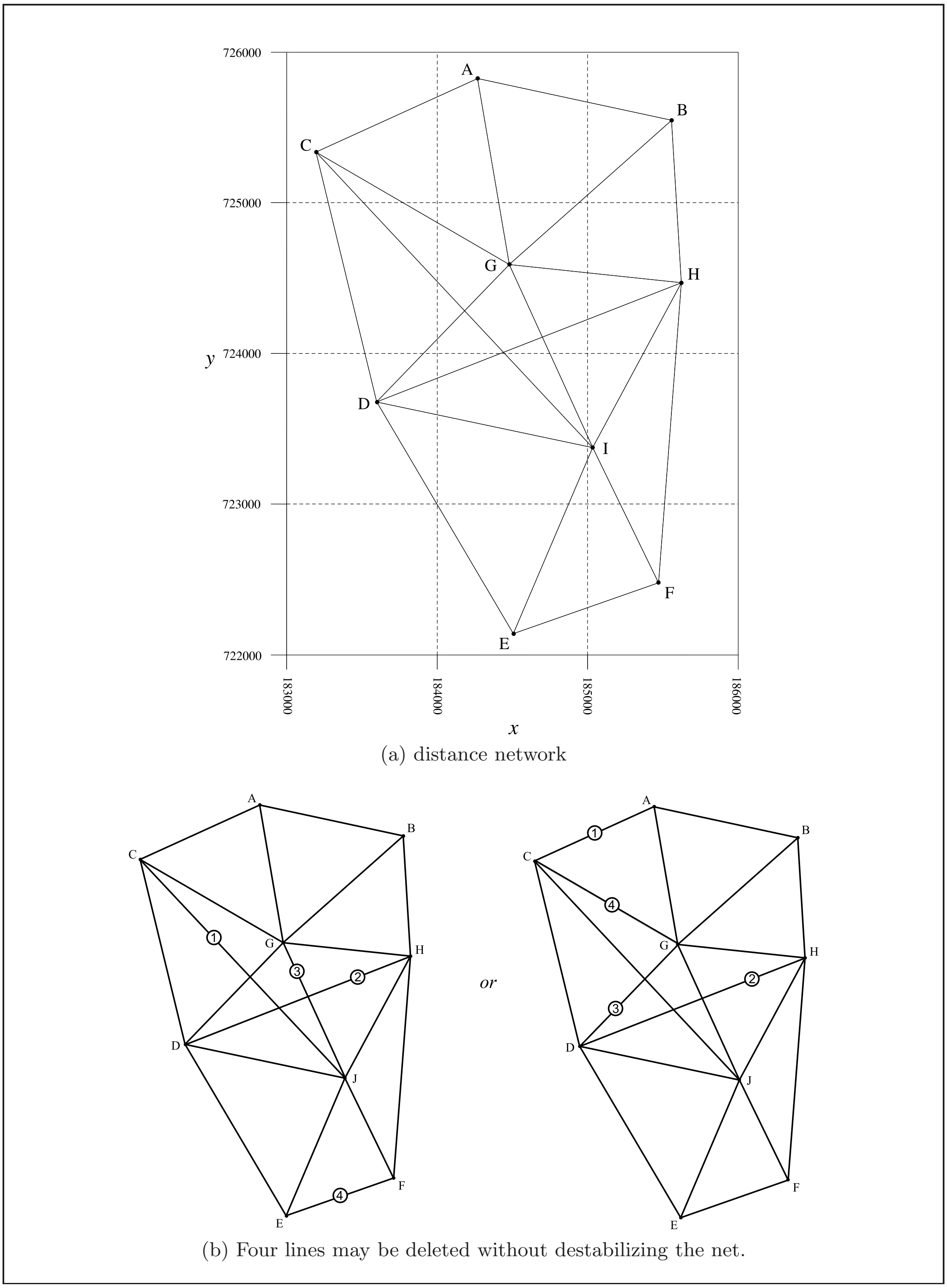
$$\begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} h_{1\mathrm{B}} \\ h_{13} \\ h_{12} \\ h_{32} \\ h_{1\mathrm{A}} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} H_\mathrm{B} \\ 0 \\ 0 \\ 0 \\ H_\mathrm{A}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

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

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



(a) distance network

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

Figure 3.6.

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

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

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:

$$
\begin{aligned}
\mathcal{L}_B(e,\lambda) &= \frac{1}{2} \underset{1\times m}{e^\mathsf{T}} \underset{m\times 1}{e} + \underbrace{\underset{1\times r}{\lambda^\mathsf{T}} (\underset{r\times m}{B^\mathsf{T}} \underset{m\times 1}{y} - \underset{r\times m}{B^\mathsf{T}} \underset{m\times 1}{e})}_{1\times 1} \longrightarrow \min_{e,\lambda} \\
\frac{\partial \mathcal{L}_B}{\partial e}(\hat{e},\hat{\lambda}) &= \underset{m\times 1}{\hat{e}} - \underset{m\times r}{B} \underset{r\times 1}{\hat{\lambda}} = \underset{m\times 1}{0} \\
\frac{\partial \mathcal{L}_B}{\partial \lambda}(\hat{e},\hat{\lambda}) &= -\underset{r\times m}{B^\mathsf{T}} \underset{m\times 1}{\hat{e}} + \underset{r\times m}{B^\mathsf{T}} \underset{m\times 1}{y} = \underset{r\times 1}{0} \qquad (w = B^\mathsf{T} y) \\
&\Longrightarrow \begin{pmatrix} \underset{m\times m}{I} & -\underset{m\times r}{B} \\ -\underset{r\times m}{B^\mathsf{T}} & \underset{r\times r}{0} \end{pmatrix} \begin{pmatrix} \hat{e} \\ \hat{\lambda} \end{pmatrix} = \begin{pmatrix} \underset{m\times 1}{0} \\ -\underset{m\times 1}{w} \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
\hat{e} = B\hat{\lambda} &\Longrightarrow B^\mathsf{T} B\hat{\lambda} = w \\
&\Longrightarrow \hat{\lambda} = (B^\mathsf{T} B)^{-1} w \qquad \operatorname{rank}(B^\mathsf{T} B) = r \\
\Longrightarrow \hat{e} &= B(B^\mathsf{T} B)^{-1} w \\
&= B(B^\mathsf{T} B)^{-1} B^\mathsf{T} y \\
&= P_B y \\
\hat{y} &= y - \hat{e} \\
&= [I - B(B^\mathsf{T} B)^{-1} B^\mathsf{T}] y \\
&= P_B^\perp y
\end{aligned}
$$

Transition parametric model $y = Ax + e$
$\longleftrightarrow$ model of condition equations $B^\mathsf{T} e = B^\mathsf{T} y$

Left multiply $y = Ax + e$ by B^T

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

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

3.5. Linearization of non-linear condition equations

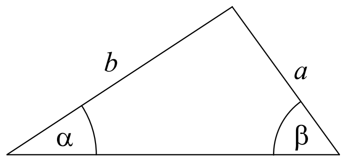


Figure 3.7.: Linearization of condition equations

Ideal situation: error free observations

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

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

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

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

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.

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

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,

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

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

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

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.

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

[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:

$$\left.\begin{matrix} y_1 \longrightarrow w_1 \\ y_2 \longrightarrow w_2 \end{matrix}\right\} \mathcal{L}_a^w = \frac{1}{2}\left[w_1(y_1 - ax)^2 + w_2(y_2 - ax)^2\right]$$

$$= \frac{1}{2}(y - ax)^{\mathsf{T}} \begin{pmatrix} w_1 & 0 \\ 0 & w_2 \end{pmatrix} (y - ax)$$

$$= \frac{1}{2}(y - ax)^{\mathsf{T}} W (y - ax)$$

$$= \frac{1}{2} y^{\mathsf{T}} W y - y^{\mathsf{T}} W a x + \frac{1}{2} x^{\mathsf{T}} a^{\mathsf{T}} W a x$$

$$= \frac{1}{2} e^{\mathsf{T}} W e$$

Necessary condition:

$$\hat{x} : \min_x \mathcal{L}_a(x) \Longrightarrow \frac{\partial \mathcal{L}_a}{\partial x}(\hat{x})$$

$$\Longrightarrow \frac{\partial \mathcal{L}}{\partial x} = -y^{\mathsf{T}} W a + a^{\mathsf{T}} W a x = -a^{\mathsf{T}} W y + a^{\mathsf{T}} W a \hat{x} = 0$$

$$\Longrightarrow a^{\mathsf{T}} W a \hat{x} = a^{\mathsf{T}} W y \quad \text{normal equation}$$

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 $\Longrightarrow$

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

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

Example

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

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

$\Longrightarrow$

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

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

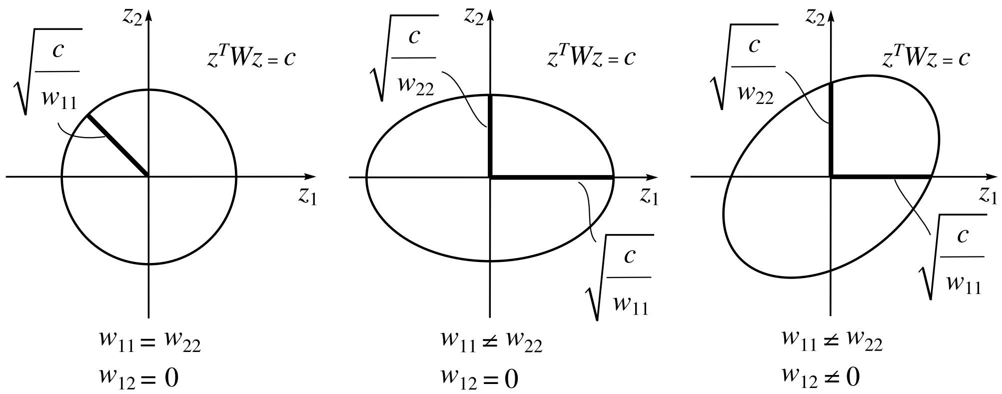
$$w_1 > w_2 \quad : \text{“}y_1 \text{ is more important than } y_2\text{”} \Longrightarrow |\hat{e}_1| < |\hat{e}_2|$$

Projectors

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

P_a idempotent matrix, oblique projector

4.1.1. Geometry



(a) Circle

(b) Ellipse aligned with coordinate axes.

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

Figure 4.1.

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

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

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

General ellipse

$\operatorname{grad} F(z_0) = 2Wz_0$ vector in z_0, orthogonal to the tangent of the ellipse in z_0

$$z - z_0 \perp Wz_0 \quad \text{or} \quad {z_0}^\mathsf{T} W(z - z_0) = 0$$

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 $(Wa)^\perp$)

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

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

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

- 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.

$m = 2$

$$\underset{2\times1}{y} = \underset{2\times1}{a}\ \underset{1\times1}{x} + \underset{2\times1}{e}$$

becomes

$$\underset{m\times1}{y} = \underset{m\times n}{A}\ \underset{n\times1}{x} + \underset{m\times1}{e}$$

Replace a by A!

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

4.2. Weighted condition equations

Geometry

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

Direction of $(Wa)^\perp$:

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

Target function to be minimized: $e^\mathsf{T}We$ 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 $e^\mathsf{T}We = c! \Longrightarrow$ line $b^\mathsf{T}y = b^\mathsf{T}e$ is tangent to $e^\mathsf{T}We = \hat{e}^\mathsf{T}W\hat{e} = c_{\min}$.

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

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

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

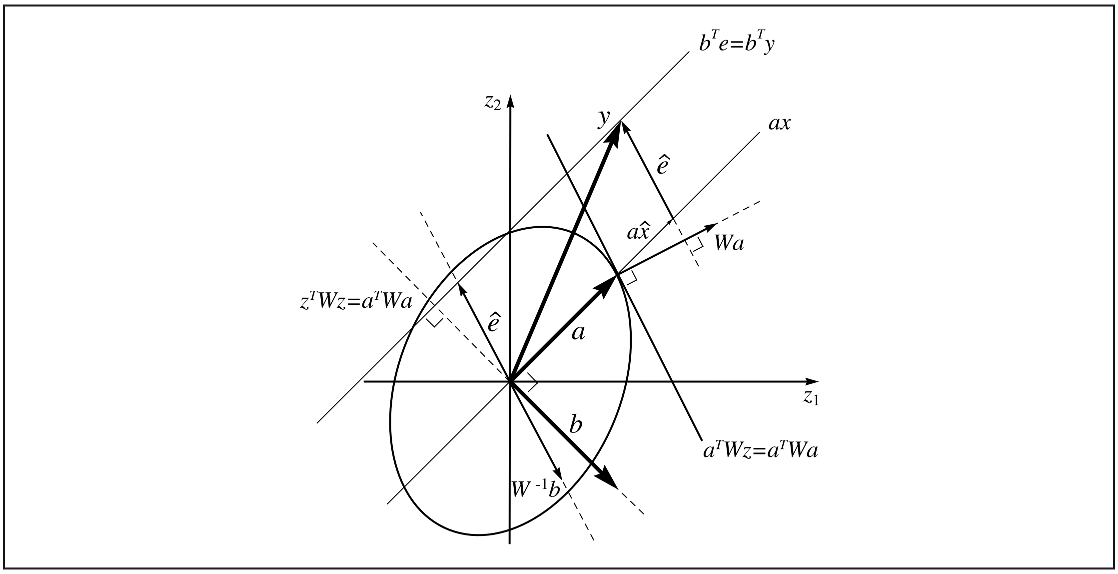


Figure 4.2.: weighted condition

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

Calculus

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

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

Lagrange:

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

Find e and λ which minimize L_b.

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

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

1. row

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

2. row

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

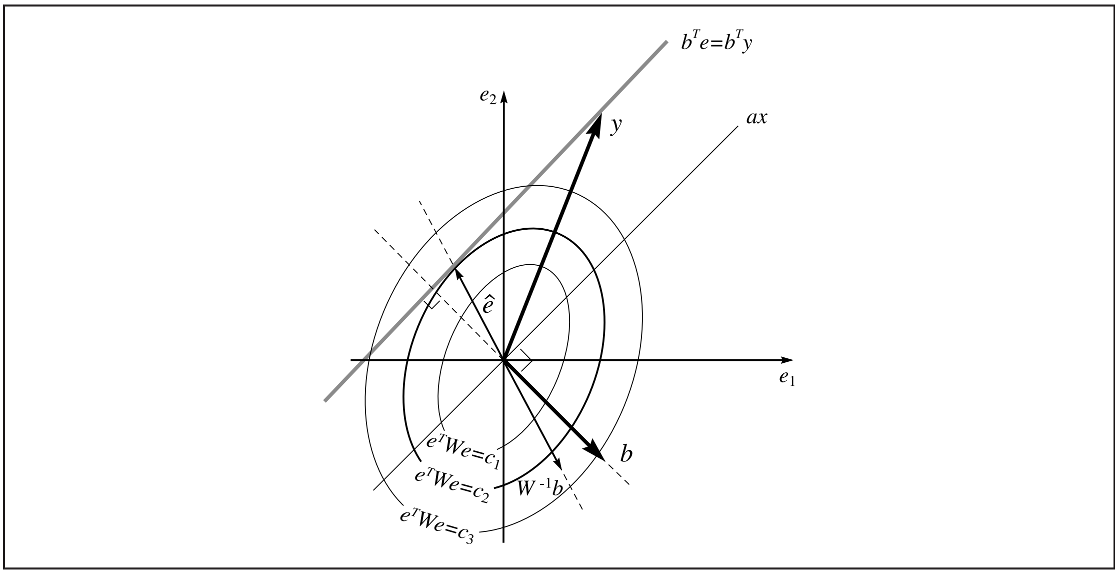


Figure 4.3.: possible ellipses

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

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

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

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

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

Constant term (RHS)

Ideal case without errors:

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

In reality:

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

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

4.3. Stochastics

Probabilistic formulation

(stochastic quantities are underlined)

	Functional model	Stochastic model: variance-covariance matrix
Version 1:	$\underline{y} = Ax + \underline{e}, \quad \mathrm{E}\{\underline{e}\} = 0$	$\mathrm{D}\{\underline{e}\} = Q_y$
Version 2:	$\mathrm{E}\{\underline{y}\} = Ax$	$\mathrm{D}\{\underline{e}\} = Q_y$

Functional model + Stochastic model = Mathematical model

Linear Variance-covariance propagation

In general:

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

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

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	$\mathrm{E}\{\underline{\hat{x}}\} = x$
Estimate	

2D-example (old)

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

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

L-property:

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

U-property:

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

B-property:

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

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

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

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

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

Higher dimensions

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

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

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

Linear Variance-covariance propagation

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

Besides:

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

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

5. Geomatics examples

5.1. A-Model: Adjustment of observation equations

5.1.1. Examples

Planar triangle

Observations: Angles $\alpha, \beta, \gamma [^\circ]$, distances $S_{12}, S_{13}, S_{23} [\mathrm{m}]$

Auxiliary quantities: Bearings $T_{12}, T_{13} [^\circ]$

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:

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

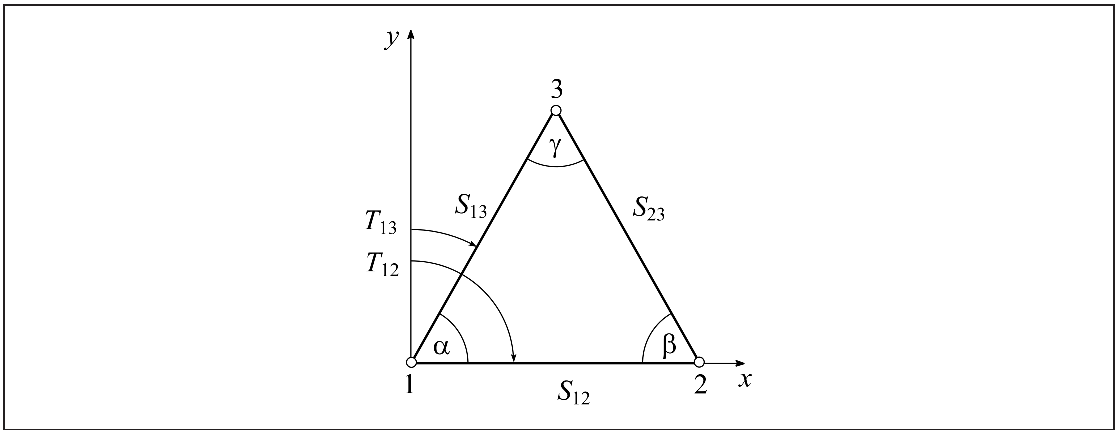


Figure 5.1.: Sketch Planar triangle

Approx. coordinates [m]			"Observations" from approx. coordinates		Observations		σ
			S^0_{12}	1 m	S_{12}	1.01 m	$\pm$ 0.01 m
			S^0_{13}	1 m	S_{13}	1.02 m	$\pm$ 0.02 m
	x_0	y_0	S^0_{23}	1 m	S_{23}	0.97 m	$\pm$ 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 [Unit]	Designmatrix A						Units	Unknowns
	$\rho := \frac{180°}{\pi}$	dx_1	dy_1	dx_2	dy_2	dx_3	dy_3		[m]
S_{12}	0.01[m]	-1	0	1	0	0	0	[–]	dx_1
S_{23}	-0.03[m]	0	0	$\frac{1}{2}$	$-\frac{\sqrt{3}}{2}$	$-\frac{1}{2}$	$\frac{\sqrt{3}}{2}$		dy_1
S_{13}	0.02[m]	$-\frac{1}{2}$	$-\frac{\sqrt{3}}{2}$	0	0	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$		dx_2
α	0[rad]	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	0	-1	$-\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	[m^{-1}]	dy_2
β	$-0.3\frac{°}{\rho}$[rad]	0	-1	$-\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$		dx_3
γ	$0.2\frac{°}{\rho}$[rad]	$-\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	0	-1		dy_3

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

Network of Points

Example 1: Monitoring situation where directions and distances to 4 points A, B, C, and D are measured from N, see below table. Coordinates of point N_0 and the orientation $\omega_N^0 = 63.5610$gon are approximately given (see Jäger 2005, Pg 241-242).

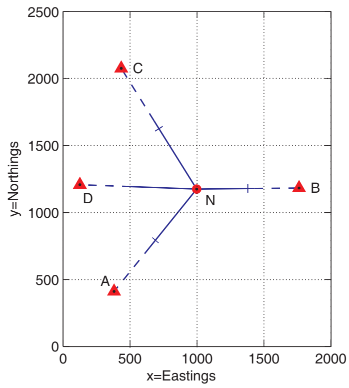


Figure 5.2.: Network of Points

Measured: From	To	Distance ($\sigma = 1$cm)	Direction ($\sigma = 0.5$mgon)
N	A	982.690 m	193.1749 gon
	B	765.000 m	337.1304 gon
	C	1063.890 m	72.0344 gon
	D	-	134.0758 gon
Given	Point	y [m]	x [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

Results: $\Longrightarrow \quad \hat{y}_N = 1175.150m, \quad \hat{x}_N = 997.722m, \quad \hat{\omega}_N = 63.5612gon,$
$S_{ND} = 873.693m, \quad \hat{e}'P\hat{e} = 0.9993218,$ (3 iterations, stop criteria $\|\widehat{\Delta\xi}\| < 10^{-10}$).

Example 2: In this example, measured distances between the network of points in figure 3.6 are adjusted. The standard deviation of observations $\sigma = 1$cm.

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	1267.52	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	184270.031	725830.033
B	185549.974	725400.000
C	183200.000	725450.000
D	183800.000	723550.000
E	184300.000	722050.000
F	185200.000	722450.000
G	184500.000	724400.000
H	185700.000	724650.000
I	184800.000	723400.000

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	0.00000	Y_E	70.03491
Y_A	0.00000	X_F	220.07441
X_B	0.00000	Y_F	12.07883
Y_B	124.26784	X_G	-44.70262
X_C	-29.04695	Y_G	176.98800
Y_C	-80.51622	X_H	-52.35495
X_D	-241.43589	Y_H	-206.79948
Y_D	143.80590	X_I	190.82430
X_E	149.15511	Y_I	-29.00163

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.

Point ID	$\hat{X}$ [m]	$\hat{Y}$ [m]
A	184270.031	725830.033
B	185549.974	725555.019
C	183185.048	725344.999
D	183598.001	723680.041
E	184499.996	722144.987
F	185469.997	722495.040
G	184480.021	724580.029
H	185625.005	724480.000
I	185030.002	723390.016

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

$$\Longrightarrow \quad \hat{e}'P\hat{e} = 0.0035, \quad \text{(6 iterations, stop criteria } \|\widehat{\Delta\xi}\| < 10^{-10}\text{)}$$

Figure 5.4 shows the convergence of estimated corrections to approximate coordinates. Fig 5.5 shows the overall convergence in adjustment iteration. Figure 5.6 shows approximate points, adjusted and datum points. Finally in figure 5.7 adjusted and datum points are shown with error ellipses.

Table 5.3 shows the A-matrix after the first iteration.

	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	-0.31848	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A-C	0	-0.94233	-0.33468	0	0	0	0	0	0	0	0	0	0	0	0
A-G	0	0	0	0	0	0	0	0	0	0.15877	-0.98731	0	0	0	0
B-G	0.68966	0	0	0	0	0	0	0	0	-0.72413	-0.68966	0	0	0	0
B-H	0.98057	0	0	0	0	0	0	0	0	0	0	0.19615	-0.98057	0	0
C-D	0	-0.30113	0.95358	0.30113	-0.95358	0	0	0	0	0	0	0	0	0	0
C-G	0	-0.77794	0.62834	0	0	0	0	0	0	0.77794	-0.62834	0	0	0	0
C-I	0	-0.61527	0.78832	0	0	0	0	0	0	0	0	0	0	0.61527	-0.78832
D-E	0	0	0	-0.31623	0.94868	0.31623	-0.94868	0	0	0	0	0	0	0	0
D-G	0	0	0	-0.63571	-0.77193	0	0	0	0	0.63571	0.77193	0	0	0	0
D-H	0	0	0	-0.86543	-0.50104	0	0	0	0	0	0	0.86543	0.50104	0	0
D-I	0	0	0	-0.98894	0.14834	0	0	0	0	0	0	0	0	0.98894	-0.14834
E-F	0	0	0	0	0	-0.91381	-0.40614	0.91381	0.40614	0	0	0	0	0	0
E-I	0	0	0	0	0	-0.34731	-0.93775	0	0	0	0	0	0	0.34731	0.93775
F-H	0	0	0	0	0	0	0	-0.22162	-0.97513	0	0	0.22162	0.97513	0	0
F-I	0	0	0	0	0	0	0	0.38806	-0.92164	0	0	0	0	-0.38806	0.92164
G-H	0	0	0	0	0	0	0	0	0	-0.97898	-0.20395	0.97898	0.20395	0	0
G-I	0	0	0	0	0	0	0	0	0	-0.28735	0.95783	0	0	0.28735	-0.95783
H-I	0	0	0	0	0	0	0	0	0	0	0	0.58430	0.81153	-0.58430	-0.81153

Figure 5.3.: A-Matrix (1^{st} iteration)

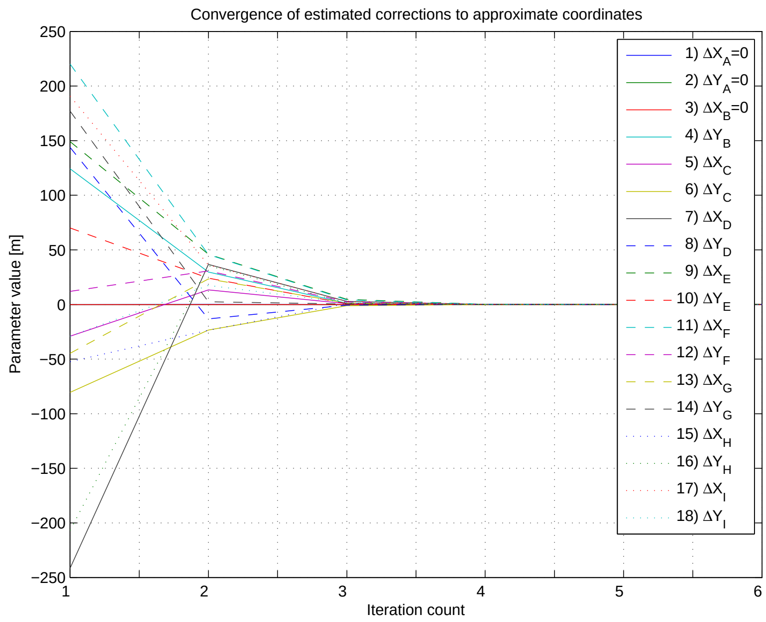


Figure 5.4.: Convergence of estimated corrections

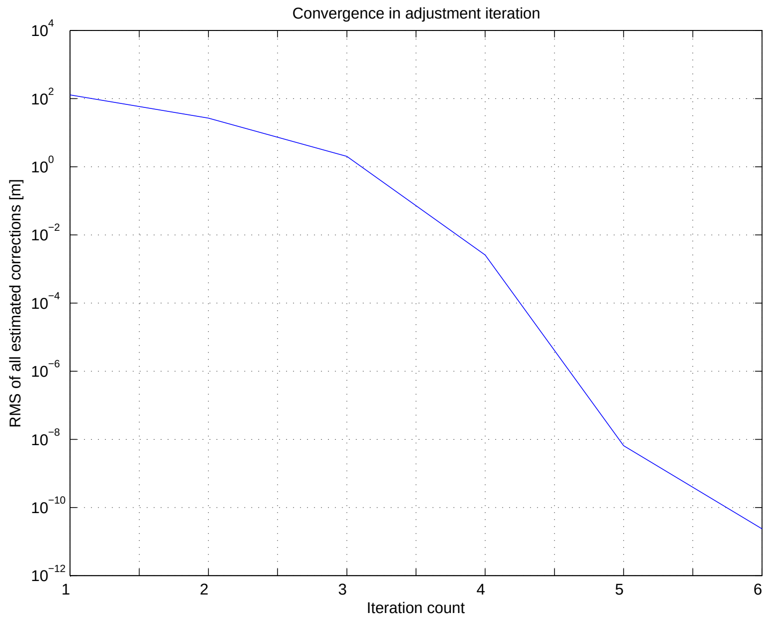


Figure 5.5.: Convergence in adjustment iteration

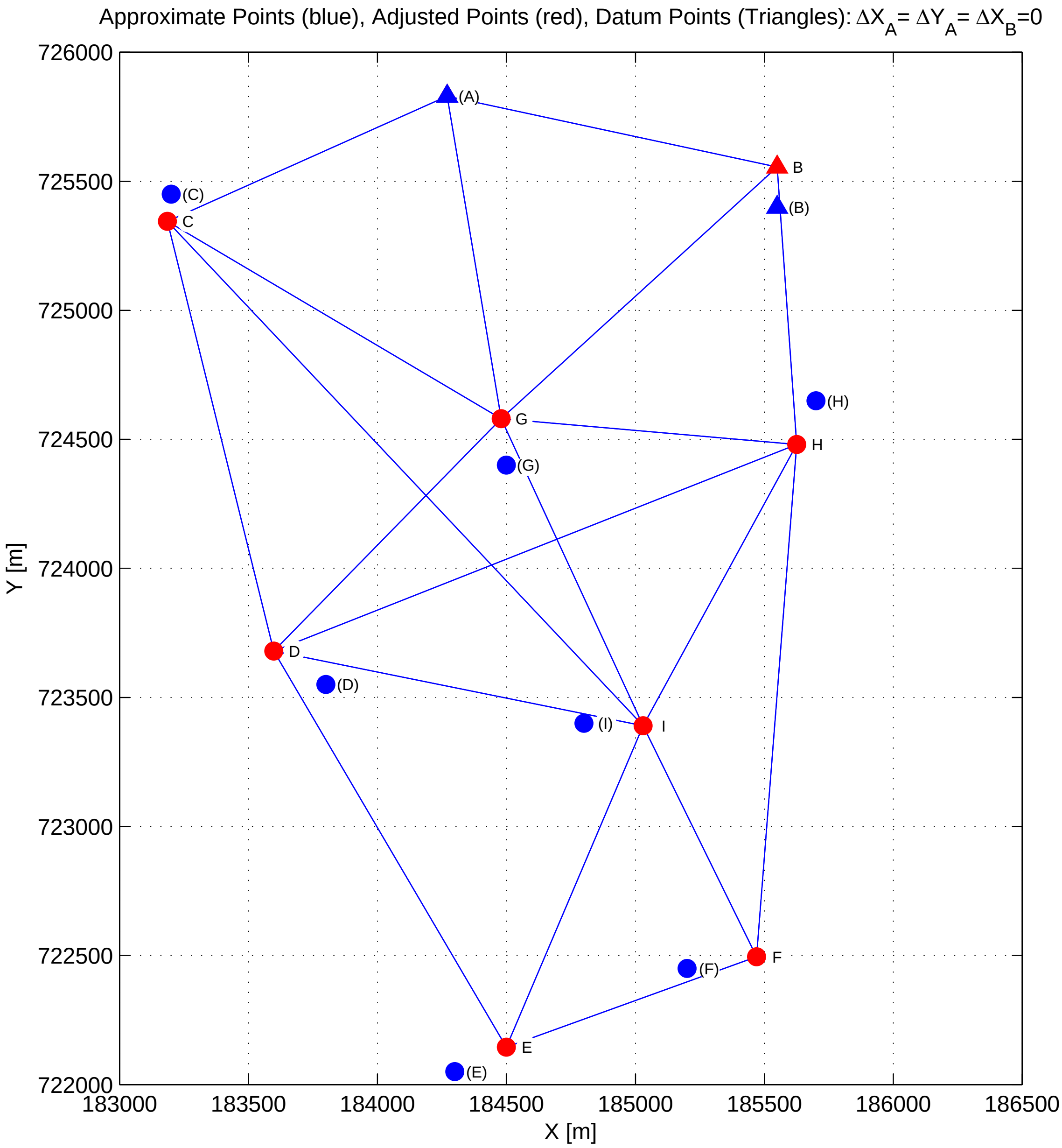


Figure 5.6.: Approximate, adjusted and datum points

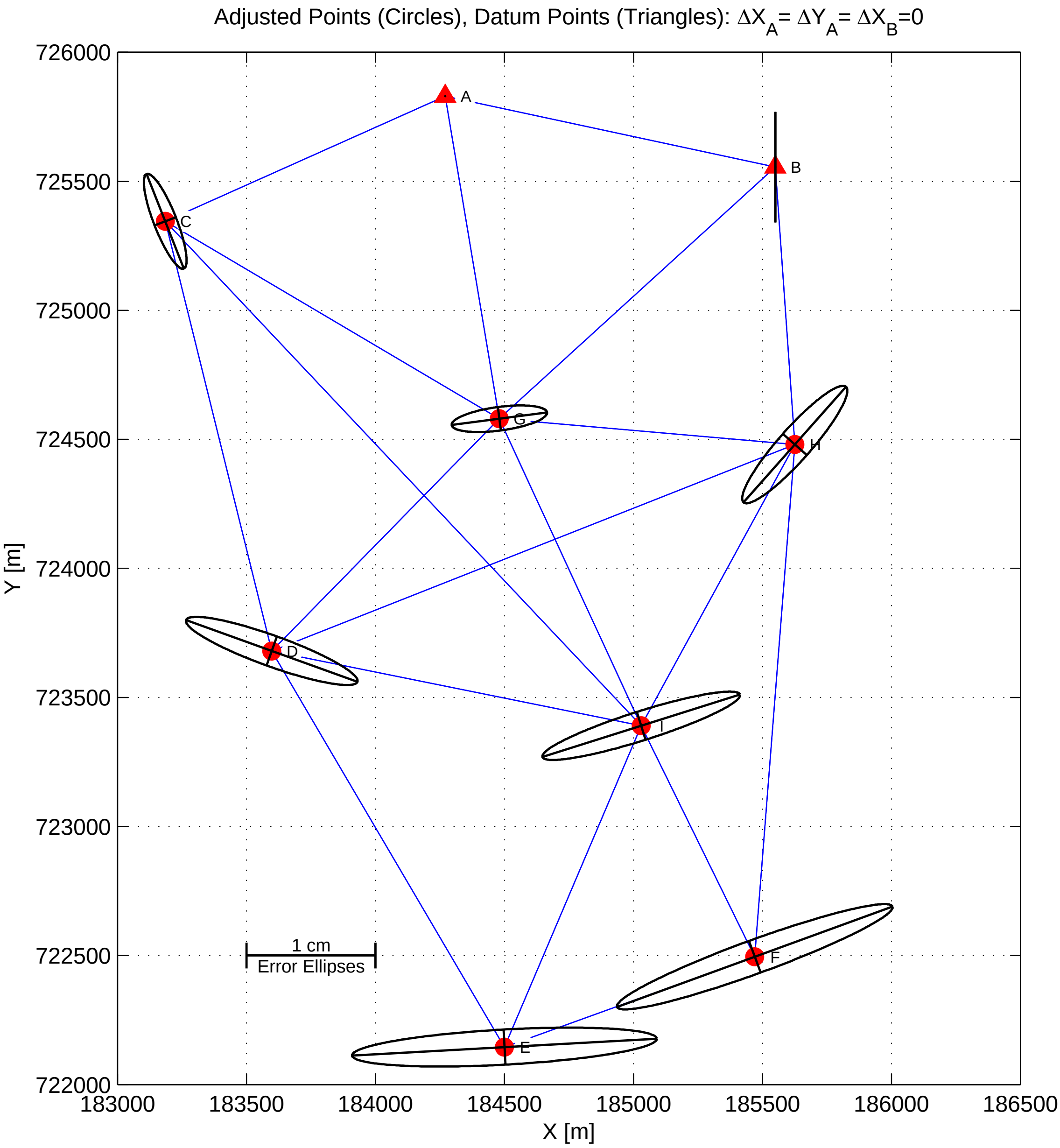


Figure 5.7.: Adjusted and datum points with error ellipses

Polynomial fit

Observations: y_i, $i = 1, \dots, m$

Given: Fixed x-coordinates x_i, $i = 1, \dots, m$

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

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

(possible additional restrictions: (a) tangent in $(x_{\mathrm{T}}, y_{\mathrm{T}})$ should pass through $(x_{\mathrm{P}}, y_{\mathrm{P}})$ or (b) fitting polynomial should pass through $(x_{\mathrm{Q}}, y_{\mathrm{Q}})$ or (c) unknown coefficient a_k shall get the numerical value $\tilde{a}_k$.

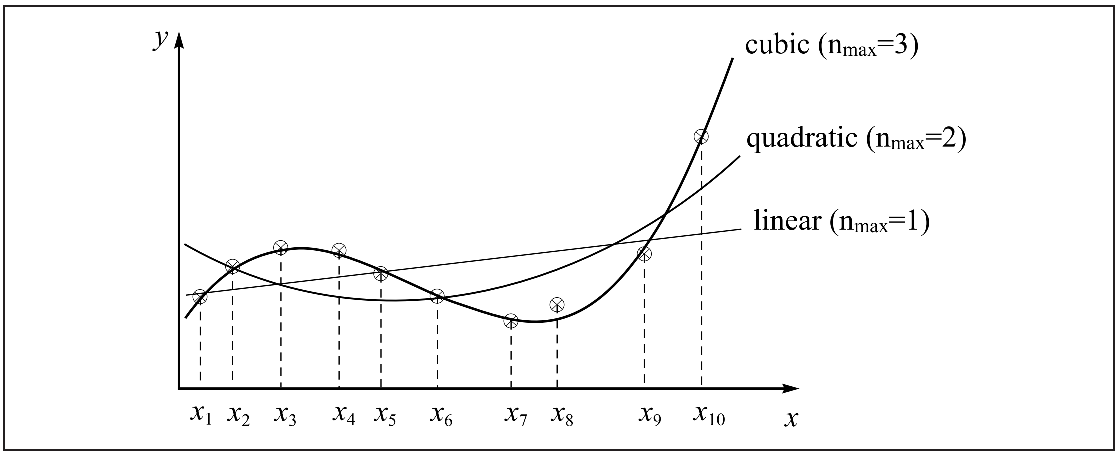


Figure 5.8.: Fitting polynomials of different degrees

Observation equation

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

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

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 larger the polynomial oscillates. Using a 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: $g(x) = f(x_{\mathrm{T}}) + f'(x_{\mathrm{T}})(x - x_{\mathrm{T}}) \Longrightarrow y_{\mathrm{P}} = g(x_{\mathrm{P}}) = f(x_{\mathrm{T}}) + f'(x_{\mathrm{T}})(x_{\mathrm{P}} - x_{\mathrm{T}})$

   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 $x_{\mathrm{P}}, y_{\mathrm{P}}$

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

   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 $n_{\max}$

$$B^\mathsf{T} = \left[\, 1 \;\; x_\mathrm{P} \;\; x_\mathrm{T}(2x_\mathrm{P} - x_\mathrm{T}) \;\; \dots \;\; x_\mathrm{T}^{n_{\max}-1}(n_{\max} x_\mathrm{P} - (n_{\max}-1)x_\mathrm{T}) \,\right]$$

$\Longrightarrow$ Tangent equation with adjusted parameters $\hat{\xi} = \left[\, \hat{a}_0, \; \dots, \; \hat{a}_{n_{\max}} \,\right]^\mathsf{T}$

$$y = a_\mathrm{T} x + b_\mathrm{T}, \quad a_\mathrm{T} := \frac{\hat{y}_\mathrm{T} - y_\mathrm{P}}{x_\mathrm{T} - x_\mathrm{P}} \qquad \text{“tangent slope”}$$

$$b_\mathrm{T} := \hat{y}_\mathrm{T} - \frac{\hat{y}_\mathrm{T} - y_\mathrm{P}}{x_\mathrm{T} - x_\mathrm{P}} x_\mathrm{T} \qquad \text{“axis intercept”}$$

$$\hat{y}_\mathrm{T} = \hat{a}_0 + \hat{a}_1 x_\mathrm{T} + \dots + \hat{a}_{n_{\max}} x_\mathrm{T}^{n_{\max}} \qquad \text{“estimated ordinate”}$$

b) adjusted polynomial shall pass through the point $(x_\mathrm{Q}, y_\mathrm{Q})$

$$y_\mathrm{Q} = \sum_{n=0}^{n_{\max}} a_n x_\mathrm{Q}^n \Longrightarrow B^\mathsf{T}\xi = y_\mathrm{Q}, \quad B^\mathsf{T} = \left(\, 1 \;\; x_\mathrm{Q} \;\; x_\mathrm{Q}^2 \;\; \dots \;\; x_\mathrm{Q}^{n_{\max}} \,\right)$$

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

$$B^\mathsf{T}\xi = \tilde{a}_k, \quad B^\mathsf{T} = \left[\, 0 \;\; \dots \;\; \underbrace{1}_{\text{position } k+1} \;\; \dots \,\right]$$

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

Examples

$$x_i = \left[\, -1, \;\; 0, \;\; 1, \;\; 2, \;\; 3, \;\; 4, \;\; 5 \,\right]^\mathsf{T}$$

$$y_i = \left[\, 1.3, \;\; 0.8, \;\; 0.9, \;\; 1.2, \;\; 2.0, \;\; 3.5, \;\; 4.1 \,\right]^\mathsf{T}$$

Example 1: no restrictions

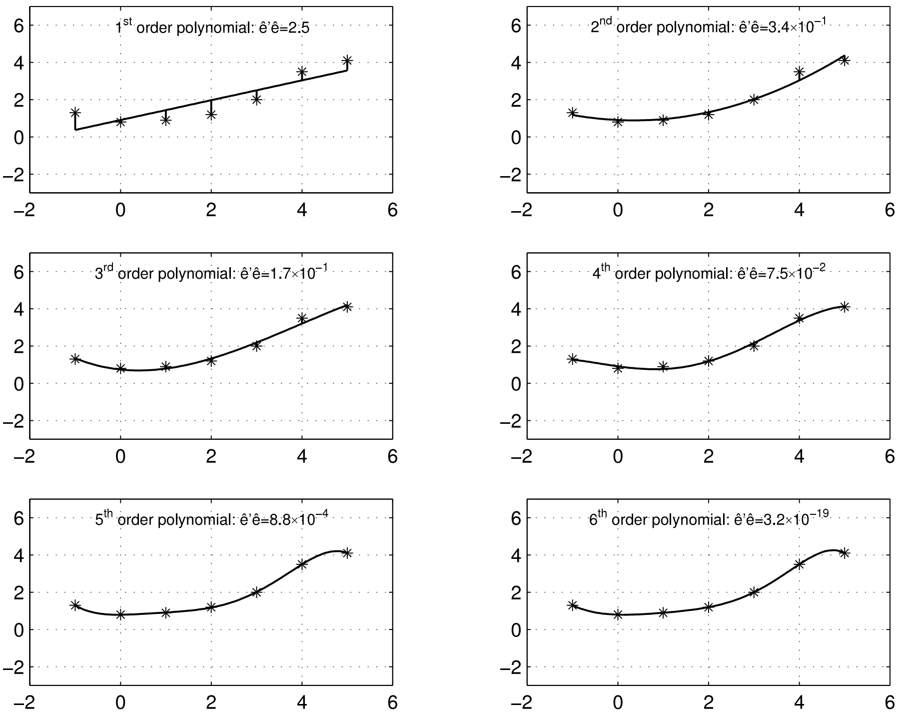


Figure 5.9.: Polynomial fit without restrictions

Example 2: Tangent in $x_{\mathrm{T}} = 1$, $\hat{y}_{\mathrm{T}}(x_{\mathrm{T}})$ shall pass through the point $x_{\mathrm{P}} = 4$, $y_{\mathrm{P}} = 2$

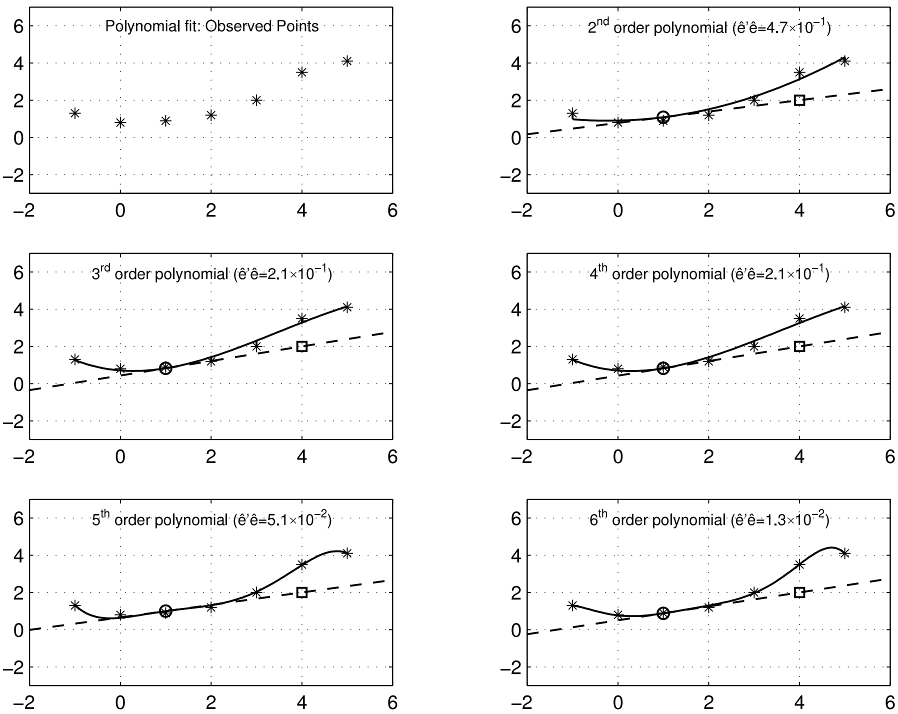


Figure 5.10.: Polynomial fit with tangent restriction

Example 3: Adjusted polynomial shall pass through the point $x_{\mathrm{Q}} = 1.5$, $y_{\mathrm{Q}} = 2$

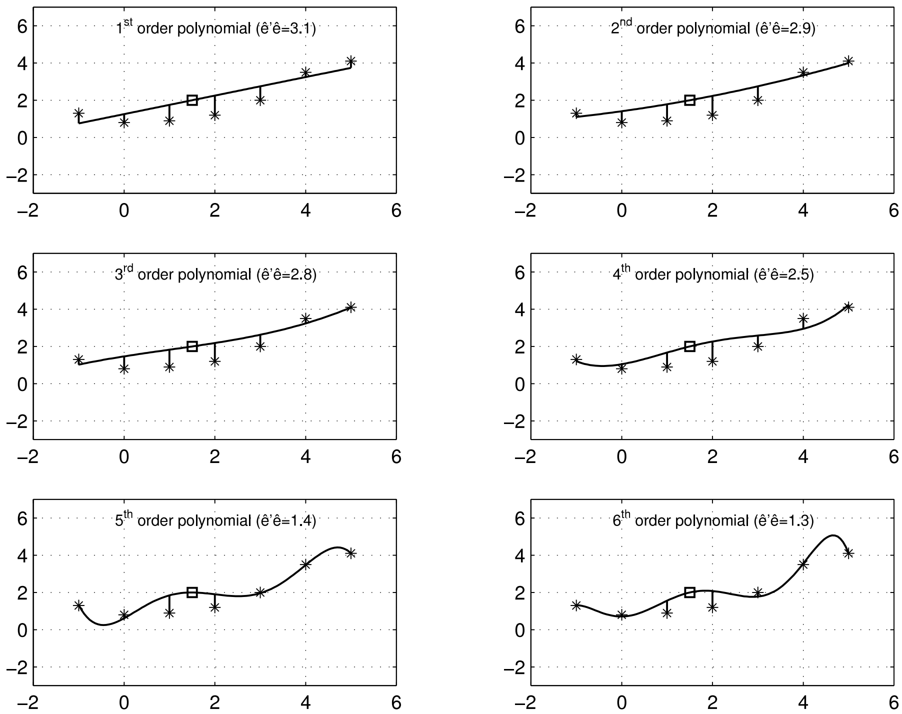


Figure 5.11.: Polynomial fit with point restriction

Example 4: Coefficient $\hat{a}_1$ shall vanish

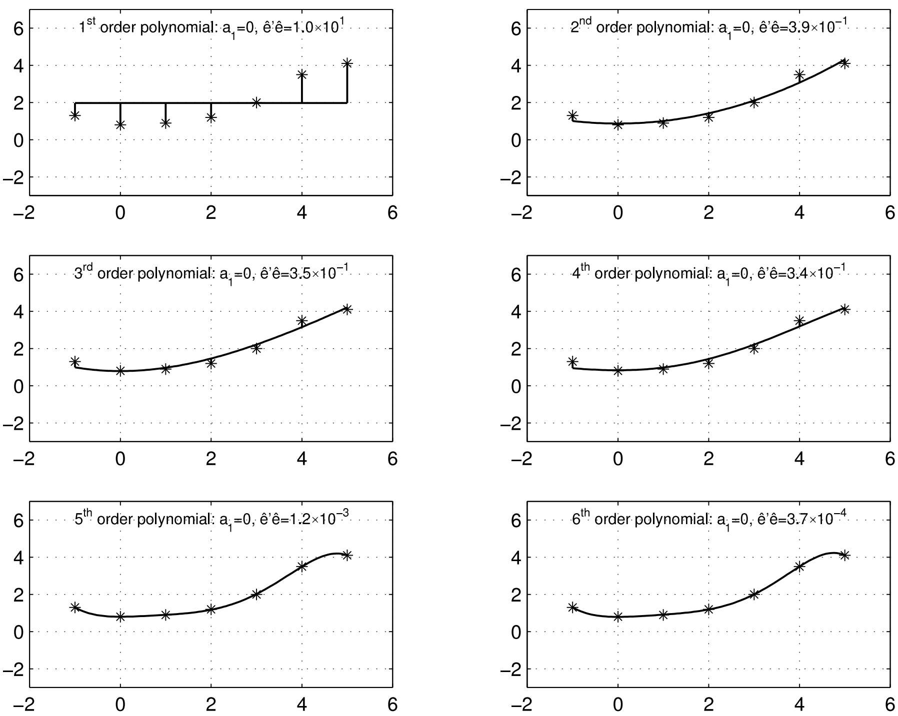


Figure 5.12.: Polynomial fit with coefficient restriction

Examples:

Various straight line fits. For the numerics, the values on page 63 have been used.

Example: **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: $\Longrightarrow \hat{a}_0 = 0.907, \quad \hat{a}_1 = 0.532, \quad \hat{e}' P \hat{e} = 2.505$

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

See figure 5.13.

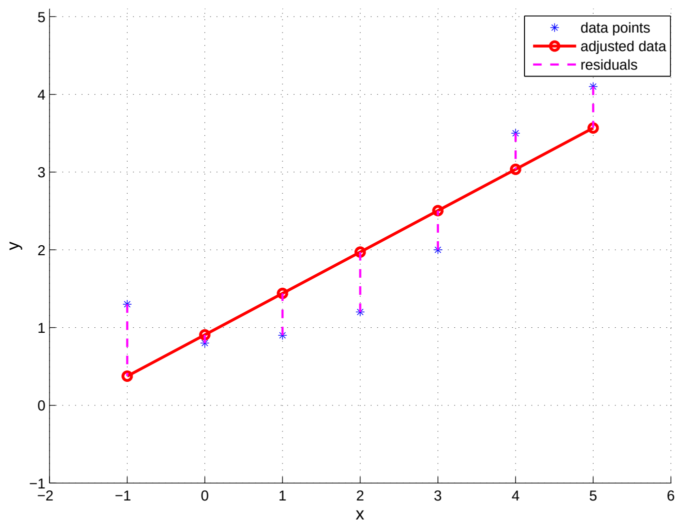


Figure 5.13.: A-model with inconsistencies in y_i, unit weight

Example: **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: $\Longrightarrow \hat{a}_0 = -0.815, \quad \hat{a}_1 = 1.428, \quad \hat{e}' P \hat{e} = 6.723$

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

See figure 5.14.

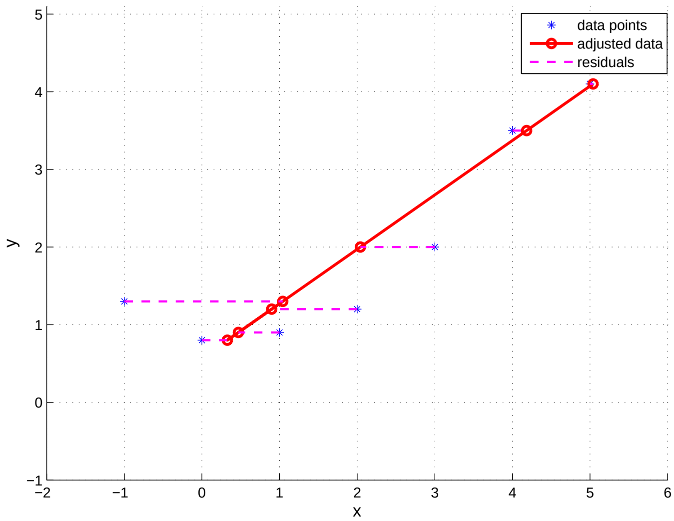


Figure 5.14.: A-model with inconsistencies in x_i, unit weight

5.2. B-Model : Adjustment of condition equations

Unknowns: Inconsistencies $\Longrightarrow$ 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" $(e_a^0, e_b^0, e_\alpha^0, e_\beta^0) =: |_0$

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

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.
$\Longrightarrow$

$$
\begin{aligned}
B^{\mathsf{T}} &= \big(\sin(\beta - e_\beta^0),\, -\sin(\alpha - e_\alpha^0),\, -(b - e_b^0)\cos(\alpha - e_\alpha^0),\, (a - e_a^0)\cos(\beta - e_\beta^0)\,\big) \\
w &= a\sin\left(\beta - e_\beta^0\right) - b\sin\left(\alpha - e_\alpha^0\right) - \left(b - e_b^0\right)\cos(\alpha - e_\alpha^0)e_\alpha^0 \\
&\quad + \left(a - e_a^0\right)\cos(\beta - e_\beta^0)e_\beta^0 \qquad \text{“Misclosure”}
\end{aligned}
$$

Example: Observations: $a = 10, b = 5, \alpha = 60°, \beta = 23.7°$

Results: $\Longrightarrow$ Initial approximate values for unknown inconsistencies: $e_a^0 = e_b^0 = 0$, $e_\alpha^0 = e_\beta^0 = 0°$ $\Longrightarrow$ (6 Iterations:) $\|\Delta\hat{e}\| < 10^{-12}$

$$
\begin{gathered}
\hat{e}_a = -5.63 \times 10^{-6}, \quad \hat{e}_b = 1.13 \times 10^{-4}, \\
\hat{e}_\alpha = 6'26''.241, \quad \hat{e}_\beta = -1°55'42''.492, \quad \hat{e}'\hat{e} = 4.017 \times 10^{-6}
\end{gathered}
$$

5.3. Mixed model

Example: **Straight Line fit** using A-Model, with inconsistencies e_{x_i} and e_{y_i} in both observations x_i and y_i ($Q_y^{-1} = Q_x^{-1} = I, \quad P = \mathrm{diag}(Q_y^{-1}, Q_x^{-1})$). For the numerics, the values on page 63 have been used.
Unknown parameters $a_0, a_1, \bar{x}_i, i = 1, ..., m$

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

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

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.

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

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

$\Longrightarrow$

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

and

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

In matrix notation;

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

$\Longrightarrow$ Where

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

Results: $\Longrightarrow$ Initial approximate values for unknown parameters: $a_0^0 = 0.800$, $a_1^0 = 0.550$, $\bar{x}_i^0 = x_i \Longrightarrow$ (20 Iterations, $\|\Delta\hat{\xi}\| < 10^{-12}$)

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

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

See figure 5.15.

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

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

Initial approximate values:

$$e_{x_i} = e_{x_i}^0 + \Delta e_{x_i}, \quad e_{y_i} = e_{y_i}^0 + \Delta e_{y_i}, \quad a_0 = a_0^0 + \Delta a_0, \quad a_1 = a_1^0 + \Delta a_1.$$

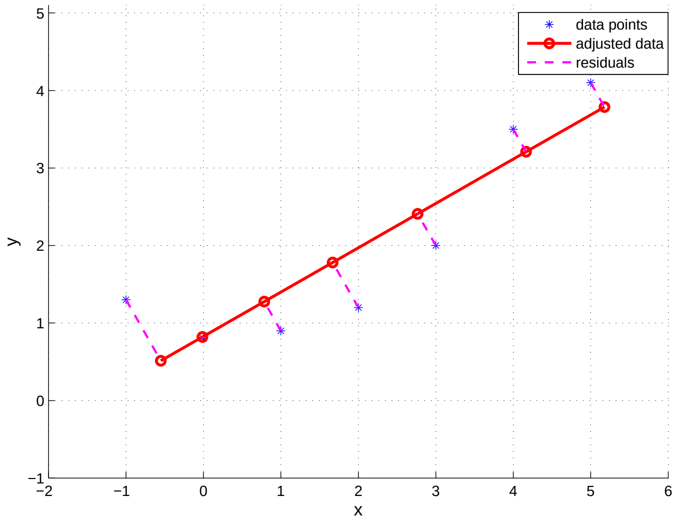


Figure 5.15.: A-model with inconsistencies in x_i and y_i, unit weights

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

$$y_i - e^0_{y_i} - [a^0_0 + a^0_1(x_i - e^0_{x_i})] - e_{y_i} + e^0_{y_i} - [\,1\ (x_i - e^0_{x_i})\,]\begin{pmatrix}\Delta a_0\\ \Delta a_1\end{pmatrix} + a^0_1 e_{x_i} - a^0_1 e^0_{x_i} = 0$$

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

$$\underset{m\times 1}{w} = y - (a^0_0 + a^0_1 x);\ \underset{m\times 2}{A} = -\begin{pmatrix}1 & x_1 - e^0_{x_1}\\ 1 & x_2 - e^0_{x_2}\\ \vdots & \vdots\\ 1 & x_m - e^0_{x_m}\end{pmatrix};$$

$$\underset{2\times 1}{\Delta\xi} = \begin{pmatrix}\Delta a_0\\ \Delta a_1\end{pmatrix};\ \underset{m\times 2m}{B^\mathsf{T}} = \begin{bmatrix}a^0_1 & & & -1 & & \\ & \ddots & & & \ddots & \\ & & a^0_1 & & & -1\end{bmatrix} = [a^0_1 I_m, -I_m];\ \underset{2m\times 1}{e} = \begin{bmatrix}e_{x_1}\\ \vdots\\ e_{x_m}\\ e_{y_1}\\ \vdots\\ e_{y_m}\end{bmatrix} = \begin{bmatrix}e_x\\ e_y\end{bmatrix}$$

Lagrangian

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

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

Results:

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

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

See figure 5.16.

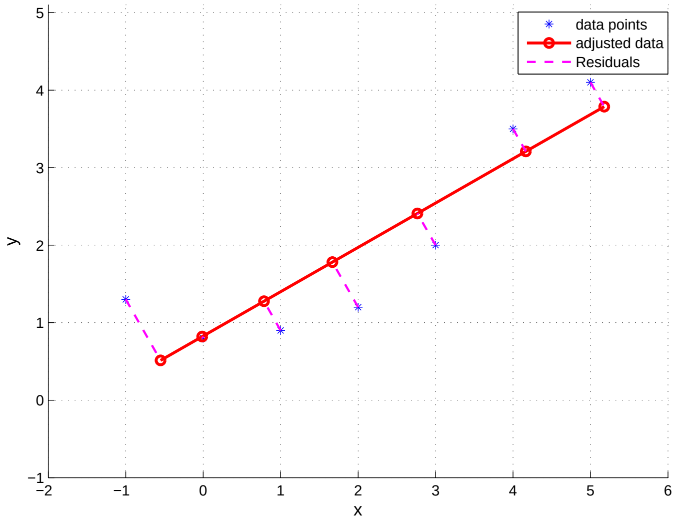


Figure 5.16.: Extended B-model with inconsistencies in x_i and y_i, unit weights

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

Dataset:

x_i	-1	0	1	2	3	4	5
y_i	1.3	0.8	0.9	1.2	2	3.5	4.1
Weights P_x	3	9	8	4	5	7	10
Weights P_y	2	8	7	5	10	8	6

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

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

$$\hat{y} = \begin{bmatrix} 0.208, & 0.620, & 1.124, & 1.633, & 2.281, & 3.288, & 3.895 \end{bmatrix}^\mathsf{T}$$

$$\hat{e}_y = \begin{bmatrix} 1.092, & 0.180, & -0.224, & -0.433, & -0.281, & 0.212, & 0.205 \end{bmatrix}^\mathsf{T}$$

$$\hat{x} = \begin{bmatrix} -0.521, & 0.105, & 0.871, & 1.644, & 2.630, & 4.159, & 5.081 \end{bmatrix}^\mathsf{T}$$

$$\hat{e}_x = \begin{bmatrix} -0.479, & -0.105, & 0.129, & 0.356, & 0.370, & -0.159, & -0.081 \end{bmatrix}^\mathsf{T}$$

See figure 5.17.

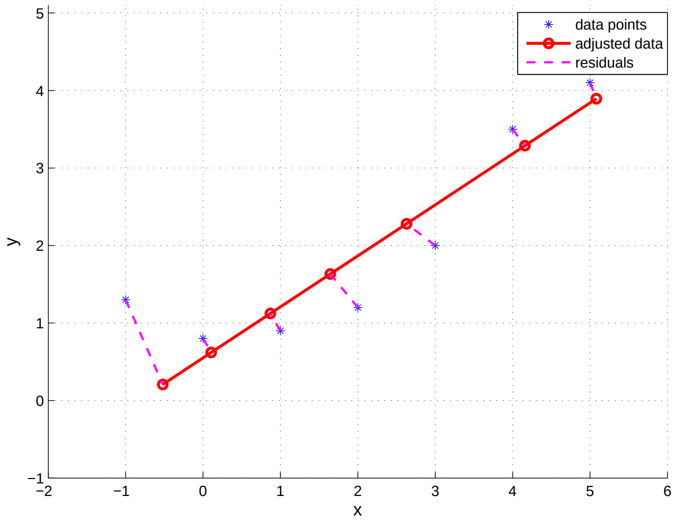


Figure 5.17.: Extended B-model with inconsistencies in x_i and y_i , non-unit weights

Least-squares adjustment - Example: 2D Coordinate Transformations.

The following two tables (See Niemeier 2008, Pg 374-375) give coordinates with respect to the “start” (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 ($P_x = Q_x^{-1} = I, P_y = Q_y^{-1} = I, P_u = Q_u^{-1} = I, P_v = Q_v^{-1} = I, P = \text{diag}(P_x, P_y, P_u, P_v)$).

Start coordinates		
Point	u[m]	v[m]
1	14029.640	12786.840
2	14914.630	12535.560
3	14771.830	11404.660
4	13221.620	11840.320
13	14735.090	12127.380
14	14253.840	11923.950
15	13603.740	11836.700
16	14291.760	12495.310
17	13931.500	12307.610

A 2D Similarity transformation (“Helmert transformation”) to transform the set of coordinates from the “Start system” to the “Target system” will be performed.

Target coordinates		
Point	x [m]	y [m]
1	19405.518	23159.823
2	20291.232	22909.817
3	20150.035	21778.202
4	18598.550	22211.755

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

Usual adjustment : only target coordinates x_i and y_i have inconsistencies

$$x_i - e_{x_i} = \lambda u_i \cos\alpha + \lambda v_i \sin\alpha + t_x$$

$$y_i - e_{y_i} = -\lambda u_i \sin\alpha + \lambda v_i \cos\alpha + t_y$$

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

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

Approximate values:

$$
\lambda = \lambda^0 + \Delta\lambda; \quad \alpha = \alpha^0 + \Delta\alpha; \quad t_x = t_x^0 + \Delta t_x; \quad t_y = t_y^0 + \Delta t_y;
$$

$$
\bar{u}_i = \bar{u}_i^0 + \Delta\bar{u}_i; \quad \bar{v}_i = \bar{v}_i^0 + \Delta\bar{v}_i.
$$

Linearisation process:

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

In matrix form:

$$\underset{4p\times 1}{l} = \begin{bmatrix} x_1 - x_1^0 \\ \vdots \\ x_p - x_p^0 \\ \cdots\cdots \\ y_1 - y_1^0 \\ \vdots \\ y_p - y_p^0 \\ \cdots\cdots \\ u_1 - \bar{u}_1^0 \\ \vdots \\ u_p - \bar{u}_p^0 \\ \cdots\cdots \\ v_1 - \bar{v}_1^0 \\ \vdots \\ v_p - \bar{v}_p^0 \end{bmatrix}; \quad \underset{4p\times 2p+4}{A} = \left[\begin{array}{cccccccccc} 1 & 0 & a_1 & b_1 & \lambda^0\cos\alpha^0 & \dots & 0 & \lambda^0\sin\alpha^0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & a_p & b_p & 0 & \dots & \lambda^0\cos\alpha^0 & 0 & \dots & \lambda^0\sin\alpha^0 \\ \hline 0 & 1 & c_1 & d_1 & -\lambda^0\sin\alpha^0 & \dots & 0 & \lambda^0\cos\alpha^0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & c_p & d_p & 0 & \dots & -\lambda^0\sin\alpha^0 & 0 & \dots & \lambda^0\cos\alpha^0 \\ \hline 0 & 0 & 0 & 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & \dots & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 1 \end{array}\right];$$

$$\underset{2p+4\times 1}{\xi} = \begin{bmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\alpha \\ \Delta\lambda \\ \Delta\bar{u}_1 \\ \vdots \\ \Delta\bar{u}_p \\ \Delta\bar{v}_1 \\ \vdots \\ \Delta\bar{v}_p \end{bmatrix}$$

Results: Initial approximate values for unknown parameters: $t_x^0 = 5500m$, $t_y^0 = 10200m$, $\lambda^0 = 1$, $\alpha^0 = 1''.5$, $\bar{u}^0 = \left[u_1, \dots, u_4\right]^\mathsf{T}$, $\bar{v}^0 = \left[v_1, \dots, v_4\right]^\mathsf{T}$
$\Longrightarrow$ (5 Iterations, $\|\Delta\hat{\xi}\| < 10^{-11}$)

Parameters: $\hat{t}_x = 5389.091m$, $\hat{t}_y = 10347.006m$, $\hat{\alpha} = -5'5''.557$, $\hat{\lambda} = 1.000409017$, $\hat{e}^\mathsf{T} P\hat{e} = 0.00128479m^2$.

Coordinates of data points in the target system

data points-target system		
Point	x [m]	y [m]
13	20112.219	22501.170
14	19631.075	22296.944
15	18980.839	22208.695
16	19668.163	22868.593
17	19308.035	22680.283

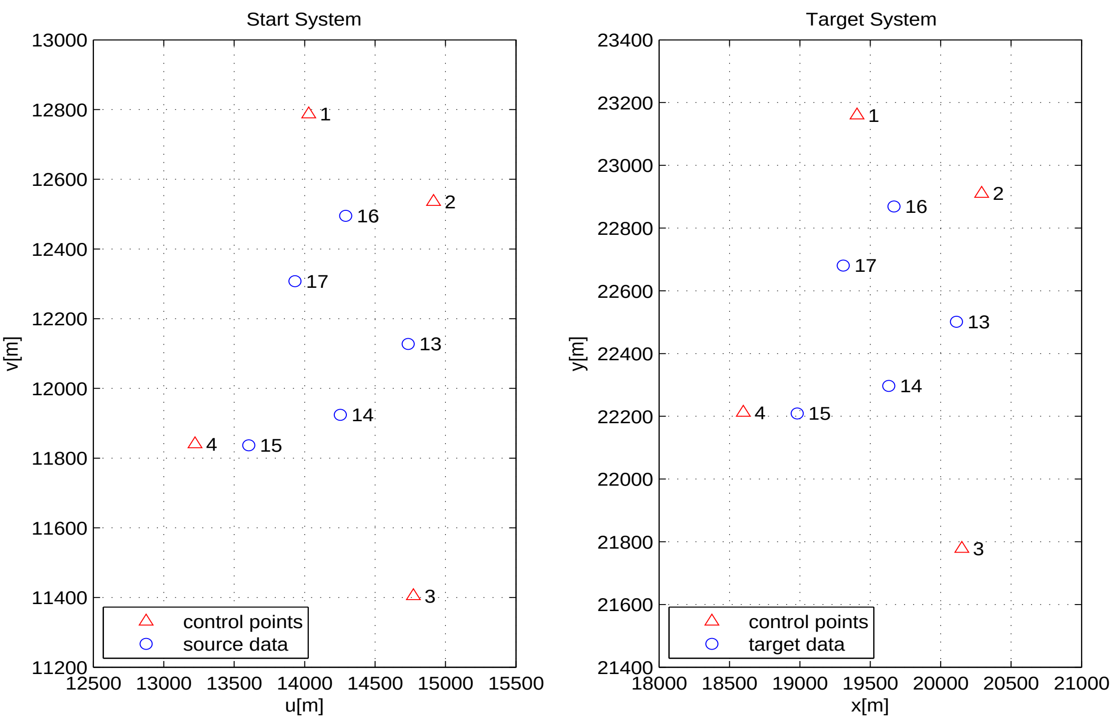


Figure 5.18.: 2D similarity transformation:Gauss Markov model; inconsistencies in both source and target systems

Mixed model approach II: **Extended B-model** with inconsistencies in both $[x, y]$ and $[u, v]$ coordinates

$$f_x := x_i - e_{x_i} - [\lambda \cos\alpha(u_i - e_{u_i}) + \lambda \sin\alpha(v_i - e_{v_i}) + t_x] = 0$$

$$f_y := y_i - e_{y_i} - [-\lambda \sin\alpha(u_i - e_{u_i}) + \lambda \cos\alpha(v_i - e_{v_i}) + t_y] = 0$$

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

Linearisation:

$$\underbrace{\begin{bmatrix} x_i - e^0_{x_i} - (\lambda^0 \cos\alpha^0 (u_i - e^0_{u_i}) + \lambda^0 \sin\alpha^0 (v_i - e^0_{v_i}) + t^0_x) \\ y_i - e^0_{y_i} - (-\lambda^0 \sin\alpha^0 (u_i - e^0_{u_i}) + \lambda^0 \cos\alpha^0 (v_i - e^0_{v_i}) + t^0_y) \end{bmatrix}}_{\underset{2p\times 1}{w}} +$$

$$\underbrace{\begin{pmatrix} \left.\frac{\partial f_{x_i}}{\partial(e_{x_i}, e_{y_i}, e_{u_i}, e_{v_i})}\right|_0 \\ \left.\frac{\partial f_{y_i}}{\partial(e_{x_i}, e_{y_i}, e_{u_i}, e_{v_i})}\right|_0 \end{pmatrix}}_{\underset{2p\times 4p}{B^\mathsf{T}}} \underset{4p\times 1}{\Delta e} + \underbrace{\begin{pmatrix} \left.\frac{\partial f_{x_i}}{\partial(t_x, t_y, \alpha, \lambda)}\right|_0 \\ \left.\frac{\partial f_{y_i}}{\partial(t_x, t_y, \alpha, \lambda)}\right|_0 \end{pmatrix}}_{\underset{2p\times 4}{A}} \underset{4\times 1}{\Delta\xi} = 0$$

In matrix notation:

$$\underbrace{-\begin{pmatrix} x_1 - e^0_{x_1} - x^0_1 \\ \vdots \\ x_p - e^0_{x_p} - x^0_p \\ \cdots\cdots\cdots \\ y_1 - e^0_{y_1} - y^0_1 \\ \vdots \\ y_p - e^0_{y_p} - y^0_p \end{pmatrix}}_{\underset{2p\times 1}{-w}} =$$

$$\underbrace{\begin{pmatrix} -1 & \dots & 0 & 0 & \dots & 0 & \lambda^0\cos\alpha^0 & \dots & 0 & \lambda^0\sin\alpha^0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & -1 & 0 & \dots & 0 & 0 & \dots & \lambda^0\cos\alpha^0 & 0 & \dots & \lambda^0\sin\alpha^0 \\ \hline 0 & \dots & 0 & -1 & \dots & 0 & -\lambda^0\sin\alpha^0 & \dots & 0 & \lambda^0\cos\alpha^0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & -1 & 0 & \dots & -\lambda^0\sin\alpha^0 & 0 & \dots & \lambda^0\cos\alpha^0 \end{pmatrix}}_{\substack{B^\mathsf{T} \\ 2p\times 4p}} \underbrace{\begin{pmatrix} \Delta e_{x_1} \\ \vdots \\ \Delta e_{x_p} \\ \Delta e_{y_1} \\ \vdots \\ \Delta e_{y_p} \\ \Delta e_{u_1} \\ \vdots \\ \Delta e_{u_p} \\ \Delta e_{v_1} \\ \vdots \\ \Delta e_{v_p} \end{pmatrix}}_{\substack{\Delta e \\ 4p\times 1}}$$

$$+ \underbrace{\begin{pmatrix} -1 & 0 & (\lambda^0\bar{u}_1\sin\alpha^0 - \lambda^0\bar{v}_1\cos\alpha^0) & -(\bar{u}_1\cos\alpha^0 + \bar{v}_1\sin\alpha^0) \\ \vdots & \vdots & \vdots & \vdots \\ -1 & 0 & (\lambda^0\bar{u}_p\sin\alpha^0 - \lambda^0\bar{v}_p\cos\alpha^0) & -(\bar{u}_p\cos\alpha^0 + \bar{v}_p\sin\alpha^0) \\ \hline 0 & -1 & (\lambda^0\bar{u}_1\cos\alpha^0 + \lambda^0\bar{v}_1\sin\alpha^0) & (\bar{u}_1\sin\alpha^0 - \bar{v}_1\cos\alpha^0) \\ \vdots & \vdots & \vdots & \vdots \\ 0 & -1 & (\lambda^0\bar{u}_p\cos\alpha^0 + \lambda^0\bar{v}_p\sin\alpha^0) & (\bar{u}_p\sin\alpha^0 - \bar{v}_p\cos\alpha^0) \end{pmatrix}}_{\substack{A \\ 2p\times 4}} \underbrace{\begin{pmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\alpha \\ \Delta\lambda \end{pmatrix}}_{\substack{\Delta\xi \\ 4\times 1}}$$

where $\bar{u}_i = u_i - e^0_{u_i}, \quad \bar{v}_i = v_i - e^0_{v_i}$

Results: Initial approximate values for unknown parameters: $t^0_x = 5500m$, $t^0_y = 10200m, \quad \lambda^0 = 1, \quad \alpha^0 = 1''.5, \quad e^0_{x_i} = e^0_{y_i} = e^0_{u_i} = e^0_{v_i} = 0\ \forall i$

$\Longrightarrow$ (7 Iterations, $\|\Delta\hat{\xi}\| < 10^{-11}$)

Parameters: $\hat{t}_x = 5389.091m, \quad \hat{t}_y = 10347.006m, \quad \hat{\alpha} = -5'5''.557, \quad \hat{\lambda} = 1.000409017$, $\hat{e}^\mathsf{T} P\hat{e} = 0.00128479m^2$.

Least squares adjustment - Example: 6-parameter affine transformation-Model I

The numerical data on page 76 (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: $t_x = t_x^0 + \Delta t_x, \quad t_y = t_y^0 + \Delta t_y, \quad \alpha = \alpha^0 + \Delta\alpha,$
$\lambda_1 = \lambda_1^0 + \Delta\lambda_1, \quad \lambda_2 = \lambda_2^0 + \Delta\lambda_2, \quad k = k^0 + \Delta k \quad \bar{u}_i = \bar{u}_i^0 + \Delta\bar{u}_i, \quad \bar{v}_i = \bar{v}_i^0 + \Delta\bar{v}_i.$

Linearization:

$$x_i - e_{x_i} = \underbrace{[\lambda_1^0 \bar{u}_i^0(\cos\alpha^0 - k^0\sin\alpha^0) + \lambda_1^0 \bar{v}_i^0(\sin\alpha^0 + k^0\cos\alpha^0) + t_x^0]}_{x_i^0} + \Delta t_x +$$

$$\underbrace{[-\lambda_1^0 \bar{u}_i^0(\sin\alpha^0 + k^0\cos\alpha^0) + \lambda_1^0 \bar{v}_i^0(\cos\alpha^0 - k^0\sin\alpha^0)]}_{a_i} \Delta\alpha +$$

$$\underbrace{[\bar{u}_i^0(\cos\alpha^0 - k^0\sin\alpha^0) + \bar{v}_i^0(\sin\alpha^0 + k^0\cos\alpha^0)]}_{b_i} \Delta\lambda_1 + \underbrace{[-\lambda_1^0 \bar{u}_i^0 \sin\alpha^0 + \lambda_1^0 \bar{v}_i^0 \cos\alpha^0]}_{f_i} \Delta k +$$

$$\underbrace{\lambda_1^0(\cos\alpha^0 - k^0\sin\alpha^0)}_{g} \Delta\bar{u}_i^0 + \underbrace{\lambda_1^0(\sin\alpha^0 + k^0\cos\alpha^0)}_{h} \Delta\bar{v}_i^0$$

$$y_i - e_{y_i} = \underbrace{[-\lambda_2^0 \bar{u}_i^0 \sin\alpha^0 + \lambda_2^0 \bar{v}_i^0 \cos\alpha^0 + t_y^0]}_{y_i^0} + \Delta t_y + \underbrace{[-\lambda_2^0(\bar{u}_i^0 \cos\alpha^0 + \bar{v}_1^0 \sin\alpha^0)]}_{c_i} \Delta\alpha +$$

$$\underbrace{(-\bar{u}_i^0 \sin\alpha^0 + \bar{v}_i^0 \cos\alpha^0)}_{d_i} \Delta\lambda_2 \underbrace{-\lambda_2^0 \sin\alpha^0}_{q} \Delta\bar{u}_i + \underbrace{\lambda_2^0 \cos\alpha^0}_{r} \Delta\bar{v}_i$$

In matrix notation:

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

Results: Initial approximate values for unknown parameters: $t_x^0 = 5500m$, $t_y^0 = 10200m$, $\lambda_1^0 = 1$, $\lambda_2^0 = 1$, $\alpha^0 = 1''.5$, $k^0 = 0$, $\bar{u}^0 = \left[u_1, \dots, u_4\right]^\mathsf{T}$, $\bar{v}^0 = \left[v_1, \dots, v_4\right]^\mathsf{T}$

$\Longrightarrow$ (5 Iterations, $\|\Delta\hat{\xi}\| < 10^{-11}$)

Parameters: $\hat{t}_x = 5388.876m$, $\hat{t}_y = 10346.871m$, $\hat{\alpha} = -5'7''.89$, $\hat{\lambda}_1 = 1.000409692$, $\hat{\lambda}_2 = 1.000406924$, $\hat{k} = 2.8233 \times 10^{-5}$ $\hat{e}^\mathsf{T} P \hat{e} = 0.0009932m^2$.

Coordinates of data points in the target system

data points-target system		
Point	x [m]	y [m]
13	20112.220	22501.176
14	19631.071	22296.945
15	18980.833	22208.689
16	19668.169	22868.593
17	19308.037	22680.279

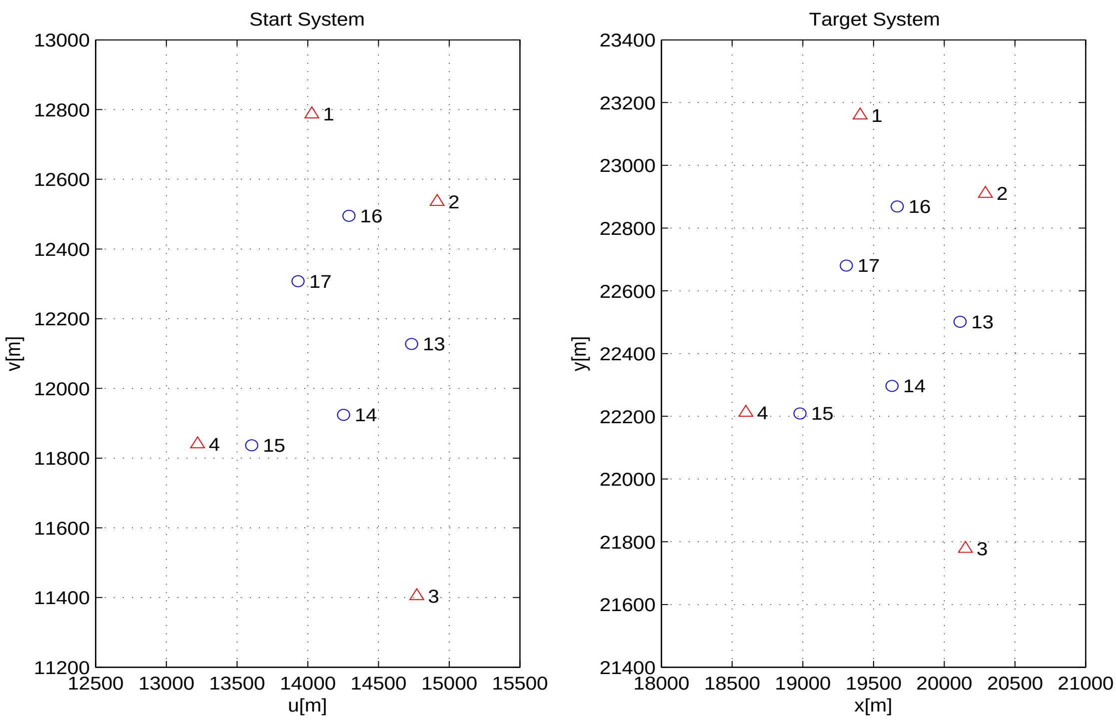


Figure 5.19.: 6 parameter affine transformation: Gauss Markov model; inconsistencies in both source and target systems

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

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

$$f_y := y_i - e_{y_i} - [-\lambda_2(u_i - e_{u_i})\sin\alpha + \lambda_2(v - e_{v_i})\cos\alpha + t_y] = 0$$

Initial approximate values: $e_{x_i} = e^0_{x_i} + \Delta e_{x_i}, \quad e_{y_i} = e^0_{y_i} + \Delta e_{y_i}, \quad e_{u_i} = e^0_{u_i} + \Delta e_{u_i},$
$e_{v_i} = e^0_{v_i} + \Delta e_{v_i}, \quad t_x = t^0_x + \Delta t_x, \quad t_y = t^0_y + \Delta t_y, \quad \alpha = \alpha^0 + \Delta\alpha,$
$\lambda_1 = \lambda^0_1 + \Delta\lambda_1, \quad \lambda_2 = \lambda^0_2 + \Delta\lambda_2, \quad k = k^0 + \Delta k$

Linearization:

$$\underbrace{\begin{pmatrix} x - e^0_x - x^0 \\ y - e^0_y - y^0 \end{pmatrix}}_{\substack{w \\ 2p\times 1}} + \underbrace{\begin{pmatrix} \left.\frac{\partial f_x}{\partial(e_x, e_y, e_u, e_v)}\right|_0 \\ \\ \left.\frac{\partial f_y}{\partial(e_x, e_y, e_u, e_v)}\right|_0 \end{pmatrix}}_{\substack{B^{\mathsf{T}} \\ 2p\times 4p}} \underset{4p\times 1}{\Delta e} + \underbrace{\begin{pmatrix} \left.\frac{\partial f_x}{\partial(t_x, t_y, \alpha, \lambda_1, \lambda_2, k)}\right|_0 \\ \\ \left.\frac{\partial f_y}{\partial(t_x, t_y, \alpha, \lambda_1, \lambda_2, k)}\right|_0 \end{pmatrix}}_{\substack{A \\ 2p\times 6}} \underset{6\times 1}{\Delta\xi} = 0$$

where $x^0_i = [\lambda^0_1 \left(u_i - e^0_{u_i}\right) \left(\cos\alpha^0 - k^0 \sin\alpha^0\right) + \lambda^0_1 \left(v_i - e^0_{v_i}\right) \left(\sin\alpha^0 + k^0\cos\alpha^0\right) + t^0_x]$ and $y^0_i = [-\lambda^0_2 \left(u_i - e_{u_i}\right)\sin\alpha^0 + \lambda^0_2 \left(v_i - e_{v_i}\right)\cos\alpha^0 + t^0_y]$

In matrix notation

$$\underbrace{-\begin{pmatrix} x_1 - e^0_{x_1} - x^0_1 \\ \vdots \\ x_p - e^0_{x_p} - x^0_p \\ \cdots\cdots\cdots \\ y_1 - e^0_{y_1} - y^0_1 \\ \vdots \\ y_p - e^0_{y_p} - y^0_p \end{pmatrix}}_{\substack{-w \\ 2p\times 1}} =$$

$$\underbrace{\left(\begin{array}{cccccccccccc} -1 & \dots & 0 & 0 & \dots & 0 & a & \dots & 0 & b & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & -1 & 0 & \dots & 0 & 0 & \dots & a & 0 & \dots & b \\ \hline 0 & \dots & 0 & -1 & \dots & 0 & q & \dots & 0 & r & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & -1 & 0 & \dots & q & 0 & \dots & r \end{array}\right)}_{\substack{B^{\mathsf{T}} \\ 2p\times 4p}} \underbrace{\begin{pmatrix} \Delta e_{x_1} \\ \vdots \\ \Delta e_{x_p} \\ \Delta e_{y_1} \\ \vdots \\ \Delta e_{y_p} \\ \Delta e_{u_1} \\ \vdots \\ \Delta e_{u_p} \\ \Delta e_{v_1} \\ \vdots \\ \Delta e_{v_p} \end{pmatrix}}_{\substack{\Delta e \\ 4p\times 1}} + \underbrace{\left(\begin{array}{cccccc} -1 & 0 & c_1 & d_1 & 0 & h_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & 0 & c_p & d_p & 0 & h_p \\ \hline 0 & -1 & f_1 & 0 & g_1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & -1 & f_p & 0 & g_p & 0 \end{array}\right)}_{\substack{A \\ 2p\times 6}} \underbrace{\begin{pmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\alpha \\ \Delta\lambda_1 \\ \Delta\lambda_2 \\ \Delta k \end{pmatrix}}_{\substack{\Delta\xi \\ 6\times 1}}$$

where $a = \lambda_1^0 \left(\cos\alpha^0 - k^0 \sin\alpha^0\right)$, $\quad b = \lambda_1^0 \left(\sin\alpha^0 + k^0 \cos\alpha^0\right)$, $\quad q = -\lambda_2^0 \sin\alpha^0$,
$r = \lambda_2^0 \cos\alpha^0$, $\quad c_i = [\lambda_1^0 \bar{u}_i^0 \left(\sin\alpha^0 + k^0 \cos\alpha^0\right) - \lambda_1^0 \bar{v}_i^0 \left(\cos\alpha^0 - k^0 \sin\alpha^0\right)]$,
$d_i = -[\bar{u}_i^0 \left(\cos\alpha^0 - k^0 \sin\alpha^0\right) + \bar{v}_i^0 \left(\sin\alpha^0 + k^0 \cos\alpha^0\right)]$, $\quad f_i = \lambda_2^0 \left(\bar{u}_i^0 \cos\alpha^0 + \bar{v}_i^0 \sin\alpha^0\right)$,
$g_i = \left(\bar{u}_i^0 \sin\alpha^0 - \bar{v}_i^0 \cos\alpha^0\right)$, $\quad h_i = -\lambda_1^0 \left(-\bar{u}_i^0 \sin\alpha^0 + \bar{v}_i^0 \cos\alpha^0\right)$

Results: Initial approximate values for unknown parameters $t_x^0 = 5500m$, $t_y^0 = 10200m$, $\quad \lambda_1^0 = 1$, $\quad \lambda_2^0 = 1$, $\quad \alpha^0 = 1''.5$, $\quad k^0 = 0$, $e_{x_i}^0 = e_{y_i}^0 = e_{u_i}^0 = e_{v_i}^0 = 0 \ \forall i$

$\Longrightarrow$ (4 Iterations, $\|\Delta\hat{\xi}\| < 10^{-11}$)

Parameters: $\hat{t}_x = 5388.876m$, $\quad \hat{t}_y = 10346.871m$, $\quad \hat{\alpha} = -5'7''.89$, $\quad \hat{\lambda}_1 = 1.000409692$, $\hat{\lambda}_2 = 1.000406924$, $\quad \hat{k} = 2.8233 \times 10^{-5}$, $\quad \hat{e}^{\mathsf{T}} P \hat{e} = 0.0009932m^2$.

Least squares adjustment - Example: 6-parameter affine transformation-Model II

The same data sets (76) 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.

$$x_i - e_{x_i} = \lambda_1 \bar{u}_i \cos\varepsilon - \lambda_2 \bar{v}_i \sin\delta + t_x$$

$$y_i - e_{y_i} = \lambda_1 \bar{u}_i \sin\varepsilon + \lambda_2 \bar{v}_i \cos\delta + ty$$

$$u_i - e_{u_i} = \bar{u}_i$$

$$v_i - e_{v_i} = \bar{v}_i$$

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

Linearization process:

$$x_i - e_{x_i} = \underbrace{[\lambda_1^0 \bar{u}_i^0 \cos\varepsilon^0 - \lambda_2^0 \bar{v}_i^0 \sin\delta^0 + t_x^0]}_{x_i^0} + \Delta t_x - \lambda_1^0 \bar{u}_i^0 \sin\varepsilon^0 \Delta\varepsilon - \lambda_2^0 \bar{v}_i^0 \cos\delta^0 \Delta\delta +$$

$$\bar{u}_i^0 \cos\varepsilon^0 \Delta\lambda_1^0 - \bar{v}_i^0 \sin\delta^0 \Delta\lambda_2^0 + \lambda_1^0 \cos\varepsilon^0 \Delta\bar{u}_i - \lambda_2^0 \sin\delta^0 \Delta\bar{v}_i$$

$$y_i - e_{y_i} = \underbrace{[\lambda_1^0 \bar{u}_i^0 \sin\varepsilon^0 + \lambda_2^0 \bar{v}_i^0 \cos\delta^0 + t_y^0]}_{y_i^0} + \Delta t_y + \lambda_1^0 \bar{u}_i^0 \cos\varepsilon^0 \Delta\varepsilon - \lambda_2^0 \bar{v}_i^0 \sin\delta^0 \Delta\delta +$$

$$\bar{u}_i^0 \sin\varepsilon^0 \Delta\lambda_1^0 + \bar{v}_i^0 \cos\delta^0 \Delta\lambda_2^0 + \lambda_1^0 \sin\varepsilon^0 \Delta\bar{u}_i + \lambda_2^0 \cos\delta^0 \Delta\bar{v}_i$$

$$u_i - e_{u_i} = \bar{u}_i^0 + \Delta\bar{u}_i^0 \quad , \quad v_i - e_{v_i} = \bar{v}_i^0 + \Delta\bar{v}_i^0$$

In matrix notation:

$$
\underset{4p\times 1}{l} = \begin{pmatrix} x_1 - x_1^0 \\ \vdots \\ x_p - x_p^0 \\ \cdots\cdots\cdots\cdots \\ y_1 - y_1^0 \\ \vdots \\ y_p - y_p^0 \\ \cdots\cdots\cdots\cdots \\ u_1 - u_1^0 \\ \vdots \\ u_p - u_p^0 \\ \cdots\cdots\cdots\cdots \\ v_1 - v_1^0 \\ \vdots \\ v_p - v_p^0 \end{pmatrix} ; \quad \underset{2p+6\times 1}{\Delta\xi} = \begin{pmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\varepsilon \\ \Delta\delta \\ \Delta\lambda_1 \\ \Delta\lambda_2 \\ \Delta\bar{u}_1 \\ \vdots \\ \Delta\bar{u}_p \\ \Delta\bar{v}_1 \\ \vdots \\ \Delta\bar{v}_p \end{pmatrix}
$$

$$
\underbrace{\left(\begin{array}{cccccccccccc}
1 & 0 & -\lambda_1^0\bar{u}_1^0 sin\varepsilon^0 & -\lambda_2^0\bar{v}_1^0 cos\delta^0 & \bar{u}_1^0 cos\varepsilon^0 & -\bar{v}_1^0 sin\delta^0 & \lambda_1^0 cos\varepsilon^0 & \dots & 0 & -\lambda_2^0 sin\delta^0 & \dots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
1 & 0 & -\lambda_1^0\bar{u}_p^0 sin\varepsilon^0 & -\lambda_2^0\bar{v}_p^0 cos\delta^0 & \bar{u}_p^0 cos\varepsilon^0 & -\bar{v}_p^0 sin\delta^0 & 0 & \dots & \lambda_1^0 cos\varepsilon^0 & 0 & \dots & -\lambda_2^0 sin\delta^0 \\
\hline
0 & 1 & \lambda_1^0\bar{u}_1^0 cos\varepsilon^0 & -\lambda_2^0\bar{v}_1^0 sin\delta^0 & \bar{u}_1^0 sin\varepsilon^0 & \bar{v}_1^0 cos\delta^0 & \lambda_1^0 sin\varepsilon^0 & \dots & 0 & \lambda_2^0 cos\delta^0 & \dots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 1 & \lambda_1^0\bar{u}_p^0 cos\varepsilon^0 & -\lambda_2^0\bar{v}_p^0 sin\delta^0 & \bar{u}_p^0 sin\varepsilon^0 & \bar{v}_p^0 cos\delta^0 & 0 & \dots & \lambda_1^0 sin\varepsilon^0 & 0 & \dots & \lambda_2^0 cos\delta^0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 1 & \dots & 0 & 0 & \dots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & 0 & \dots & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 1 & \dots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 1
\end{array}\right)}_{\underset{4p\times 2p+6}{A}}
$$

Results: Initial approximate values for unknown parameters: $t_x^0 = 5500m$, $t_y^0 = 10200m$, $\epsilon^0 = 1''.5$, $\delta^0 = 3''.5$, $\lambda_1^0 = 1$, $\lambda_2^0 = 1$, $\bar{u}^0 = \left[u_1, \ldots, u_4\right]^\mathsf{T}$, $\bar{v}^0 = \left[v_1, \ldots, v_4\right]^\mathsf{T}$
$\Longrightarrow$ (7 Iterations, $\|\Delta\hat{\xi}\| < 10^{-11}$)

Parameters: $\hat{t}_x = 5388.876m$, $\hat{t}_y = 10346.871m$, $\hat{\epsilon} = 5'7''.89$, $\hat{\delta} = 5'2''.06$, $\hat{\lambda}_1 = 1.000409734$, $\hat{\lambda}_2 = 1.000406883$, $\hat{e}^\mathsf{T} P \hat{e} = 0.0009932m^2$.

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

$$f_x := x_i - e_{x_i} - \left[\lambda_1 \left(u_i - e_{u_i}\right)\cos\varepsilon - \lambda_2 \left(v_i - e_{v_i}\right)\sin\delta + t_x\right] = 0$$

$$f_y := y_i - e_{y_i} - \left[\lambda_1 \left(u_i - e_{u_i}\right)\sin\varepsilon + \lambda_2 \left(v_i - e_{v_i}\right)\cos\delta + t_y\right] = 0$$

Initial approximate values: $e_{x_i} = e_{x_i}^0 + \Delta e_{x_i}$, $e_{y_i} = e_{y_i}^0 + \Delta e_{y_i}$, $e_{u_i} = e_{u_i}^0 + \Delta e_{u_i}$, $e_{v_i} = e_{v_i}^0 + \Delta e_{v_i}$, $t_x = t_x^0 + \Delta t_x$, $t_y = t_y^0 + \Delta t_y$, $\varepsilon = \varepsilon^0 + \Delta\varepsilon$, $\delta = \delta^0 + \Delta\delta$, $\lambda_1 = \lambda_1^0 + \Delta\lambda_1$, $\lambda_2 = \lambda_2^0 + \Delta\lambda_2$

Linearisation:

$$\underbrace{\begin{pmatrix} x - e_x^0 - x^0 \\ y - e_y^0 - y^0 \end{pmatrix}}_{\substack{w \\ 2p\times 1}} + \underbrace{\begin{pmatrix} \left.\frac{\partial f_x}{\partial(e_x, e_y, e_u, e_v)}\right|_0 \\ \\ \left.\frac{\partial f_y}{\partial(e_x, e_y, e_u, e_v)}\right|_0 \end{pmatrix}}_{\substack{B^\mathsf{T} \\ 2p\times 4p}} \underset{4p\times 1}{\Delta e} + \underbrace{\begin{pmatrix} \left.\frac{\partial f_x}{\partial(t_x, t_y, \varepsilon, \delta, \lambda_1, \lambda_2)}\right|_0 \\ \\ \left.\frac{\partial f_y}{\partial(t_x, t_y, \varepsilon, \delta, \lambda_1, \lambda_2)}\right|_0 \end{pmatrix}}_{\substack{A \\ 2p\times 6}} \underset{6\times 1}{\Delta\xi} = 0$$

where $x_i^0 = \lambda_1^0 \left(u_i - e_{u_i}\right)\cos\varepsilon^0 - \lambda_2^0 \left(v_i - e_{v_i}\right)\sin\delta^0 + t_x^0$ and $y_i^0 = \lambda_1^0 \left(u_i - e_{u_i}\right)\sin\varepsilon^0 + \lambda_2^0 \left(v_i - e_{v_i}\right)\cos\delta^0 + t_y^0$

In matrix notation

$$\underbrace{-\begin{pmatrix} x_1 - e_{x_1}^0 - x_1^0 \\ \vdots \\ x_p - e_{x_p}^0 - x_p^0 \\ \cdots\cdots\cdots\cdots \\ y_1 - e_{y_1}^0 - y_1^0 \\ \vdots \\ y_p - e_{y_p}^0 - y_p^0 \end{pmatrix}}_{\substack{-w \\ 2p\times 1}} =$$

$$\underbrace{\left(\begin{array}{cccccccccccc} -1 & \dots & 0 & 0 & \dots & 0 & \lambda_1^0 \cos\varepsilon^0 & \dots & 0 & -\lambda_2^0 \sin\delta^0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & -1 & 0 & \dots & 0 & 0 & \dots & \lambda_1^0 \cos\varepsilon^0 & 0 & \dots & -\lambda_2^0 \sin\delta^0 \\ \hline 0 & \dots & 0 & -1 & \dots & 0 & \lambda_1^0 \sin\varepsilon^0 & \dots & 0 & \lambda_2^0 \cos\delta^0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & -1 & 0 & \dots & \lambda_1^0 \sin\varepsilon^0 & 0 & \dots & \lambda_2^0 \cos\delta^0 \end{array}\right)}_{\underset{2p\times 4p}{B^\mathsf{T}}} \underbrace{\begin{pmatrix} \Delta e_{x_1} \\ \vdots \\ \Delta e_{x_p} \\ \Delta e_{y_1} \\ \vdots \\ \Delta e_{y_p} \\ \Delta e_{u_1} \\ \vdots \\ \Delta e_{u_p} \\ \Delta e_{v_1} \\ \vdots \\ \Delta e_{v_p} \end{pmatrix}}_{\underset{4p\times 1}{\Delta e}} +$$

$$\underbrace{\left(\begin{array}{cccccc} -1 & 0 & \lambda_1^0 \bar{u}_1 \sin\varepsilon^0 & \lambda_2^0 \bar{v}_1 \cos\delta^0 & -\bar{u}_1 \cos\varepsilon^0 & \bar{v}_1 \sin\delta^0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & 0 & \lambda_p^0 \bar{u}_p \sin\varepsilon^0 & \lambda_2^0 \bar{v}_p \cos\delta^0 & -\bar{u}_p \cos\varepsilon^0 & \bar{v}_p \sin\delta^0 \\ \hline 0 & -1 & -\lambda_1^0 \bar{u}_1 \cos\varepsilon^0 & \lambda_2^0 \bar{v}_1 \sin\delta^0 & -\bar{u}_1 \sin\varepsilon^0 & -\bar{v}_1 \cos\delta^0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & -1 & -\lambda_1^0 \bar{u}_p \cos\varepsilon^0 & \lambda_2^0 \bar{v}_p \sin\delta^0 & -\bar{u}_p \sin\varepsilon^0 & -\bar{v}_p \cos\delta^0 \end{array}\right)}_{\underset{2p\times 6}{A}} \underbrace{\begin{pmatrix} \Delta t_x \\ \Delta t_y \\ \Delta\varepsilon \\ \Delta\delta \\ \Delta\lambda_1 \\ \Delta\lambda_2 \end{pmatrix}}_{\underset{6\times 1}{\Delta\xi}}$$

where $\bar{u}_i = u_i - e^0_{u_i}$ and $\bar{v}_i = v_i - e^0_{u_i}$

Results: Initial approximate values for unknown parameters: $t^0_x = 5500m$, $t^0_y = 10200m$, $\epsilon^0 = 1''.5$, $\delta^0 = 3''.5$, $\lambda^0_1 = 1$, $\lambda^0_2 = 1$, $e^0_{x_i} = e^0_{y_i} = e^0_{u_i} = e^0_{v_i} = 0 \ \forall i$

$\Longrightarrow$ (20 Iterations, $\|\Delta\hat{\xi}\| < 10^{-11}$)

Parameters: $\hat{t}_x = 5388.876m$, $\hat{t}_y = 10346.871m$, $\hat{\epsilon} = 5'7''.89$, $\hat{\delta} = 5'2''.06$, $\hat{\lambda}_1 = 1.000409734$, $\hat{\lambda}_2 = 1.000406882$, $\hat{e}^\mathsf{T} P \hat{e} = 0.0009932m^2$.

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

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

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

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

Linearization (with $e^0_{x_i} = e^0_{y_i} = 0$ in the first iteration and given x_0)

$$x = x_0 + \Delta x, \qquad e_{x_i} = e^0_{x_i} + \Delta e_{x_i}, \qquad e_{y_i} = e^0_{y_i} + \Delta e_{y_i}$$

$$\begin{aligned} f(x_0 + \Delta x, y - e) &= f(x_0, y) + \left.\frac{\partial f}{\partial x}\right|_{x_0,y} \Delta x + \left.\frac{\partial f}{\partial e}\right|_{x_0,y} e + \text{ terms of higher order} = 0 \\ &\doteq w + A\Delta x + B^{\mathsf{T}} e = 0 \\ g(x_0 + \Delta x) &= g(x_0) + \left.\frac{\partial g}{\partial x}\right|_{x_0} \Delta x + \text{ terms of higher order} = 0 \\ &\doteq w_R + R\Delta x = 0 \end{aligned}$$

Linear model and adjustment principle

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

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

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

Necessary condition

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

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

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

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

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

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

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

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 use partitioning formula:

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

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

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

$$\begin{aligned} \hat{\lambda}_R &= -Q_{21}A^\mathsf{T}M^{-1}w - Q_{22}w_R \\ &= Q_{22}[R^\mathsf{T}(A^\mathsf{T}M^{-1}A)^{-1}A^\mathsf{T}M^{-1}w - w_R] \\ &= [R(A^\mathsf{T}M^{-1}A)^{-1}R^\mathsf{T}][w_R - R^\mathsf{T}(A^\mathsf{T}M^{-1}A)^{-1}A^\mathsf{T}M^{-1}w] \\ -w &= -M\hat{\lambda} + A\Delta\hat{x} \\ \Longrightarrow \hat{\lambda} &= M^{-1}(A\Delta\hat{x} + w) \\ \hat{e} &= W^{-1}B\hat{\lambda} = W^{-1}BM^{-1}(A\Delta\hat{x} + w) \end{aligned}$$

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

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

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

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

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

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

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

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

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

N is symmetric

$H \cdot (5.3)$

$$\Longrightarrow H\underbrace{NR}_{0} + HR^\mathsf{T}S = H \Longrightarrow S = (HR^\mathsf{T})^{-1}H$$

$H \cdot (5.4)$

$$\Longrightarrow H\underbrace{NS^\mathsf{T}}_{0} + HR^\mathsf{T}Q = 0 \Longrightarrow HR^\mathsf{T}Q = 0$$

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

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

$$\Longrightarrow R = (N + R^\mathsf{T}R)^{-1}[I - R^\mathsf{T}(HR^\mathsf{T})^{-1}H]$$

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

if $w_R = 0$:

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

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

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

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

$$\implies \quad \begin{pmatrix} \frac{2(x_i - e^0_{x_i} - x^0_\mathrm{M})}{a_0^2} & \frac{2(y_i - e^0_{y_i} - y^0_\mathrm{M})}{b_0^2} \end{pmatrix} \begin{pmatrix} e_{x_i} \\ e_{y_i} \end{pmatrix}$$

$$+ \begin{pmatrix} -\frac{2(x_i - e^0_{x_i} - x^0_\mathrm{M})}{a_0^2} & -\frac{2(y_i - e^0_{y_i} - y^0_\mathrm{M})}{b_0^2} & -\frac{2(x_i - e^0_{x_i} - x^0_\mathrm{M})^2}{a_0^3} & -\frac{2(y_i - e^0_{y_i} - y^0_\mathrm{M})^2}{b_0^3} \end{pmatrix} \begin{pmatrix} \Delta x_\mathrm{M} \\ \Delta y_\mathrm{M} \\ \Delta a \\ \Delta b \end{pmatrix}$$

$$+ \frac{(x_i - e^0_{x_i} - x^0_\mathrm{M})^2}{a_0^2} + \frac{(y_i - e^0_{y_i} - y^0_\mathrm{M})^2}{b_0^2} - 1 + \frac{2(x_i - e^0_{x_i} - x^0_\mathrm{M})}{a_0^2} e^0_{x_i} + \frac{2(y_i - e^0_{y_i} - y^0_\mathrm{M})}{b_0^2} e^0_{y_i} = 0$$

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

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

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

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

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

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

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

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

$$\left.\begin{aligned} &W\hat{e} + B\hat{\lambda} = 0 \quad \Longrightarrow \quad \hat{e} = -W^{-1}B\hat{\lambda} \\ &B^{\mathsf{T}}\hat{e} + A\Delta\hat{x} = -w \end{aligned}\right\} \Longrightarrow -B^{\mathsf{T}}W^{-1}B\hat{\lambda} + A\Delta\hat{x} = -w$$

$$\begin{aligned} \Longrightarrow \quad \hat{\lambda} &= (B^{\mathsf{T}}W^{-1}B)^{-1}(A\Delta\hat{x} + w) \\ \Longrightarrow \quad & A^{\mathsf{T}}(B^{\mathsf{T}}W^{-1}B)^{-1}A\Delta\hat{x} + A^{\mathsf{T}}(B^{\mathsf{T}}W^{-1}B)^{-1}w = 0 \\ \Longrightarrow \quad \Delta\hat{x} &= -[A^{\mathsf{T}}(B^{\mathsf{T}}W^{-1}B)^{-1}A]^{-1}A^{\mathsf{T}}(B^{\mathsf{T}}W^{-1}B)^{-1}w \\ \Longrightarrow \quad \hat{e} &= -W^{-1}B(B^{\mathsf{T}}W^{-1}B)^{-1}(A\Delta\hat{x} + w) \\ &= W^{-1}B(B^{\mathsf{T}}W^{-1}B)^{-1}\{A[A^{\mathsf{T}}(B^{\mathsf{T}}W^{-1}B)^{-1}A]^{-1}A^{\mathsf{T}}(B^{\mathsf{T}}W^{-1}B)^{-1} - I\}w \end{aligned}$$

Numerics

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

$$\begin{aligned} x_{\mathrm{M}}^0 &= y_{\mathrm{M}}^0 = 0 \\ a_0 &= b_0 = 120 \end{aligned}$$

$\Longrightarrow$ (10 Iterations: $\|\Delta\hat{x}\| < 10^{-12}, \|\Delta\hat{e}\| < 10^{-12}$):

$$\hat{x}_{\mathrm{M}} = -0.598, \quad \hat{y}_{\mathrm{M}} = -1.942, \quad \hat{a} = 131.087, \quad \hat{b} = 115.131, \quad \hat{e}'W\hat{e} = 523.208$$

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

See figure 5.20.

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

$$\begin{aligned} g(x) = g(x_m, y_m, a, b) &= a - b = 0 \\ \Longrightarrow \qquad R = [0, 0, 1, -1], &\quad w_R = 0 \end{aligned}$$

$\Longrightarrow$ (10 Iterations: $\|\Delta\hat{x}\| < 10^{-12}, \|\Delta\hat{e}\| < 10^{-12}$):

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

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

See figure 5.21.

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

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

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

$\Longrightarrow$ (10 Iterations: $\|\Delta\hat{x}\| < 10^{-12}, \|\Delta\hat{e}\| < 10^{-12}$):

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

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

See figure 5.22.

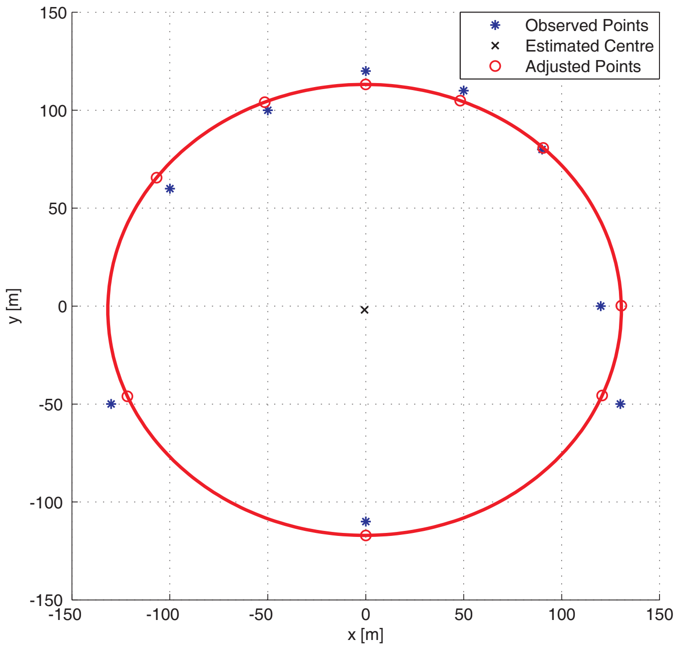


Figure 5.20.: Ellipse fit (mixed model), no restriction

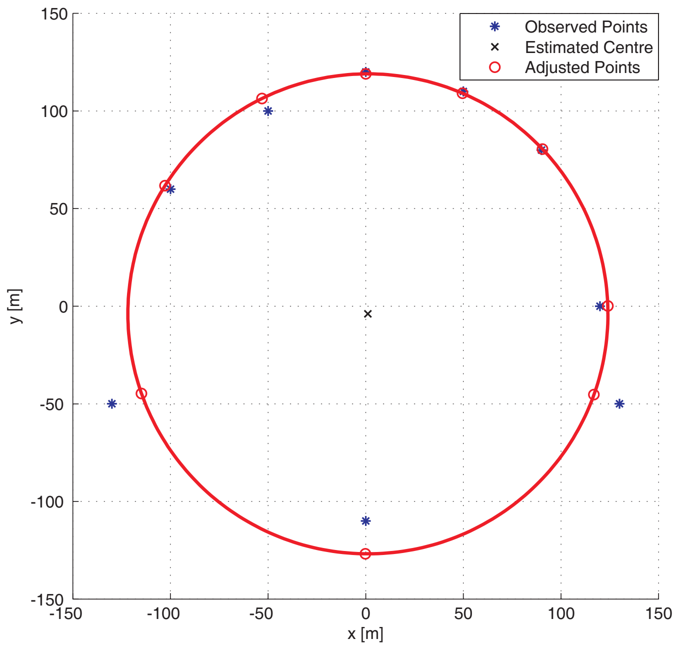


Figure 5.21.: Ellipse fit (mixed model), circle restriction

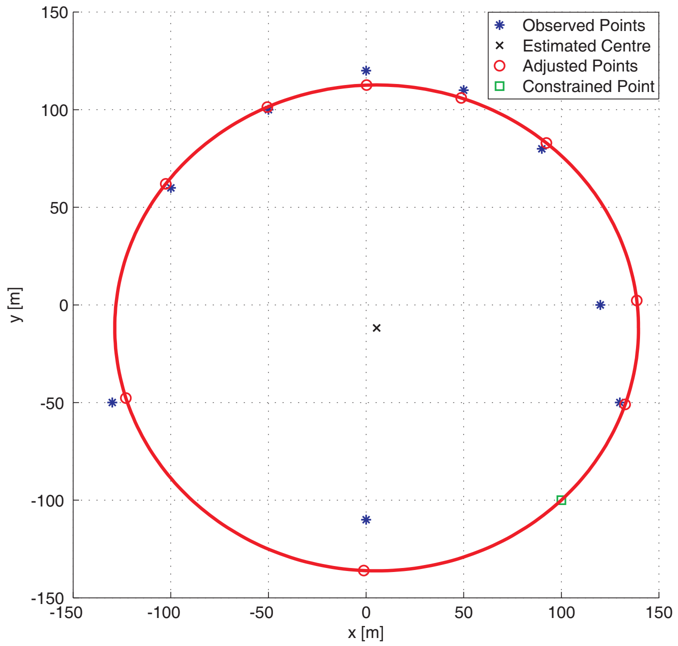


Figure 5.22.: Ellipse fit (mixed model), point restriction

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.

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

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

Linear algebra:

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

Q: Eigenvalues of a projector?

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

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

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

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

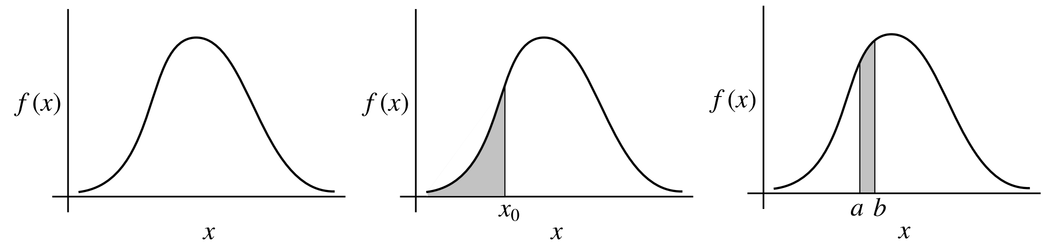
probability density function (PDF) Wahrscheinlichkeitsdichte

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

Note: not necessarily normal distribution.

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

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



(a) probability density function (b) probability calculations by integrating over the PDF (c) intervall

Figure 6.1.

Probability calculations by integrating over the PDF.

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

Cumulative density function

Verteilungsfunktion

cumulative distribution or density function (CDF)

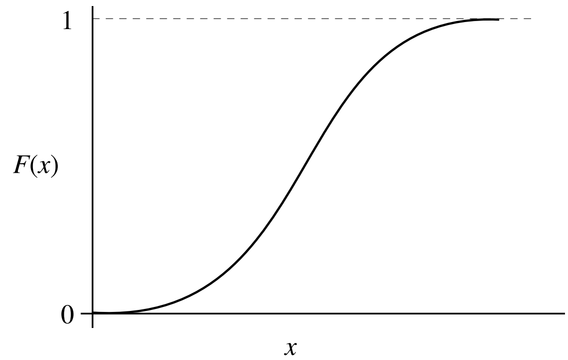


Figure 6.2.: cumulative distribution or density function

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

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.

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

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

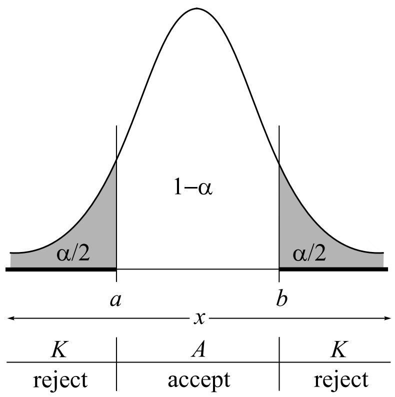


Figure 6.3.: Confidence and significance level

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

Sicherheitswahrscheinlichkeit
Irrtumswahrscheinlichkeit
Konfidenzbereich (Annahmebereich)
kritischer Bereich (Ablehnungs-, Verwerfungsbereich)

Now: given a realization x of $\underline{x}$

If $a \leq x \leq b$, there is no reason to reject the hypothesis.

Otherwise: reject hypothesis.

e. g.

$$\begin{aligned} \hat{\underline{e}} &= P_a^{\perp} \underline{y} \\ Q_{\hat{e}} &= P_a^{\perp} Q_y = Q_y - Q_{\hat{y}} \end{aligned}$$

Example Normal distribution

define a, b: determine α

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

MATLAB: `normpdf`

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

kritischer Wert $k =$ *critical Value*

define α: determine a, b

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

MATLAB: `norminv`

Rejection of hypothesis

$\Longrightarrow$ an alternative hypothesis must hold

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

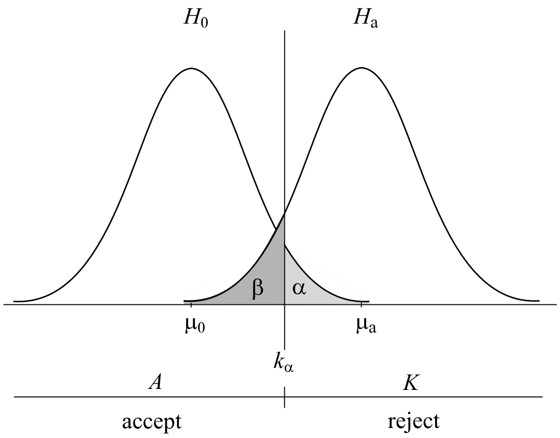


Figure 6.4.: accept or reject hypothesis?

	H_0 true	H_0 false
$x \in K$ $\Longrightarrow$ reject H_0	Wrong $\Longrightarrow$ Type I error (false alarm) $P(x \in K \mid H_0) = \alpha$	OK
$x \notin K$ $\Longrightarrow$ accept H_0	OK	Wrong $\Longrightarrow$ Type II error (failed Alarm) $P(x \notin K \mid H_a) = \beta$

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

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

Testg"ute

6.4. Distributions

Standard normal distribution (univariate)

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

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

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

Standard normal (multivariate) → Chi-square distribution

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

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

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

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

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

Non-standard normal → central Chi-square distribution

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

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

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

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

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

The same is true when

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

Non-standard normal → non-central Chi-square distribution

$$\underline{x} \sim N(\mu, I)$$

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

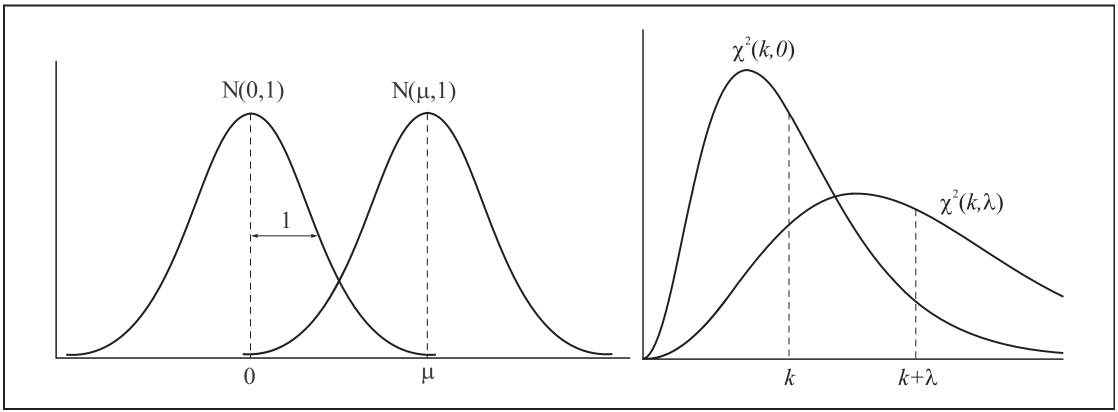


Figure 6.5.: Central/Non-central normal and χ^2-distribution

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

General case

$$
\underline{x} \sim N(\mu, Q_x)
$$

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

$$
\mathrm{E}\left\{\underline{x}\right\} = \mu
$$

$$
\mathrm{D}\left\{\underline{x}\right\} = Q_x
$$

$$
\mathrm{E}\left\{\underline{x}^{\mathsf{T}}Q_x^{-1}\underline{x}\right\} = k+\lambda, \qquad = \mu^{\mathsf{T}}Q_x^{-1}\mu
$$

7. Statistical Testing

7.1. Global model test: a first approach

Statistics of estimated residuals

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

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 $\underline{T}$:

$$
\begin{aligned}
\underline{T} &= \underline{\hat{e}}^{\mathsf{T}} Q_y^{-1} \underline{\hat{e}} \sim \chi^2(m - n, 0) \\
\mathrm{E}\{\underline{T}\} &= m - n
\end{aligned}
$$

Thus, we have a test statistic and we know its distribution. This is the starting point for global model testing.

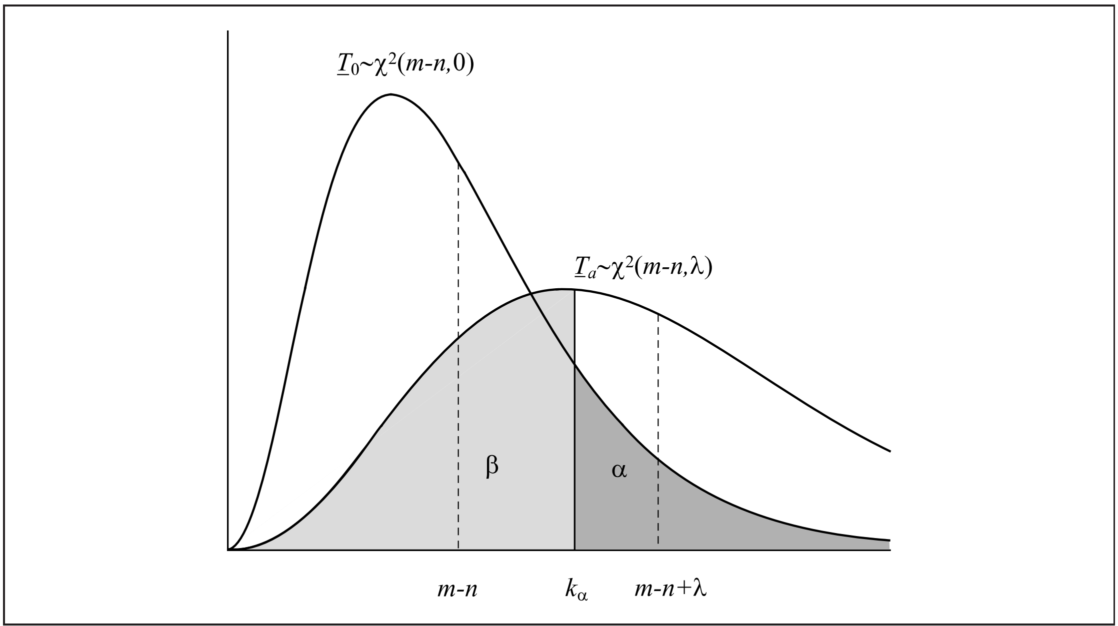


Figure 7.1.: Distribution of the test statistic $\underline{T}$ under the null and alternative hypotheses. (Non-centrality parameter λ to be explained later)

$T > k_\alpha$: 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:

- error in the (deterministic) observation model A,
- measurement error: $\mathrm{E}\{\underline{e}\} \neq 0$,
- wrong assumptions in the stochastic model: $\mathrm{D}\{\underline{e}\} \not\sim Q_y$.

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:

$$\underline{\hat{\sigma}}^2 = \frac{\underline{\hat{e}}^{\mathsf{T}} Q^{-1} \underline{\hat{e}}}{m - n} \Longrightarrow \mathrm{E}\left\{ \underline{\hat{\sigma}}^2 \right\} = \sigma^2 ,$$

unverzerrte Sch"atzung

which shows that $\underline{\hat{\sigma}}^2$ is an *unbiased estimate* of σ^2.

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.

$$\mathrm{E}\left\{\frac{\hat{\underline{\sigma}}^2}{\sigma^2}\right\} = 1$$

7.2. Testing procedure

Null hypothesis and alternative hypothesis

If the null hypothesis is described by $\mathrm{E}\{\underline{y}\} = Ax$, $\mathrm{D}\{\underline{y}\} = Q_y$, 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
$\mathrm{E}\{\underline{y}\} = Ax;\quad \mathrm{D}\{\underline{y}\} = Q_y$	$\mathrm{E}\{\underline{y}\} = Ax + C\nabla;\quad \mathrm{D}\{\underline{y}\} = Q_y$ $= \begin{pmatrix} A & C \end{pmatrix}\begin{pmatrix} x \\ \nabla \end{pmatrix}$
$\downarrow$	$\downarrow$
$\hat{\underline{x}}_0$	$\hat{\underline{x}}_\mathrm{a}, \hat{\underline{\nabla}}$
$\downarrow$	$\downarrow$
$\hat{\underline{y}}_0 = A\hat{\underline{x}}_0$	$\hat{\underline{y}}_\mathrm{a} = A\hat{\underline{x}}_\mathrm{a} + C\hat{\underline{\nabla}}$
$\downarrow$	$\downarrow$
$\hat{\underline{e}}_0$	$\hat{\underline{e}}_\mathrm{a}$
$\downarrow$	$\downarrow$
$\hat{\underline{e}}_0^\mathsf{T} Q_y^{-1}\hat{\underline{e}}_0 \sim \chi^2(m-n)$	$\hat{\underline{e}}_\mathrm{a}^\mathsf{T} Q_y^{-1}\hat{\underline{e}}_\mathrm{a} \sim \chi^2(m-n-q)$

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

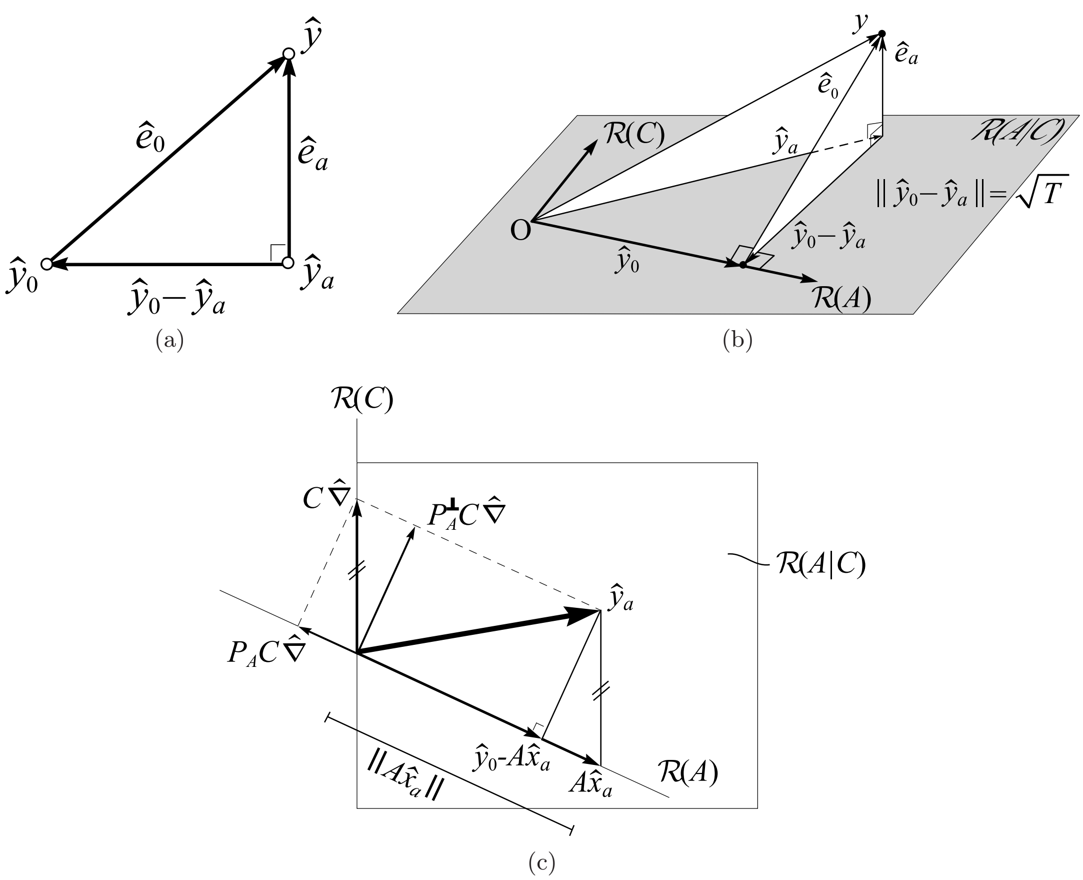


Figure 7.2.: Hypotheses H_0 and H_a

$$
\begin{aligned}
\underline{\hat{y}}_{\mathrm{a}} &= A\underline{\hat{x}}_{\mathrm{a}} + C\underline{\hat{\nabla}} \\
&= \underbrace{A\underline{\hat{x}}_{\mathrm{a}} + P_A C\underline{\hat{\nabla}}}_{\underline{\hat{y}}_0} + P_A^{\perp} C\underline{\hat{\nabla}} \\
\Longrightarrow \underline{\hat{y}}_{\mathrm{a}} - \underline{\hat{y}}_0 &= P_A^{\perp} C\underline{\hat{\nabla}} \\
\Longrightarrow \underline{T} &= P_A^{\perp} C\underline{\hat{\nabla}}^{\mathsf{T}} Q_y^{-1} P_A^{\perp} C\underline{\hat{\nabla}} = \underline{\hat{\nabla}}^{\mathsf{T}} C^{\mathsf{T}} \underbrace{P_A^{\perp\,\mathsf{T}} Q_y^{-1} P_A^{\perp}}_{=Q_y^{-1} P_A^{\perp} = Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1}} C\underline{\hat{\nabla}} \\
&= \underline{\hat{\nabla}}^{\mathsf{T}} C^{\mathsf{T}} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C\underline{\hat{\nabla}} \qquad \text{Second version} \\
\underline{T} &= \underline{\hat{e}}_0^{\mathsf{T}} Q_y^{-1} \underline{\hat{e}}_0 - \underline{\hat{e}}_{\mathrm{a}}^{\mathsf{T}} Q_y^{-1} \underline{\hat{e}}_{\mathrm{a}} \\
&= (\underline{\hat{y}}_0 - \underline{\hat{y}}_{\mathrm{a}})^{\mathsf{T}} Q_y^{-1} (\underline{\hat{y}}_0 - \underline{\hat{y}}_{\mathrm{a}}) \qquad \text{Third version} \\
&= \underline{\hat{\nabla}}^{\mathsf{T}} C^{\mathsf{T}} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C\underline{\hat{\nabla}}
\end{aligned}
$$

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}

$$
\begin{aligned}
H_0 \quad &: \quad A^{\mathsf{T}} Q_y^{-1} A \underline{\hat{x}}_0 = A^{\mathsf{T}} Q_y^{-1} \underline{y} \\
H_{\mathrm{a}} \quad &: \quad \begin{pmatrix} A^{\mathsf{T}} \\ C^{\mathsf{T}} \end{pmatrix} Q_y^{-1} \begin{pmatrix} A & C \end{pmatrix} \begin{pmatrix} \underline{\hat{x}} \\ \underline{\hat{\nabla}} \end{pmatrix} = \begin{pmatrix} A^{\mathsf{T}} \\ C^{\mathsf{T}} \end{pmatrix} Q_y^{-1} \underline{y} \\
\Longleftrightarrow &\begin{pmatrix} \underset{n\times n}{A^{\mathsf{T}} Q_y^{-1} A} & \underset{n\times q}{A^{\mathsf{T}} Q_y^{-1} C} \\ \underset{q\times n}{C^{\mathsf{T}} Q_y^{-1} A} & \underset{q\times q}{C^{\mathsf{T}} Q_y^{-1} C} \end{pmatrix} \begin{pmatrix} \underline{\hat{x}}_{\mathrm{a}} \\ \underline{\hat{\nabla}} \end{pmatrix} = \begin{pmatrix} A^{\mathsf{T}} Q_y^{-1} \underline{y} \\ C^{\mathsf{T}} Q_y^{-1} \underline{y} \end{pmatrix}
\end{aligned}
$$

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
$\underline{z} \sim N(0, Q_z)$	$\underline{z} \sim N(Q_z \underline{\hat{\nabla}}, Q_z)$
$\underline{T} \sim \chi^2(q, 0)$	$\underline{T} \sim \chi^2(q, \lambda)$
	$\lambda = \nabla^\mathsf{T} Q_z Q_z^{-1} Q_z \nabla = \nabla^\mathsf{T} C^\mathsf{T} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C \nabla$

Summary

Test quantity $\underline{T} = \underline{\hat{e}}_0^\mathsf{T} Q_y^{-1} \underline{\hat{e}}_0 - \underline{\hat{e}}_\mathrm{a}^\mathsf{T} Q_y^{-1} \underline{\hat{e}}_\mathrm{a}$ 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 $\underline{T}$

$$
\begin{aligned}
&(1)\quad \underline{\hat{e}}_0^\mathsf{T} Q_y^{-1} \underline{\hat{e}}_0 - \underline{\hat{e}}_\mathrm{a}^\mathsf{T} Q_y^{-1} \underline{\hat{e}}_\mathrm{a} \\
&(2)\quad (\underline{\hat{y}}_0 - \underline{\hat{y}}_\mathrm{a})^\mathsf{T} Q_y^{-1} (\underline{\hat{y}}_0 - \underline{\hat{y}}_\mathrm{a}) \\
&(3)\quad \underline{\hat{\nabla}}^\mathsf{T} C^\mathsf{T} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C \underline{\hat{\nabla}} \\
&(4)\quad \underline{\hat{e}}_0^\mathsf{T} Q_y^{-1} C \left(C^\mathsf{T} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C \right)^{-1} C^\mathsf{T} Q_y^{-1} \underline{\hat{e}}_0 \\
&(5)\quad \underline{z}^\mathsf{T} Q_z^{-1} \underline{z}; \quad \underline{z} := C^\mathsf{T} Q_y^{-1} \underline{\hat{e}}_0; \quad Q_z = C^\mathsf{T} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C
\end{aligned}
$$

Versions (1)–(3) explicitely involve the computation of H_a while cases (4) and (5) require only $\underline{\hat{e}}_0$ and some C.

For the reason that

$$
\begin{aligned}
\underline{z} &\sim N(0, Q_z) \qquad \text{under } H_0 \\
\underline{z} &\sim N(Q_z \nabla, Q_z) \qquad \text{under } H_\mathrm{a}
\end{aligned}
$$

test quantity $\underline{T}$ is distributed

$$
\begin{aligned}
\underline{T} &\sim \chi^2(q, 0) \qquad \text{under } H_0 \\
\underline{T} &\sim \chi^2(q, \lambda), \quad \lambda = \nabla^\mathsf{T} Q_z \nabla \qquad \text{under } H_\mathrm{a}
\end{aligned}
$$

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

$$
\begin{aligned}
\operatorname{rank}(A|C) &\overset{!}{=} n + q = n + (m - n) = m \\
&\Longrightarrow \quad o(A|C) = m \times n + q = m \times m \quad \text{“quadratic”} \\
&\Longrightarrow \quad \text{redundancy} = m - n - q = 0 \\
&\Longrightarrow \quad \underline{\hat{e}}_\mathrm{a} = 0 \\
&\Longrightarrow \quad \underline{\hat{y}}_\mathrm{a} = \underline{y} \\
&\Longrightarrow \quad \underline{T} = \underline{\hat{e}}_0^\mathsf{T} Q_y^{-1} \underline{\hat{e}}_0
\end{aligned}
$$

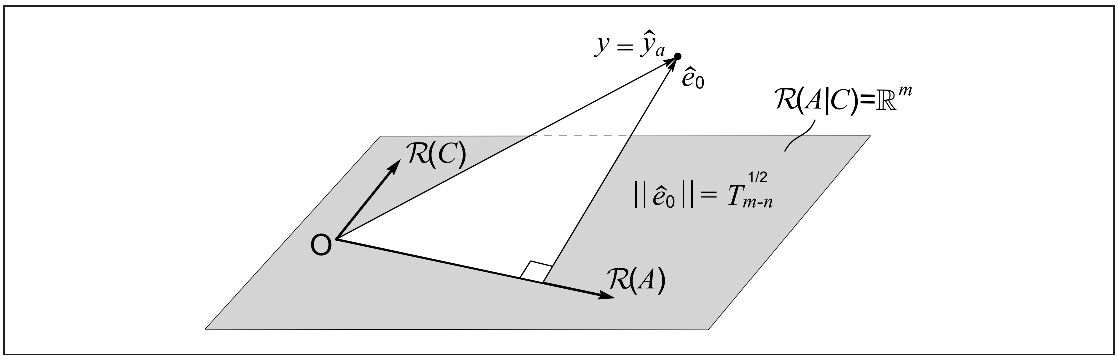


Figure 7.3.: Test quantity $\underline{T}$

$$H_0 : \mathrm{E}\left\{\underline{y}\right\} = Ax \quad \text{versus} \quad H_\mathrm{a} : \mathrm{E}\left\{\underline{y}\right\} \in \mathbf{R}^m$$

$$\underline{T} \sim \chi^2_{m-n,0}$$
$$\underline{T} \sim \chi^2_{m-n,\lambda}, \quad \lambda = \nabla^\mathsf{T} Q_z \nabla$$

For the reason that $\hat{\underline{e}}_\mathrm{a} = 0$, 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

$$\begin{aligned}
\underline{T} &= \hat{\underline{e}}_0^\mathsf{T} Q_y^{-1} C \left(C^\mathsf{T} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C \right)^{-1} C^\mathsf{T} Q_y^{-1} \hat{\underline{e}}_0 \\
&= \frac{\left(\hat{\underline{e}}_0^\mathsf{T} Q_y^{-1} C \right)^2}{C^\mathsf{T} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C} \\
&= \frac{\hat{\underline{\nabla}}^2}{\left(C^\mathsf{T} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C \right)^{-1}} \\
&= \frac{\hat{\underline{\nabla}}^2}{\sigma^2_{\hat{\nabla}}} \\
\hat{\underline{\nabla}}^2 &= \frac{C^\mathsf{T} Q_y^{-1} \hat{\underline{e}}}{C^\mathsf{T} Q_y^{-1} Q_{\hat{e}_0} Q_y^{-1} C}
\end{aligned}$$

Important application:

Detection of a gross error (outlier, blunder) in the observations, which leads to a wrong model specification.

$$H_0 : \mathrm{E}\{\underline{y}\} = Ax \qquad H_\mathrm{a} : \mathrm{E}\{\underline{y}\} = Ax + C\nabla$$

$$C = \left[0, 0, \ldots, \underbrace{1}_{i\text{-te Pos.}}, 0, \ldots, 0\right]^\mathsf{T}$$

Reject H_0 if $\underline{T} = \frac{\hat{\underline{\nabla}}^2}{\sigma^2_{\hat{\nabla}}} > k_\alpha$ or if $\sqrt{\underline{T}} = \frac{\hat{\underline{\nabla}}}{\sigma_{\hat{\nabla}}} < -\sqrt{k_\alpha}$ and $\sqrt{\underline{T}} = \frac{\hat{\underline{\nabla}}}{\sigma_{\hat{\nabla}}} > \sqrt{k_\alpha}$ ($\hat{\underline{\nabla}}$ 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

$$\sqrt{\underline{T}} = \frac{\hat{\underline{e}}_i}{\sigma_{\hat{e}_i}}$$

$$H_0 : \sqrt{\underline{T}} \sim N(0, 1) \qquad H_\mathrm{a} : \sqrt{\underline{T}} \sim N(\nabla\sqrt{\underline{T}}, 1)$$

$$\text{with } \nabla\sqrt{\underline{T}} = \sqrt{C^\mathsf{T} Q_y^{-1} Q_{\hat{e}} Q_y^{-1} C}\nabla$$

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.

$$\lambda_0 = \lambda\left(\alpha, q = m - n, \gamma = \gamma_0\right) = \lambda\left(\alpha_1, q = 1, \gamma = \gamma_0\right)$$

$q = 1$:

$$\left.\begin{array}{l}\gamma_0 = 1 - \beta_0 \\ \alpha_1\end{array}\right\} \Longrightarrow \lambda(\longrightarrow \mu) = \lambda_0$$

$q = m - n$:

$$\left.\begin{array}{l}\lambda_0 \\ \gamma_0 = 1 - \beta_0\end{array}\right\} \Longrightarrow \alpha = \alpha_{m-n}$$

e. g.: $\alpha_1 = 1\%$ (usually small), $\beta_0 = 20\% \Longrightarrow \alpha_{m-n} \approx 30\%$

7.4. Internal reliability

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*.

interne Zuverl"assigkeit

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.

Quastion: 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

$$\begin{aligned}
\underline{\hat{e}} &= R_A^{\perp} \underline{y} \\
&= \left[I - A \left(A^{\mathsf{T}} Q_y^{-1} A \right)^{-1} A^{\mathsf{T}} Q_y^{-1} \right] \underline{y} \\
&= R\underline{y} \qquad R = \text{redundancy matrix} \\
\underline{\hat{e}}_i &= R_{ij} \underline{y}_j \\
&= r_i \underline{y}_i + \ldots \\
\Longrightarrow \delta \underline{\hat{e}} &= r_i \delta \underline{y}_i
\end{aligned}$$

$\Longrightarrow$ 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

$$q = 1, c_i, Q_y = \text{diagonal}$$

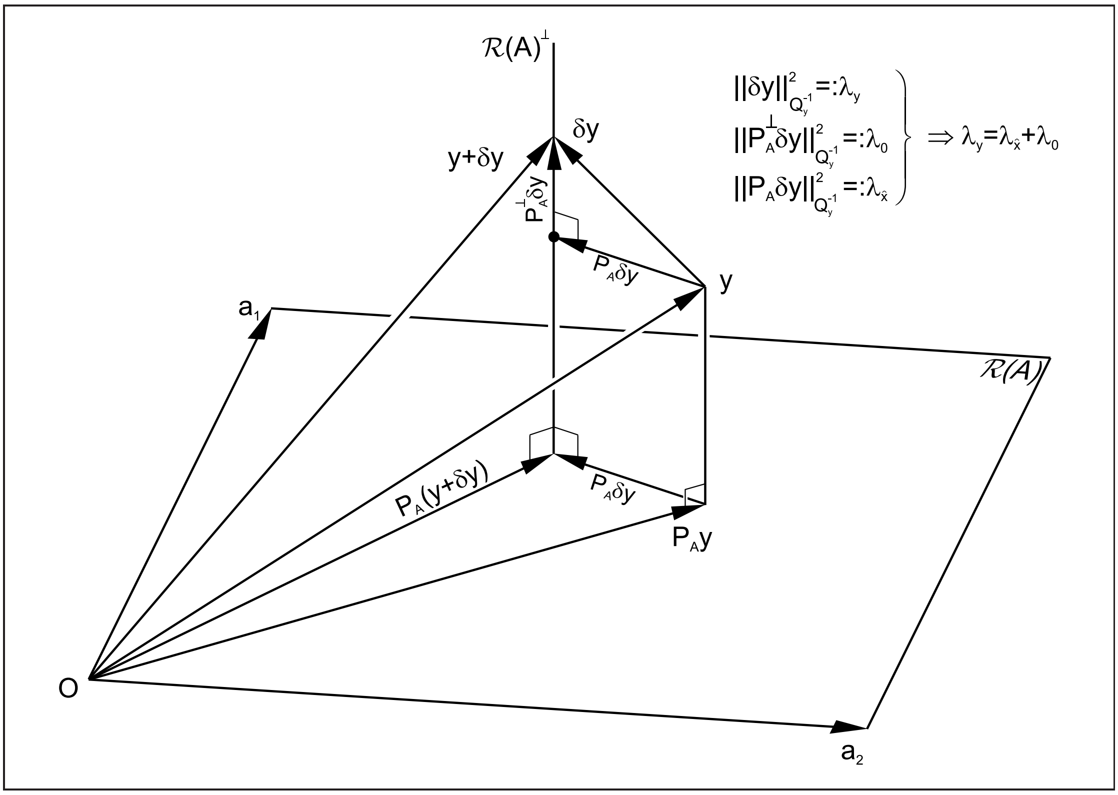


Figure 7.4.: Decomposition of λ_y

$$
\begin{aligned}
\Longrightarrow \underline{\lambda}_{\hat{x}} &= \underline{\lambda}_{y_i} - \underline{\lambda}_0 \\
&= \frac{1}{r_i}\underline{\lambda}_0 - \underline{\lambda}_0 \\
&= \frac{1 - r_i}{r_i}\underline{\lambda}_0 \\
&= \frac{\sigma^2_{\hat{y}_i}\sigma^2_{y_i}}{\sigma^2_{y_i}(\sigma^2_{y_i} - \sigma^2_{\hat{y}_i})}\underline{\lambda}_0 \\
&= \frac{\sigma^2_{\hat{y}_i}}{\sigma^2_{y_i} - \sigma^2_{\hat{y}_i}}\underline{\lambda}_0 \\
&= \frac{1}{\frac{\sigma^2_{y_i}}{\sigma^2_{\hat{y}_i}} - 1}\underline{\lambda}_0
\end{aligned}
$$

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 $\underline{\hat{x}}_{(2)}=\underline{\hat{x}}_{(1)}+\dots$.

$\Longrightarrow$ Solve

$$Q^{-1}_{\hat{x}_{(2)}}=Q^{-1}_{\hat{x}_{(1)}}+{A_2}^{\mathsf{T}}Q_2^{-1}A_2$$

for $Q^{-1}_{\hat{x}_{(1)}}$

$$\Longrightarrow Q^{-1}_{\hat{x}_{(1)}}=Q^{-1}_{\hat{x}_{(2)}}-{A_2}^{\mathsf{T}}Q_2^{-1}A_2$$

Substitute the result in $\underline{\hat{x}}_{(2)}=\left(Q^{-1}_{\hat{x}_{(1)}}+{A_2}^{\mathsf{T}}Q_2^{-1}A_2\right)^{-1}\left(Q^{-1}_{\hat{x}_{(1)}}\underline{\hat{x}}_{(1)}+{A_2}^{\mathsf{T}}Q_2^{-1}\underline{y}_2\right)$

$$\begin{aligned}\Longrightarrow \underline{\hat{x}}_{(2)}&=Q_{\hat{x}_{(2)}}\left(Q^{-1}_{\hat{x}_{(2)}}\underline{\hat{x}}_{(1)}-{A_2}^{\mathsf{T}}Q_2^{-1}A_2\underline{\hat{x}}_{(1)}+{A_2}^{\mathsf{T}}Q_2^{-1}\underline{y}_2\right)\\&=\underline{\hat{x}}_{(1)}+Q_{\hat{x}_{(2)}}{A_2}^{\mathsf{T}}Q_2^{-1}\left(\underline{y}_2-A_2\underline{\hat{x}}_{(1)}\right)\\&=\underline{\hat{x}}_{(1)}+K\underline{v}_2\\ \underline{v}_2&=\underline{y}_2-A_2\underline{\hat{x}}_{(1)}\\ K&=Q_{\hat{x}_{(2)}}{A_2}^{\mathsf{T}}Q_2^{-1}\end{aligned}$$

$A_2\underline{\hat{x}}_{(1)}\dots$ predicted observation
$\underline{v}_2\dots$ predicted residual (attention!)
$K\dots$ *gain matrix*

Verst"arkungs-matrix

disadvantage: too many matrix inversions

$$\underset{n\times n}{Q^{-1}_{\hat{x}_{(1)}}},\qquad \underset{n\times n}{\left(Q^{-1}_{\hat{x}_{(1)}}+{A_2}^{\mathsf{T}}Q_2^{-1}A_2\right)^{-1}},\qquad \underset{m_2\times m_2}{Q_2^{-1}}$$

8.1.4. Formulation using condition equations

$$B^{\mathsf{T}}A=0$$

$$\Longrightarrow\begin{pmatrix}-A_2 & I\end{pmatrix}\begin{pmatrix}I\\ A_2\end{pmatrix}=0$$

$$\left(-A_2 \;\; I \right) \mathrm{E}\left\{ \begin{pmatrix} \underline{\hat{x}}_{(1)} \\ \underline{y}_2 \end{pmatrix} \right\} = 0; \qquad \mathrm{D}\left\{ \begin{pmatrix} \underline{\hat{x}}_{(1)} \\ \underline{y}_2 \end{pmatrix} \right\} = \begin{pmatrix} Q_{\hat{x}_{(1)}} & 0 \\ 0 & Q_2 \end{pmatrix}$$

$$\left. \begin{aligned} B^\mathsf{T}\, \mathrm{E}\left\{\underline{y}\right\} &= 0 \\ \mathrm{D}\left\{\underline{y}\right\} &= Q_y \end{aligned} \right\} \Longrightarrow \underline{\hat{y}} = \left[I - Q_y B \left(B^\mathsf{T} QyB \right)^{-1} B^\mathsf{T} \right] \underline{y}$$

$$\Longrightarrow \begin{pmatrix} \underline{\hat{x}}_{(2)} \\ \underline{\hat{y}}_2 \end{pmatrix} = \left[\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} - \begin{pmatrix} -Q_{\hat{x}_{(1)}} {A_2}^\mathsf{T} \\ Q_2 \end{pmatrix} \left(Q_2 + A_2 Q_{\hat{x}_{(1)}} {A_2}^\mathsf{T} \right)^{-1} \left(-A_2 \;\; I \right) \right] \begin{pmatrix} \underline{\hat{x}}_1 \\ \underline{y}_2 \end{pmatrix}$$

$\Longrightarrow$ Measurement update

$$\begin{aligned} \underline{\hat{x}}_{(2)} &= \underline{\hat{x}}_{(1)} + Q_{\hat{x}_{(1)}} {A_2}^\mathsf{T} \left(Q_2 + A_2 Q_{\hat{x}_{(1)}} {A_2}^\mathsf{T} \right)^{-1} \left(\underline{y}_2 - A_2 \underline{\hat{x}}_{(1)} \right) \\ &= \underline{\hat{x}}_{(1)} + K \underline{v}_2 \\ \underset{m_2 \times m_2}{K} &= Q_{\hat{x}_{(1)}} {A_2}^\mathsf{T} \left(Q_2 + A_2 Q_{\hat{x}_{(1)}} {A_2}^\mathsf{T} \right)^{-1} \end{aligned}$$

$\Longrightarrow$ 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:

$$\begin{aligned}
\underline{\hat{x}}_{(k)} &= \underline{\hat{x}}_{(k-1)} + K_k \underline{v}_k \\
\underline{v}_k &= \underline{y}_k - A_k \underline{\hat{x}}_{(k-1)} \\
K_k &= Q_{\hat{x}_{(k-1)}} A_k{}^{\mathsf{T}} \left(Q_k + A_k Q_{\hat{x}_{(k-1)}} A_k{}^{\mathsf{T}} \right)^{-1} \\
Q_{\hat{x}_{(k)}} &= \left[I - K_k A_k \right] Q_{\hat{x}_{(k-1)}}
\end{aligned}$$

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. $-b^{\mathsf{T}}B = I$

1. $\underbrace{-b^\mathsf{T}A}_{=0} - b^\mathsf{T}bC = -b^\mathsf{T} \Longrightarrow C = -(b^\mathsf{T}b)^{-1}b^\mathsf{T}$

2. $-b^\mathsf{T}B + b^\mathsf{T}bD = 0 \Longrightarrow I + b^\mathsf{T}bD = 0 \Longrightarrow D = -(b^\mathsf{T}b)^{-1}$

1. $A + b(b^\mathsf{T}b)^{-1}b^\mathsf{T} = I \Longrightarrow A = I - b(b^\mathsf{T}b)^{-1}b^\mathsf{T}$

2. $B + b(b^\mathsf{T}b)^{-1} = 0 \Longrightarrow B = -b(b^\mathsf{T}b)^{-1}$

$$\begin{bmatrix} I & -b \\ -b^\mathsf{T} & 0 \end{bmatrix}^{-1} = \begin{bmatrix} I - b(b^\mathsf{T}b)^{-1}b^\mathsf{T} & -b(b^\mathsf{T}b)^{-1} \\ -(b^\mathsf{T}b)^{-1}b^\mathsf{T} & -(b^\mathsf{T}b)^{-1} \end{bmatrix}$$

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

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

$$HDQ = 0 \Longrightarrow Q = 0$$

$$\begin{aligned} \text{A.1} + D \cdot \text{A.3} &\Longrightarrow (A^{\mathsf{T}}A)R + D(HD)^{-1}H + DD^{\mathsf{T}}R = I \\ &\Longrightarrow (A^{\mathsf{T}}A + DD^{\mathsf{T}})R = I - D(HD)^{-1}H \\ &\Longrightarrow R = (A^{\mathsf{T}}A + DD^{\mathsf{T}})^{-1}(I - D(HD)^{-1}H) \end{aligned}$$

Inserting R and S into the normal equations

$$\begin{cases} \hat{x} = RA^{\mathsf{T}}y = (A^{\mathsf{T}}A + DD^{\mathsf{T}})^{-1}A^{\mathsf{T}}y - (A^{\mathsf{T}}A + DD^{\mathsf{T}})^{-1}D(HD)^{-1}\underbrace{HA^{\mathsf{T}}}_{0}y \\ \hat{\lambda} = SA^{\mathsf{T}}y = (HD)^{-1}\underbrace{HA^{\mathsf{T}}}_{0}y = 0 \end{cases}$$

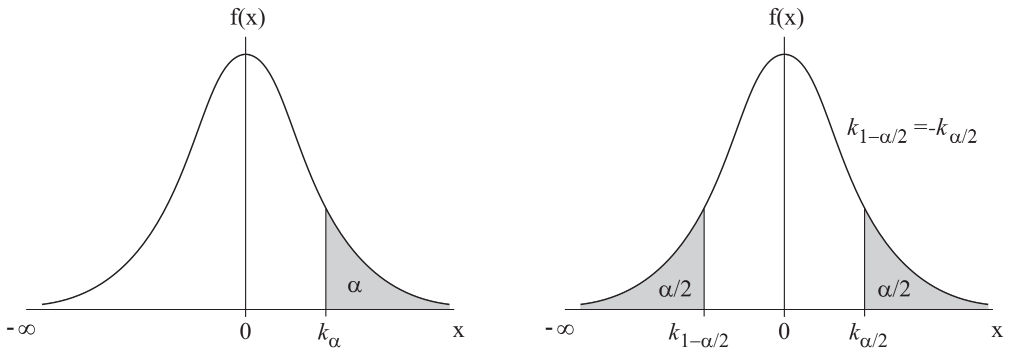
$$\begin{aligned} &\Longrightarrow \hat{\lambda} = 0 \\ &\Longrightarrow \hat{x} = (A^{\mathsf{T}}A + DD^{\mathsf{T}})^{-1}A^{\mathsf{T}}y + H(H^{\mathsf{T}}D)^{-1}c \\ \hat{e} \;\; = \;\; & y - A\hat{x} = (I - A(A^{\mathsf{T}}A + DD^{\mathsf{T}})^{-1}A^{\mathsf{T}})y \end{aligned}$$

B. Statistical Tables

B.1. Standard Normal Distribution

$$\text{Computation of one-sided level of significance } \alpha = 1 - \int_{-\infty}^{k_\alpha} f(x)\,dx \text{ and}$$

$$\text{two-sided level of significance } \alpha = 2 \int_{-\infty}^{k_{1-\alpha/2}} f(x)\,dx$$



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 $\alpha = 1 - \int_{-\infty}^{k_\alpha} f(x)\,dx$ and

two-sided level of significance $\alpha = 2 \int_{-\infty}^{k_{1-\alpha/2}} f(x)\,dx$ (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:

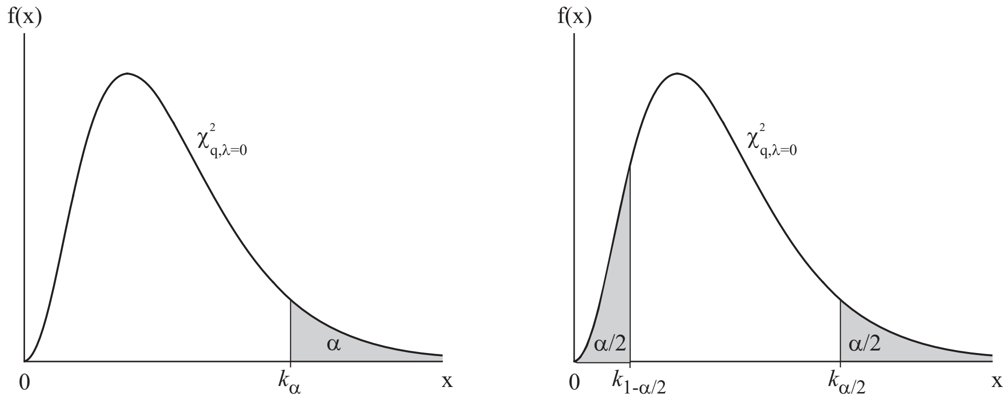
α=1-normcdf(k_α) k_α=norminv(1-α)

Example (k_α=0.87, one-sided):

α=0.1922=1-normcdf(0.87) k_α=0.87=norminv(1-0.1922)

B.2. Central χ^2-Distribution

Computation of critical value $k_\alpha = \chi^2_{1-\alpha}(q, \lambda = 0)$



q\α	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 $k_\alpha = \chi^2_{1-\alpha}(q, \lambda = 0)$ (continued)

q\α	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

q\α	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 $k_\alpha = \chi^2_{1-\alpha}(q, \lambda = 0)$ (continued)

q\α	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_α=chi2inv(1-α,q) α=1-chi2cdf(k_α,q)

Example (α=0.95, one-sided):

k_α=11.59=chi2inv(1-0.95,21) α=0.95=1-chi2cdf(11.59,21)

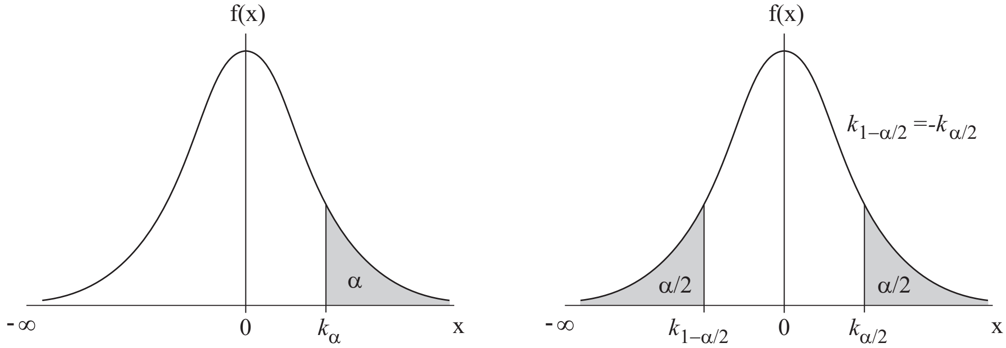
Example (α=0.01, two-sided):

$k_{\alpha/2}$=41.4=chi2inv(1-0.01/2,21)

$k_{1-\alpha/2}$=8.034=chi2inv(0.01/2,21)

B.3. Central t-Distribution

Computation of critical value $k_\alpha = t_{1-\alpha}(q)$



q\α	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 $k_\alpha = t_{1-\alpha}(q)$

q\\α	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_α=tinv(1-α,q) α=1-tcdf(k_α,q)

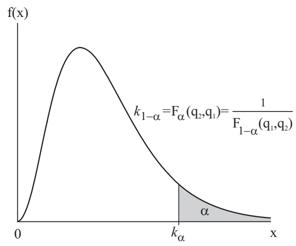
Example (α=0.005, one-sided):

k_α=2.898=tinv(1-0.005,17) α=0.005=1-tcdf(2.898,17)

.

B.4. Central F-Distribution

Computation of critical value $k_\alpha = F_{1-\alpha}(q_1, q_2, \lambda = 0)$



α=0.10, 1-α=0.90

$q_2 \backslash q_1$	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 $k_\alpha = F_{1-\alpha}(q_1, q_2, \lambda = 0)$ (continued)

α=0.10, 1-α=0.90

$q_2 \backslash q_1$	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 $k_\alpha = F_{1-\alpha}(q_1, q_2, \lambda = 0)$

α=0.05, 1-α=0.95

$q_2 \backslash q_1$	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 $k_\alpha = F_{1-\alpha}(q_1, q_2, \lambda = 0)$ (continued)

α=0.05, 1-α=0.95

$q_2 \backslash q_1$	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 $k_\alpha = F_{1-\alpha}(q_1, q_2, \lambda = 0)$

α=0.025, 1-α=0.975

$q_2 \backslash q_1$	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 $k_\alpha = F_{1-\alpha}(q_1, q_2, \lambda = 0)$ (continued)

α=0.025, 1-α=0.975

$q_2 \backslash q_1$	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 $k_\alpha = F_{1-\alpha}(q_1, q_2, \lambda = 0)$

α=0.01, 1-α=0.99

$q_2 \backslash q_1$	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 $k_\alpha = F_{1-\alpha}(q_1, q_2, \lambda = 0)$ (continued)

α=0.01, 1-α=0.99

$q_2 \backslash q_1$	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_α=finv(1-α, q_1, q_2)

Example (α=0.05, one-sided):

k_α=2.774=finv(1-0.05,20,10)

$$k_{1-\alpha}=0.360=\frac{1}{\text{finv(1-0.05,20,10)}}=\text{finv(0.05,10,20)}$$

C. Book recommendations and other material

C.1. Scientific books

- *Anderson, James M. and Edward M. Mikhail*
  Surveying. Theory and Practice
  7^{th} edition
  McGraw-Hill, 1998
  ISBN 0-07-015914-9
- *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
- *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 Aus-

bildung 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
  Addision-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$^{\text{th}}$ 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

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

- *Strang Gilbert*
  Linear Algebra, MIT Course 18.06
  http://ocw.mit.edu/courses/mathematics/18-06-linear-algebra-spring-2010/