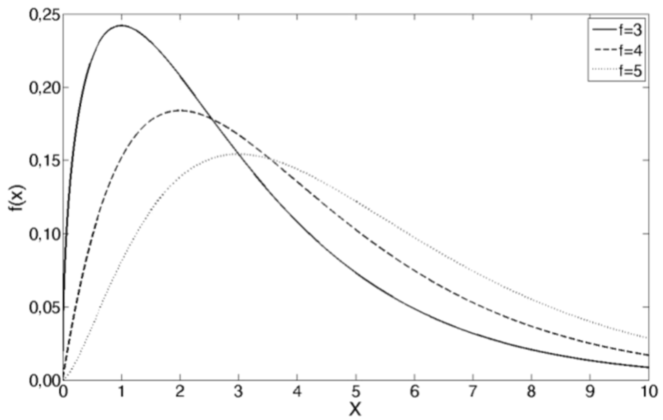


Adjustment Theory I

General Information

Prof. Dr.-Ing Frank Neitzel | Institute of Geodesy and Geoinformation Science

Version: 16.10.2024

Prof. Dr.-Ing. Frank Neitzel, Dr.-Ing. Sven Weisbrich

Leaflet for Module: Adjustment Theory I

Content and timing

Module type: compulsory optional subject

Workload: 4 semester periods per week (SPW)

- 2 SPW lecture
- 2 SPW exercise

Schedule:

Monday	10:15 am – 11:45 am:	lecture (H 6131)
Friday	8:15 am – 9:45 am:	exercise, Group A (H 6134)
	10:15 pm – 11:45 pm:	exercise, Group B (H 6134)

Contents

- Definitions of basic statistical terms
- Random variables
- The random vector
- Propagation of observation errors
- The Gaussian or Normal Distribution
- Introduction to least squares adjustment
- Applications of least squares adjustment
- Least squares adjustment with constraints for the unknown parameters

Methods

- Lectures, blackboard writings and presentations
- Exercises in the computer lab (H6134)

Presentations, lecture notes and further material will be provided on the ISIS system. https://isis.tu-berlin.de/

Examination

Examination: written exam at the end of the semester in **March 2025**.

Homework

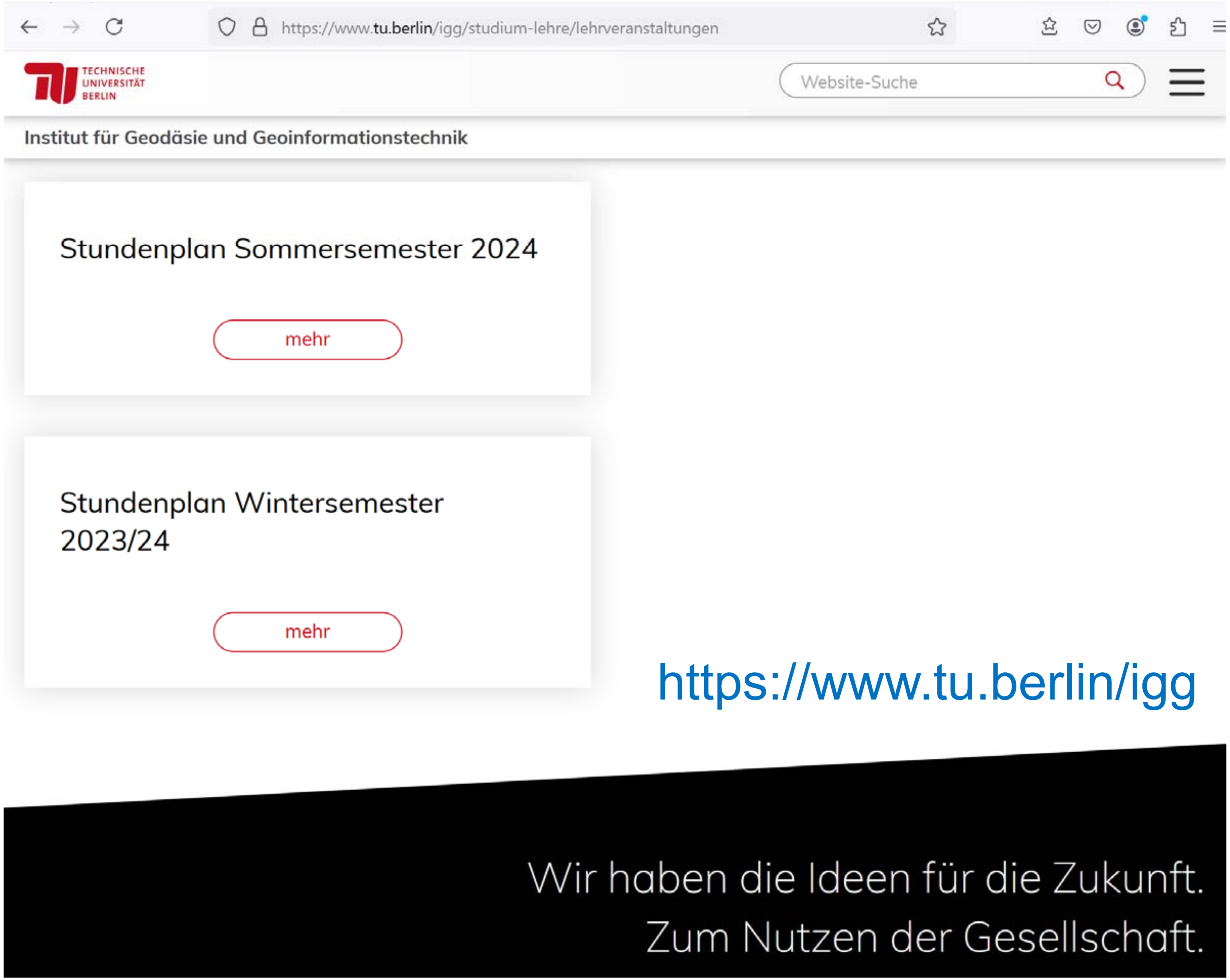
Every homework should be prepared by a **group** of 5 students (min. 4 students) and submitted **in time**.

Every homework will be corrected individually and will be evaluated according to the following scheme:

+ +	Correct solution
+	Some minor mistakes
-	Too many errors, reconsideration highly recommended

Credit Points

The module "Adjustment Theory I" is considered to be passed if the written exam has been marked sufficiently. Successful participants will receive 6 ECTS point.



https://www.tu.berlin/igg

o´clock	Monday 1st semester	Monday 3rd semester	Monday 3rd semester	Tuesday 1st semester	Tuesday 3rd semester	Wednesday 1st semester	Wednesday 3rd semester	Wednesday 3rd semester
8–10		IV 3633 L 9079 Kada **Semantic 3D/4[illegible]dels**					IV 3633 L xxx Ge **Real-Time Multi-Sensor Precise Positioning** Start: Block course from 6 - 17 Jan. 2025 More information on ISIS	
10–12	IV 3633 L 201 Neitzel **Adjustment Theory I Lecture** Start: 21.10.2024 H6131	**Stereo Image Processing** Start: 14.10.2024 LF 05-28		VL 3633 L 9070 Kada **Geoinformatics** Start: 15.10.2024 H 6131	IV 3633 L 9079 Löwner / Wamhoff **Semantic 3D/4D City Models** Start: 15.10.2024 H 6134		VL 3633 L 218 Oberst / Neumann **Geodesy and Dynamics of Outer Planets** Start: 23.10.2024 H 6131	
12–14	VL 0433 L 120 Hellwich **Photogrammetric Computer Vision**	VL 3633 L 90xx Kada **Deep Learning for Geographical Data** Start: 14.10.2024 H 6131		VL 3633 L 90xx Kada **Geodatabases** Start: 15.10.2024 H 6131	IV 3633 L 246 Galas **GNSS Signal Processing and Data Communication** Start: 15.10.2024 H 6134	IV 3633 L 202 Flechtner / Schuh / Oberst Galas / Wickert **Introduction to Space Geodesy** Start: 16.10.2024 H 6131		PJ 0433 L 160 Hellwich **Project Hot Topics in Computer Vision A** Start: 16.10.204 MAR 0.015
14–16	Start: 14.10.2024 H 107	UE 3633 L 90xx Kaufhold **Deep Learning for Geographical Data Exercise** Start: 14.10.2024 H 6131		UE 3633 L 90xx Löwner **Geodatabases Exercise, Group B** Start: 15.10.2024 H 6134	IV 3633 L 299 Wickert **Geoscientific Aspects of Geodesy** Start: 15.10.2024 H 6131	IV 3633 L 202 Heinkelmann **Introduction to Space Geodesy Exercise** Start: 23.10.2024 H 6131	IV 3633 L 241 Schuh / Männel / Galas **Space Geodetic Techniques** Start: 23.10.2024 H 6134	
16–18	UE 3633 L 90xx Löwner **Geodatabases Exercise, Group A** Start: 21.10.2024 H 6134	IV 3633 L 241 Schuh / Männel / Galas **Space Geodetic Techniques** Start: 21.10.2024 H 6131	VL 0433 L 110 Hellwich **Digital Image Processing** Start: 14.10.2024 A 151				PJ 3633 247 Schuh / Wickert / Flechtner Oberst / Galas / Ge **Space Geodesy and Navigation Project** Start: 16.10.2024 H 6131	
18–20								

o´clock	Thursday		Friday		
	1st semester	3rd semester	1st semester	3rd semester	
8–10			IV 3633 L 201 Neitzel / Weisbrich **Adjustment Theory I** **Exercise, Group A** Start: 18.10.2024 H 6134	**Seminar** **Geoinformation Technology** [illegible]	PJ 3633 L 90xx Kada **Project** **Geoinformation Technology** Start: 18.10.2024 H 6131
10–12	UE 3633 L 9071 Kaufhold **Geoinformatics** **Exercise, Group A** Start: 17.10.2024 H 6134	IV 3633 L 231 Neitzel / Weisbrich **Analysis of Stochastic Processes** Start: 17.10.2024 H 6131	IV 3633 L 201 Neitzel / Weisbrich **Adjustment Theory I** **Exercise, Group B** Start: 18.10.2024 H 6134	**Digital Image Processing** Start: 18.10.2024 H 3010	
12–14	UE 3633 L 9071 Kaufhold **Geoinformatics** **Exercise, Group B** Start: 17.10.2024 H 6134	IV 3633 L 253 Neitzel / Weisbrich **Transformation of** **Geodetic Networks** Start: 17.10.2024 H 6131	UE 0433 L 121 Hellwich **Photogrammetric** **Computer Vision** Start: 18.10.2024 Room: To be announced	PJ 3633 L 255 Neitzel / Weisbrich **Engineering Geodesy and** **Adjustment Calculation Project** Start: 18.10.2024 H 6131 / Geodätenstand	
14–16		PJ 3633 L 250 Neitzel / Weisbrich **Engineering Geodesy B**			
16–18		Start: 17.10.2024 H 6131 / Geodätenstand Colloquium 3633 L 990 Neitzel / Sorge **Geodetic Colloquium**			
18–20		look at board H 6131			

https://www.tu.berlin/igg/studium-lehre/studienangebot/module

TECHNISCHE UNIVERSITÄT BERLIN

Anmelden Webmail Website-Suche und Direktzugang Fakultäten / Zentralinstitute DE / EN

Institut für Geodäsie und Geoinformationst STUDIUM & LEHRE FORSCHUNG EINRICHTUNGEN & SERVICES ÜBER UNS PODCASTS

Sie befinden sich hier: Technische Universität Berlin » Fakultät VI - Planen Bauen Umwelt » Institut für Geodäsie und Geoinformationstechnik » Studium & Lehre » Studienangebot » Module

Modulübersicht

Modulkatalog

Das Masterstudium umfasst Module mit einem Gesamtumfang von mindestens 120 Leistungspunkten (LP).
Die Module sind wie folgt zu wählen:

1. Semester:

Es müssen Module im Umfang von 30 LP aus dem Basisbereich (FOU Foundation Section) gewählt werden.

2./3. Semester:

Aus vier möglichen Themenblöcken (GIS, SGN, EGA, CV) ist ein Vertiefungsbereich auszuwählen, in dem mindestens 21 LP gesammelt werden müssen.
In diesem Vertiefungsbereich ist das Projektseminar (mind. 6 LP) ein Pflichtfach. Aus den anderen drei Themenblöcken sind jeweils Module im Umfang von 9 LP zu wählen.

Eine beliebige Anzahl von Modulen aus dem fächerübergreifenden Angebot der TU Berlin und anderen nat./int. Universitäten ist im Bereich "freie Wahl" zu treffen. Hier müssen insgesamt mindestens 12 LP gesammlt werden. Die Wahl eines Sprachmoduls sowie von Veranstaltungen aus dem FÜS-Angebot der TU Berlin wird empfohlen es können aber auch Kurse aus dem Angebot des Masterprogramms "M.Sc. Geodesy and Geoinformation Science" gewählt werden.

Adjustment Theory I

Module title:	Credits:	Responsible person:
Adjustment Theory I	6	Neitzel, Frank
	Office:	**Contact person:**
	No information	*No information*
Website:	**Display language:**	**E-mail address:**
No information	Englisch	ega@geodesy.tu-berlin.de

Learning Outcomes

Students will gain a profound theoretical and methodical knowledge in variance-covariance propagation and parameter estimation via least-squares adjustment. They are able to design functional and stochastic models and apply these on geodetic and general engineering tasks.

Content

Adjustment Theory I [IV 3633 L 201]

Content

- Distribution of a random variable
- Mean value, expectation and true value of a random variable
- Variance and standard deviation
- Random vectors
- Propagation of observation errors
- Propagation of variances and covariances
- The Gaussian or Normal Distribution
- Introduction to least squares adjustment
- Functional model and stochastic model
- Observation equations
- Linear functional models
- Non-linear functional models
- Variance-Covariance Matrices (VCM) for the results
- Least squares adjustment with constraints for the unknown parameters
- Applications of least squares adjustment

Didactic concept

-Lecture 50%
-Exercise 50%

Module Components

Course Name	Type	Number	Cycle	SWS
Adjustment Theory I	IV	3633 L 201	WS	4

Workload and Credit Points

Adjustment Theory I (Integrierte Veranstaltung)	Multiplier	Hours	Total
Overall attendance	15.0	4.0h	60.0h
Preparation and planning	15.0	2.0h	30.0h
Homework	4.0	15.0h	60.0h
Examination preparation	1.0	30.0h	30.0h
			180.0h

The Workload of the module sums up to 180.0 Hours. Therefore the module contains 6 Credits.

Description of Teaching and Learning Methods

- Lecture 50%
- Exercise 50%

Requirements for participation and examination

Desirable prerequisites for participation in the courses:

Profound knowledge of Linear Algebra, basic knowledge of Applied Geodesy, basic knowledge of MatLab / Octave

Mandatory requirements for the module test application:

No information

Module completion

Grading:	**Type of exam:**	**Language:**	**Duration/Extent:**
graded	Written exam	English	120 Minuten

Duration of the Module

This module can be completed in one semester.

Maximum Number of Participants

This module is not limited to a number of students.

Registration Procedures

no information.

Recommended reading, Lecture notes

Lecture notes:
unavailable

Electronical lecture notes :
available

Recommended literature:

Ghilani C. & Wolf P. (2006): Adjustment Computations. New Jersey, John Wiley

Harvey, B.R. (2016): Practical Least Squares and Statistics for Surveyors. Monograph 13, Third Edition reprinted, Surveying, CVEN, The University of New South Wales, Australia, 332 + x pp., ISBN 0-7334-2339-6

Mikhail, E. & Ackerman, F. (1976): Observations and Least Squares. New York: University Press of America

Niemeier, W. (2008): Ausgleichungsrechnung. Walter de Gruyter, New York

Assigned Degree Programs

This moduleversion is used in the following modulelists:

Geodesy and Geoinformation Science (Master of Science)
StuPO (21.03.2007)
Modullisten der Semester: WS 2021/22

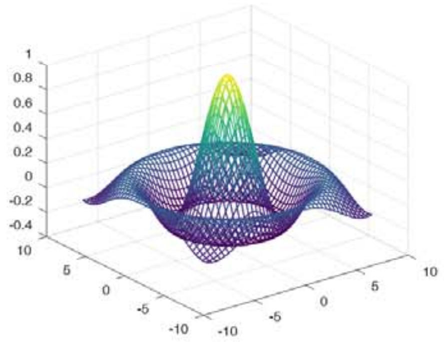
Miscellaneous

No information

https://octave.org

GNU Octave

About Donate Download Get Involved Support



GNU Octave

Scientific Programming Language

- Powerful mathematics-oriented syntax with built-in 2D/3D plotting and visualization tools
- Free software, runs on GNU/Linux, macOS, BSD, and Microsoft Windows
- Drop-in compatible with many Matlab scripts

https://octave.org

</> Syntax Examples

The Octave syntax is largely compatible with Matlab. The Octave interpreter can be run in GUI mode, as a console, or invoked as part of a shell script. More Octave examples can be found in the Octave wiki.

Solve systems of equations with linear algebra operations on **vectors** and **matrices**.

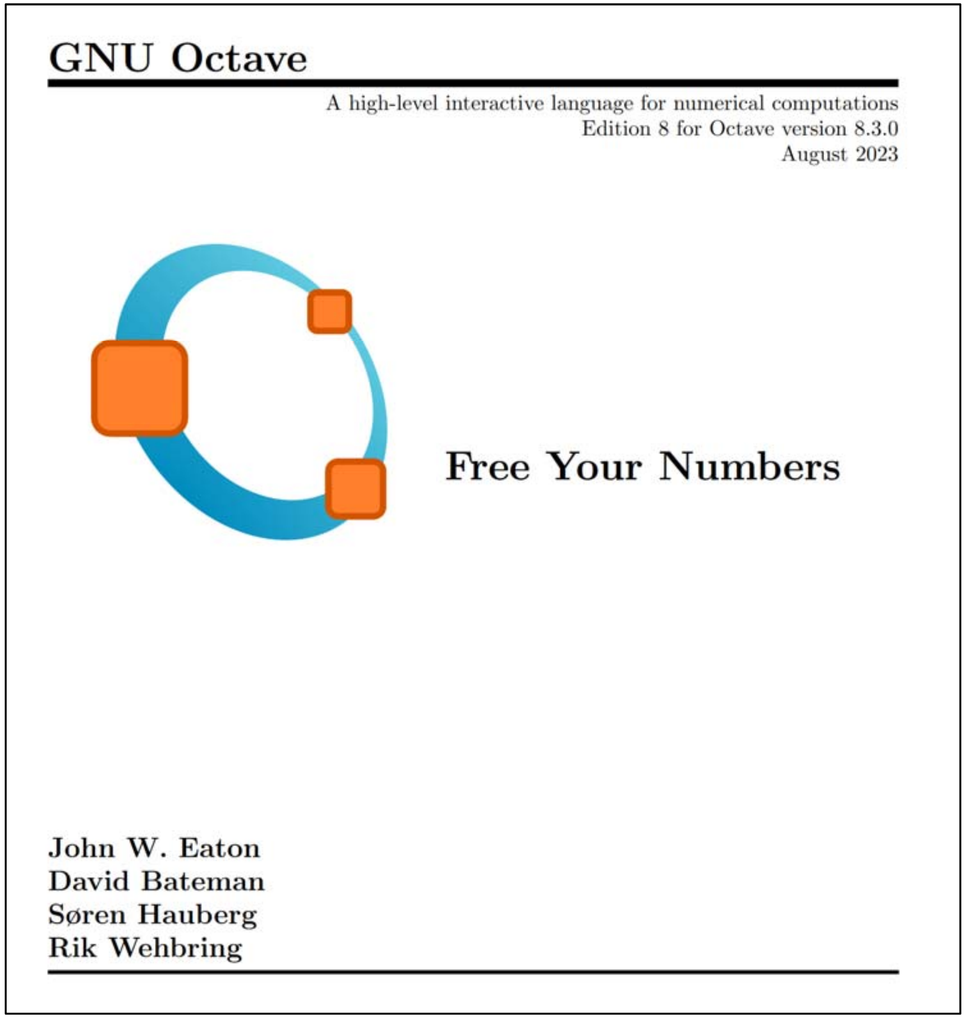
```
b = [4; 9; 2] # Column vector
A = [ 3 4 5;
      1 3 1;
      3 5 9 ]
x = A \ b     # Solve the system Ax = b
```

Visualize data with **high-level plot commands** in 2D and 3D.

```
x = -10:0.1:10; # Create an evenly-spaced vector from -10..10
y = sin (x);    # y is also a vector
plot (x, y);
title ("Simple 2-D Plot");
xlabel ("x");
ylabel ("sin (x)");
```

Click here to see the plot output

Octave Manual


GNU Octave

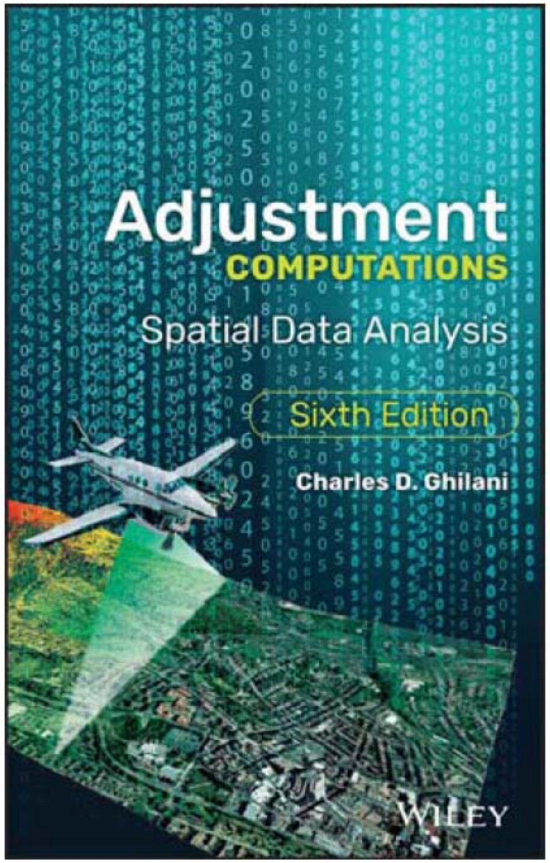
A high-level interactive language for numerical computations
Edition 8 for Octave version 8.3.0
August 2023

Free Your Numbers

John W. Eaton
David Bateman
Søren Hauberg
Rik Wehbring

https://octave.org/octave.pdf

Literature

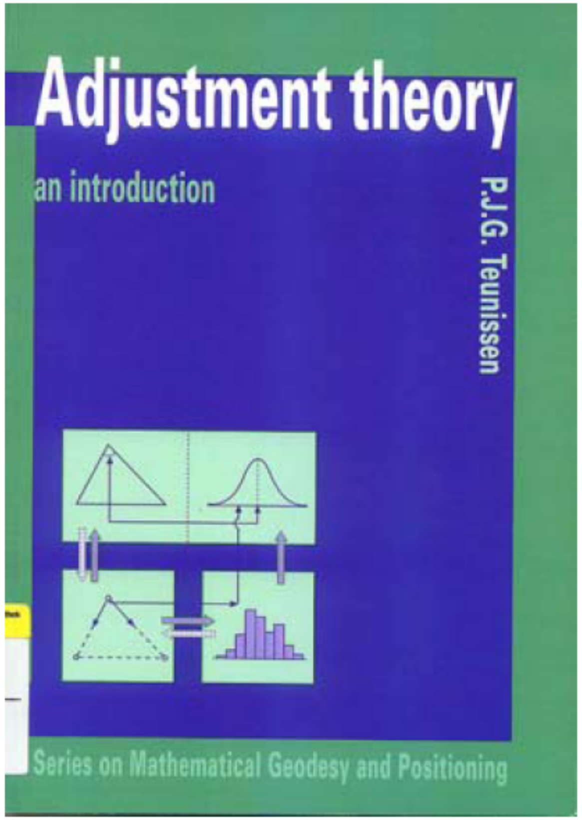


Adjustment Computations: Spatial Data Analysis

by Charles D. Ghilani

- **Publisher:** John Wiley & Sons; 6th edition (2017)
- **Language:** English
- **ASIN:** B07666XDCC
- **Pages:** 721

Literature

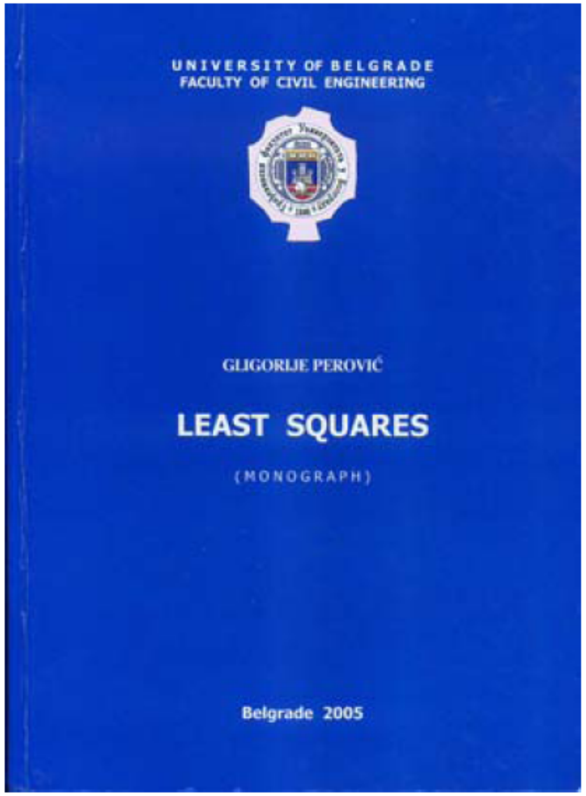


Adjustment theory, an introduction

by P.J.G. Teunissen

- **Publisher:** VSSD;
  1st edition (2003)
- **Language: English**
- **ISBN-13:** 978-9040719745
- **Pages:** 202

Literature

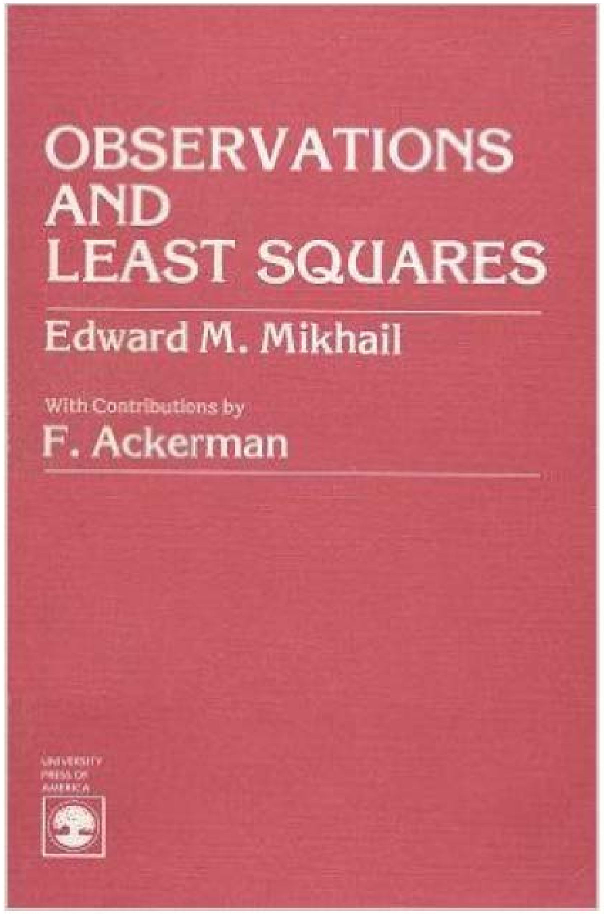


Least Squares

by Gligorije Perovic

- **Publisher:** Faculty of Civil Engineering, University of Belgrade (2005)
- **Language: English**
- **ISBN:** 9788690740901
- **Pages:** 648

Literature

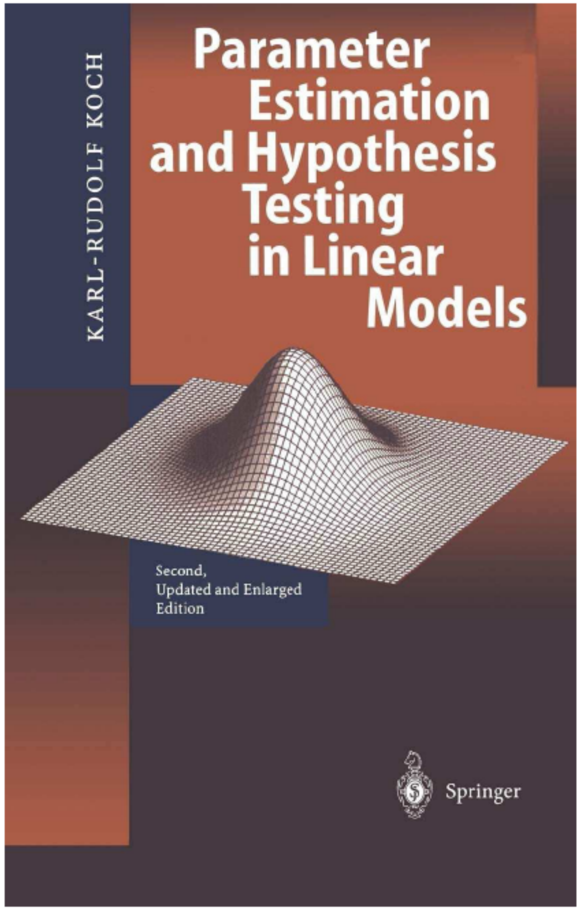


Observations and Least Squares

by Edward M. Mikhail

- **Publisher:** Rowman & Littlefield; illustrated edition (1983)
- **Language: English**
- **ISBN-10:** 0819123978
- **ISBN-13:** 978-0819123978
- **Pages:** 497

Literature



Parameter Estimation and Hypothesis Testing in Linear Models

by Karl-Rudolf Koch

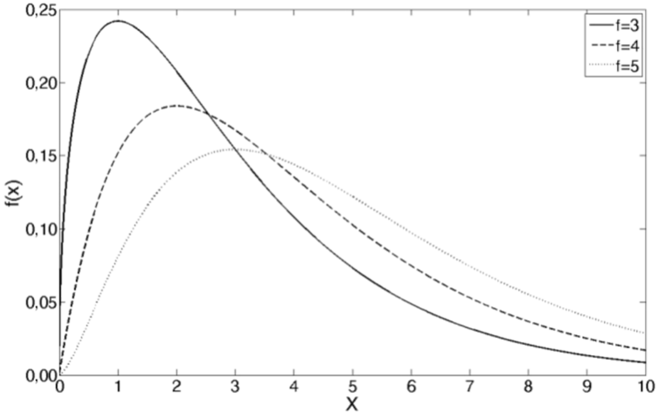
- **Publisher:** Springer; 2nd edition (1999)
- **Language: English**
- **ISBN-13:** 978-3540652571
- **Pages:** 354

Literature

Ausgleichungsrechnung

by Wolfgang Niemeier

- **Publisher:** De Gruyter; 2^{nd} edition (2008)
- **Language:** German
- **ISBN-10:** 3110190559
- **ISBN-13:** 978-3110190557
- **Pages:** 508

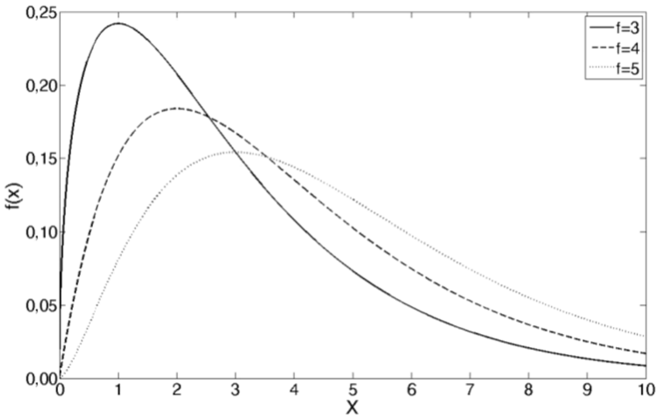


Technische Universität Berlin

Adjustment Theory I

General Information

Prof. Dr.-Ing Frank Neitzel | Institute of Geodesy and Geoinformation Science



Institute of Geodesy and Geoinformation Science

Chair of Geodesy and Adjustment Theory

Adjustment Theory I

Chapter 1: Definitions

Prof. Dr.-Ing. Frank Neitzel

Last update: 16 October 2024

Motivation

- Description of parts of the geometrical-physical reality
- Basis: Measurements
- Linking measurements with desired quantities (modeling):

$$\boxed{\mathbf{L} = \mathbf{F}(\mathbf{x})}$$

$\mathbf{L}$: Measurements

$\mathbf{X}$: Unknowns

$\mathbf{F}(\mathbf{x})$: Functional model

Contents

1. Definitions

1.1 Statistics

From New Latin „statisticus“ $\rightarrow$ of state affairs,
e.g. registration of births and deaths.

First known use of the term „statistics“ in 1770

Definition

> „Statistics is the study of the collection, organization, analysis and interpretation of data.
>
> It deals with all aspects of this including the planning of data collection in terms of the design of surveys and experiments“
>
> (Source: The Oxford Dictionary of Statistical Terms)

Task: Describe → Estimate → Decide

Descriptive statistics summarizes the population data by describing what was observed in the sample numerically.

Typical questions in Geodesy:

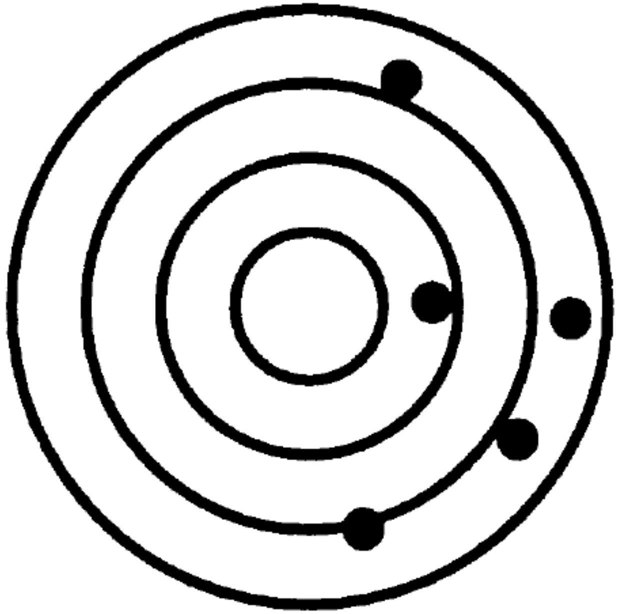
- What is the precision of a distance measurement?
- Is the precision of an angular measurement sufficient to determine coordinates (y, x) with a precision of 2 cm?
- How can I obtain the precision of a function of my observations?
- What is the probability that a measured value exeeds a given threshold value?

Accuracy versus Precision

- Geodesists strive for both **accuracy** and **precision**.
- Many people use the terms “accuracy” and “precision” interchangeably.
- However, in geodesy (as well as other technical and scientific fields), these words have **different meanings**.
- To geodesists, “accuracy” refers to how closely a measurement or observation comes to measuring a "true value," since measurements and observations are always subject to error.
- “Precision” refers to how closely repeated measurements or observations come to duplicating measured or observed values.
- Using five cases of rifle shots fired at a bull’s eye target, each with different results, helps to distinguish the meaning of these two terms.

Case 1: Not accurate, not precise

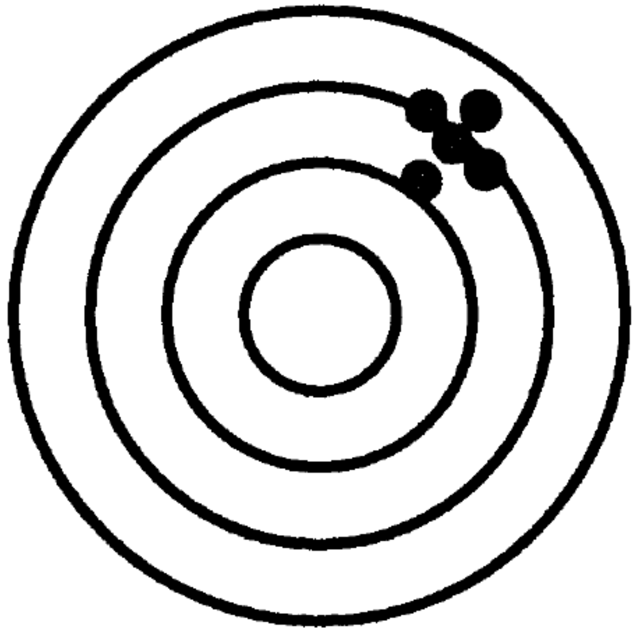
- A shooter stands, aims through the rifle's telescopic sight, and fires five shots at a target.
- Upon examining the target, the shooter sees that all five shots are high, low or right and scattered all around that part of the target.
- These shots were neither accurate (not close to the center) nor precise (not close to each other).



Not Accurate
Not Precise

Case 2: Precise, not accurate

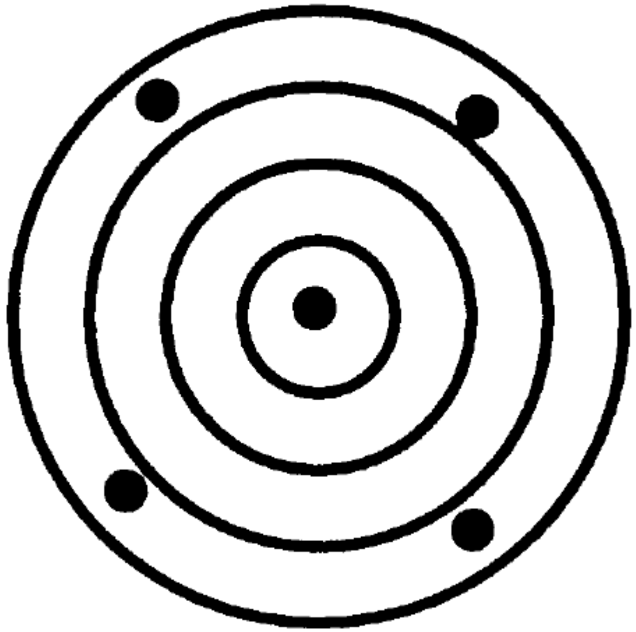
- The shooter assumes a prone position, rests the barrel of the rifle on a support, takes careful aim, holds his breath, and gently squeezes the trigger.
- The target shows that these five shots are very close together, but all five are high and to the right of the bull's eye.
- These shots are precise (close together), but not accurate (not close to the center of the target).



Precise
Not Accurate

Case 3: Accurate, not precise

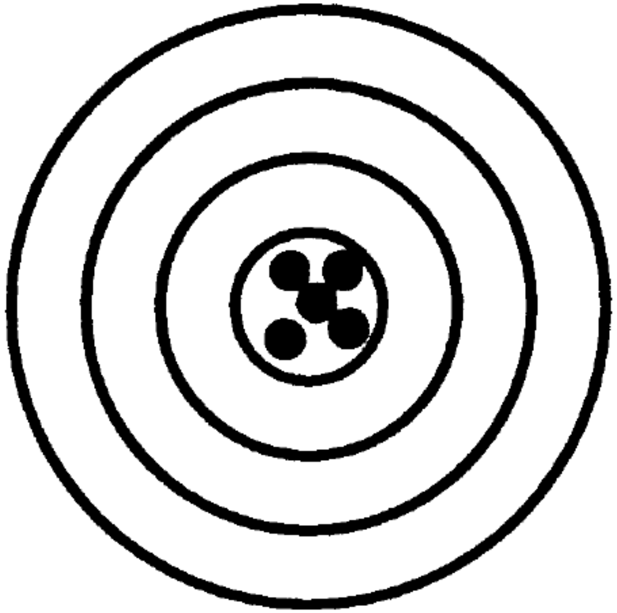
- The shooter adjusts the rifle's telescopic sight and, full of confidence that the problem of inaccuracy has been solved, stands and quickly fires five shots.
- Upon studying the target, the five holes are scattered across the target, but the location of each of the five is very close to the bull's eye.
- These shots are accurate, but not precise.



Accurate
Not Precise

Case 4: Accurate, precise

- The shooter again assumes a prone position, rests the barrel of the rifle on a support, takes careful aim, holds his breath, and gently squeezes the trigger five times.
- This time, the five holes are very close to the center of the target (accurate) and very close together (precise).



Accurate
Precise

A surveying example

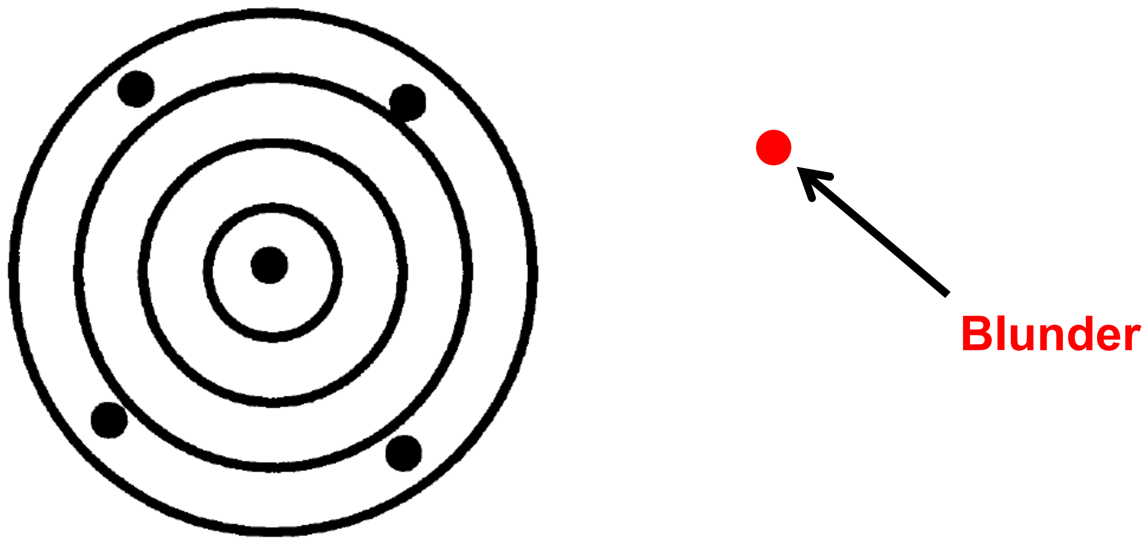
- Imagine surveyors very carefully measuring the distance between two survey points about 30 meters apart 10 times with a measuring tape.
- All 10 of the results agree with each other to within two millimeters. These would be very precise measurements.
- However, suppose the tape they used was too long by 10 millimeters. Then the measurements, even though very precise, would not be accurate.
- Other factors that could affect the accuracy or precision of tape measurements include:
  - incorrect spacing of the marks on the tape,
  - use of the tape at a temperature different from the temperature at which it was calibrated,
  - and use of the tape without the correct tension to control the amount of sag in the tape.

When ever we perform **measurements**, errors are present.

They are divided into blunders, systematic errors, random errors

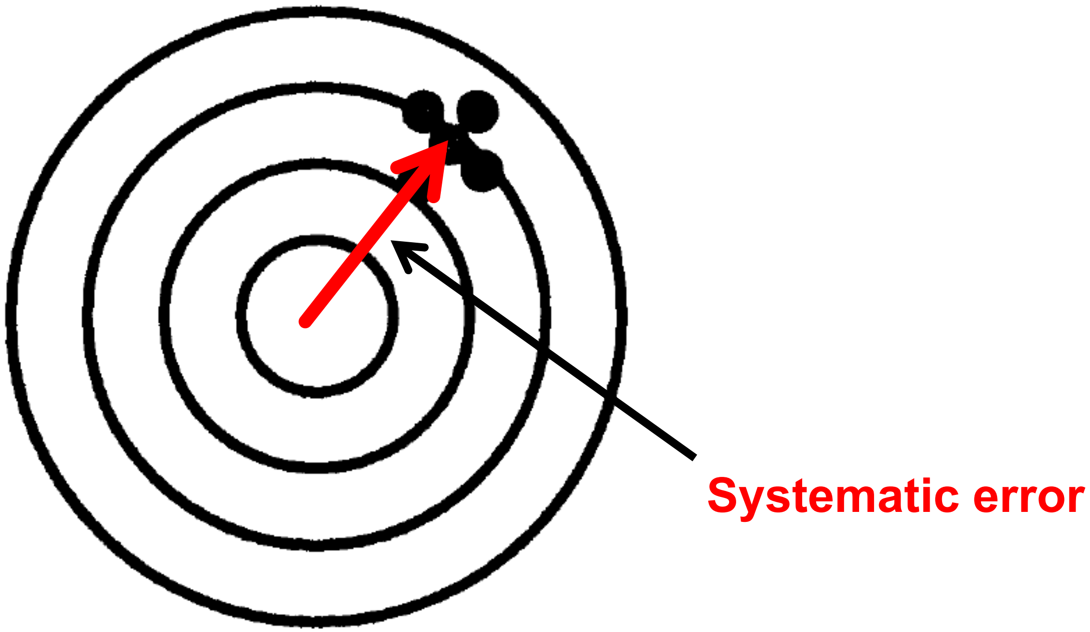
Blunders (or gross errors)

- Caused by mistakes in reading measuring instruments, misidentification of targets
- Can be spotted by checking measurements and then eliminated



Systematic errors influence the result always in the **same sense**

- Caused by
  - Poor calibration
  - one sided use of measuring instruments
  - external influences on the instrument or measuring object, such as temperature, pressure, etc.



Systematic errors

- Can be eliminated to a large extend by
  - instrument calibration
  - proper selection of measuring procedure
  - mathematical compensation

Random errors are all remaining unknown „elementary errors“ after elimination of blunders and systematic errors

- Caused by
  - limitation of human senses
  - limitation of the measuring instruments
  - uncontrollable changes of the environment or the object to be measured

Random errors will equally often be positive or negative.
In mathematical statistics they are considered as independent stochastic variables $\rightarrow$ They have a probability attached to them.

The Nature of Measurement

Evaluation of measurements

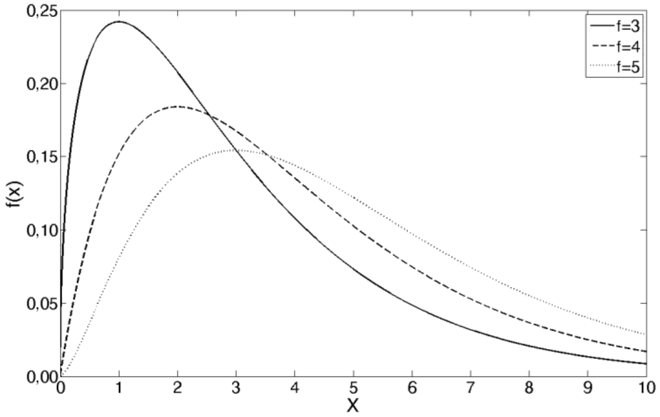
Tasks

1. Derive the most probable value of the desired unknown quantity
2. Provide a value for the precision of a single measurement
3. Estimate the precision of the mean and of its confidence region

Adjustment Theory I

Chapter 1: Definitions

Prof. Dr.-Ing. Frank Neitzel



Institute of Geodesy and Geoinformation Science

Chair of Geodesy and Adjustment Theory

Adjustment Theory I

Chapter 2: Random Variables

Prof. Dr.-Ing. Frank Neitzel

Last update: 23 October 2023

Contents

2.1 Definitions

Random Variable X

- Outcome of an experiment expressed by a number

- Can have different values (measurement or observation x_i) within a certain interval (sample space)

- For each measurement x_i we can specify the corresponding probability $P(x_i)$

Discrete Random Variable

- A variable which can only take a countable number of values

$$W(X) = \{x_1, x_2, x_3, \ldots, x_n\}$$

- e.g possible result of rolling two dice

(1, 1), (1, 2), …

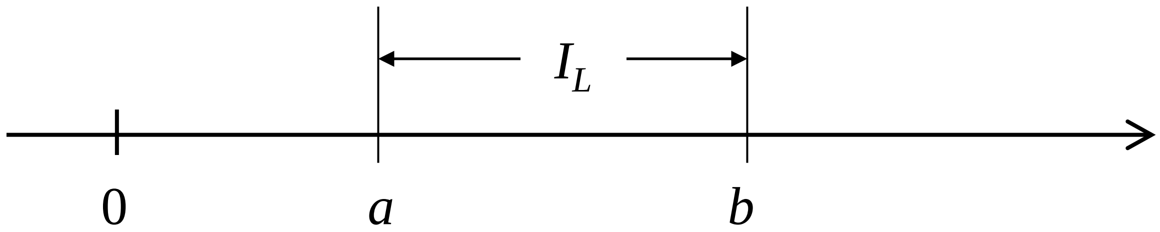
- Roulette

Discrete Random Variable

⇒ Discrete random variables are not considered in geodesy

Continous Random Variable

The sample space is an interval I_L on the number line



$$W(L) = \langle a, b \rangle$$

$$l_j \in I_L \quad with \quad j = 1, 2, 3, \ldots, n$$

n realizations l_j (observations) of the random variable L

Example

For the result of a distance measurement between two given points is (assuming a sufficiently fine resolution) every value possible (within certain borders).

$\Rightarrow$ Such random variables are denoted as **continous random variables**

Random Variable L

Realization

The observations l_j with $j = 1, 2, 3, \ldots, n$

are realizations of the random variable L

Sample

- A sample is a subset of a population, e.g. sample of observations (realizations)
- Sample size n (number of realizations)
- Statistical population ($n \rightarrow \infty$)

Examples

Example 1: Height of all students

- Sample space: $W(L) = \langle 0, \infty \rangle$, e.g. $W(L) = \langle 1.00 \text{ m}, 2.00 \text{ m} \rangle$
- Statistical population: All students of a certain population
- Sample: n students of this semester
- Observation l_j: Height of student j

Example 2: Distance between two points

- Sample space: $W(L) = \langle 0, \infty \rangle$, e.g. $W(L) = \langle 99.000 \text{ m}, 100.000 \text{ m} \rangle$
- Statistical population: All possible measuring results in the interval
- Sample: n observations l_j for the distance
- Observation l_j: Distance j

Difference between Example 1 and Example 2?

- Distance L has a true value $\tilde{L}$ and a mean value $\bar{L}$, even if we don‘t know the true value
- Body height L has no true value $\tilde{L}$, only a mean value $\bar{L}$

2.2.1 Frequency function and probability density

Problem: Large series of observations are often confusing

Goal: Extraction of information

Solution: Classification of observations

Random variable L has been measured n times

$\rightarrow$ Sample with a sample size n

$$l_j\,,\ \ j = 1, 2, 3, \ldots, n$$

Observations l_j from repeated measurements will (in most of the cases) not yield the same value.

Following the rules of probability theory the observations will scatter around the (theoretical) mean value (the expected value) μ.

Observations l_j are obtained from a certain interval $[a, b]$

$$\boxed{a \le l_j \le b\,, \quad \forall l_j}$$

Interval $[a, b]$ can be divided in m discrete intervals (bins) K_i with width Δx

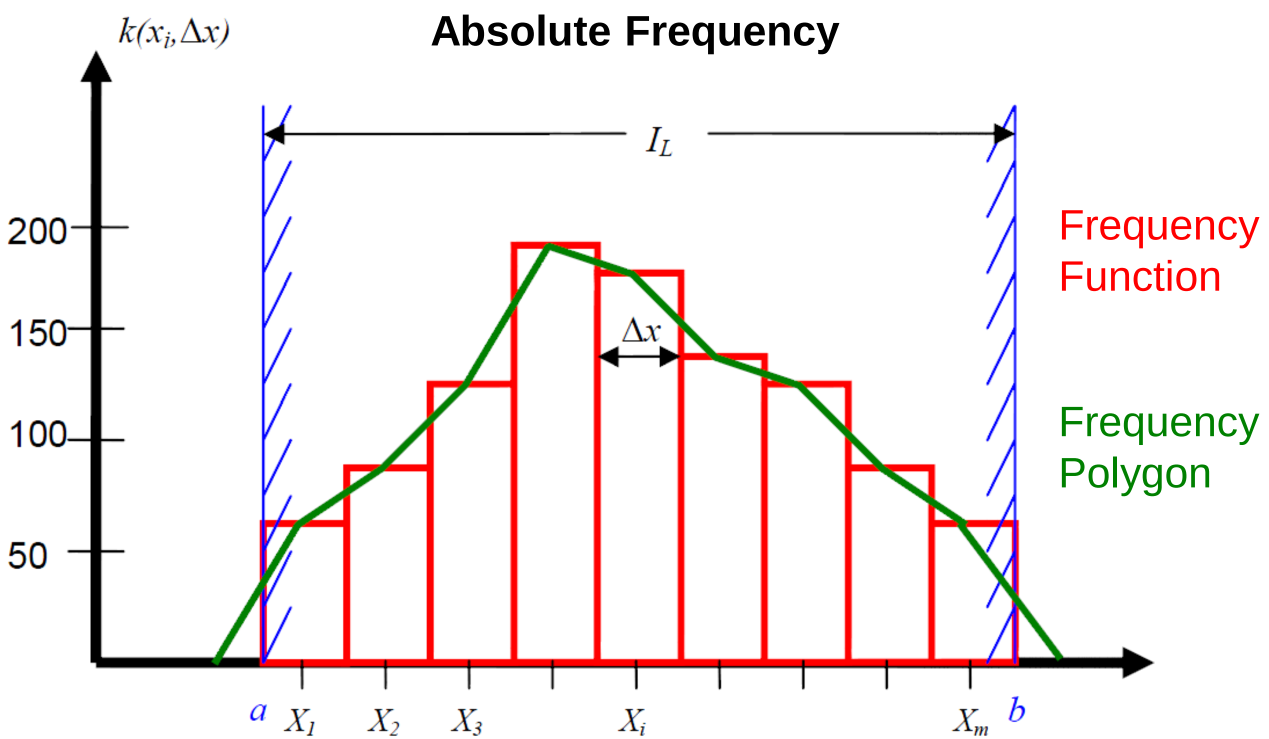
$$\boxed{\Delta x = \frac{b-a}{m}}$$

Absolute frequency k_i: Number k_i of observations in bin K_i

For k_i it holds $\boxed{0 \le k_i \le n}$, as $\sum_{i=1}^{m} k_i = n$

Graphical representation as **histogram** with m rectangles with height k_i erected over discrete intervals (bins)

Also possible:
Graphical representation with polygon that connects k_i with straight line segments



How many bins? $m \approx \sqrt{n}$

Example: Age of employees in a factory

21	22	24	24	25	26	27	27	27	27
28	28	28	28	29	31	31	31	31	32
32	32	33	33	33	34	34	35	35	35
35	35	36	36	36	36	36	37	37	37
37	37	37	37	38	38	39	39	39	39
40	40	41	41	41	41	41	42	42	42
43	43	44	45	45	45	46	47	47	47
47	48	48	48	49	49	49	52	52	52
53	53	53	54	54	55	55	55	56	57
58	60	60	61	64	65	65	65	65	65

Relative frequency h_i: Ratio of k_i to number of observations n

$$\boxed{h_i = h\left(x_i, \Delta x\right) = \frac{k_i}{n} \qquad i = 1, 2, \ldots, m}$$

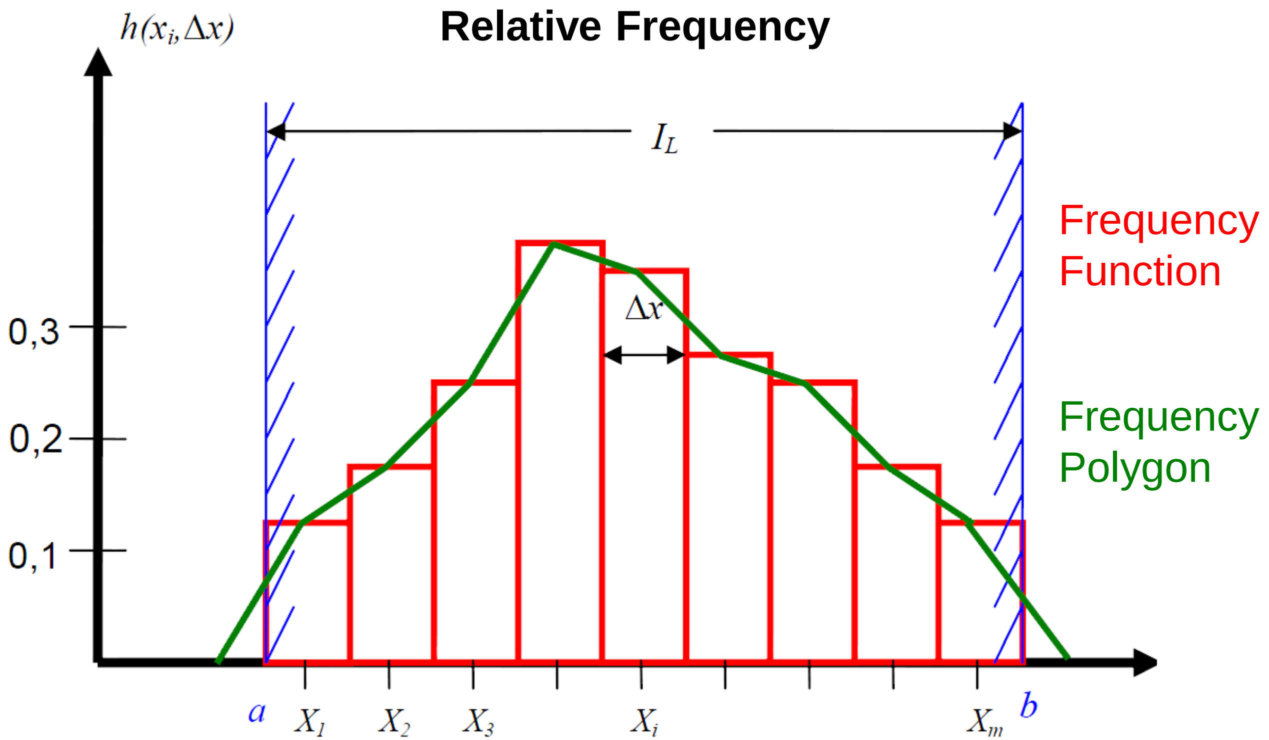
For h_i it holds $\boxed{0 \le h\left(x_i, \Delta x\right) \le 1}$,

as $$\sum_{i=1}^{m} h\left(x_i, \Delta x\right) = \sum_{i=1}^{m} \frac{k_i}{n} = \frac{1}{n} \sum_{i=1}^{m} k_i = 1$$

n

All bins considered together build the **frequency function**

Graphical representation as **histogram** with m rectangles with height h_i or as frequency polygon that connects h_i with straight line segments



Disadvantages of frequency function:

a) not very representative for small n (small number of observations)
b) not a continuous function

Solution for a): Transition to the population ($n \rightarrow \infty$)

A measured series with $n = \infty$ is denoted as **population**.
For $n \rightarrow \infty$ the relative frequency converges to its limit, the **probability** $p(x_i, \Delta x)$:

$$\boxed{\lim_{n\to\infty}\{h(x_i, \Delta x)\} = \lim_{n\to\infty}\left(\frac{k_i}{n}\right) = p(x_i, \Delta x)}$$

Probability Function

For $n \rightarrow \infty$ this function is representative

Solution for b): Transition to the probability density

The form of the probability function depends on the position of the discrete intervals (bins) on the number line which is denoted with x_i and on the width Δx of the bin.

$\Rightarrow$ A representation independent of x_i and Δx is desirable.

To overcome this drawback we divide the probabilty function by the width Δx and let the width become differentially small ($\Delta x \rightarrow dx$).

$\Rightarrow$ As limit we obtain the **probability density function**

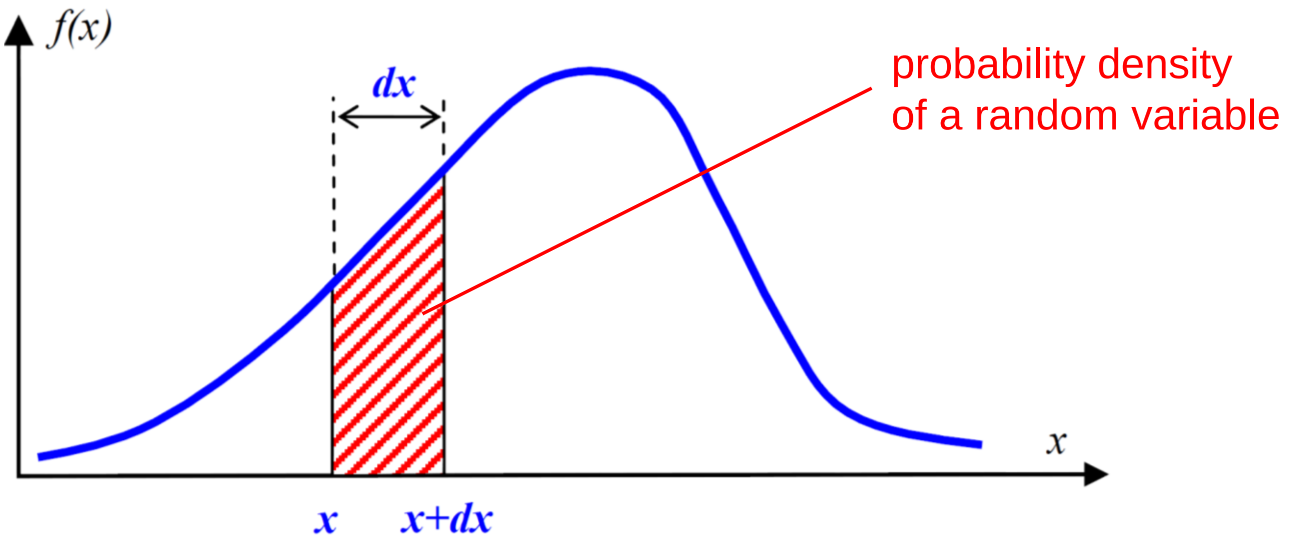
$$\boxed{\lim_{\Delta x \rightarrow dx} \left\{ \frac{p\left(x_i, \Delta x\right)}{\Delta x} \right\} = f\left(x\right)}$$

Probability *Density* Function

This is a continous function now

Continous function $f(x)$ of the parameter x that yields the **probability** P that a measurement L ranges between x and $x + dx$.

$$\boxed{p(x_i, dx) = P(x \leq L \leq x + dx) = f(x)\,dx}$$



$$\boxed{\sum p(x_i, dx) = \int_{-\infty}^{+\infty} f(x)\,dx = 1}$$

Necessary condition for a probability density

2.2.2 Cumulative frequency function and distribution function

The frequency function from 2.2.1 yields the relative frequency of observations in certain bins of the sample space.

All observations fall into the interval $[a, b]$ that is divided in m bins with the width Δx. (*not very useful for practical use*)

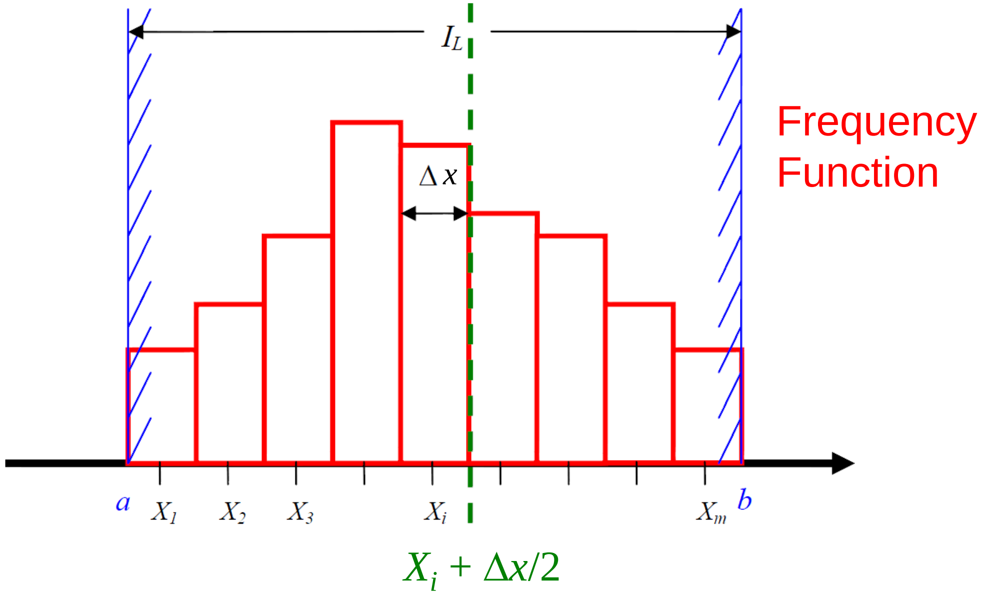
The ratio of k_i to number of observations n was introduced as frequency function h_i

$$\boxed{h_i = h\left(x_i, \Delta x\right) = \frac{k_i}{n} \qquad i = 1, 2, \ldots, m}$$ with m = number of bins

The cumulative frequency function yields the relative frequency of observations that are not larger than a certain threshold $x_i + \dfrac{\Delta x}{2}$.

(*often of interest in practical use*)

2.2.2 Cumulative frequency function and distribution function



Transition from the representation as histogram

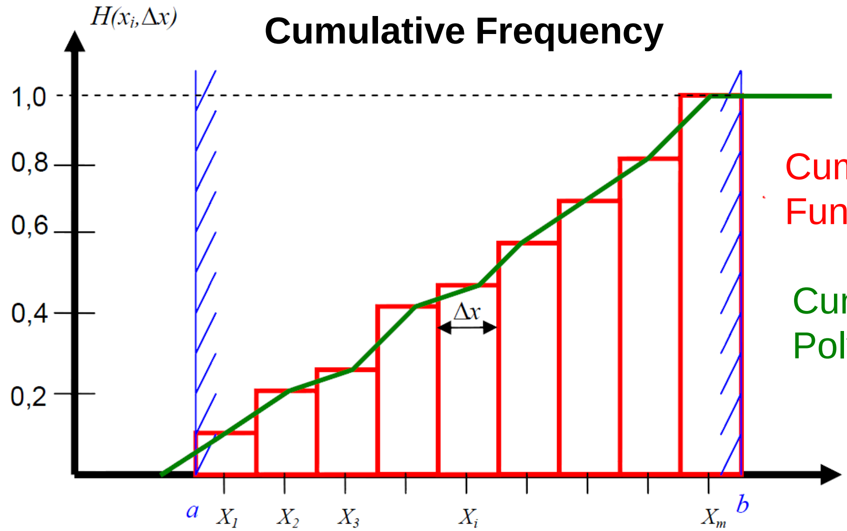
$$\boxed{h(x_i, \Delta x) = \frac{k_i}{n} = \frac{\text{"number of values in the } i^{\text{th}} \text{ bin"}}{\text{"sample size"}} \qquad i = 1, 2, \ldots, m}$$

to representation as cumulative frequency function

$$\boxed{H(x_i, \Delta x) = \sum_{k=1}^{i} h(x_k, \Delta x)}$$

For the cumulative frequency function it holds $\boxed{0 \le H(x_i, \Delta x) \le 1}$

Graphical representation of cumulative frequency function usually as **cumulative frequency polygon**



Distribution Function: As in 2.2.1 we can consider the population to obtain continuous representation $n \rightarrow \infty$, $\Delta x \rightarrow dx$

Instead of cumulative frequency we obtain the probability that an observation is below a certain threshold b.

$$\boxed{F(b) = P(L \le b) = P(-\infty \le L \le b) = \int_{-\infty}^{b} f(x)\,dx}$$

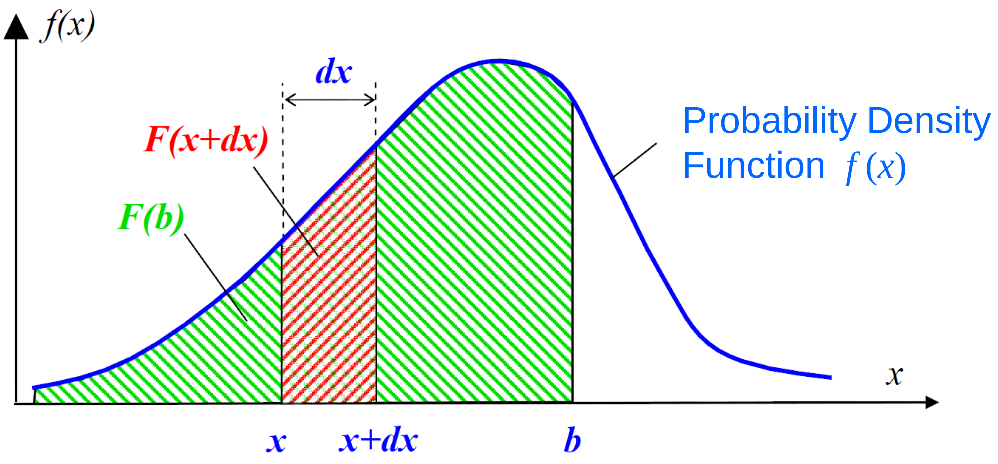
Distribution Function

Necessary condition for a probability density

$$\boxed{\int_{-\infty}^{+\infty} f(x)\,dx = 1}$$

Graphical representation of distribution function $F(b)$ and probability density function $f(x)$

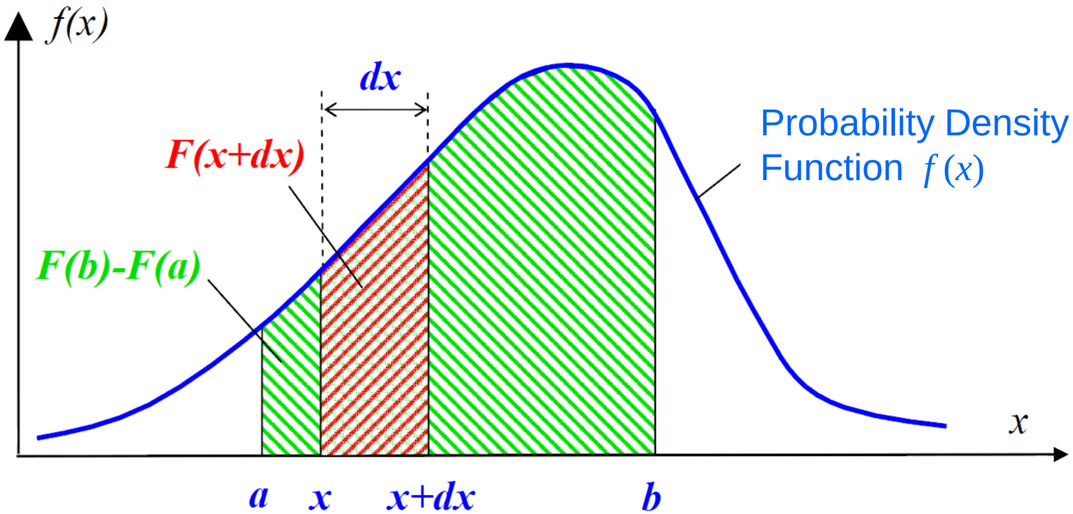
Graphical representation of the probability that an observation is below a certain threshold b.



$$\boxed{F(b) = P(L \le b) = P(-\infty \le L \le b) = \int_{-\infty}^{b} f(x)\,dx}$$

2.2.2 Cumulative frequency function and distribution function

Graphical representation of the probability that a random variable falls into the interval $[a, b]$ as area (distribution function $F(b) - F(a)$) under the probability density function $f(x)$.



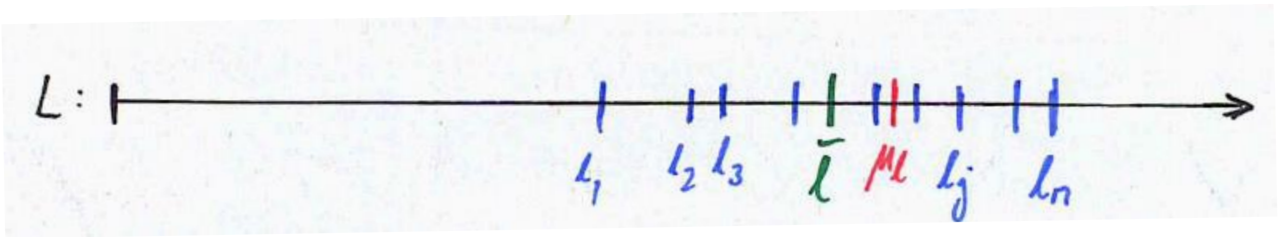
If we know the distribution function of an observation L (random variable) we can compute the probability that L falls between the limits a and b.

$$\boxed{P(a \le L \le b) = P(L \le b) - P(L \le a) = F(b) - F(a) = \int_a^b f(x)\,dx}$$

2.3.1 Definitions

Given: Random variable L with its realizations $l_j\,,\; j = 1, 2, 3, \dots, n$

sample → „observation vector“



$$\mathbf{l} = \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{bmatrix}$$ n observations of the same random variable L

observation vector

Wanted: - Extraction of information

- „Best realization“ close to the „true value“

Empirical mean or arithmetic mean $\bar{l}$

$$\boxed{\bar{l} = \frac{1}{n}\sum_{j=1}^{n} l_j = \frac{1}{n} \cdot \mathbf{e}^{\mathrm{T}} \cdot \mathbf{l}} \quad with \;\; \mathbf{e}^{\mathrm{T}} = [1 \quad 1 \quad \cdots \quad 1]$$

Unbiased estimation of the expectation value

Expectation value μ_L

For $n \to \infty$, empirical mean converges to expectation μ_L

$$\boxed{\mu_L = E(L) = \lim_{n\to\infty}(\bar{l}) = \lim_{n\to\infty}\left\{\frac{1}{n}\sum_{j=1}^{n} l_j\right\} = \lim_{n\to\infty}\left\{\frac{1}{n} \cdot \mathbf{e}^{\mathrm{T}} \cdot \mathbf{l}\right\}}$$

Definition: $\mu_L = \int_{-\infty}^{\infty} l \cdot f(l) dl$ with probability density function $f(l)$

In general

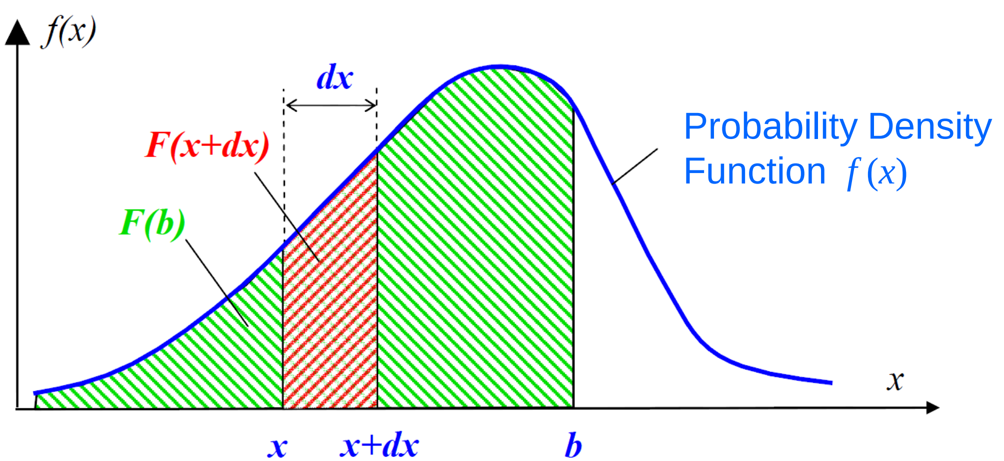
$$\boxed{\mu_X = E(X) = \lim_{n\to\infty}\left\{\frac{1}{n}\sum_{j=1}^{n} x_j\right\}}$$ Expectation operator

Definition for arbitrary $\delta > 0$: $\boxed{\lim_{n\to\infty} P(|l-\mu_l| > \delta) = 0}$

Expectation for given probability density $f(x)$

From section 2.2.2: Probability that a random variable X falls into a differential interval

$$\boxed{P(x \le X \le x+dx) = f(x)dx = p_x}$$



From n observations in total (with $n \to \infty$) n_x fall into this interval

$$\boxed{n_x = n \cdot p_x = n \cdot f(x)dx}$$

The mean of all these values is

$$\boxed{\frac{1}{n}\sum_{x=-\infty}^{+\infty} x \cdot n_x = \frac{1}{n}\sum_{x=-\infty}^{+\infty} x \cdot n \cdot f(x)dx = \mu_x}$$

Replace sigma-sign by an integral

$$\boxed{\mu_x = \int_{-\infty}^{+\infty} x \cdot f(x)dx}$$

Expectation

Expectation μ_x for a known probability density $f(x)$

True Value $\tilde{L}$

True Value $\tilde{L}$ is in general not equal to the expectation μ_L: $\boxed{\tilde{L} \neq \mu_L}$

True Value $\tilde{L}$ is in general unknown!

Example for a known true value:
Sum of angles in a triangle = 200 gon

True Value $\boxed{\tilde{L} = \mu_L - \Delta}$ with Δ as systematic deviation

True deviation $\boxed{\eta_j = \varepsilon_j + \Delta_j = l_j - \tilde{L}}$

Random Deviation (with respect to the expectation)

$\boxed{\varepsilon_j = l_j - \mu_L}$ as vectors $\boxed{\boldsymbol{\varepsilon} = \mathbf{l} - \mathbf{e} \cdot \mu_L}$ with $\mathbf{e} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$ "unit vector"

Residual ("correction") (with respect to the mean value)

$\boxed{v_i = \bar{l} - l_j}$ as vectors $\boxed{\mathbf{v} = \mathbf{e} \cdot \bar{l} - \mathbf{l}}$

In geodesy we try to avoid the occurrence of systematic errors by applying

- appropriate measurement set-ups
- appropriate methods for data assessment

If, and only if, no systematic errors are present with $\Delta = 0$

it holds $\tilde{L} = \mu_L$ and $\eta = \varepsilon$.

2.3.2 Calculation rules for expectations

Given: One or more random variables X, resp. a vector of random variables $\mathbf{X}$ with probability densities $f(x)$ and expectations μ_X.

A random variable Y is obtained by applying X (or $\mathbf{X}$) in a given function g:

$$\boxed{Y = g(X)}$$

1. Probability density $f(y)$ of Y

$$y + dy = g(x + dx)$$

$$P\{x \leq X \leq x + dx\} = P\{y \leq Y \leq y + dy\}$$

Probability densities

$$f(x)dx = f(y)dy \quad |: dy$$

$$f(y) = f(x)\frac{dx}{dy}$$

$$= f(x)\frac{1}{\frac{dy}{dx}} \quad \text{with} \quad \frac{dy}{dx} = g'(x)$$

$$\boxed{f(y) = \frac{f(x)}{g'(x)} = f(x)\frac{dx}{dy}}$$

2.3.2 Calculation rules for expectations

2. Expectation μ_Y of Y

$$\boxed{E(Y) = \mu_Y = \int_{-\infty}^{+\infty} y \cdot f(y)dy = \int_{-\infty}^{+\infty} g(x)f(x)dx}$$

3. Expectation of a linear function $Y = g(X) = a + bX$ with a and b as arbitrary constant values

$$E(Y) = \int_{-\infty}^{+\infty} g(x)f(x)dx = \int_{-\infty}^{+\infty}(a + bx)f(x)dx$$

$$= a \underbrace{\int_{-\infty}^{+\infty} f(x)dx}_{1} + b \underbrace{\int_{-\infty}^{+\infty} x \cdot f(x)dx}_{\mu_X}$$

$$\boxed{\mu_Y = E(Y) = a + b \cdot \mu_X = a + b \cdot E(X)}$$

4. Expectation of a sum $Y = g(X) = \sum_{i=1}^{m} X_i = X_1 + X_2 + \ldots + X_m$ of random variables X_i

$$\boxed{\mu_Y = E(Y) = E\left\{\sum_{i=1}^{m} X_i\right\} = \sum_{i=1}^{m} E(X_i)}$$

Addition rule for expectations $E(X_1 + X_2) = E(X_1) + E(X_2)$
„Expectation of a sum = Sum of expectations“

5. Expectation of a product $Y = g(X) = \prod_{i=1}^{m} X_i = X_1 \cdot X_2 \cdot \ldots \cdot X_m$ of random variables X_i

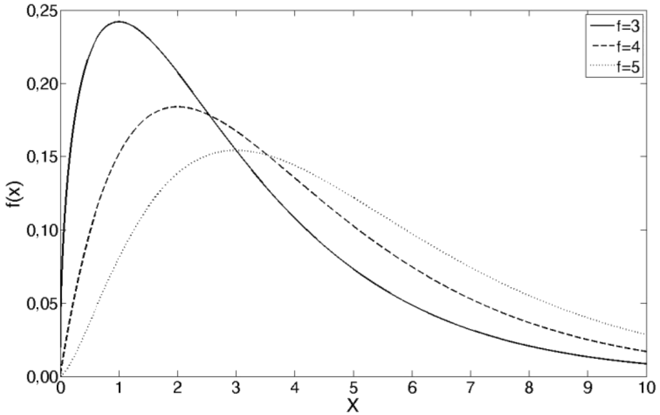
$$\boxed{\mu_Y = E(Y) = E\left\{\prod_{i=1}^{m} X_i\right\} = \prod_{i=1}^{m} E(X_i)}$$

Multiplication rule for expectations (for <u>stochastic independent</u> random variables) $E(X_1 \cdot X_2) = E(X_1) \cdot E(X_2)$
„Expectation of a product = Product of expectations“

Adjustment Theory I

Chapter 2: Random Variables

Prof. Dr.-Ing. Frank Neitzel



Institute of Geodesy and Geoinformation Science

Chair of Geodesy and Adjustment Theory

Adjustment Theory I

Chapter 2: Random Variables

Prof. Dr.-Ing. Frank Neitzel

Last update: 23 October 2023

Contents

2.1 Definitions

Random Variable X

- Outcome of an experiment expressed by a number
- Can have different values (measurement or observation x_i) within a certain interval (sample space)
- For each measurement x_i we can specify the corresponding probability $P(x_i)$

Discrete Random Variable

- A variable which can only take a countable number of values

$$W(X) = \{x_1, x_2, x_3, \dots, x_n\}$$

- e.g possible result of rolling two dice

(1, 1), (1, 2), …

- Roulette

2. Random Variables

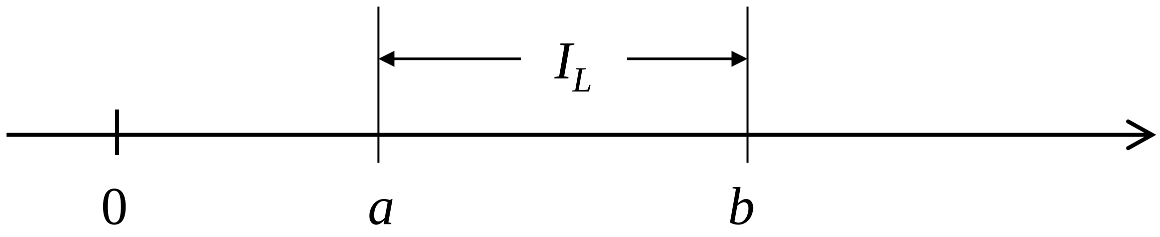
Discrete Random Variable

$\Rightarrow$ Discrete random variables are not considered in geodesy

2.1 Definitions

Continous Random Variable

The sample space is an interval I_L on the number line



$$W(L) = \langle a, b \rangle$$

$$l_j \in I_L \quad with \quad j = 1, 2, 3, \ldots, n$$

n realizations l_j (observations) of the random variable L

Example

For the result of a distance measurement between two given points is (assuming a sufficiently fine resolution) every value possible (within certain borders).

$\Rightarrow$ Such random variables are denoted as **continous random variables**

Random Variable L

Realization

The observations l_j with $j = 1, 2, 3, \ldots, n$

are realizations of the random variable L

Sample

- A sample is a subset of a population, e.g. sample of observations (realizations)
- Sample size n (number of realizations)
- Statistical population ($n \rightarrow \infty$)

Examples

Example 1: Height of all students

- Sample space: $W(L) = \langle 0, \infty \rangle$, e.g. $W(L) = \langle 1.00\ \mathrm{m}, 2.00\ \mathrm{m} \rangle$
- Statistical population: All students of a certain population
- Sample: n students of this semester
- Observation l_j: Height of student j

Example 2: Distance between two points

- Sample space: $W(L) = \langle 0, \infty \rangle$, e.g. $W(L) = \langle 99.000\ \mathrm{m}, 100.000\ \mathrm{m} \rangle$
- Statistical population: All possible measuring results in the interval
- Sample: n observations l_j for the distance
- Observation l_j: Distance j

Difference between Example 1 and Example 2?

- Distance L has a true value $\tilde{L}$ and a mean value $\bar{L}$, even if we don‘t know the true value
- Body height L has no true value $\tilde{L}$, only a mean value $\bar{L}$

2.2.1 Frequency function and probability density

Problem: Large series of observations are often confusing

Goal: Extraction of information

Solution: Classification of observations

Random variable L has been measured n times

$\rightarrow$ Sample with a sample size n

$$l_j\,,\;\; j = 1, 2, 3, \ldots, n$$

Observations l_j from repeated measurements will (in most of the cases) not yield the same value.

Following the rules of probability theory the observations will scatter around the (theoretical) mean value (the expected value) μ.

Observations l_j are obtained from a certain interval $[a, b]$

$$\boxed{a \le l_j \le b\,, \quad \forall l_j}$$

Interval $[a, b]$ can be divided in m discrete intervals (bins) K_i with width Δx

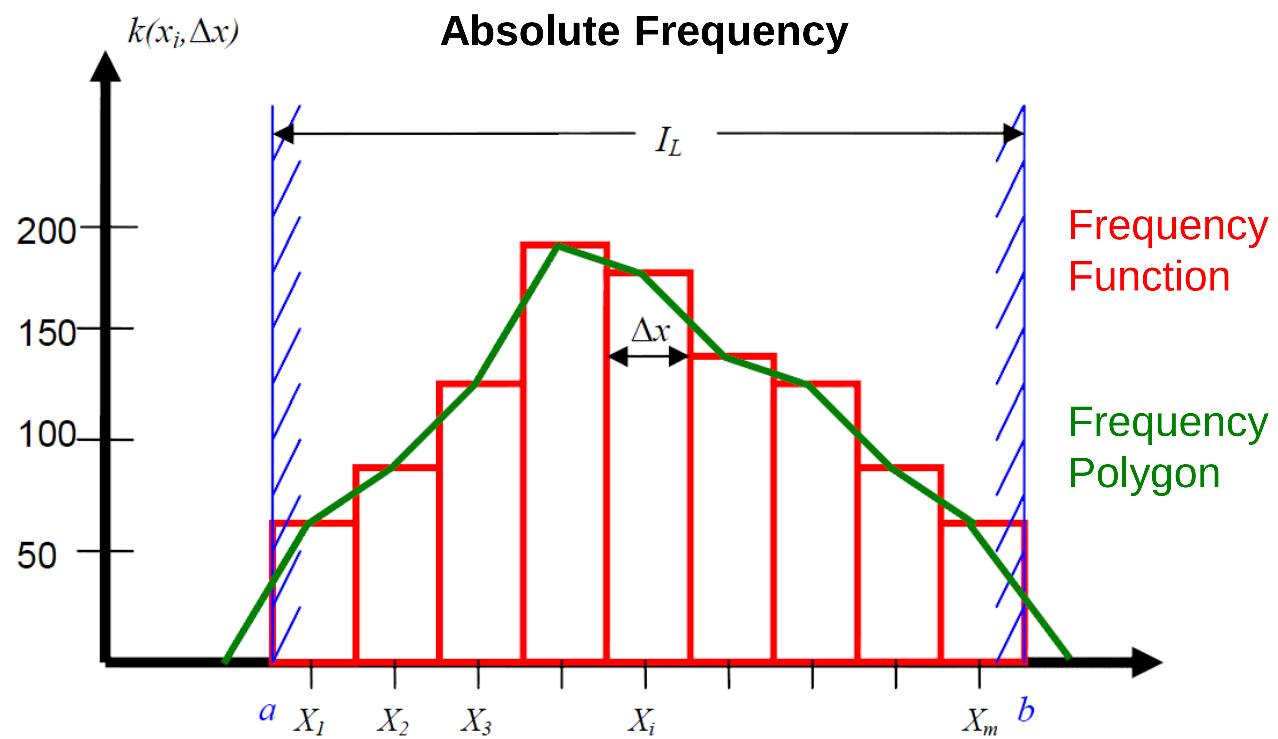
$$\boxed{\Delta x = \frac{b-a}{m}}$$

Absolute frequency k_i: Number k_i of observations in bin K_i

For k_i it holds $\boxed{0 \le k_i \le n}$, as $\sum_{i=1}^{m} k_i = n$

Graphical representation as **histogram** with m rectangles with height k_i erected over discrete intervals (bins)

Also possible:
Graphical representation with polygon that connects k_i with straight line segments



How many bins? $m \approx \sqrt{n}$

Example: Age of employees in a factory

21	22	24	24	25	26	27	27	27	27
28	28	28	28	29	31	31	31	31	32
32	32	33	33	33	34	34	35	35	35
35	35	36	36	36	36	36	37	37	37
37	37	37	37	38	38	39	39	39	39
40	40	41	41	41	41	41	42	42	42
43	43	44	45	45	45	46	47	47	47
47	48	48	48	49	49	49	52	52	52
53	53	53	54	54	55	55	55	56	57
58	60	60	61	64	65	65	65	65	65

2.2.1 Frequency function and probability density

Relative frequency h_i: Ratio of k_i to number of observations n

$$\boxed{h_i = h(x_i, \Delta x) = \frac{k_i}{n} \qquad i = 1, 2, \ldots, m}$$

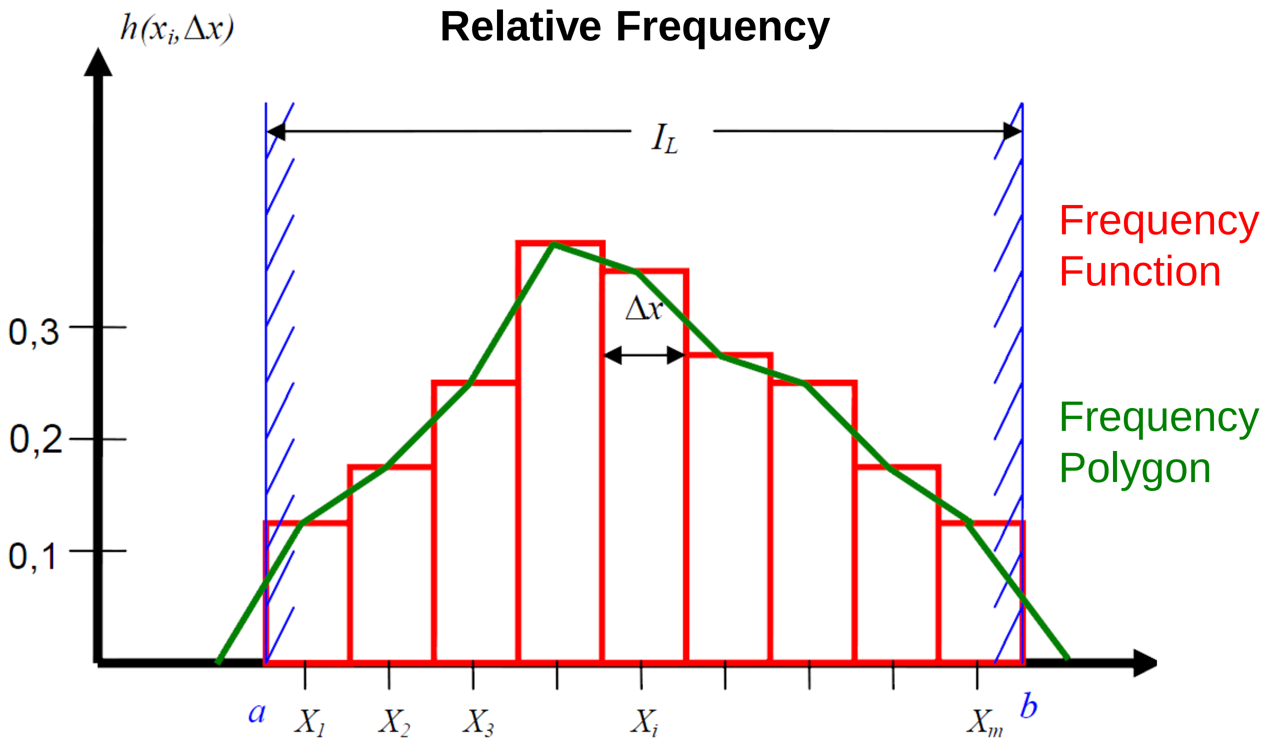
For h_i it holds $\boxed{0 \le h(x_i, \Delta x) \le 1}$,

as
$$\sum_{i=1}^{m} h(x_i, \Delta x) = \sum_{i=1}^{m} \frac{k_i}{n} = \frac{1}{n} \sum_{i=1}^{m} k_i = 1$$

n

All bins considered together build the **frequency function**

Graphical representation as **histogram** with m rectangles with height h_i or as frequency polygon that connects h_i with straight line segments



Disadvantages of frequency function:

a) not very representative for small n (small number of observations)

b) not a continuous function

Solution for a): Transition to the population ($n \rightarrow \infty$)

A measured series with $n = \infty$ is denoted as **population**.
For $n \rightarrow \infty$ the relative frequency converges to its limit,
the **probability** $p(x_i, \Delta x)$:

$$\boxed{\lim_{n\to\infty}\left\{h\left(x_i, \Delta x\right)\right\} = \lim_{n\to\infty}\left(\frac{k_i}{n}\right) = p\left(x_i, \Delta x\right)}$$

Probability Function

For $n \rightarrow \infty$ this function is representative

Solution for b): Transition to the probability density

The form of the probability function depends on the position of the discrete intervals (bins) on the number line which is denoted with x_i and on the width Δx of the bin.

$\Rightarrow$ A representation independent of x_i and Δx is desirable.

To overcome this drawback we divide the probabilty function by the width Δx and let the width become differentially small ($\Delta x \rightarrow dx$).

$\Rightarrow$ As limit we obtain the **probability density function**

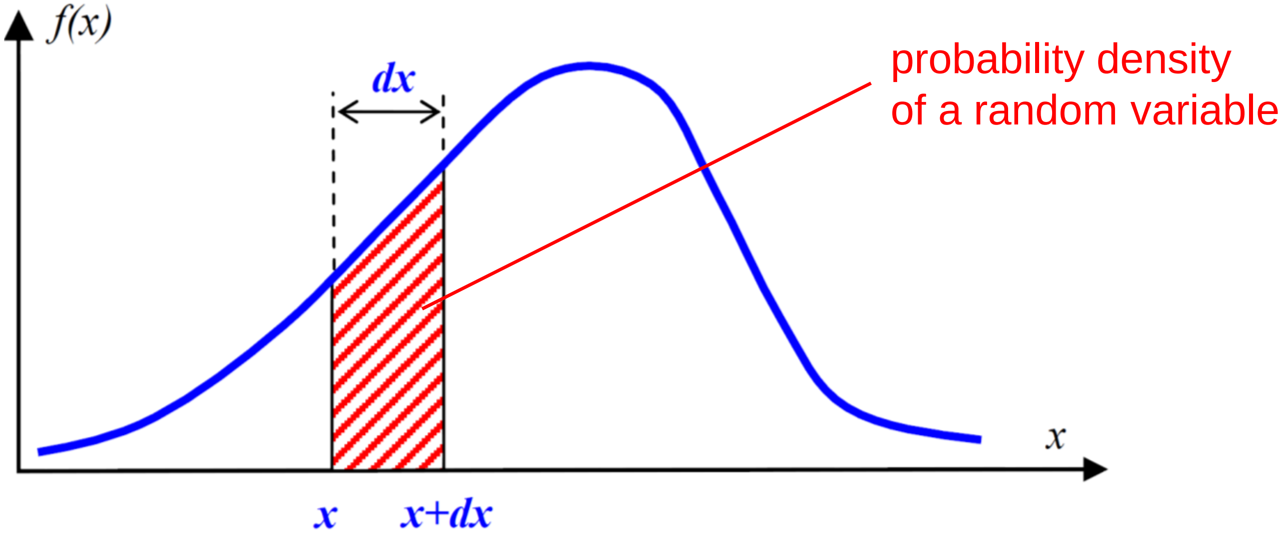
$$\boxed{\lim_{\Delta x \rightarrow dx} \left\{ \frac{p(x_i, \Delta x)}{\Delta x} \right\} = f(x)}$$ **Probability *Density* Function**

This is a continous function now

2.2.1 Frequency function and probability density

Continous function $f(x)$ of the parameter x that yields the **probability** P that a measurement L ranges between x and $x + dx$.

$$\boxed{p(x_i, dx) = P(x \leq L \leq x + dx) = f(x)dx}$$



$$\boxed{\sum p(x_i, dx) = \int_{-\infty}^{+\infty} f(x)dx = 1}$$

Necessary condition for a probability density

The frequency function from 2.2.1 yields the relative frequency of observations in certain bins of the sample space.

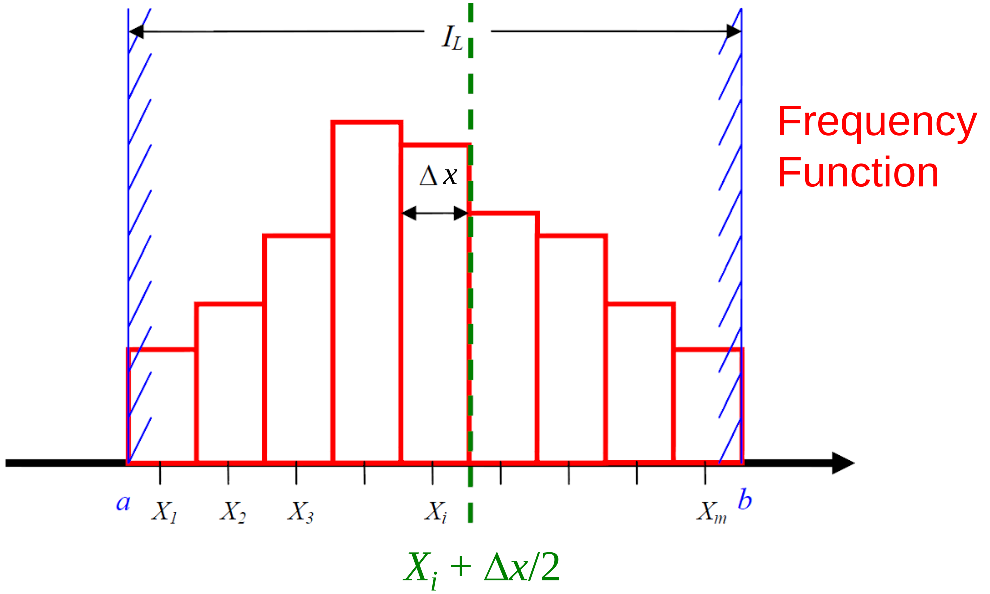
All observations fall into the interval $[a, b]$ that is divided in m bins with the width Δx. (*not very useful for practical use*)

The ratio of k_i to number of observations n was introduced as frequency function h_i

$$\boxed{h_i = h\left(x_i, \Delta x\right) = \frac{k_i}{n} \qquad i = 1, 2, \ldots, m}$$ with m = number of bins

The cumulative frequency function yields the relative frequency of observations that are not larger than a certain threshold $x_i + \dfrac{\Delta x}{2}$.

(*often of interest in practical use*)



Transition from the representation as histogram

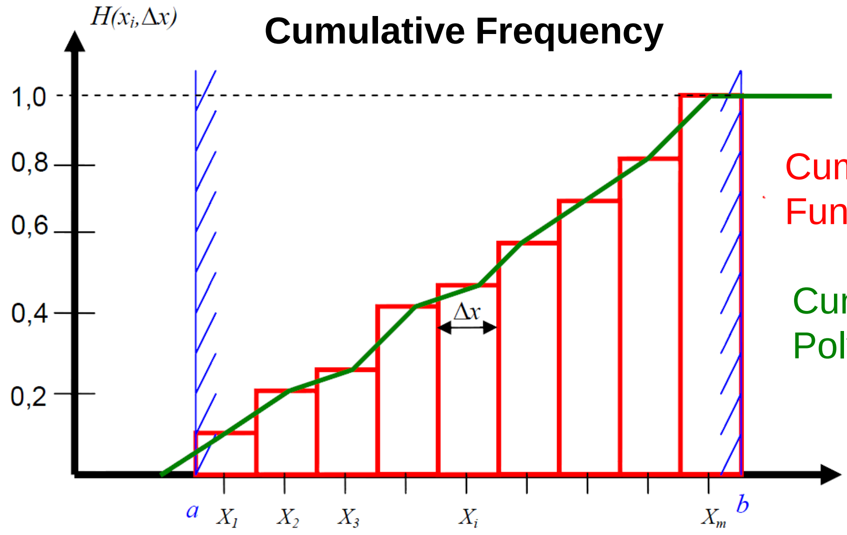
$$h(x_i, \Delta x) = \frac{k_i}{n} = \frac{\text{"number of values in the } i^{\text{th}} \text{ bin"}}{\text{"sample size"}} \qquad i = 1, 2, \ldots, m$$

to representation as cumulative frequency function

$$H(x_i, \Delta x) = \sum_{k=1}^{i} h(x_k, \Delta x)$$

For the cumulative frequency function it holds $\boxed{0 \le H(x_i, \Delta x) \le 1}$

Graphical representation of cumulative frequency function usually as **cumulative frequency polygon**



Distribution Function: As in 2.2.1 we can consider the population to obtain continuous representation $n \to \infty\ , \Delta x \to dx$

Instead of cumulative frequency we obtain the probability that an observation is below a certain threshold b.

$$\boxed{F(b) = P(L \le b) = P(-\infty \le L \le b) = \int_{-\infty}^{b} f(x)\,dx}$$

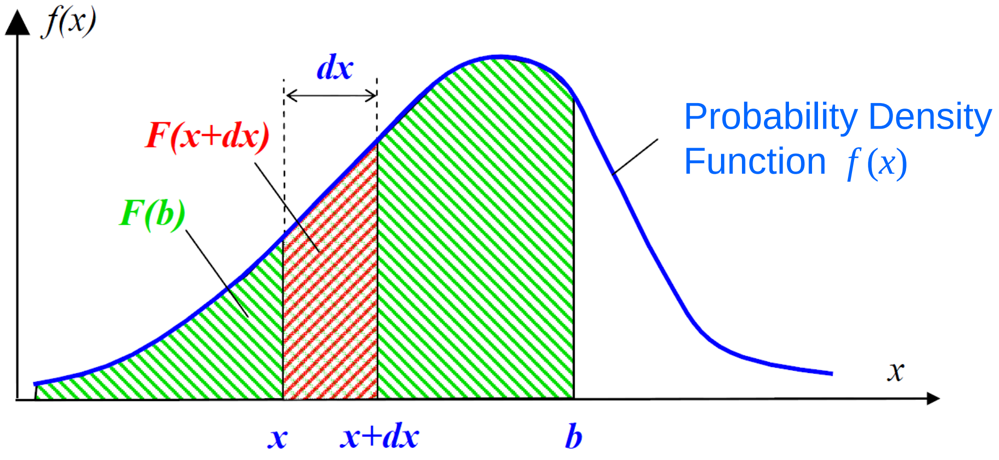
Distribution Function

Necessary condition for a probability density

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

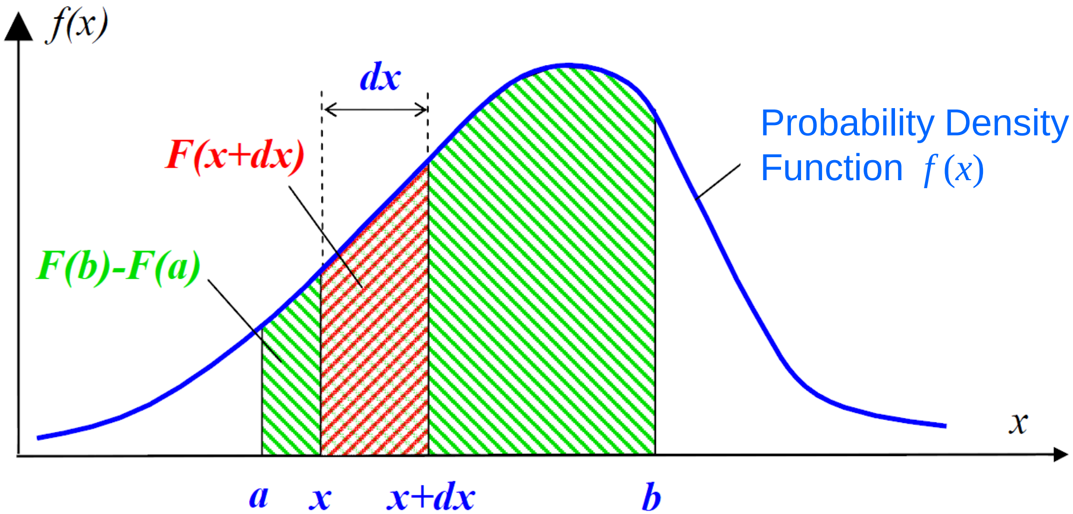
Graphical representation of distribution function $F(b)$ and probability density function $f(x)$

Graphical representation of the probability that an observation is below a certain threshold b.



$$F(b) = P(L \le b) = P(-\infty \le L \le b) = \int_{-\infty}^{b} f(x)\,dx$$

Graphical representation of the probability that a random variable falls into the interval $[a, b]$ as area (distribution function $F(b) - F(a)$) under the probability density function $f(x)$.



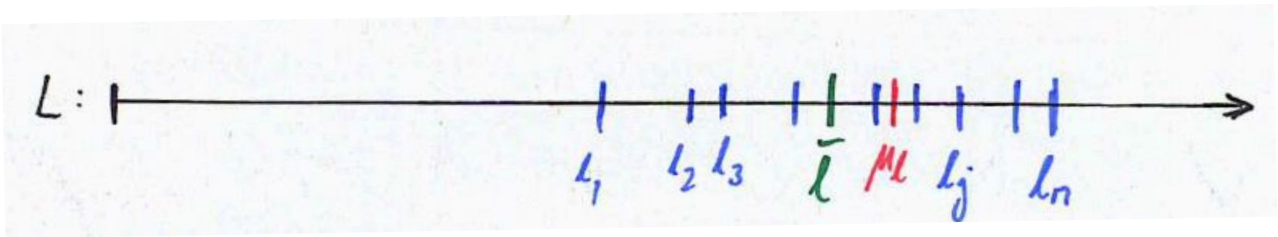
If we know the distribution function of an observation L (random variable) we can compute the probability that L falls between the limits a and b.

$$\boxed{P(a \leq L \leq b) = P(L \leq b) - P(L \leq a) = F(b) - F(a) = \int_a^b f(x)\,dx}$$

2.3.1 Definitions

Given: Random variable L with its realizations l_j , $j = 1, 2, 3, \ldots, n$

sample → „observation vector“



$$\mathbf{l} = \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{bmatrix}$$ n observations of the same random variable L

observation vector

Wanted: - Extraction of information

- „Best realization“ close to the „true value“

Empirical mean or arithmetic mean $\bar{l}$

$$\boxed{\bar{l} = \frac{1}{n}\sum_{j=1}^{n} l_j = \frac{1}{n} \cdot \mathbf{e}^{\mathrm{T}} \cdot \mathbf{l}} \quad with \;\; \mathbf{e}^{\mathrm{T}} = [1 \quad 1 \quad \cdots \quad 1]$$

Unbiased estimation of the expectation value

Expectation value μ_L

For $n \to \infty$, empirical mean converges to expectation μ_L

$$\boxed{\mu_L = E(L) = \lim_{n\to\infty}(\bar{l}) = \lim_{n\to\infty}\left\{\frac{1}{n}\sum_{j=1}^{n} l_j\right\} = \lim_{n\to\infty}\left\{\frac{1}{n} \cdot \mathbf{e}^{\mathrm{T}} \cdot \mathbf{l}\right\}}$$

Definition: $\mu_L = \int_{-\infty}^{\infty} l \cdot f(l)dl$ with probability density function $f(l)$

In general

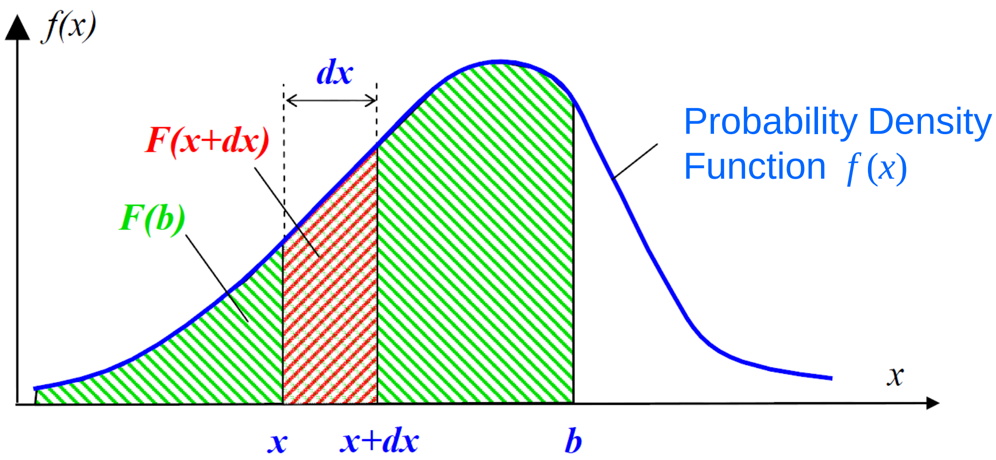
$$\boxed{\mu_X = E(X) = \lim_{n\to\infty}\left\{\frac{1}{n}\sum_{j=1}^{n} x_j\right\}}$$ Expectation operator

Definition for arbitrary $\delta > 0$: $\boxed{\lim_{n\to\infty} P(|l-\mu_l| > \delta) = 0}$

Expectation for given probability density $f(x)$

From section 2.2.2: Probability that a random variable X falls into a differential interval

$$\boxed{P(x \le X \le x+dx) = f(x)dx = p_x}$$



From n observations in total (with $n \to \infty$) n_x fall into this interval

$$\boxed{n_x = n \cdot p_x = n \cdot f(x)dx}$$

The mean of all these values is

$$\boxed{\frac{1}{n}\sum_{x=-\infty}^{+\infty} x \cdot n_x = \frac{1}{n}\sum_{x=-\infty}^{+\infty} x \cdot n \cdot f(x)dx = \mu_x}$$

Replace sigma-sign by an integral

$$\boxed{\mu_x = \int_{-\infty}^{+\infty} x \cdot f(x)dx}$$ Expectation

Expectation μ_x for a known probability density $f(x)$

True Value $\tilde{L}$

True Value $\tilde{L}$ is in general not equal to the expectation μ_L: $\boxed{\tilde{L} \neq \mu_L}$

True Value $\tilde{L}$ is in general unknown!

Example for a known true value:
Sum of angles in a triangle = 200 gon

True Value $\boxed{\tilde{L} = \mu_L - \Delta}$ with Δ as systematic deviation

True deviation $\boxed{\eta_j = \varepsilon_j + \Delta_j = l_j - \tilde{L}}$

Random Deviation (with respect to the expectation)

$\boxed{\varepsilon_j = l_j - \mu_L}$ as vectors $\boxed{\boldsymbol{\varepsilon} = \mathbf{l} - \mathbf{e} \cdot \mu_L}$ with $\mathbf{e} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$ "unit vector"

Residual ("correction") (with respect to the mean value)

$\boxed{v_i = \bar{l} - l_j}$ as vectors $\boxed{\mathbf{v} = \mathbf{e} \cdot \bar{l} - \mathbf{l}}$

In geodesy we try to avoid the occurrence of systematic errors by applying

- appropriate measurement set-ups
- appropriate methods for data assessment

If, and only if, no systematic errors are present with $\Delta = 0$

it holds $\tilde{L} = \mu_L$ and $\eta = \varepsilon$.

2.3.2 Calculation rules for expectations

Given: One or more random variables X, resp. a vector of random variables $\mathbf{X}$ with probability densities $f(x)$ and expectations μ_X.

A random variable Y is obtained by applying X (or $\mathbf{X}$) in a given function g:

$$\boxed{Y = g(X)}$$

1. Probability density $f(y)$ of Y

$$y + dy = g(x + dx)$$

$$P\{x \le X \le x + dx\} = P\{y \le Y \le y + dy\}$$

Probability densities

$$f(x)dx = f(y)dy \quad |: dy$$

$$f(y) = f(x)\frac{dx}{dy}$$

$$= f(x)\frac{1}{\frac{dy}{dx}} \quad \text{with} \quad \frac{dy}{dx} = g'(x)$$

$$\boxed{f(y) = \frac{f(x)}{g'(x)} = f(x)\frac{dx}{dy}}$$

2.3.2 Calculation rules for expectations

2. Expectation μ_Y of Y

$$\boxed{E(Y) = \mu_Y = \int_{-\infty}^{+\infty} y \cdot f(y)dy = \int_{-\infty}^{+\infty} g(x)f(x)dx}$$

3. Expectation of a linear function $Y = g(X) = a + bX$ with a and b as arbitrary constant values

$$E(Y) = \int_{-\infty}^{+\infty} g(x)f(x)dx = \int_{-\infty}^{+\infty} (a + bx)f(x)dx$$

$$= a \underbrace{\int_{-\infty}^{+\infty} f(x)dx}_{1} + b \underbrace{\int_{-\infty}^{+\infty} x \cdot f(x)dx}_{\mu_X}$$

$$\boxed{\mu_Y = E(Y) = a + b \cdot \mu_X = a + b \cdot E(X)}$$

4. Expectation of a sum $Y = g(X) = \sum_{i=1}^{m} X_i = X_1 + X_2 + \ldots + X_m$ of random variables X_i

$$\boxed{\mu_Y = E(Y) = E\left\{\sum_{i=1}^{m} X_i\right\} = \sum_{i=1}^{m} E(X_i)}$$

Addition rule for expectations $E(X_1 + X_2) = E(X_1) + E(X_2)$
„Expectation of a sum = Sum of expectations“

5. Expectation of a product $Y = g(X) = \prod_{i=1}^{m} X_i = X_1 \cdot X_2 \cdot \ldots \cdot X_m$ of random variables X_i

$$\boxed{\mu_Y = E(Y) = E\left\{\prod_{i=1}^{m} X_i\right\} = \prod_{i=1}^{m} E(X_i)}$$

Multiplication rule for expectations (for <u>stochastic independent</u> random variables) $E(X_1 \cdot X_2) = E(X_1) \cdot E(X_2)$
„Expectation of a product = Product of expectations“

Adjustment Theory I

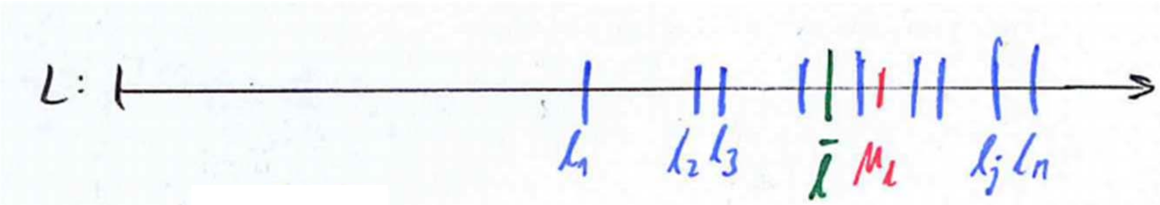
Chapter 2: Random Variables

Prof. Dr.-Ing. Frank Neitzel

2.4.1 Definition of dispersion measures

Given: Random variable L with its realizations l_j , $j = 1, 2, 3, \dots, n$

sample → „observation vector“



$$\mathbf{l} = \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{bmatrix}$$ n observations for the same random variable L

All observations originate from the same population ($n \to \infty$)
$\Rightarrow$ Expectation $E(L) = \mu_L$

Wanted: Dispersion measure

$\boxed{\varepsilon_j = l_j - \mu_L \quad \forall j}$ Random Deviations, „random errors“

„actual – nominal“

$$\boxed{\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} = \begin{bmatrix} l_1 - \mu_L \\ l_2 - \mu_L \\ \vdots \\ l_n - \mu_L \end{bmatrix} = \mathbf{l} - \mathbf{e} \cdot \mu_L}$$

Vector of deviations “error vector”

Values ε_j contain information about dispersion of single observation

Variance, if probability density is known (theoretical variance)

Definition of variance as measure of dispersion of a random variable (accuracy measure):

$$\sigma_l^2 = E\left(\varepsilon^2\right) = E\{(l-\mu_L)^2\} = E\left(l^2\right) - \mu_L^2$$

$$\boxed{\sigma_l^2 = E\left(\varepsilon^2\right) = \lim_{n\to\infty}\left\{\frac{1}{n}\sum_{j=1}^{n}\varepsilon_j^2\right\} = \lim_{n\to\infty}\left\{\frac{1}{n}\cdot\boldsymbol{\varepsilon}^{\mathrm{T}}\cdot\boldsymbol{\varepsilon}\right\}}$$

Theoretical Variance

Variance σ^2 is the mean of the squared ε_j

$$\boxed{\sigma_l = +\sqrt{\sigma_l^2}}$$

Theoretical Standard deviation

Standard deviation σ is the (positive) square root of the variance σ^2

2.4.1 Definition of dispersion measures

If we know the probability density function $f(x)$, we can compute the variance of a random variable directly (without observations):

$$\sigma_x^2 = E\left(\varepsilon^2\right) = E\{(x-\mu_X)^2\}$$

$$\boxed{\sigma_x^2 = \int_{+\infty}^{-\infty} (x-\mu_X)^2 f(x)\,dx}$$ Theoretical Variance

with $f(x) =$ density of x

2.4.2 Empirical dispersion measures

If we consider n observations (measurements, empirical values), we can compute the empirical standard deviation s as estimation for the theoretical standard deviation σ.

Attention: We have to distinguish between two cases

Case A	Case B
Expectation μ_L known	Expectation μ_L unknown

2.4.2.1 Expectation μ_L of the random variable is known (CASE A)

Given:

- Random variable L with its realizations l_j , $j = 1, 2, 3, \dots, n$
- Observation vector $\mathbf{l} = [l_1\ l_2\ \dots\ l_n]^\mathrm{T}$ with $n \ll \infty$
- Known expectation $E(L) = \mu_L$

$\Rightarrow$ Vector of random deviations $\boxed{\boldsymbol{\varepsilon} = \mathbf{l} - \mathbf{e} \cdot \mu_L}$

2.4.2.1 Expectation μ_L of the random variable is known (CASE A)

Wanted: Estimation s_l^2 for the theoretical variance σ_l^2

$$\boxed{s_l^2 = \frac{\varepsilon_1^2+\varepsilon_2^2+\ldots+\varepsilon_n^2}{n} = \frac{1}{n}\sum_{j=1}^{n}\varepsilon_j^2 = \frac{1}{n}\boldsymbol{\varepsilon}^{\mathrm{T}}\cdot\boldsymbol{\varepsilon}}$$

Empirical Variance of a single observation

$$\boxed{s_l = +\sqrt{s_l^2}}$$

Empirical Standard Deviation of a single observation

It holds: $E\left(s_l^2\right) = \sigma_l^2$ Empirical variance s_l^2 is an unbiased estimate of the theoretical variance σ_l^2.

but: $E(s_l) \neq \sigma_l$ Empirical standard deviation s_l is not an unbiased estimate of the theoretical standard deviation σ_l !

Usually $E(s_l) < \sigma_l$

2.4.2.2 Expectation μ_L of the random variable is unknown (CASE B)

Given:
- Random variable L with its realizations l_j , $j = 1, 2, 3, \dots, n$
- Observation vector $\mathbf{l} = [l_1 \; l_2 \; \dots \; l_n]^{\mathrm{T}}$ with $n \ll \infty$

Not known: Expectation $E(L) = \mu_L$

Wanted: Estimation s_l^2 for the theoretical variance σ_l^2

2.4.2.2 Expectation μ_L of the random variable is unknown (CASE B)

Solution: Replace expectation μ_L by the mean value $\bar{l}$

$$\boxed{\bar{l} = \frac{1}{n}\sum_{j=1}^{n} l_j = \frac{1}{n}\cdot \mathbf{e}^{\mathrm{T}} \cdot \mathbf{l}}$$

Empirical mean or arithmetic mean

with $\mathbf{e} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$

$$\boxed{v_j = \bar{l} - l_j \quad \forall j}$$

Residuals

Check: $\sum_{j=1}^{n} v_j \overset{!}{=} 0$

$$\boxed{\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} \bar{l} - l_1 \\ \bar{l} - l_2 \\ \vdots \\ \bar{l} - l_n \end{bmatrix} = \mathbf{e}\cdot \bar{l} - \mathbf{l}}$$

Vector of residuals

2.4.2.2 Expectation μ_L of the random variable is unknown (CASE B)

$$\boxed{s_l^2 = \frac{v_1^2+v_2^2+\ldots+v_n^2}{\underbrace{(n-1)}_{f}} = \frac{1}{(n-1)}\sum_{i=1}^{n} v_i^2 = \frac{1}{n-1}\mathbf{v}^{\mathrm{T}}\cdot\mathbf{v}}$$

Empirical Variance of a single observation

f: degree of freedom

"redundancy"

$$\boxed{s_l = +\sqrt{s_l^2}}$$

Empirical Standard Deviation of a single observation

2.4.2.2 Expectation μ_L of the random variable is unknown (CASE B)

It holds: $E(s_l^2) = \sigma_l^2$ Empirical variance s_l^2 is an unbiased estimate of the theoretical variance σ_l^2.

but: $E(s_l) \neq \sigma_l$ Empirical standard deviation s_l is <u>not</u> an unbiased estimate of the theoretical standard deviation σ_l .

Usually $E(s_l) < \sigma_l$.

One “information” (one degree of freedom f) ”number of redundant observations” is used for the estimation of the mean $\bar{l} \Rightarrow$ we have to divide by $f = n - 1$.

2.4.2.3 Standard Deviation of an arithmetic mean

Given:
- Standard deviation of a single observation s_l from random deviations (CASE A from 2.4.2.1) or from residuals (CASE B form 2.4.2.2)
- Arithmetic mean $\bar{l}$ from n observations

Definitions: Variance of an arithmetic mean $\bar{l}$ from n observations

$$\boxed{s_{\bar{l}}^2 = \frac{s_l^2}{n}}$$

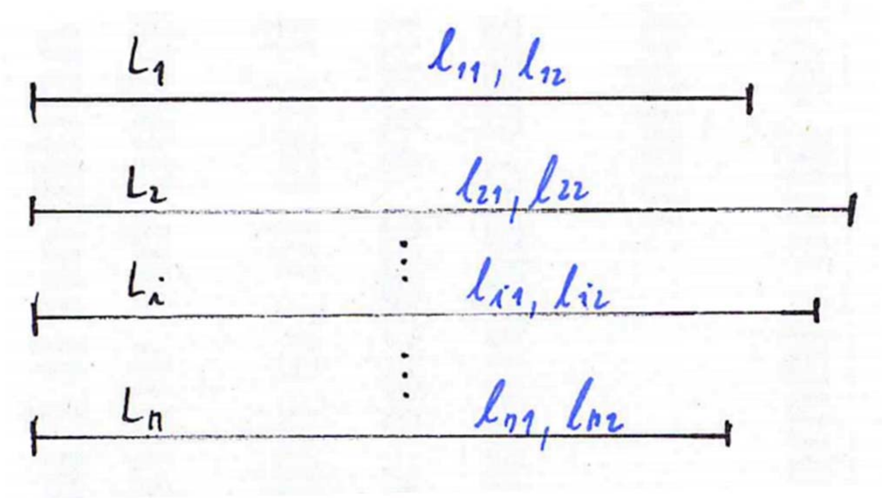
Standard deviation of an arithmetic mean $\bar{l}$ from n observations

$$\boxed{s_{\bar{l}} = \frac{s_l}{\sqrt{n}}}$$

Standard deviation of a single observation s_l can be computed from random deviations or from residuals.

2.4.2.4 Standard Deviation for double measurements of same precision

Given: Several random variables have each been measured 2 times



$\mathbf{L}^{\mathrm{T}} = [L_1\ L_2\ \cdots\ L_n]$ Random Vector

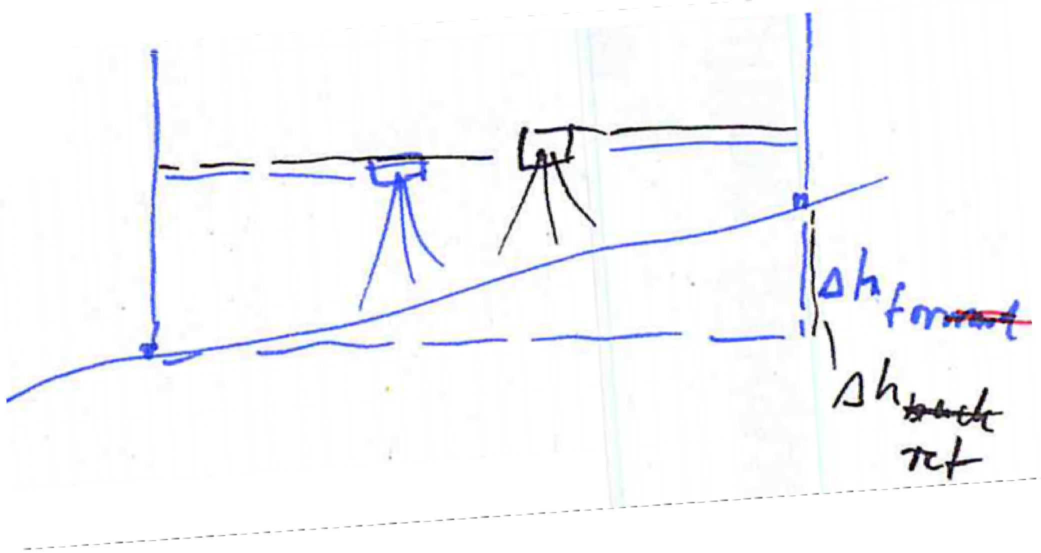
$\mathbf{l}_1^{\mathrm{T}} = [l_{11}\ l_{21}\ \cdots\ l_{n1}]$ first series of observations

$\mathbf{l}_2^{\mathrm{T}} = [l_{12}\ l_{22}\ \cdots\ l_{n2}]$ second series of observations

Wanted: Estimation s_l^2 for the theoretical variance σ_l^2

Solution: Computation of differences from double measurements

Example: Difference in the height difference from forward survey and return survey ($d = \Delta h_{for} - \Delta h_{ret}$) in differential levelling



Observation differences d_j for each double measurement

$$\boxed{d_j = l_{j2} - l_{j1}} \text{ for } j = 1, 2, \dots, n$$

$$\boxed{\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} = \mathbf{l}_2 - \mathbf{l}_1 = \begin{bmatrix} l_{12} - l_{11} \\ l_{22} - l_{21} \\ \vdots \\ l_{n2} - l_{n1} \end{bmatrix}}$$

2.4.2.4 Standard Deviation for double measurements of same precision

Standard deviation of a single observation l_{j1} or l_{j2}

$$\boxed{s_l = \sqrt{\frac{1}{2n}\mathbf{d}^{\mathrm{T}} \cdot \mathbf{d}} = \sqrt{\frac{d_1^2 + d_2^2 + \ldots + d_n^2}{2n}} = \sqrt{\frac{\sum d^2}{2n}}}$$

Empirical Standard Deviation of a single observation

Standard Deviation $s_{\bar{l}}$ of the arithmetic mean from both observations

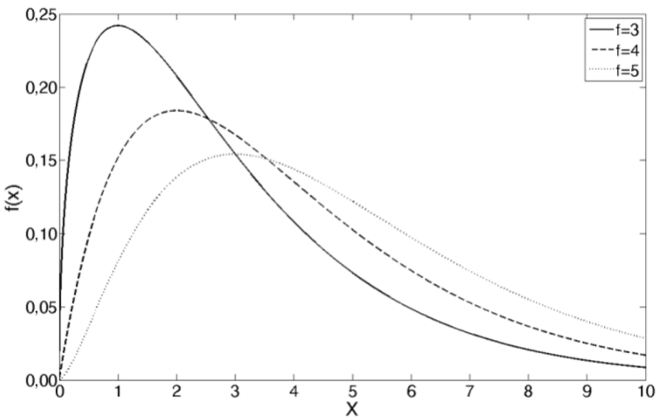
$$\boxed{s_{\bar{l}} = \frac{s_l}{\sqrt{2}} = \frac{1}{2}\sqrt{\frac{\sum d^2}{n}}}$$

Empirical Standard Deviation of the arithmetic mean

Adjustment Theory I

Chapter 2: Random Variables

Prof. Dr.-Ing. Frank Neitzel



Technische Universität Berlin

Adjustment Theory I

Chapter 3 - The Random Vector

Prof. Dr.-Ing Frank Neitzel | Institute of Geodesy and Geoinformation Science

Version: 27.10.2024

Contents

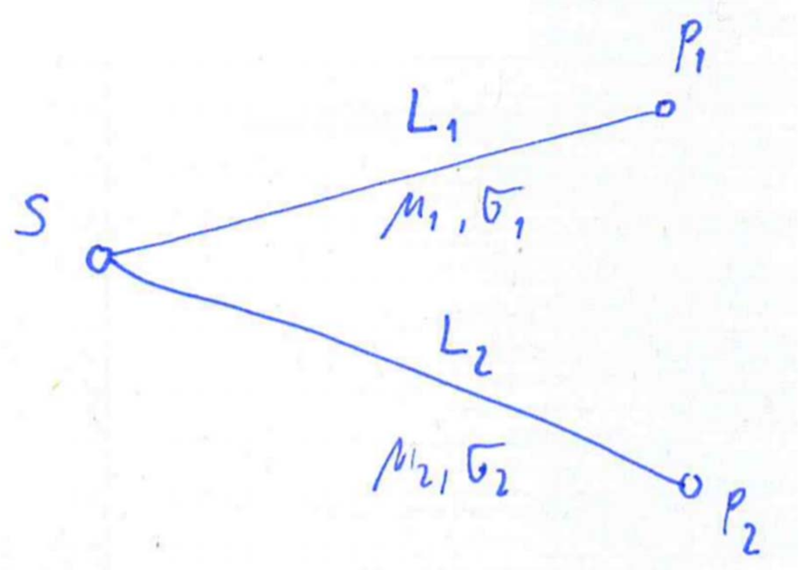
3. The Random Vector

3.1 The two-dimensional random vector

3.1.1 Theoretical variance and theoretical correlation coefficient

▶ Given: 2D random vector $\mathbf{L}_{2\times 1} = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix}$ with the random variables L_1 and L_2

Example: Measurement of two distances from the same point



3.1.1 Theoretical Variance and Theoretical Correlation Coefficient

Technische Universität Berlin

- Vector of expectations

$$\boldsymbol{\mu}_{2\times1} = \begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix} = E(\mathbf{L}_{2\times1}) = \begin{bmatrix}E(L_1)\\ E(L_2)\end{bmatrix}$$

- Vector of random deviations

$$\boldsymbol{\varepsilon}_{L\,2\times1} = \begin{bmatrix}\varepsilon_1\\ \varepsilon_2\end{bmatrix} = \mathbf{L}_{2\times1} - \boldsymbol{\mu}_{2\times1} = \begin{bmatrix}L_1 - \mu_1\\ L_2 - \mu_2\end{bmatrix}$$

- as in 2.4.1

$$E\left(\boldsymbol{\varepsilon}_{L\,2\times1}\right) = E\{\mathbf{L}_{2\times1} - \mathbf{e}_{2\times1}\cdot\mu_L\} = \underbrace{E(\mathbf{L}_{2\times1})}_{\boldsymbol{\mu}_L} - \boldsymbol{\mu}_L = \mathbf{0}$$

$$E\left\{\boldsymbol{\varepsilon}_{L\,2\times1}\cdot\boldsymbol{\varepsilon}^{\mathrm{T}}_{L\,1\times2}\right\} = E\left\{\begin{bmatrix}\varepsilon_1\\ \varepsilon_2\end{bmatrix}_{2\times1}\cdot\begin{bmatrix}\varepsilon_1 & \varepsilon_2\end{bmatrix}_{1\times2}\right\} = E\left\{\begin{bmatrix}\varepsilon_1^2 & \varepsilon_1\cdot\varepsilon_2\\ \varepsilon_2\cdot\varepsilon_1 & \varepsilon_2^2\end{bmatrix}\right\}$$

$$= \begin{bmatrix}E(\varepsilon_1^2) & E(\varepsilon_1\cdot\varepsilon_2)\\ E(\varepsilon_2\cdot\varepsilon_1) & E(\varepsilon_2^2)\end{bmatrix}$$

- Theoretical Variance of $\mathbf{L}$

$$E\left(\varepsilon_i^2\right) = \sigma_i^2 \qquad \text{for } i = 1, 2$$

3.1.1 Theoretical Variance and Theoretical Correlation Coefficient

▶ Definitions

- Theoretical Covariances between L_1 and L_2

$$E(\varepsilon_1 \cdot \varepsilon_2) = \lim_{n\to\infty} \left\{ \frac{1}{n} \cdot \sum_{j=1}^{n} \varepsilon_{1j} \cdot \varepsilon_{2j} \right\} = \sigma_{12}$$

$$\sigma_{12} = \sigma_{21}$$

- Theoretical Variance-Covariance Matrix (VCM) of $\mathbf{L}$

$$\mathbf{\Sigma}_{LL_{2\times 2}} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} = E\left\{ \boldsymbol{\varepsilon}_{L_{2\times 1}} \cdot \boldsymbol{\varepsilon}_{L_{1\times 2}}^{\mathrm{T}} \right\}$$

- Variances are always positive (“+”)
- Covariances are a “measure of the dependency” between L_1 and L_2

→ For stochastic independent values: $\sigma_{12} = 0$

3.1.1 Theoretical Variance and Theoretical Correlation Coefficient

▶ Definitions

- Theoretical Correlation Coefficient between L_1 and L_2

$$\rho_{12} = \frac{\sigma_{12}}{\sigma_1 \cdot \sigma_2}$$
$$\rho_{12} = \rho_{21}$$

- Fixed limits $-1 \leq \rho_{12} \leq +1$ **!**
- Stochastic independent (no correlation): $\rho_{12} = 0$
- Maximum correlation in the “same direction”: $\rho_{12} = +1$
- Maximum correlation in “opposite direction”: $\rho_{12} = -1$

$$\rho_{12} = \frac{\sigma_{12}}{\sigma_1 \cdot \sigma_2} \Rightarrow \sigma_{12} = \rho_{12} \cdot \sigma_1 \cdot \sigma_2$$

- Another possible representation for $\mathbf{\Sigma}_{LL}$

$$\mathbf{\Sigma}_{LL_{2\times2}} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \rho_{12} \cdot \sigma_1 \cdot \sigma_2 \\ \rho_{12} \cdot \sigma_2 \cdot \sigma_1 & \sigma_2^2 \end{bmatrix}$$

3.1.2 Empirical Variance and Empirical Correlation Coefficient

▶ Given: 2D random vector $\mathbf{L}_{2\times 1} = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix}$ with random variables L_1, L_2 and its realisations in the observation matrix $\mathbf{l}$

$$\mathbf{l}_{n\times 2} = \begin{bmatrix} l_{11} & l_{21} \\ l_{12} & l_{22} \\ \vdots & \vdots \\ l_{1n} & l_{2n} \end{bmatrix}$$

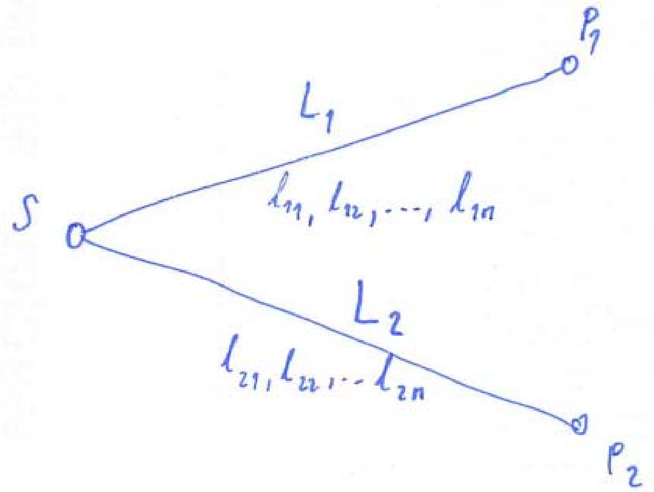
n realisations of two random variables

Corresponding observations l_{1j} and l_{2j} are obtained “pairwise”

→ Measurements are performed in a small time interval
under almost the same conditions

3.1.2 Empirical Variance and Empirical Correlation Coefficient

▶ <u>Example:</u> Measurement of two distances from the same point at the same time.



▶ Wanted: Empirical estimation r_{12} for the theoretical correlation coefficient ρ_{12}

→ We have to consider two cases, see 2.4.2

CASE A	CASE B
Expectation μ_L is known	Expectation μ_L is unknown

3.1.2.1 Empirical Variance-Covariance for known expectation μ_L (CASE A)

Technische Universität Berlin

- Vector of Expectations

$$\boldsymbol{\mu}_{2\times 1} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = E(\mathbf{L}_{2\times 1}) = \begin{bmatrix} E(L_1) \\ E(L_2) \end{bmatrix}$$

- Matrix of Random Deviations

$$\boldsymbol{\varepsilon}_{L_{n\times 2}} = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{21} \\ \varepsilon_{12} & \varepsilon_{22} \\ \vdots & \vdots \\ \varepsilon_{1n} & \varepsilon_{2n} \end{bmatrix} = \mathbf{L}_{n\times 2} - \mathbf{e}_{n\times 1} \cdot \boldsymbol{\mu}^{\mathrm{T}}_{1\times 2} = \begin{bmatrix} l_{11} - \mu_1 & l_{21} - \mu_2 \\ l_{12} - \mu_1 & l_{22} - \mu_2 \\ \vdots & \vdots \\ l_{1n} - \mu_1 & l_{2n} - \mu_2 \end{bmatrix}$$

3.1.2.1 Empirical Variance-Covariance for known expectation μ_L (CASE A)

Technische Universität Berlin

- Definitions

$$\mathbf{S}_{LL_{2\times 2}} = \frac{1}{n} \cdot \boldsymbol{\varepsilon}^{\mathrm{T}}_{L_{2\times n}} \cdot \boldsymbol{\varepsilon}_{L_{n\times 2}} = \frac{1}{n} \cdot \begin{bmatrix} \sum_{j=1}^{n} \varepsilon_{1j}^2 & \sum_{j=1}^{n} \varepsilon_{1j} \cdot \varepsilon_{2j} \\ \sum_{j=1}^{n} \varepsilon_{2j} \cdot \varepsilon_{1j} & \sum_{j=1}^{n} \varepsilon_{2j}^2 \end{bmatrix}$$

- Empirical Variances of L_1 and L_2

$$s_i^2 = \frac{1}{n} \cdot \sum_{j=1}^{n} \varepsilon_{ij}^2 \qquad \text{for } i = 1, 2$$

- Empirical Covariances between L_1 and L_2

$$s_{12} = \frac{1}{n} \cdot \sum_{j=1}^{n} \varepsilon_{1j} \cdot \varepsilon_{2j} \qquad s_{12} = s_{21}$$

- Empirical Variance-Covariance Matrix VCM of $\mathbf{L}$

$$\mathbf{S}_{LL_{2\times 2}} = \frac{1}{n} \cdot \boldsymbol{\varepsilon}^{\mathrm{T}}_{L_{2\times n}} \cdot \boldsymbol{\varepsilon}_{L_{n\times 2}} = \begin{bmatrix} s_1^2 & s_{12} \\ s_{21} & s_2^2 \end{bmatrix}$$

3.1.2.2 Empirical Variance-Covariance for unknown expectation μ_L (CASE B)

▶ Vector of expectations $\boldsymbol{\mu}_L$ is not known → has to be replaced by

- Vector of mean values

$$\bar{\mathbf{l}}_{2\times 1} = \begin{bmatrix} \bar{l}_1 \\ \bar{l}_2 \end{bmatrix} = \frac{1}{n} \cdot \mathbf{l}_{2\times n}^{\mathrm{T}} \cdot \mathbf{e}_{n\times 1} = \frac{1}{n} \cdot \begin{bmatrix} l_{11} + l_{12} + \cdots + l_{1n} \\ l_{21} + l_{22} + \cdots + l_{2n} \end{bmatrix}$$

→ Matrix of residuals

$$\mathbf{v}_{n\times 2} = \mathbf{e}_{n\times 1} \cdot \bar{\mathbf{l}}_{1\times 2}^{\mathrm{T}} - \mathbf{l}_{n\times 2} = \begin{bmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \\ \vdots & \vdots \\ v_{1n} & v_{2n} \end{bmatrix} = \begin{bmatrix} \bar{l}_1 - l_{11} & \bar{l}_2 - l_{21} \\ \bar{l}_1 - l_{12} & \bar{l}_2 - l_{22} \\ \vdots & \vdots \\ \bar{l}_1 - l_{1n} & \bar{l}_2 - l_{2n} \end{bmatrix}$$

3.1.2.2 Empirical Variance-Covariance for unknown expectation μ_L (CASE B)

▶ Definitions

- Empirical Variances of L_1 and L_2

$$s_i^2 = \frac{1}{(n-1)} \cdot \sum_{j=1}^{n} v_{ij}^2 \qquad \text{for } i = 1, 2$$

- Empirical Covariances between L_1 and L_2

$$s_{12} = \frac{1}{(n-1)} \cdot \sum_{j=1}^{n} v_{1j} \cdot v_{2j}$$

$$s_{12} = s_{21}$$

- Empirical Variance-Covariance Matrix VCM of $\mathbf{L}$

$$\mathbf{S}_{LL_{2\times 2}} = \frac{1}{(n-1)} \cdot \mathbf{v}_{2\times n}^{\mathrm{T}} \cdot \mathbf{v}_{n\times 2} = \begin{bmatrix} s_1^2 & s_{12} \\ s_{21} & s_2^2 \end{bmatrix}$$

3.1.2.3 Fusion of CASE A and CASE B yields Empirical Correlation

Technische Universität Berlin

▶ Definitions

- Empirical Correlation Coefficient between L_1 and L_2

$$r_{12} = \frac{s_{12}}{s_1 \cdot s_2}$$

$$s_{12} = s_{21}$$

- Fixed limits $-1 \leq r_{12} \leq +1$ **!**
- stochastic independent (no correlation): $r_{12} = 0$
- Maximum correlation in the “same direction”: $r_{12} = +1$
- Maximum correlation in “opposite direction”: $r_{12} = -1$

$$r_{12} = \frac{s_{12}}{s_1 \cdot s_2} \Rightarrow s_{12} = r_{12} \cdot s_1 \cdot s_2$$

- Another possible representation for $\mathbf{S}_{LL}$

$$\mathbf{S}_{LL_{2\times 2}} = \begin{bmatrix} s_1^2 & s_{12} \\ s_{21} & s_2^2 \end{bmatrix} = \begin{bmatrix} s_1^2 & r_{12} \cdot s_1 \cdot s_2 \\ r_{12} \cdot s_2 \cdot s_1 & s_2^2 \end{bmatrix}$$

Summary

- Theoretical Variance-Covariance Matrix

$$\boldsymbol{\Sigma}_{ll} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}$$

If l_1, l_2 stochastic independent: Covariances $\sigma_{12} = \sigma_{21} = 0$

- Empirical Variance Covariance Matrix

$$\mathbf{S}_{ll} = \begin{bmatrix} s_1^2 & s_{12} \\ s_{21} & s_2^2 \end{bmatrix}$$

Summary

► Computation of empirical VCM for known expectation μ

$$\mathbf{S}_{ll} = \frac{\boldsymbol{\varepsilon}^{\mathrm{T}}\boldsymbol{\varepsilon}}{n} = \frac{1}{n} \cdot \begin{bmatrix} \sum_{j=1}^{n} \varepsilon_{1j}^{2} & \sum_{j=1}^{n} \varepsilon_{1j} \cdot \varepsilon_{2j} \\ \sum_{j=1}^{n} \varepsilon_{1j} \cdot \varepsilon_{2j} & \sum_{j=1}^{n} \varepsilon_{2j}^{2} \end{bmatrix}$$

► Computation of empirical VCM for unknown expectation μ

$$\mathbf{S}_{ll} = \frac{\mathbf{v}^{\mathrm{T}}\mathbf{v}}{n-1} = \frac{1}{n-1} \cdot \begin{bmatrix} \sum_{j=1}^{n} v_{1j}^{2} & \sum_{j=1}^{n} v_{1j} \cdot v_{2j} \\ \sum_{j=1}^{n} v_{1j} \cdot v_{2j} & \sum_{j=1}^{n} v_{2j}^{2} \end{bmatrix}$$

Summary

Technische Universität Berlin

- Theoretical Correlation Coefficient

$$\rho_{12} = \frac{\sigma_{12}}{\sigma_1 \sigma_2} \;, \quad -1 \leq \rho \leq +1$$

- Empirical Correlation Coefficient

$$r_{12} = \frac{s_{12}}{s_1 s_2} \;, \quad -1 \leq r \leq +1$$

Correlation can originate from

1. Mathematical correlation (functional correlation, algebraic correlation)

   If we apply a functional relationship between two or more realisations of a random variable, we obtain a correlation between the resulting estimations.
   → Originates on a purely mathematical basis

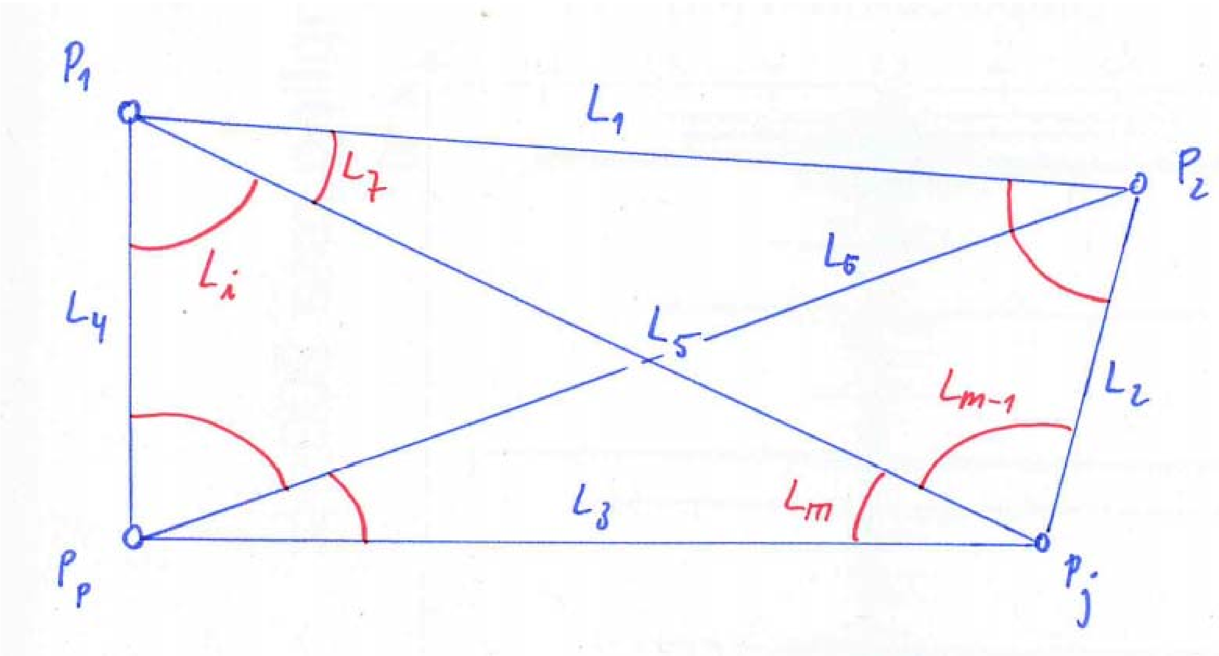
2. Physical correlation

   Correlation between two realisations of a random variable (→ originates from the measurement) due to small systematic deviations, that are not (or not sufficiently) considered in the functional model.

3.2 The m-dimensional random vector

3.2.1 Theoretical expectation and theoretical covariance matrix

<u>Example:</u> Measurement of directions and distances in a geodetic network



3.2.1 Theoretical expectation and theoretical covariance matrix

- Given: m random variables $L_1, L_2, \dots, L_m$

- Random vector

$$\mathbf{L}_{m\times 1} = \begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ L_m \end{bmatrix}$$

- Vector of expectations

$$\boldsymbol{\mu}_{m\times 1} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{bmatrix} = E\{\mathbf{L}_{m\times 1}\} = \begin{bmatrix} E(L_1) \\ E(L_2) \\ \vdots \\ E(L_m) \end{bmatrix}$$

3.2.1 Theoretical expectation and theoretical covariance matrix

- Vector of random deviations

$$\boldsymbol{\varepsilon}_{m\times 1} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{bmatrix} = \mathbf{L}_{m\times 1} - \boldsymbol{\mu}_{m\times 1} = \begin{bmatrix} L_1 - \mu_1 \\ L_2 - \mu_2 \\ \vdots \\ L_m - \mu_m \end{bmatrix}$$

- as in 2.4.1

$$E(\boldsymbol{\varepsilon}_{m\times 1}) = E\{\mathbf{L}_{m\times 1} - \boldsymbol{\mu}_{L_{m\times 1}}\} = \underbrace{E(\mathbf{L}_{m\times 1})}_{\boldsymbol{\mu}_L} - \boldsymbol{\mu}_{L_{m\times 1}} = \mathbf{0}$$

$$E\{\boldsymbol{\varepsilon}_{L_{m\times 1}} \cdot \boldsymbol{\varepsilon}^{\mathrm{T}}_{L_{1\times m}}\} = \begin{bmatrix} E(\varepsilon_1^2) & E(\varepsilon_1 \cdot \varepsilon_2) & \cdots & E(\varepsilon_1 \cdot \varepsilon_m) \\ E(\varepsilon_2 \cdot \varepsilon_1) & E(\varepsilon_2^2) & \cdots & E(\varepsilon_2 \cdot \varepsilon_m) \\ \vdots & \vdots & \ddots & \vdots \\ E(\varepsilon_m \cdot \varepsilon_1) & E(\varepsilon_m \cdot \varepsilon_2) & \cdots & E(\varepsilon_m \cdot \varepsilon_m) \end{bmatrix} = \boldsymbol{\Sigma}_{LL_{m\times m}}$$

3.2.1 Theoretical expectation and theoretical covariance matrix

▶ Theoretical VCM of L

$$\mathbf{\Sigma}_{LL_{m\times m}} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1m} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1} & \sigma_{m2} & \cdots & \sigma_m^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \rho_{12}\cdot\sigma_1\cdot\sigma_2 & \cdots & \sigma_{1m} \\ \rho_{21}\cdot\sigma_2\cdot\sigma_1 & \sigma_2^2 & \cdots & \sigma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{m1}\cdot\sigma_m\cdot\sigma_1 & \rho_{m2}\cdot\sigma_m\cdot\sigma_2 & \cdots & \sigma_m^2 \end{bmatrix}$$

▶ Theoretical Variance of L_i

$$E\left(\varepsilon_i^2\right) = \sigma_i^2$$

3.2.2 Empirical expectation and empirical covariance matrix

- Given: m random variables $\mathbf{L}_{m\times 1} = \begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ L_m \end{bmatrix}$

- $m \times n$-dimensional observation matrix

$$\mathbf{l}_{m\times n} = \begin{bmatrix} l_{11} & l_{21} & \cdots & l_{m1} \\ l_{12} & l_{22} & \cdots & l_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ l_{1n} & l_{2n} & \cdots & l_{mn} \end{bmatrix}$$

m different random variables were measured n times

- m-dimensional vector of mean values respective expectations

$$\bar{\mathbf{l}} = \begin{bmatrix} \bar{l}_1 \\ \bar{l}_2 \\ \vdots \\ \bar{l}_m \end{bmatrix} \qquad \text{resp.} \qquad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{bmatrix}$$

3.2.2 Empirical expectation and empirical covariance matrix

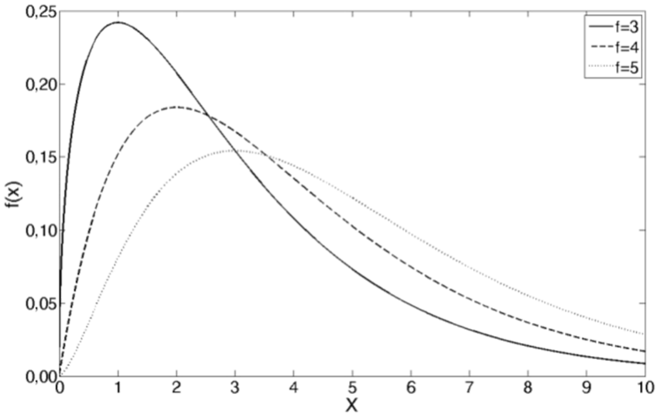
▶ m-dimensional VCM

$$\mathbf{S}_{ll_{m\times m}} = \begin{bmatrix} s_1^2 & s_{12} & \cdots & s_{1m} \\ s_{21} & s_2^2 & \cdots & s_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ s_{m1} & s_{m2} & \cdots & s_m^2 \end{bmatrix}$$

▶ Correlation coefficient for the m-dimensional case

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \cdot \sigma_j} \qquad \text{respective} \qquad r_{ij} = \frac{s_{ij}}{s_i \cdot s_j} \qquad \text{for } i, j = 1,2, \ldots, m \text{ and } i \neq j$$

→ All computations (e.g. residuals, empirical variances and covariances and correlation coefficients as in 2-dimensional case)

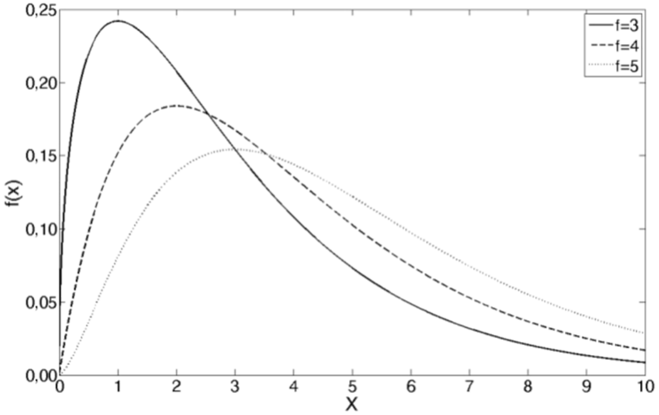


Technische Universität Berlin

Adjustment Theory I

Chapter 3 - The Random Vector

Prof. Dr.-Ing Frank Neitzel | Institute of Geodesy and Geoinformation Science



Technische Universität Berlin

Adjustment Theory I

Chapter 4 – Propagation of Observation Errors

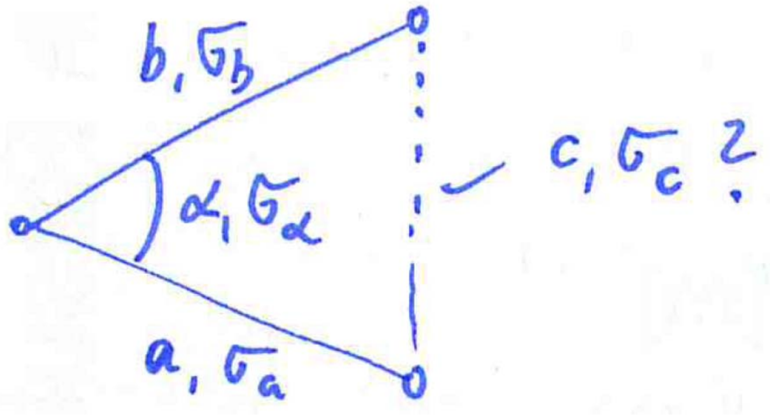
Prof. Dr.-Ing. Frank Neitzel | Institute of Geodesy and Geoinformation Science

Version: 01 November 2024

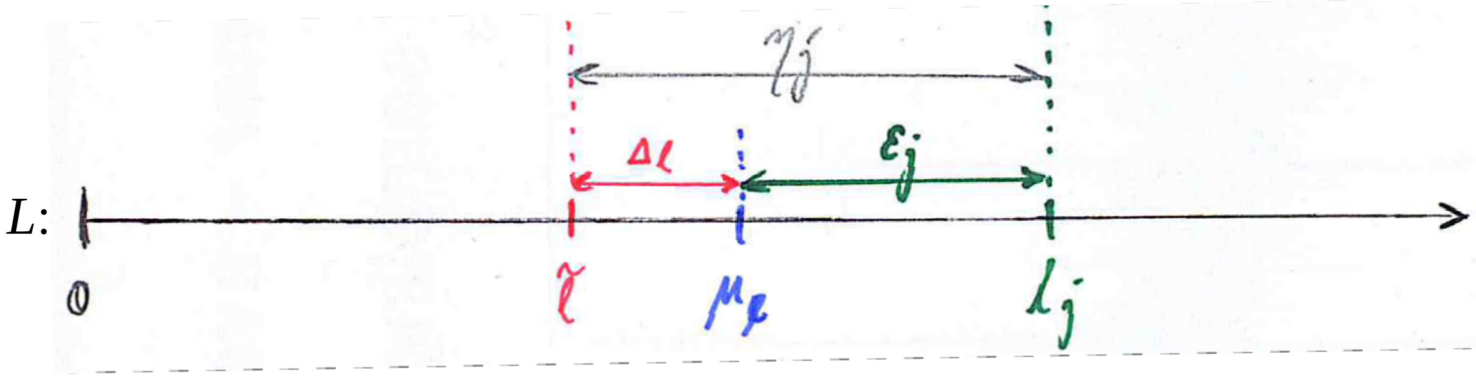
Contents

4. Propagation of Observation Errors

Example:



4.1 True, systematic and random deviations

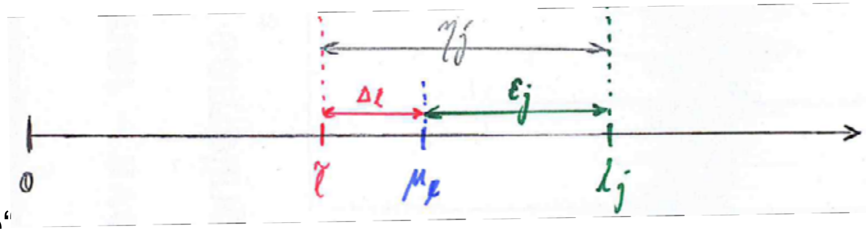


4.1 Introduction

▶ For a single random variable L

- Observation l_j
- True value of the random variable $\tilde{l}$
- Expectation $\mu_L = E(L) = \lim_{n\to\infty} \bar{l} = \lim_{n\to\infty} (\frac{1}{n}\sum_{j=1}^{n} l_j) = \lim_{n\to\infty} (\frac{1}{n} \cdot \mathbf{e}_{1\times n}^{\mathrm{T}} \cdot \mathbf{l}_{n\times 1})$
- Random deviations: $\varepsilon_j = l_j - \mu_L \quad \forall j$

  („random errors“) „actual – nominal“



- Systematic deviations: $\Delta_l = \mu_L - \tilde{l}$

  („systematic errors“) „expectation – true value“

- True deviations: $\eta_j = l_j - \tilde{l} = \Delta_l + \varepsilon_j \quad \forall j$

  („true errors“)

$$\boxed{l_j = \tilde{l} + \eta_j = \tilde{l} + \Delta_l + \varepsilon_j = \mu_L + \varepsilon_j} \quad \forall j$$

4.1 Introduction

► For random vectors $\mathbf{L}$

- Random vector:

$$\mathbf{L}_{1\times n}^{\mathrm{T}} = [L_1 \quad L_2 \quad \cdots \quad L_n]$$

- Observation vector:

$$\mathbf{l}_{1\times n}^{\mathrm{T}} = [l_1 \quad l_2 \quad \cdots \quad l_n]$$

- Vector of true values:

$$\tilde{\mathbf{l}}_{1\times n}^{\mathrm{T}} = [\tilde{l}_1 \quad \tilde{l}_2 \quad \cdots \quad \tilde{l}_n]$$

- Expectation vector:

$$\boldsymbol{\mu}_{L\,1\times n}^{\;\mathrm{T}} = [\mu_1 \quad \mu_2 \quad \cdots \quad \mu_n] = E(\mathbf{L}_{1\times n}^{\mathrm{T}})$$

4.1 Introduction

▶ For random vectors $\mathbf{L}$

- Random deviation:

$$\boldsymbol{\varepsilon}_{l_{n\times1}} = \mathbf{l}_{n\times1} - \boldsymbol{\mu}_{L_{n\times1}}$$

- Systematic deviations:

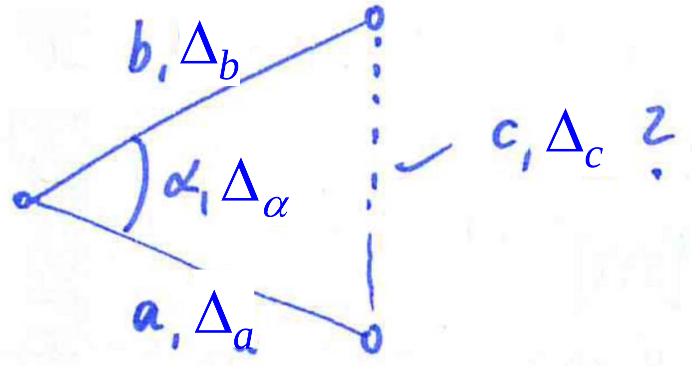
$$\Delta_{l_{n\times1}} = \boldsymbol{\mu}_{L_{n\times1}} - \tilde{\mathbf{l}}_{n\times1}$$

- True deviations: $\boldsymbol{\eta}_{l_{n\times1}} = \mathbf{l}_{n\times1} - \tilde{\mathbf{l}}_{n\times1} = \Delta_{l_{n\times1}} + \boldsymbol{\varepsilon}_{l_{n\times1}}$

$$\boxed{\mathbf{l}_{n\times1} = \tilde{\mathbf{l}}_{n\times1} + \boldsymbol{\eta}_{l_{n\times1}} = \tilde{\mathbf{l}}_{n\times1} + \Delta_{l_{n\times1}} + \boldsymbol{\varepsilon}_{L_{n\times1}} = \boldsymbol{\mu}_{l_{n\times1}} + \boldsymbol{\varepsilon}_{l_{n\times1}}}$$

4.2 Propagation of systematic deviations

<u>Example</u>:



- Given: Observation vector

$$\mathbf{l}_{n\times1} = \tilde{\mathbf{l}}_{n\times1} + \Delta_{l_{n\times1}} + \boldsymbol{\varepsilon}_{l_{n\times1}}$$

- Wanted: Quantities $\mathbf{f}_{u\times1}$ that cannot be directly measured, but can be calculated from $\mathbf{l}_{n\times1}$ as $\mathbf{f}_{u\times1} = \boldsymbol{\Phi}_{u\times1}(\mathbf{l}_{n\times1})$.

"$\mathbf{f}$ can be derived from $\mathbf{l}$ via a functional model $\boldsymbol{\Phi}$"

4.2 Propagation of systematic deviations

► Question: What is the influence $\Delta\mathbf{f}_{u\times 1}$ of systematic errors Δl_j in $\mathbf{l}_{n\times 1}$ on the derived quantities $\mathbf{f}_{u\times 1}$?

► Relation between $\mathbf{f}$ and $\mathbf{l}$ $\rightarrow$ functional model

$$\mathbf{f}_{u\times 1} = \boldsymbol{\Phi}_{u\times 1}(\mathbf{l}_{n\times 1}) = \begin{bmatrix} \varphi_1(\mathbf{l}_{n\times 1}) \\ \varphi_2(\mathbf{l}_{n\times 1}) \\ \vdots \\ \varphi_{\mathrm{u}}(\mathbf{l}_{n\times 1}) \end{bmatrix}$$

derived quantities

$\rightarrow$ unknowns

functional relation

contains arbitrary (non-linear) differentiable functions

4.2.1 Linear Functional Models

If relation between $\mathbf{f}$ and $\mathbf{l}$ linear $\rightarrow$ linear functional model

$$\boxed{\mathbf{f}_{u\times 1} = \mathbf{F}_{u\times n} \cdot \mathbf{l}_{n\times 1}}$$

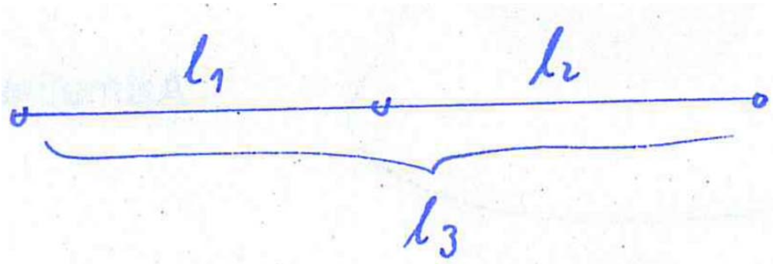
"Design Matrix" observation vector

contains the coefficients

of the linear functional model

4.2.1 Linear Functional Models

Example: Zero correction of EDM



$$a = l_1 + l_2 - l_3$$

Observation vector: $\mathbf{l}_{3\times1} = \begin{bmatrix} l_1 \\ l_2 \\ l_3 \end{bmatrix}$, design matrix $\mathbf{F}_{1\times3} = [1 \quad 1 \quad -1]$

Propagation of true deviations („Propagation law for true deviations“)

$$\boxed{\boldsymbol{\eta}_{f_{u\times1}} = \mathbf{F}_{u\times n} \cdot \boldsymbol{\eta}_{l_{n\times1}}}$$

4.2.1 Linear Functional Models

▶ Now we consider systematic deviations

- True deviation of $\mathbf{l}_{n\times1}$:

$$\boldsymbol{\eta}_{l_{n\times1}} = \Delta_{l_{n\times1}} + \boldsymbol{\varepsilon}_{l_{n\times1}}$$

- True deviation of $\mathbf{f}_{u\times1}$:

$$\boldsymbol{\eta}_{f_{u\times1}} = \Delta_{f_{u\times1}} + \boldsymbol{\varepsilon}_{f_{u\times1}}$$

$$\boldsymbol{\eta}_f = \Delta_f + \boldsymbol{\varepsilon}_f = \mathbf{F}\cdot(\Delta_l + \boldsymbol{\varepsilon}_l) = \mathbf{F}\cdot\Delta_l + \mathbf{F}\cdot\boldsymbol{\varepsilon}_l$$

- Expectations:

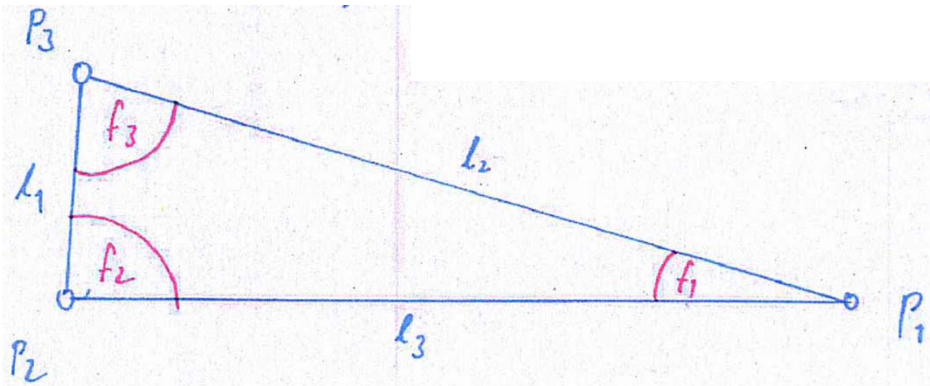
$$E(\boldsymbol{\eta}_f) = \Delta_f + \underbrace{E(\boldsymbol{\varepsilon}_f)}_{=0} = \mathbf{F}\cdot\Delta_l + \mathbf{F}\cdot\underbrace{E(\boldsymbol{\varepsilon}_l)}_{=0}$$

- Propagation law for systematic deviations:

$$\Rightarrow \quad \boxed{\Delta_{f_{u\times1}} = \mathbf{F}_{u\times n}\cdot\Delta_{l_{n\times1}}}$$

4.2.2 Non-linear Functional Models

- Problem: Function of the observations $\mathbf{f}_{u\times 1} = \mathbf{\Phi}_{u\times 1}(\mathbf{l}_{n\times 1})$ oftentimes not linear
- Example:
  - 3 measured distances in a triangle
  - compute the three angles $\mathbf{f}_{u\times 1} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}$
  - and the corresponding systematic deviations $\Delta\mathbf{f}_{u\times 1} = \begin{bmatrix} \Delta f_1 \\ \Delta f_2 \\ \Delta f_3 \end{bmatrix}$



4.2.2 Non-linear Functional Models

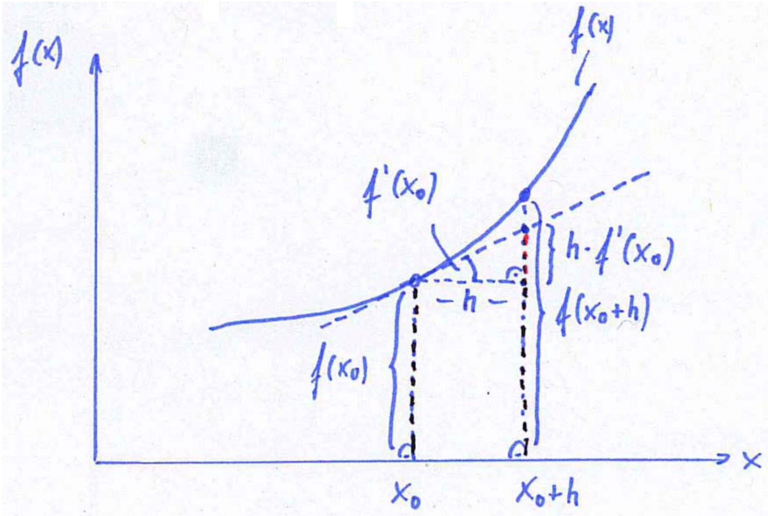
- Solution: Linearisation of the function

$$\mathbf{f}_{u\times1} = \tilde{\mathbf{f}}_{u\times1} + \boldsymbol{\eta}_{f_{u\times1}} = \boldsymbol{\Phi}_{u\times1}(\mathbf{l}_{n\times1}) = \boldsymbol{\Phi}_{u\times1}(\tilde{\mathbf{l}}_{n\times1} + \boldsymbol{\eta}_{l_{n\times1}})$$

  - $\boldsymbol{\eta}_{l_{n\times1}}$ true deviation of $\mathbf{l}_{n\times1}$
  - $\tilde{\mathbf{f}}_{u\times1}$ true value of $\mathbf{f}_{u\times1}$
  - $\boldsymbol{\eta}_{f_{u\times1}}$ true deviation of $\mathbf{f}_{u\times1}$

- Taylor Series introduced by Brook Taylor in 1715
- We know from mathematics: Taylor series is a representation of a function as an infinite sum of terms ($n \to \infty$) that are calculated from the values of the function's derivatives at a single point
- In practice: Approximation of a function by using a finite number n of terms of its Taylor Series

4.2.2 Non-linear Functional Models



▶ Function at the point x_0: $f(x_0)$ with $x_0 \mathrel{\widehat{=}}$ “starting value” or “approximate value”

▶ Function at the point $x_0 + h$:

“factorial 3” $= 1 \cdot 2 \cdot 3 = 6$

$$f(x_0 + h) = f(x_0) + \frac{h}{1!} \cdot f'(x_0) \underbrace{+ \frac{h^2}{2!} \cdot f''(x_0) + \frac{h^3}{3!} \cdot f'''(x_0) + \cdots + \frac{h^n}{n!} \cdot f^{(n)}(x_0)}$$

Terms of higher order can be neglected if $h \ll x_0$!

$$\boxed{f(x_0 + h) = f(x_0) + h \cdot f'(x_0)}$$ Linearisation of a function

4.2.2 Non-linear Functional Models

Technische Universität Berlin

Example 1:

$$f(x) = x^2,\ x_0 = 19,\ h = 1$$

$$f(x_0 + h) = x_0^2 + 1 \cdot 2x_0 = 19^2 + 2 \cdot 19 = 399$$

Direct solution: $f(x_0 + h) = (19 + 1)^2 = 400$

→ Difference, because we have assumed that terms of higher order $= 0$, but $h \ll x_0$ is not the case in this example

→ Difference is called "linearisation error"

▶ Question: How can we avoid a large linearisation error?

▶ Answer: h must be a small value

→ Starting value must be "close" to the solution

4.2.2 Non-linear Functional Models

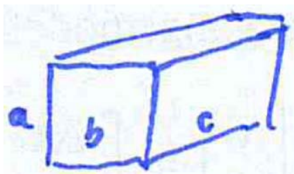
Example 2:

$$f(x) = x^2,\ x_0 = 19.9,\ h = 0.1$$

$$f(x_0 + h) = 19.9^2 + 0.1 \cdot 2 \cdot 19.9 = 396.01 + 3.98 = 399.99$$

- Function of more variables

  e.g. computation of volume $V = a \cdot b \cdot c$



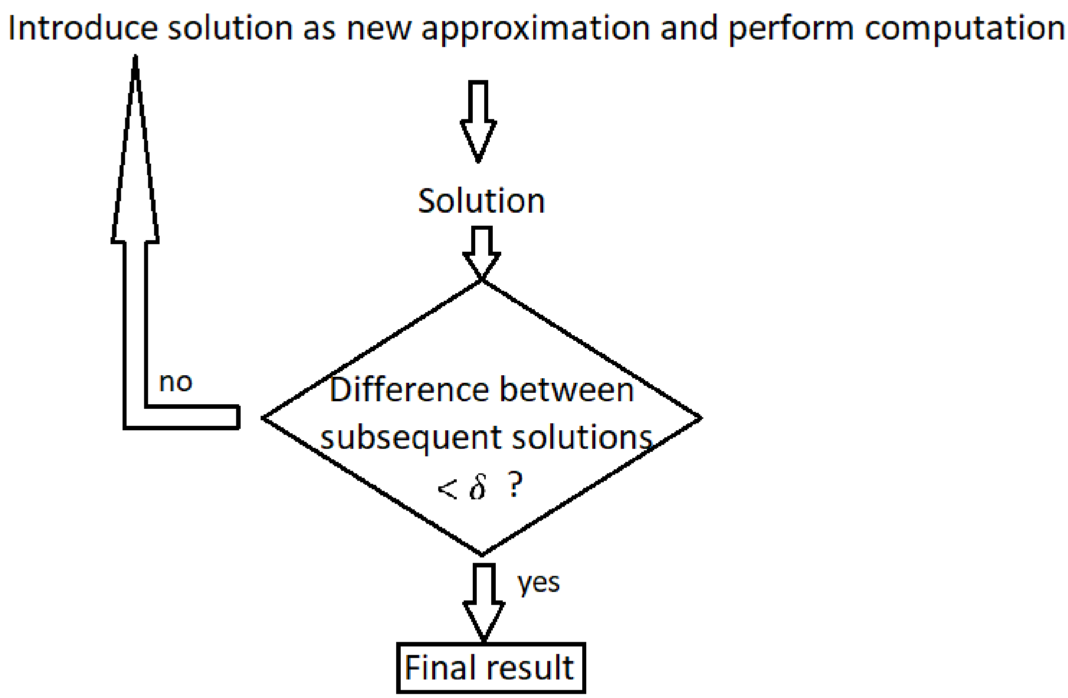
$$f(a_0 + h, b_0 + k, c_0 + m) = f(a_0, b_0, c_0) + h \cdot \left(\frac{\partial f}{\partial a}\right)_0 + k \cdot \left(\frac{\partial f}{\partial b}\right)_0 + m \cdot \left(\frac{\partial f}{\partial c}\right)_0$$

Partial derivative at the place of starting values

4.2.2 Non-linear Functional Models

- General problem: How can we be sure that we have reached an appropriate solution?

  → Iterative computing!



4.2.2 Non-linear Functional Models

$$\tilde{\mathbf{f}}_{u\times 1} + \boldsymbol{\eta}_{f\,u\times 1} = \boldsymbol{\Phi}_{u\times 1}(\tilde{\mathbf{l}}_{n\times 1} + \boldsymbol{\eta}_{l\,n\times 1})$$

contains arbitrary (non-linear)

differentiable functions

$$\boxed{\tilde{\mathbf{f}} + \boldsymbol{\eta}_f = \boldsymbol{\Phi}_{u\times 1}\left(\tilde{\mathbf{l}}_{n\times 1}\right) + \frac{\partial\boldsymbol{\Phi}(\mathbf{l})}{\partial\mathbf{l}}_{u\times n} \cdot \boldsymbol{\eta}_{l\,n\times 1} + \frac{\partial^2\boldsymbol{\Phi}(\mathbf{l})}{2\cdot\partial\mathbf{l}^2}_{u\times n} \cdot \boldsymbol{\eta}^2_{l\,n\times 1} + \cdots}$$

Taylor Series

4.2.2 Non-linear Functional Models

Under the assumption that

$$\|\boldsymbol{\eta}_l\|_{n\times1} \ll \|\mathbf{l}\|_{n\times1}$$

Euclidean norm

$$\sqrt{\sum_{i=1}^{n} {\eta_{l_i}}^2}$$

$$\sqrt{\sum_{i=1}^{n} {l_i}^2}$$

we can truncate the Taylor Series after the first term

$$\tilde{\mathbf{f}} + \boldsymbol{\eta}_f = \boldsymbol{\Phi}(\tilde{\mathbf{l}}) + \frac{\partial \boldsymbol{\Phi}(\mathbf{l})}{\partial \boldsymbol{l}} \cdot \boldsymbol{\eta}_l$$

$$\text{with } \tilde{\mathbf{f}} = \boldsymbol{\Phi}(\tilde{\mathbf{l}})$$

$$\Rightarrow \boldsymbol{\eta}_{f_{u\times1}} = \underbrace{\frac{\partial \boldsymbol{\Phi}(\mathbf{l})}{\partial \mathbf{l}}}_{\mathbf{J}_{u\times n}} \cdot \boldsymbol{\eta}_{l_{n\times1}}$$

4.2.2 Non-linear Functional Models

Jacobian Matrix

$$\mathbf{J}_{u\times n} = \begin{bmatrix} \frac{\partial \varphi_1(\mathbf{l})}{\partial l_1} & \frac{\partial \varphi_1(\mathbf{l})}{\partial l_2} & \cdots & \frac{\partial \varphi_1(\mathbf{l})}{\partial l_n} \\ \frac{\partial \varphi_2(\mathbf{l})}{\partial l_1} & \frac{\partial \varphi_2(\mathbf{l})}{\partial l_2} & \cdots & \frac{\partial \varphi_2(\mathbf{l})}{\partial l_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \varphi_u(\mathbf{l})}{\partial l_1} & \frac{\partial \varphi_u(\mathbf{l})}{\partial l_2} & \cdots & \frac{\partial \varphi_u(\mathbf{l})}{\partial l_n} \end{bmatrix} = \begin{bmatrix} j_{11} & j_{12} & \cdots & j_{1n} \\ j_{21} & j_{22} & \cdots & j_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ j_{u1} & j_{u2} & \cdots & j_{un} \end{bmatrix}$$

<u>Now</u>: We take elements of matrix $\mathbf{J}$ and insert them into matrix $\mathbf{F}$

$$\mathbf{F}_{u\times n} = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{u1} & f_{u2} & \cdots & f_{un} \end{bmatrix} = \begin{matrix} f_{11} = j_{11}, f_{12} = j_{12} \;\ldots f_{1n} = j_{1n} \\ \vdots \\ f_{u1} = j_{u1}, f_{u2} = j_{u2} \;\ldots f_{un} = j_{un} \end{matrix}$$

4.2.2 Non-linear Functional Models

- Advantage of linearisation of the functional model:

  We can apply the simple formulas for the linear case!

- Propagation law for true deviations:

$$\boxed{\boldsymbol{\eta}_{f_{u\times 1}} = \mathbf{F}_{u\times n} \cdot \boldsymbol{\eta}_{l_{n\times 1}}}$$

- Propagation law for systematic deviations:

$$\boxed{\Delta_{f_{u\times 1}} = \mathbf{F}_{u\times 1} \cdot \Delta_{l_{n\times 1}}}$$

4.3 Propagation of random deviations

$$\boldsymbol{\eta}_{f\,u\times 1} = \Delta_{f\,u\times 1} + \boldsymbol{\varepsilon}_{f\,u\times 1} = \mathbf{F}_{u\times n} \cdot \Delta_{l\,n\times 1} + \mathbf{F}_{u\times n} \cdot \boldsymbol{\varepsilon}_{l\,n\times 1}$$

▶ Propagation of random deviations

$$\Rightarrow \quad \boxed{\boldsymbol{\varepsilon}_{f\,u\times 1} = \mathbf{F}_{u\times n} \cdot \boldsymbol{\varepsilon}_{l\,n\times 1}}$$

Problem: Random deviations $\boldsymbol{\varepsilon}_{l\,n\times 1}$ of $\mathbf{l}_{n\times 1}$ are usually not known in practice

Solution: Transition to variances and covariances $\Rightarrow$ Variance-Covariance Matrices

4.3 Propagation of random deviations

▶ Given:

$$\boldsymbol{\Sigma}_{LL_{n\times n}} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix} \text{ of } \mathbf{l}_{n\times 1}$$

▶ Wanted:

$$\boldsymbol{\Sigma}_{ff_{u\times u}} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1u} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2u} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{u1} & \sigma_{u2} & \cdots & \sigma_u^2 \end{bmatrix} \text{ of } \mathbf{f}_{u\times 1}$$

4.3 Propagation of random deviations

▶ Definition:

$$\boldsymbol{\Sigma}_{LL_{n\times n}} = E\{\boldsymbol{\varepsilon}_{L_{n\times 1}} \cdot \boldsymbol{\varepsilon}^{\mathrm{T}}_{L_{1\times n}}\}$$

$$\boldsymbol{\Sigma}_{ff_{u\times u}} = E\left\{\boldsymbol{\varepsilon}_{f_{u\times 1}} \cdot \boldsymbol{\varepsilon}^{\mathrm{T}}_{f_{1\times u}}\right\}$$

$$\boldsymbol{\varepsilon}_{f_{u\times 1}} \cdot \boldsymbol{\varepsilon}^{\mathrm{T}}_{f} = \mathbf{F}_{u\times n} \cdot \boldsymbol{\varepsilon}_{L_{n\times 1}} \cdot \boldsymbol{\varepsilon}^{\mathrm{T}}_{L_{1\times n}} \cdot \mathbf{F}^{\mathrm{T}}_{n\times u}$$

$$\Rightarrow \underbrace{E\left\{\boldsymbol{\varepsilon}_{f_{u\times 1}} \cdot \boldsymbol{\varepsilon}^{\mathrm{T}}_{f_{1\times u}}\right\}}_{\boldsymbol{\Sigma}_{ff}} = \mathbf{F}_{u\times n} \underbrace{E\{\boldsymbol{\varepsilon}_{L_{n\times 1}} \cdot \boldsymbol{\varepsilon}^{\mathrm{T}}_{L_{1\times n}}\}}_{\boldsymbol{\Sigma}_{LL}} \mathbf{F}^{\mathrm{T}}_{n\times u}$$

▶ Propagation law of variances and covariances (Variance-covariance propagation)

$$\boldsymbol{\Sigma}_{ff_{u\times u}} = \mathbf{F}_{u\times n} \cdot \boldsymbol{\Sigma}_{ll_{n\times n}} \cdot \mathbf{F}^{\mathrm{T}}_{n\times u}$$

General case of variance-covariance propagation

4.4 Propagation of variances and covariances

Technische Universität Berlin

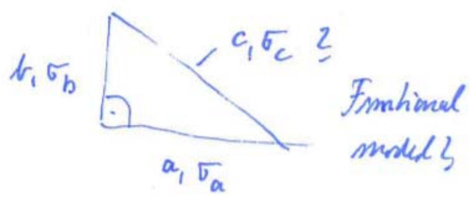
4.4.1 General case: n correlated observations, u correlated unknowns

$$\mathbf{\Sigma}_{ff_{u\times u}} = \mathbf{F}_{u\times n} \cdot \mathbf{\Sigma}_{ll_{n\times n}} \cdot \mathbf{F}^{\mathrm{T}}_{u\times n}$$

▶ Step 1: Set up the functional model for u functional relations $\mathbf{\Phi}(\mathbf{l})$ with n observations (measurements) l_i:

$$\begin{aligned} x_1 &= \Phi_1(\mathbf{l}) = \varphi_1(l_1, l_2, \ldots, l_n) \\ x_2 &= \Phi_2(\mathbf{l}) = \varphi_2(l_1, l_2, \ldots, l_n) \\ &\vdots \\ x_u &= \Phi_u(\mathbf{l}) = \varphi_u(l_1, l_2, \ldots, l_n) \end{aligned}$$

unknowns as functions of observations



$c = \sqrt{a^2 + b^2}$

4.4.1 General case: n correlated observations, u correlated unknowns

▶ Step 2: Set up the stochastic model for n observations:

$$\boldsymbol{\Sigma}_{LL_{n\times n}} = \begin{bmatrix} \sigma_{l_1}^2 & \sigma_{l_1 l_2} & \cdots & \sigma_{1n} \\ \sigma_{l_2 l_1} & \sigma_{l_2}^2 & \cdots & \sigma_{l_2 l_n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{l_n l_1} & \sigma_{l_n l_2} & \cdots & \sigma_{l_n}^2 \end{bmatrix}$$

<u>Question</u>: Functional model linear or nonlinear?

- if functional model linear, design matrix $\mathbf{F}$ directly given
- if functional model nonlinear, proceed with step 3

4.4.1 General case: n correlated observations, u correlated unknowns

▶ Step 3: Linearisation of the functional model

→ computation of partial derivatives $\frac{\partial \varphi_j}{\partial l_i}$ with $j = 1, 2, \dots, u$ and $i = 1, 2, \dots, n$

Arrange partial derivatives in Jacobian matrix $\mathbf{J}$

$$\mathbf{J}_{u \times n} = \begin{bmatrix} \frac{\partial \varphi_1}{\partial l_1} & \frac{\partial \varphi_1}{\partial l_2} & \cdots & \frac{\partial \varphi_1}{\partial l_n} \\ \frac{\partial \varphi_2}{\partial l_1} & \frac{\partial \varphi_2}{\partial l_2} & \cdots & \frac{\partial \varphi_2}{\partial l_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \varphi_u}{\partial l_1} & \frac{\partial \varphi_u}{\partial l_2} & \cdots & \frac{\partial \varphi_u}{\partial l_n} \end{bmatrix} = \begin{bmatrix} j_{11} & j_{12} & \cdots & j_{1n} \\ j_{21} & j_{22} & \cdots & j_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ j_{u1} & j_{u2} & \cdots & j_{un} \end{bmatrix}$$

→ Insert elements of $\mathbf{J}$ into matrix $\mathbf{F}$

4.4.1 General case: n correlated observations, u correlated unknowns

▶ Step 4: Computation of VCM for the unknowns

- With theoretical VCM $\mathbf{\Sigma}_{LL}$:

$$\boxed{\mathbf{\Sigma}_{xx} = \mathbf{F} \cdot \mathbf{\Sigma}_{LL} \cdot \mathbf{F}^{\mathrm{T}}}$$

$$\mathbf{\Sigma}_{xx} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1u} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2u} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{u1} & \sigma_{u2} & \cdots & \sigma_u^2 \end{bmatrix}$$

- Theoretical standard deviation of the unknowns

$$\sigma_i = \sqrt{\sigma_i^2} \qquad i = 1, \dots, u$$

- With empirical VCM $\mathbf{S}_{LL}$:

$$\boxed{\mathbf{S}_{xx} = \mathbf{F} \cdot \mathbf{S}_{LL} \cdot \mathbf{F}^{\mathrm{T}}}$$

$$\mathbf{S}_{xx} = \begin{bmatrix} s_1^2 & s_{12} & \cdots & s_{1u} \\ s_{21} & s_2^2 & \cdots & s_{2u} \\ \vdots & \vdots & \ddots & \vdots \\ s_{u1} & s_{u2} & \cdots & s_u^2 \end{bmatrix}$$

- Empirical standard deviation of the unknowns

$$s_i = \sqrt{s_i^2} \qquad i = 1, \dots, u$$

4.4.2 Special case: n uncorrelated observations, 1 unknown value

$$\mathbf{\Sigma}_{LL_{n\times n}} = \begin{bmatrix} \sigma_{l_1}^2 & & & 0 \\ & \sigma_{l_2}^2 & & \\ & & \ddots & \\ 0 & & & \sigma_{l_n}^2 \end{bmatrix}$$

$$x = \Phi(\mathbf{l}) = \varphi(l_1, l_2, \dots, l_n)$$

- Step 1:

  Set up the functional model for one functional relation $\Phi(\mathbf{l})$ with n observations (measurements) l_i

$$x = \Phi(\mathbf{l}) = \varphi(l_1, l_2, \dots, l_n)$$

4.4.2 Special case: n uncorrelated observations, 1 unknown value

▶ Step 2: Set up the stochastic model for n uncorrelated observations

$$\mathbf{\Sigma}_{LL} = \begin{bmatrix} \sigma_{l_1}^2 & & & 0 \\ & \sigma_{l_2}^2 & & \\ & & \ddots & \\ 0 & & & \sigma_{l_n}^2 \end{bmatrix}$$

Question: Functional model linear or nonlinear?

- if functional model linear, design matrix $\mathbf{F}$ directly given
- if functional model nonlinear, proceed with step 3

▶ Step 3: Linearisation of the functional model

→ computation of partial derivatives $\frac{\partial \varphi}{\partial l_i}$ $\quad i = 1, 2, \ldots, n$

$$\mathbf{J}_{1\times n} = \begin{bmatrix} \frac{\partial \varphi}{\partial l_1} & \frac{\partial \varphi}{\partial l_2} & \frac{\partial \varphi}{\partial l_3} & \cdots & \frac{\partial \varphi}{\partial l_n} \end{bmatrix}$$

→ Insert elements of $\mathbf{J}$ into matrix $\mathbf{F}$

4.4.2 Special case: n uncorrelated observations, 1 unknown value

▶ Step 4: Computation of „VCM“ for the unknowns

$$\boldsymbol{\Sigma}_{xx_{1\times 1}} = \begin{bmatrix} \frac{\partial\varphi}{\partial l_1} & \frac{\partial\varphi}{\partial l_2} & \frac{\partial\varphi}{\partial l_3} & \cdots & \frac{\partial\varphi}{\partial l_n} \end{bmatrix}_{1\times n} \begin{bmatrix} \sigma_{l_1}^2 & & & 0 \\ & \sigma_{l_2}^2 & & \\ & & \ddots & \\ 0 & & & \sigma_{l_n}^2 \end{bmatrix}_{n\times n} \begin{bmatrix} \frac{\partial\varphi}{\partial l_1} \\ \frac{\partial\varphi}{\partial l_1} \\ \vdots \\ \frac{\partial\varphi}{\partial l_n} \end{bmatrix}_{n\times 1}$$

- Theoretical variance:

$$\sigma_x^2 = \left(\frac{\partial\varphi}{\partial l_1}\right)^2 \cdot \sigma_{l_1}^2 + \left(\frac{\partial\varphi}{\partial l_2}\right)^2 \cdot \sigma_{l_2}^2 + \ldots + \left(\frac{\partial\varphi}{\partial l_n}\right)^2 \cdot \sigma_{l_n}^2$$

- Empirical variance:

$$s_x^2 = \left(\frac{\partial\varphi}{\partial l_1}\right)^2 \cdot s_{l_1}^2 + \left(\frac{\partial\varphi}{\partial l_2}\right)^2 \cdot s_{l_2}^2 + \ldots + \left(\frac{\partial\varphi}{\partial l_n}\right)^2 \cdot s_{l_n}^2$$

Variance propagation for uncorrelated observations

4.4.2 Special case: n uncorrelated observations, 1 unknown value

- Theoretical standard deviation:

$$\sigma_x = \sqrt{\sigma_x^2}$$

- Empirical standard deviation:

$$s_x = \sqrt{s_x^2}$$

→ σ_x respective s_x can easily be computed without applying matrix calculus

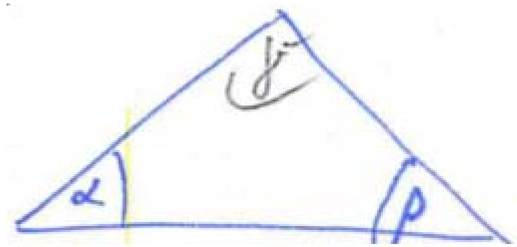
Please note: General case includes the special case, but the special case does not include the general case.

4.4.3 Variance – Covariance propagation: Some examples

Example 1:

▶ Angles α and β have been measured with a standard deviation of

$$s_\alpha = s_\beta = 1 \text{ mgon}$$



▶ Compute the standard deviation s_γ

▶ General or special case?

▶ Step 1: Functional model $\gamma = 200 \text{ gon} - \alpha - \beta$

$$\underbrace{\gamma}_{\mathbf{f}} = \underbrace{\begin{array}{c} \begin{array}{cc} \alpha & \beta \end{array} \\ \begin{bmatrix} -1 & -1 \end{bmatrix} \\ \begin{array}{cc} f_{11} & f_{12} \end{array} \end{array}}_{\mathbf{F}} \underbrace{\begin{bmatrix} \alpha \\ \beta \end{bmatrix}}_{\mathbf{l}} + 200 \text{ gon}$$

4.4.3 Variance – Covariance propagation: Some examples

- Step 2: Stochastic model

$$\mathbf{S}_{LL} = \begin{bmatrix} (1\ \text{mgon})^2 & 0 \\ 0 & (1\ \text{mgon})^2 \end{bmatrix}$$

- Steps 3 and 4:

$$s_\gamma^2 = (f_{11})^2 \cdot s_\alpha^2 + (f_{12})^2 \cdot s_\beta^2$$

with $f_{11} = -1$ and $f_{12} = -1$

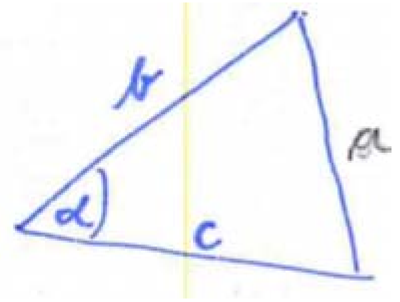
$$\Rightarrow s_\gamma^2 = 1 \cdot (1\ \text{mgon})^2 + 1 \cdot (1\ \text{mgon})^2$$

$$s_\gamma = \sqrt{2}\ \text{mgon}$$

$$s_\gamma = 1.41\ \text{mgon}$$

4.4.3 Variance – Covariance propagation: Some examples

Example 2:



- Given: $b = 20.29$ m, $s_b = 1.0$ cm

  $c = 75.75$ m, $s_c = 2.0$ cm

  $\alpha = 27.292$ gon, $s_\alpha = 0.015$ gon

- Wanted:

$$a, s_a$$

- General or special case?

- Step 1: Functional model

$$a = \sqrt{b^2 + c^2 - 2bc \cdot \cos\alpha} \quad \rightarrow \quad \text{non-linear}$$

4.4.3 Variance – Covariance propagation: Some examples

▶ Step 2: Stochastic model

$$\mathbf{S}_{LL} = \begin{bmatrix} (s_b)^2 & & 0 \\ & (s_c)^2 & \\ 0 & & (s_\alpha)^2 \end{bmatrix} = \begin{bmatrix} (0.01)^2 & & 0 \\ & (0.02)^2 & \\ 0 & & \left(\frac{0.015}{\rho}\right)^2 \end{bmatrix}$$

with $\rho = \frac{200}{\pi}$

- Attention: Units must coincide!
- Solution: Choose $[\mathrm{m}]$ and $[\mathrm{rad}]$ in $\mathbf{S}_{LL}$

4.4.3 Variance – Covariance propagation: Some examples

$$a = \sqrt{b^2 + c^2 - 2bc \cdot \cos\alpha}$$

- Steps 3 and 4:

$$s_a^2 = \left(\frac{\partial a}{\partial b}\right)^2 \cdot (s_b)^2 + \left(\frac{\partial a}{\partial c}\right)^2 \cdot (s_c)^2 + \left(\frac{\partial a}{\partial \alpha}\right)^2 \cdot (s_\alpha)^2$$

with:

$$\frac{\partial a}{\partial b} = \frac{1}{2\sqrt{}} \cdot (2b - 2c \cdot \cos\alpha) = \frac{b - c \cdot \cos\alpha}{\sqrt{}} = \frac{b - c \cdot \cos\alpha}{a}$$

$$\frac{\partial a}{\partial c} = \frac{1}{2\sqrt{}} \cdot (2c - 2b \cdot \cos\alpha) = \frac{c - b \cdot \cos\alpha}{a}$$

$$\frac{\partial a}{\partial \alpha} = \frac{1}{2\sqrt{}} \cdot (2bc \cdot \sin\alpha) = \frac{bc \cdot \sin\alpha}{\sqrt{}} = \frac{bc \cdot \sin\alpha}{a}$$

$$s_a^2 = \left(\frac{b - c \cdot \cos\alpha}{a}\right)^2 \cdot (s_b)^2 + \left(\frac{c - b \cdot \cos\alpha}{a}\right)^2 \cdot (s_c)^2 + \left(\frac{bc \cdot \sin\alpha}{a}\right)^2 \cdot (s_\alpha)^2$$

$$\Rightarrow s_a = 0.022\ \mathrm{m}$$

4.4.3 Variance – Covariance propagation: Some examples

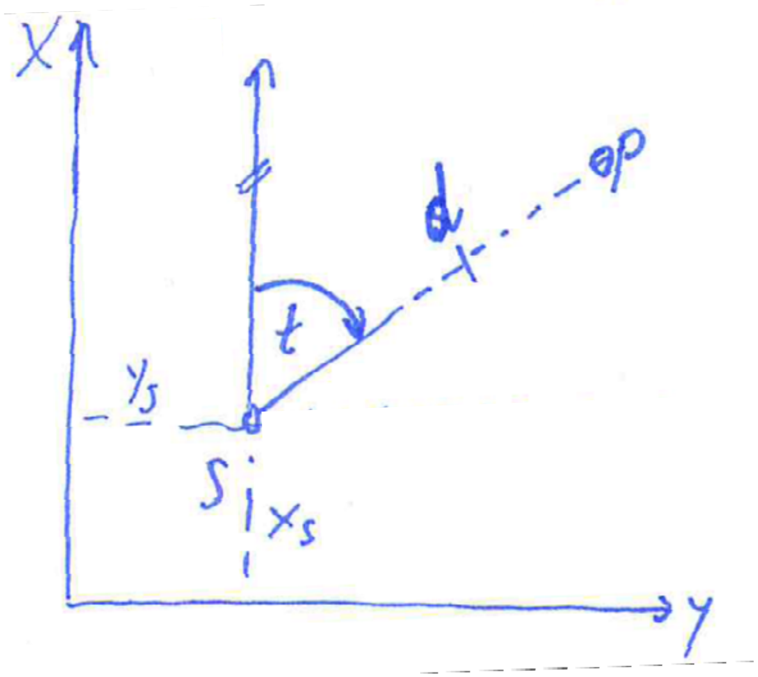
Example 3: Computation of Cartesian Coordinates

- Given:

  $Y_S = 1000.000 \text{ m}, X_S = 1000.000 \text{ m}$
  fixed values (error-free)

  $t = 77.1234 \text{ gon}, \sigma_t = 3.0 \text{ mgon}$

  $d = 987.654 \text{ m}, \quad \sigma_d = 20.0 \text{ mm}$



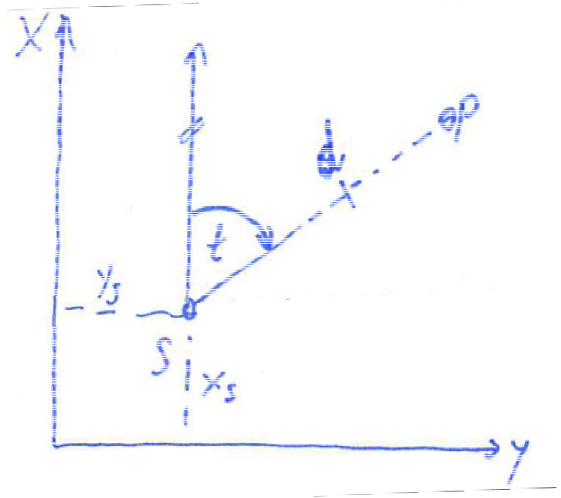
- Wanted:

  Y_P, X_P and $\sigma_{Y_P}, \sigma_{X_P}$ and ρ_{Y_P,X_P}

- General or special case?

4.4.3 Variance – Covariance propagation: Some examples

- Step 1: Functional model



$$\begin{bmatrix} Y_p \\ X_p \end{bmatrix} = \begin{bmatrix} \Phi_1(\mathbf{l}) \\ \Phi_2(\mathbf{l}) \end{bmatrix} = \begin{bmatrix} \varphi_1(d,t) \\ \varphi_2(d,t) \end{bmatrix}$$

$$\begin{bmatrix} Y_P \\ X_P \end{bmatrix} = \begin{bmatrix} Y_S + d \cdot \sin t \\ X_S + d \cdot \cos t \end{bmatrix} = \begin{bmatrix} 1924.5700 \text{ m} \\ 1347.3193 \text{ m} \end{bmatrix}$$

- Step 2:
  - Observation vector

$$\mathbf{l} = \begin{bmatrix} d \\ t \end{bmatrix} = \begin{bmatrix} 987.654 \text{ m} \\ 77.1234 \text{ gon} \end{bmatrix}$$

  - Stochastic model

$$\mathbf{\Sigma}_{LL} = \begin{bmatrix} (\sigma_\alpha)^2 & 0 \\ 0 & (\sigma_t)^2 \end{bmatrix} = \begin{bmatrix} (0.02)^2 & 0 \\ 0 & \left(\dfrac{0.003}{\rho}\right)^2 \end{bmatrix}$$

4.4.3 Variance – Covariance propagation: Some examples

▶ Step 3: Linearisation, computation of partial derivatives

$$\frac{\partial Y_P}{\partial d} = \sin t\,; \quad \frac{\partial Y_P}{\partial t} = d \cdot \cos t$$

$$\begin{bmatrix} Y_P \\ X_P \end{bmatrix} = \begin{bmatrix} Y_S + d \cdot \sin t \\ X_S + d \cdot \cos t \end{bmatrix}$$

$$\frac{\partial X_P}{\partial d} = \cos t; \quad \frac{\partial X_P}{\partial t} = -d \cdot \sin t$$

Design matrix $\mathbf{F} = \mathbf{J}$:

$$\mathbf{J} = \begin{array}{c c} & \begin{array}{cc} d \quad & \quad t \end{array} \\ \begin{array}{c} Y_P \\ \\ X_P \end{array} & \begin{bmatrix} \dfrac{\partial Y_P}{\partial d} & \dfrac{\partial Y_P}{\partial t} \\ \dfrac{\partial X_P}{\partial d} & \dfrac{\partial X_P}{\partial t} \end{bmatrix} \end{array} = \begin{bmatrix} \sin t & d \cdot \cos t \\ \cos t & -d \cdot \sin t \end{bmatrix}$$

4.4.3 Variance – Covariance propagation: Some examples

▶ Step 4: VCM of the unknowns

$$\mathbf{\Sigma}_{xx} = \mathbf{F} \cdot \mathbf{\Sigma}_{LL} \cdot \mathbf{F}^{\mathrm{T}}$$

$$\ldots$$

$$\mathbf{\Sigma}_{xx} = \begin{bmatrix} 618.4138 & -581.4212 \\ -581.4212 & 1947.7531 \end{bmatrix}_{[\mathrm{mm}^2]}$$

$$\sigma_{Y_P} = \sqrt{618.4138} = 24.8679 \text{ mm}$$

$$\sigma_{X_P} = \sqrt{1947.7531} = 44.1334 \text{ mm}$$

$$\rho_{Y_P,X_P} = \frac{-581.4212}{24.8679 \cdot 44.1334} = -0.53$$

4.4.4 Variance – Covariance propagation for complex functional relationships

Technische Universität Berlin

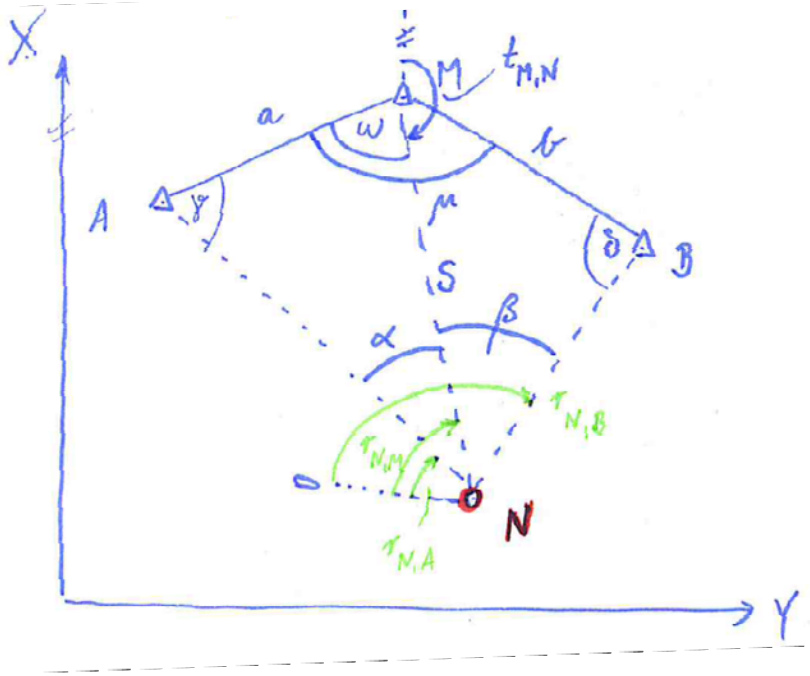
Example: Resection

▶ Given:

- Coordinates of points A, M, B as fixed values (error-free)
- Measurements: Horizontal directions $r_{N,A}, r_{N,M}, r_{N,B}$ with their standard deviations $\sigma_{r_{N,A}}, \sigma_{r_{N,M}}, \sigma_{r_{N,B}}$

▶ Wanted:

- Coordinates Y_N, X_N
- Standard deviations $\sigma_{Y_N}, \sigma_{X_N}$
- Correlation ρ_{Y_N,X_N}

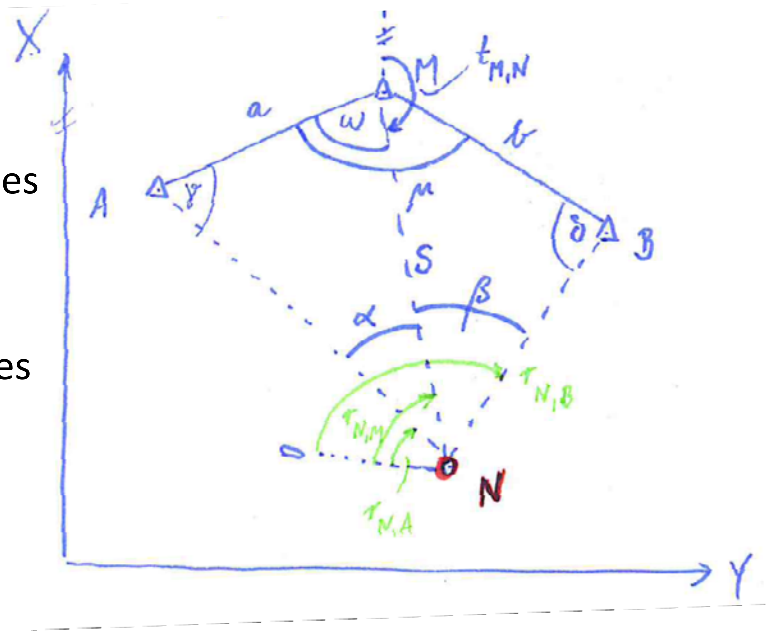


4.4.4 Variance – Covariance propagation for complex functional relationships

▶ Functional relationships:

1. $\alpha = r_{N,M} - r_{N,A}$
2. $\beta = r_{N,B} - r_{N,M}$
3. $\varepsilon = 2\pi - \alpha - \beta - \mu$ — computed from coordinates $(\varepsilon = \gamma + \delta)$
4. $a_1 = \dfrac{\sin\alpha}{a}$
5. $a_2 = \dfrac{\sin\beta}{b}$ — distances computed from coordinates
6. $\gamma = \arctan\dfrac{a_1 \cdot \sin\varepsilon}{a_2 + a_1 \cdot \cos\varepsilon}$
7. $s = \dfrac{\sin\gamma}{a_1}$
8. $t_{M,N} = t_{M,A} - \pi + \alpha + \gamma$ — computed from coordinates $\quad (t_{M,N} = t_{M,A} - \omega)$
9. $\boxed{Y_N = Y_M + s \cdot \sin t_{M,N}}$
10. $\boxed{X_N = X_M + s \cdot \cos t_{M,N}}$

Computation of Cartesian Coordinates



4.4.4 Variance – Covariance propagation for complex functional relationships

- Question: How can we compute $\sigma_{Y_N}, \sigma_{X_N}$?
- First idea: Insert formulas into one another to obtain a formula for Y_N resp. X_N that depends only on the original measurements $r_{N,A}, r_{N,M}, r_{N,B}$
- We obtain:

$$Y_N = \varphi_1(r_{N,A}, r_{N,M}, r_{N,B})$$
$$X_N = \varphi_2(r_{N,A}, r_{N,M}, r_{N,B})$$

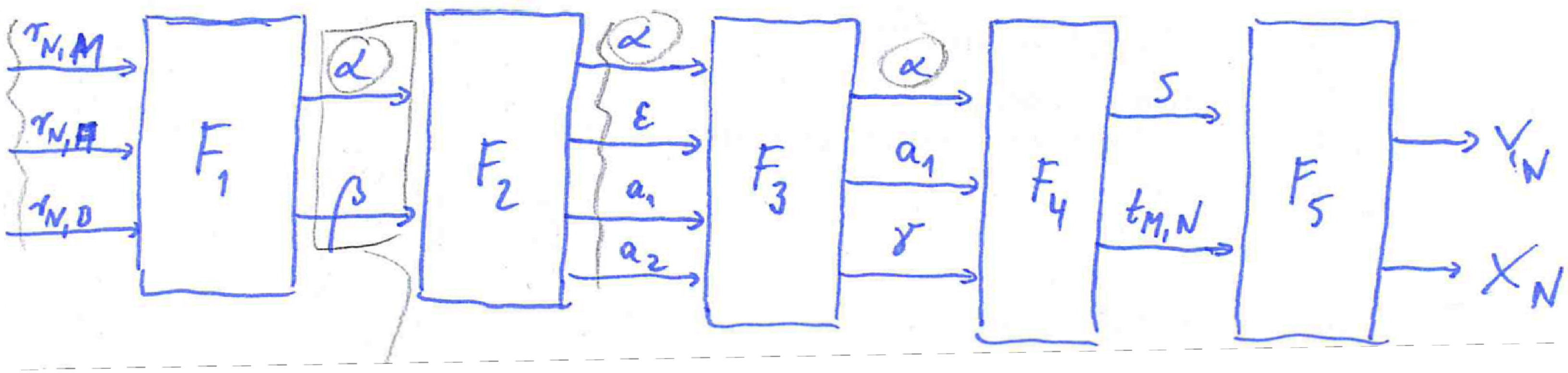
- Problem 1: We obtain very complicated equations, very complex part. derivatives
- Problem 2: Y_N, X_N are correlated! If we want to use Y_N, X_N for further computation, we need the full VCM → We have to apply the general case of VC propagation
- Solution: We keep the decomposition in partial steps and apply chain rule of matrix multiplication

Remark: If (and only if) we are only interested in $\sigma_{Y_N}, \sigma_{X_N}$ and not in the correlation ρ_{Y_N,X_N} we can compute $\sigma_{Y_N}, \sigma_{X_N}$ separately by applying the special case two times

4.4.4 Variance – Covariance propagation for complex functional relationships

We have to consider ten functional relationships

Flowchart



VC propagation: $\boldsymbol{\Sigma}_{xx} = \mathbf{F} \cdot \boldsymbol{\Sigma}_{LL} \cdot \mathbf{F}^{\mathrm{T}}$

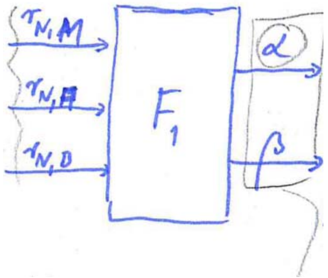
with $\mathbf{F} = \mathbf{F}_k \cdot \mathbf{F}_{k-1} \cdot \mathbf{F}_{k-2} \cdot \dots \cdot \mathbf{F}_1$

in our example $\mathbf{F} = \mathbf{F}_5 \cdot \mathbf{F}_4 \cdot \mathbf{F}_3 \cdot \mathbf{F}_2 \cdot \mathbf{F}_1$

4.4.4 Variance – Covariance propagation for complex functional relationships

$\mathbf{F}_1$:

- Functional model:



$$\alpha = r_{N,M} - r_{N,A}$$
$$\beta = r_{N,B} - r_{N,M}$$

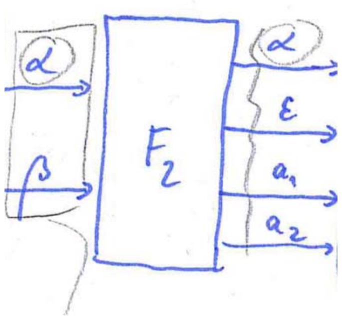
- Design matrix:

$$\mathbf{F}_1 = \begin{array}{c c} & \begin{array}{ccc} r_{N,M} & r_{N,A} & r_{N,B} \end{array} \\ \begin{array}{c} \alpha \\ \beta \end{array} & \begin{bmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \end{array}$$

4.4.4 Variance – Covariance propagation for complex functional relationships

$\mathbf{F}_2$:

- Functional model:



$$\alpha = \alpha$$

(α needed for further computation
→ we have to introduce "identity equation")

$$\varepsilon = 2\pi - \alpha - \beta - \mu$$

$$a_1 = \frac{\sin\alpha}{a}$$

$$a_2 = \frac{\sin\beta}{b}$$

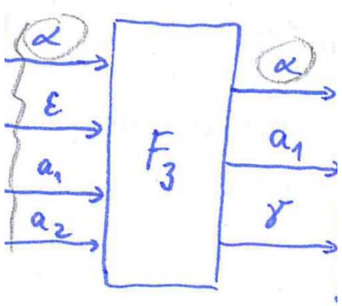
- Design matrix:

$$\mathbf{F}_2 = \begin{array}{c} \\ \alpha \\ \varepsilon \\ a_1 \\ a_2 \end{array} \begin{array}{c} \begin{array}{cc} \alpha & \beta \end{array} \\ \begin{bmatrix} 1 & 0 \\ -1 & -1 \\ \dfrac{\cos\alpha}{a} & 0 \\ 0 & \dfrac{\cos\beta}{b} \end{bmatrix} \end{array}$$

4.4.4 Variance – Covariance propagation for complex functional relationships

$\mathbf{F}_3$:

- Functional model:



$$\alpha = \alpha$$

$$a_1 = a_1$$

$$\gamma = \arctan \frac{a_1 \cdot \sin \varepsilon}{a_2 + a_1 \cdot \cos \varepsilon}$$

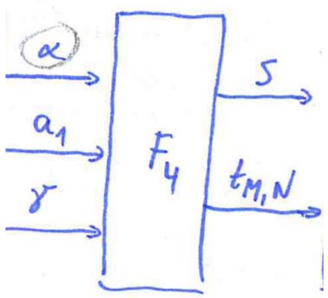
- Design matrix:

$$\mathbf{F}_3 = \begin{array}{c} \\ \alpha \\ a_1 \\ \gamma \end{array} \begin{array}{cccc} \alpha & \varepsilon & a_1 & a_2 \\ \multicolumn{4}{c}{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \dfrac{\partial \gamma}{\partial \varepsilon} & \dfrac{\partial \gamma}{\partial a_1} & \dfrac{\partial \gamma}{\partial a_2} \end{bmatrix}} \end{array}$$

4.4.4 Variance – Covariance propagation for complex functional relationships

$\mathbf{F}_4$:

- Functional model:



$$s = \frac{\sin\gamma}{a_1}$$

$$t_{M,N} = t_{M,A} - \pi + \alpha + \gamma$$

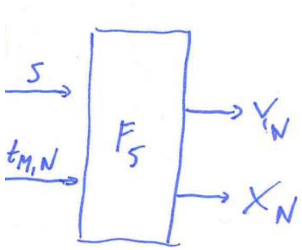
- Design matrix:

$$\mathbf{F}_4 = \begin{array}{c} \\ s \\ t_{M,N} \end{array} \begin{array}{c} \begin{array}{ccc} \alpha & a_1 & \gamma \end{array} \\ \begin{bmatrix} 0 & -\dfrac{\sin\gamma}{(a_1)^2} & \dfrac{\cos\gamma}{a_1} \\ 1 & 0 & 1 \end{bmatrix} \end{array}$$

4.4.4 Variance – Covariance propagation for complex functional relationships

$\mathbf{F}_5$:

- Functional model:



$$Y_N = Y_M + s \cdot \sin t_{M,N}$$

$$X_N = X_M + s \cdot \cos t_{M,N}$$

- Design matrix:

$$\mathbf{F}_5 = \begin{array}{c c} & \begin{array}{cc} \quad s \quad & \quad t_{M,N} \quad \end{array} \\ \begin{array}{c} Y_N \\ X_N \end{array} & \begin{bmatrix} \sin t_{M,N} & s \cdot \cos t_{M,N} \\ \cos t_{M,N} & -s \cdot \sin t_{M,N} \end{bmatrix} \end{array}$$

4.4.4 Variance – Covariance propagation for complex functional relationships

VC propagation:

$$\mathbf{F} = \mathbf{F}_5 \cdot \mathbf{F}_4 \cdot \mathbf{F}_3 \cdot \mathbf{F}_2 \cdot \mathbf{F}_1$$

$$\mathbf{\Sigma}_{xx} = \mathbf{F} \cdot \mathbf{\Sigma}_{LL} \cdot \mathbf{F}^{\mathrm{T}}$$

with $$\mathbf{\Sigma}_{LL} = \begin{bmatrix} \left(\sigma_{r_{N,M}}\right)^2 & 0 & 0 \\ 0 & \left(\sigma_{r_{N,A}}\right)^2 & 0 \\ 0 & 0 & \left(\sigma_{r_{N,B}}\right)^2 \end{bmatrix}$$

→ Attention: Same order as columns in $\mathbf{F}_1$

Result:

$$\mathbf{\Sigma}_{xx_{2\times 2}} = \begin{bmatrix} \left(\sigma_{Y_N}\right)^2 & \mathrm{Cov}(Y_N, X_N) \\ \mathrm{Cov}(X_N, Y_N) & \left(\sigma_{X_N}\right)^2 \end{bmatrix}$$

Correlation: $$\rho_{Y_N,X_N} = \frac{\mathrm{Cov}(Y_N,X_N)}{\sigma_{Y_N} \cdot \sigma_{X_N}}$$

4.4.5 Variance – Covariance propagation: Standard cases that often occur in practice

Technische Universität Berlin

Scaling of an observation with a constant factor

- Given: measurement l, standard deviation s_l
- Functional model:

$$f(l) = a \cdot l \qquad \text{with} \qquad a = \text{const.}$$

- Coefficient:

$$f_l = a$$

- VC propagation:

$$s_f^2 = f_l^2 \cdot s_l^2 \quad \Rightarrow \quad s_f^2 = a^2 \cdot s_l^2$$

Standard deviation $\quad s_f = \sqrt{s_f^2} \quad \Rightarrow \quad \boxed{s_f = a \cdot s_l}$

4.4.5 Variance – Covariance propagation: Standard cases that often occur in practice

Addition or subtraction of observations

- Given:

n uncorrelated measurements $l_1, l_2, \ldots, l_n$ with their standard deviations $s_{l_1}, s_{l_2}, \ldots, s_{l_n}$

- Functional model:

$$f(l_1, l_2, \ldots, l_n) = l_1 \pm l_2 \pm \cdots \pm l_n$$

- Coefficients:

$$f_{l_1} = +1, \qquad f_{l_2} = \pm 1, \qquad \ldots, \qquad f_{l_n} = \pm 1$$

- VC propagation:

$$s_f^2 = s_{l_1}^2 + s_{l_2}^2 + \cdots + s_{l_n}^2 \qquad \text{(for addition and subtraction!)}$$

Standard deviation $\boxed{s_f = \sqrt{s_{l_1}^2 + s_{l_2}^2 + \cdots + s_{l_n}^2}}$

4.4.5 Variance – Covariance propagation: Standard cases that often occur in practice

▶ In the case that all observations have the same precision

→ Same standard deviation s_l for all observations l_i

$$s_l = s_{l_1} = s_{l_2} = \cdots = s_{l_n}$$

→ Standard deviation

$$s_f = \sqrt{s_l^2 + s_l^2 + \cdots + s_l^2}$$

$$s_f = \sqrt{n \cdot s_l^2}$$

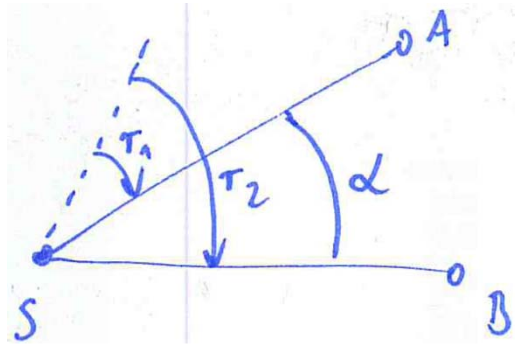
$$\boxed{s_f = \sqrt{n} \cdot s_l}$$

<u>Rule</u>: Standard deviation grows with the square root if the number of summands.

4.4.5 Variance – Covariance propagation: Standard cases that often occur in practice

Technische Universität Berlin

Example:



▶ Given:

- Horizontal directions r_1, r_2 (observations)
- Standard deviations $s_{r_1} = 1 \text{ mgon}, s_{r_2} = 1 \text{ mgon}$

▶ Wanted:

- Horizontal angle α
- Standard deviation s_α

▶ Solution:

$$s_\alpha = \sqrt{2} \cdot 1 \text{ mgon}$$
$$s_\alpha = 1.41 \text{ mgon}$$

4.4.5 Variance – Covariance propagation: Standard cases that often occur in practice

Arithmetic mean of observations

▶ Given:

- n uncorrelated measurements $l_1, l_2, \dots, l_n$
- Standard deviations $s_{l_1}, s_{l_2}, \dots, s_{l_n}$

▶ Functional model:

$$\bar{x} = f(l_1, l_2, \dots, l_n) = \frac{l_1 + l_2 + \cdots + l_n}{n} = \frac{1}{n} l_1 + \frac{1}{n} l_2 + \cdots + \frac{1}{n} l_n$$

▶ Coefficients:

$$f_{l_1} = \frac{1}{n} \;\;,\;\; f_{l_2} = \frac{1}{n} \;\;,\;\; \dots,\;\; f_{l_n} = \frac{1}{n}$$

4.4.5 Variance – Covariance propagation: Standard cases that often occur in practice

▶ VC propagation:

$$s_{\bar{x}}^2 = \left(\frac{1}{n}\right)^2 s_{l_1}^2 + \left(\frac{1}{n}\right)^2 s_{l_2}^2 + \cdots + \left(\frac{1}{n}\right)^2 s_{l_n}^2 = \frac{1}{n^2}(s_{l_1}^2 + s_{l_2}^2 + \cdots + s_{l_n}^2)$$

In the case that all observations have the same precision

→ same standard deviations s_l for all observations l_i

$$s_l = s_{l_1} = s_{l_2} = \cdots = s_{l_n}$$

$$s_{\bar{x}}^2 = \frac{1}{n^2}(s_l^2 + s_l^2 + \cdots + s_l^2)$$

$$s_{\bar{x}}^2 = \frac{1}{n^2} \cdot n \cdot s_l^2$$

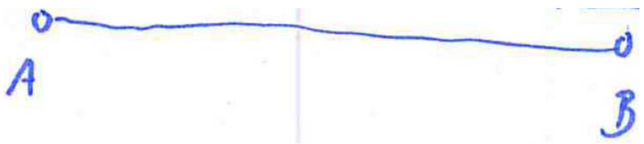
Standard deviation $\boxed{s_{\bar{x}} = \frac{s_l}{\sqrt{n}}}$

Rule: If we perform measurements for the same random variable n times, the standard deviation for the arithmetic mean decreases by the factor $1/\sqrt{n}$

4.4.5 Variance – Covariance propagation: Standard cases that often occur in practice

Technische Universität Berlin

Example:



▶ Given:

- Distance $A - B$ has been measured 10 times
- Standard deviation $s_l = 1$ cm

▶ Wanted:

- Standard deviation of the arithmetic mean

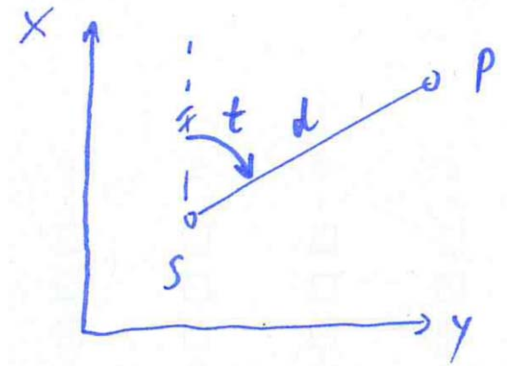
$$s_{\bar{x}} = 1\ \text{cm} \cdot \frac{1}{\sqrt{10}}$$
$$s_{\bar{x}} = 3.16\ \text{mm}$$

4.4.5 Variance – Covariance propagation: Standard cases that often occur in practice

Technische Universität Berlin

Some more remarks:

► $$\left.\begin{aligned} Y_P &= Y_s + d \cdot \sin t \\ X_P &= X_s + d \cdot \cos t \end{aligned}\right\}$$ General case, because we have two correlated unknowns



► If (and only if) we are interested in s_{Y_P}, s_{X_P} and not in the correlation r_{Y_P,X_P}, we can compute s_{Y_P}, s_{X_P} separately by applying the special case two times:

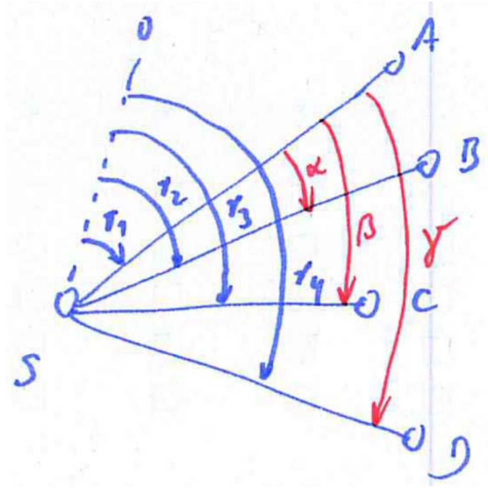
$$s_{Y_P}^2 = \left(\frac{\partial Y_P}{\partial d}\right)^2 \cdot s_d^2 + \left(\frac{\partial Y_P}{\partial t}\right)^2 \cdot s_t^2$$

$$s_{X_P}^2 = \left(\frac{\partial X_P}{\partial d}\right)^2 \cdot s_d^2 + \left(\frac{\partial X_P}{\partial t}\right)^2 \cdot s_t^2$$

► But: If we use our results (here Y_P, X_P) for further computations, we must consider the covariances and therefore apply the general case of VC propagation

4.4.5 Variance – Covariance propagation: Standard cases that often occur in practice

Example:



▶ Given:

- Horizontal directions r_1, r_2, r_3, r_4 (observations)
- Standard deviations $\sigma_{r_1} = \sigma_{r_2} = \sigma_{r_3} = \sigma_{r_4} = 0.5 \text{ mgon}$

▶ Wanted:

- Horizontal angles α, β, γ
- Standard deviations $\sigma_\alpha, \sigma_\beta, \sigma_\gamma$

▶ If we are only interested in $\sigma_\alpha, \sigma_\beta, \sigma_\gamma$ as a final result, we can apply the special case of VC propagation:

$$\alpha = r_2 - r_1 \ \Rightarrow \ \sigma_\alpha = \sqrt{2} \cdot 0.5 \text{ mgon} = 0.71 \text{ mgon}$$

$$\beta = r_3 - r_1 \ \Rightarrow \ \sigma_\beta = 0.71 \text{ mgon}$$

$$\gamma = r_4 - r_1 \ \Rightarrow \ \sigma_\gamma = 0.71 \text{ mgon}$$

4.4.5 Variance – Covariance propagation: Standard cases that often occur in practice

▶ If we want to use α, β, γ for further computation (which usually is the case), we have to apply the general case of VC propagation:

- Step 1: Functional model

$$\mathbf{X} = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} \Phi_1(\mathbf{l}) \\ \Phi_2(\mathbf{l}) \\ \Phi_3(\mathbf{l}) \end{bmatrix} = \begin{bmatrix} r_2 - r_1 \\ r_3 - r_1 \\ r_4 - r_1 \end{bmatrix}$$

- Step 2: Observation vector and stochastic model

$$\mathbf{l} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{bmatrix}, \qquad \mathbf{\Sigma}_{LL} = \begin{bmatrix} 0.5^2 & & 0 & \\ & 0.5^2 & & \\ & 0 & 0.5^2 & \\ & & & 0.5^2 \end{bmatrix}$$

4.4.5 Variance – Covariance propagation: Standard cases that often occur in practice

- Step 3: Design matrix (coefficients of the linear functional model)

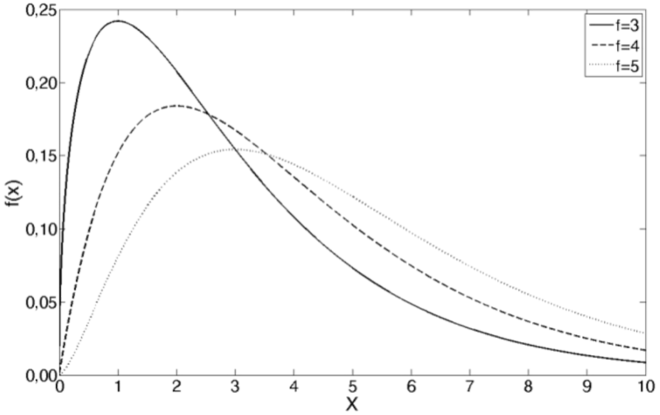
$$\mathbf{F} = \begin{array}{c} \\ \alpha \\ \beta \\ \gamma \end{array} \begin{array}{c} \begin{array}{cccc} r_1 & r_2 & r_3 & r_4 \end{array} \\ \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

- Step 4: VCM of the unknowns

$$\mathbf{\Sigma}_{xx} = \mathbf{F} \cdot \mathbf{\Sigma}_{LL} \cdot \mathbf{F}^{\mathrm{T}} = \begin{bmatrix} 0.50 & 0.25 & 0.25 \\ 0.25 & 0.50 & 0.25 \\ 0.25 & 0.25 & 0.50 \end{bmatrix}$$

Complete matrix $\mathbf{\Sigma}_{xx}$ must be considered for further computations!

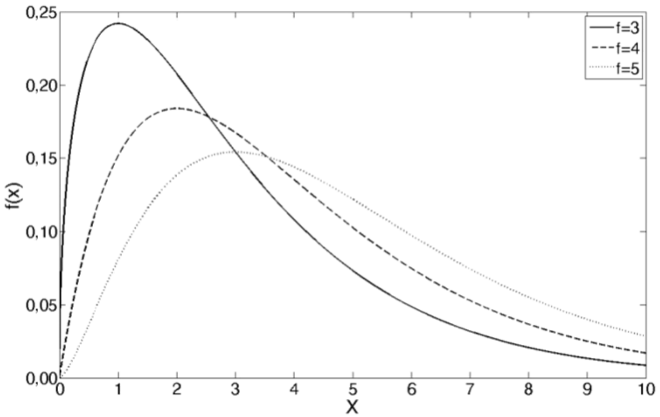
Standard deviations: $\sigma_\alpha = \sigma_\beta = \sigma_\gamma = \sqrt{0.5} = 0.71\ \mathrm{mgon}$



Adjustment Theory I

Chapter 4 – Propagation of Observation Errors

Prof. Dr.-Ing. Frank Neitzel | Institute of Geodesy and Geoinformation Science



Institute of Geodesy and Geoinformation Science

Chair of Geodesy and Adjustment Theory

Adjustment Theory I

Chapter 5: The Gaussian or Normal Distribution

Prof. Dr.-Ing. Frank Neitzel

Last update: 22 November 2024

Contents

5. The Gaussian or Normal Distribution

5.1 Introduction

– In previous sections we have discussed random variables with arbitrary distribution

– Now: we consider random variables with certain distribution functions

– Amongst all distribution functions the Normal Distribution plays a special role because random variables in geodesy (almost) follow this distribution

– For the derivation of normal distribution we apply „theory of elementary errors“.

5.2 Theory of Elementary Errors

Theory: Was developed by BESSEL and HAGEN in 1837

Random deviation ε of an observation L can be described by a sum of q very small „elementary errors" Δ_i.

$$\varepsilon = \Delta_1 + \Delta_2 + \Delta_3 + ... + \Delta_q \quad ; \qquad \varepsilon = \sum_{i=1}^{q} \Delta_i$$

under the following conditions:

– All values Δ_i have the same absolute value $\Delta_i = |\delta| \quad \forall_i$

– All values Δ_i differ only in their sign $\varepsilon = \pm\delta \pm \delta \pm \delta + ...$

– Positive and negative signs have same probability

$$P(\Delta_i = +\delta) = P(\Delta_i = -\delta) = 0.5$$

For $q \rightarrow \infty \;\Rightarrow\; \delta \rightarrow 0$

(for ∞ number of elementary errors the magnitude of the errors is almost zero)

Expectation $E\left(\Delta_h\right) = 0$

$$\Rightarrow \quad E\left(\varepsilon\right) = \sum_{h=1}^{q} E\left(\Delta_h\right) = 0 \qquad\qquad \sigma^2 = E\left(\varepsilon^2\right) = \sum_{h=1}^{q} E\left(\Delta_h^2\right) = q \cdot \delta^2$$

Examples: $\varepsilon = \Delta_1 + \Delta_2 + \Delta_3 + \ldots.. + \Delta_q$

Probability: $P_i(\varepsilon) = \frac{k_i}{\sum k_i}$

$\Rightarrow q = 1 \quad : \quad \varepsilon = \Delta_1$

Case number	Δ_1	ε	Frequency (k)	Relative Frequency = Probability P
1	+	$+\delta$	1	½ = 0.5
2	-	$-\delta$	1	½ = 0.5

$\Rightarrow q = 2 \quad : \quad \varepsilon = \Delta_1 + \Delta_2$

Case number	Δ_1	Δ_2	ε	Frequency (k)	Relative Frequency = Probability P
1	+	+	$+2\delta$	1	¼ = 0.25
2	+	-	0	2	2/4 = 0.5
3	-	+	0		
4	-	-	-2δ	1	¼ = 0.25

5.2 Theory of Elementary Errors

$$\Rightarrow \quad q = 4 \quad : \quad \varepsilon = \Delta_1 + \Delta_2 + \Delta_3 + \Delta_4$$

Case number	Δ_1 Δ_2 Δ_3 Δ_4	ε	Frequency (k)	Relative Frequency = Probability P
1	+ + + +	4δ	1	1/16 = 0.0625
2	+ + + -	2δ	4	4/16 = 0.25
3	+ + - +			
4	+ - + +			
5	- + + +			
6	+ + - -	0	6	6/16 = 0.375
7	+ - - +			
8	- - + +			
9	- + + -			
10	- + - +			
11	+ - + -			

$$\Rightarrow \; q = 4 \quad : \quad \varepsilon = \Delta_1 + \Delta_2 + \Delta_3 + \Delta_4$$

Case number	Δ_1 Δ_2 Δ_3 Δ_4	ε	Frequency (k)	Relative Frequency = Probability P
12	- - - +	-2δ	4	4/16 = 0.25
13	- - + -			
14	- + - -			
15	+ - - -			
16	- - - -	-4δ	1	1/16 = 0.0625

$\longrightarrow$ **The larger the deviation (“error”) the smaller the probability**

Pascal Triangle

$\varepsilon = \sum\Delta_i$	-6δ	-5δ	-4δ	-3δ	-2δ	$-\delta$	0	$+\delta$	$+2\delta$	$+3\delta$	$+4\delta$	$+5\delta$	$+6\delta$	$\sum k$
$q=1$						1		1						2
$q=2$					1		2		1					4
$q=3$				1		3		3		1				8
$q=4$			1		4		6		4		1			16
$q=5$		1		5		10		10		5		1		32
$q=6$	1		6		15		20		15		6		1	64

Formula for all cases $\boxed{\varepsilon = (q - 2 \cdot w) \cdot \delta}$ with $w = 0, 1, 2, \ldots, q$

$q = 1$

$w = 0$	$\varepsilon = (1 - 2 \cdot 0) \cdot \delta$	$+\delta$
$w = 1$	$\varepsilon = (1 - 2 \cdot 1) \cdot \delta$	$-\delta$

$q = 2$

$w = 0$	$\varepsilon = (2 - 2 \cdot 0) \cdot \delta$	$+2\delta$
$w = 1$	$\varepsilon = (2 - 2 \cdot 1) \cdot \delta$	0
$w = 2$	$\varepsilon = (2 - 2 \cdot 2) \cdot \delta$	-2δ

General Law

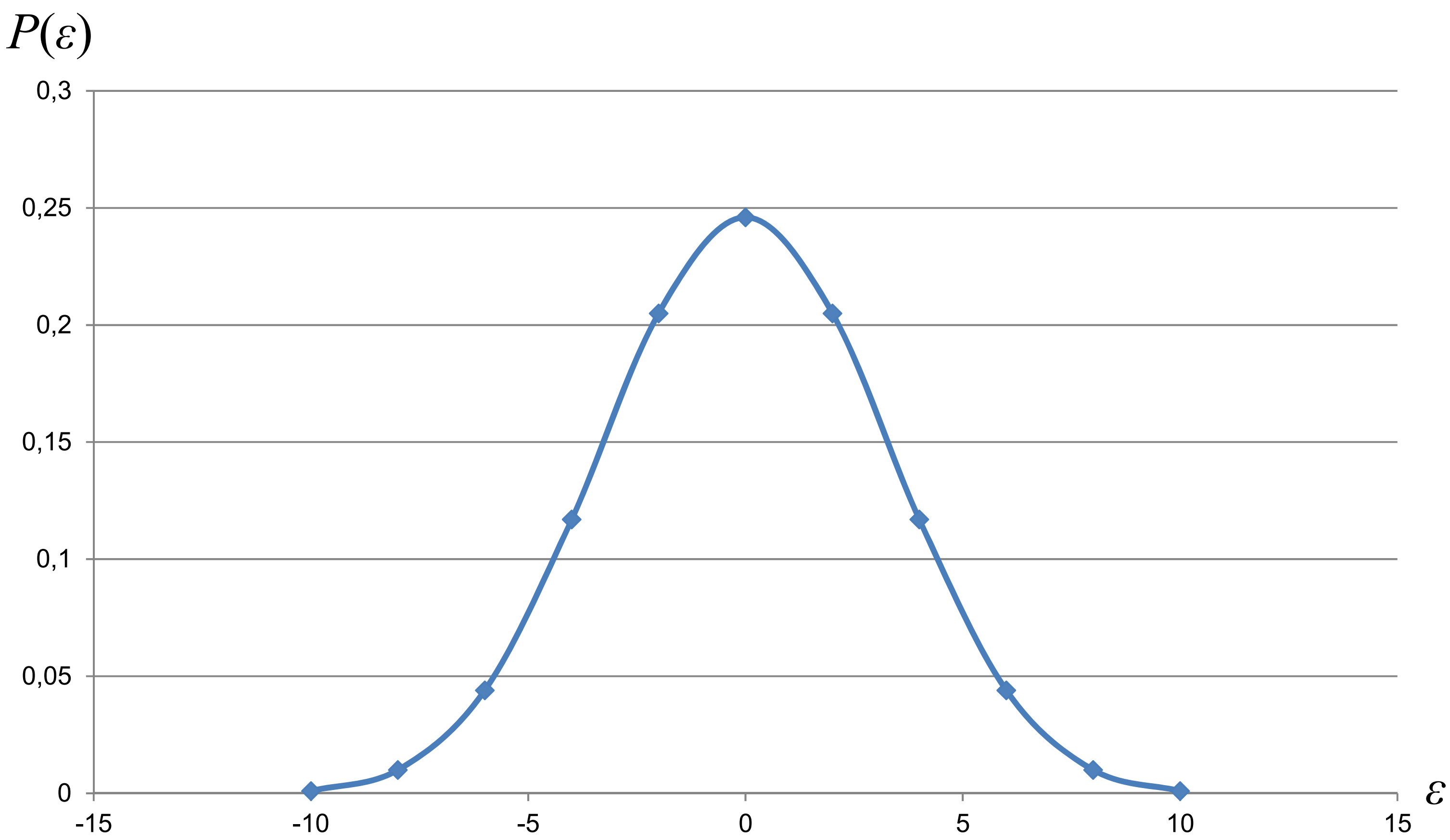
$$P(\varepsilon) = P\{\varepsilon = (q-2w)\delta\} = \frac{1}{2^q}\binom{q}{w}, \quad w = 0,1,2,\ldots,q \qquad \text{with} \quad \binom{q}{w} = \frac{q!}{(q-w)!\,w!}$$

Please note: $\sum_{w=0}^{q}\binom{q}{w} = 2^q$ and $\sum_{w=0}^{q} P(\varepsilon) = \frac{1}{2^q}\sum_{w=0}^{q}\binom{q}{w} = 1$

q = 10 We can now directly compute ε_i and $P(\varepsilon_i)$, w = 0, 1, 2, …, 10

w	0	1	2	3	4	5	6	7	8	9	10
ε	10δ	8δ	6δ	4δ	2δ	0	-2δ	-4δ	-6δ	-8δ	-10δ
$P(\varepsilon)$	0.001	0.010	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.010	0.001

Graphical representation for q = 10



Now: $q \rightarrow \infty \;\Rightarrow\; \delta \rightarrow 0$

$$P\left(\xi \leq \varepsilon < \xi + d\xi\right) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\xi^2}{2\sigma^2}\right) d\xi = f\left(\xi\right) d\xi$$

Definition

A random variable is normally distributed if it can be described by a sum of a large amount of “elementary errors” with (more or less) same magnitude.

It is: $\varepsilon = \Delta_1 + \Delta_2 + \ldots + \Delta_q$

With: ε random deviation

Δ_i elementary components of the deviation

$E(\Delta_i) = 0 \quad \forall_i$

$q \rightarrow \infty$

no Δ_i is dominating

Comparison with BESSEL and HAGEN:

– All values Δ_i have the same absolute value $\Delta_i = |\delta| \quad \forall_i$

– All values Δ_i differ only in their sign $\varepsilon = \pm\delta \pm \delta \pm \delta \pm \ldots$

– Positive and negative signs have same probability $P(\Delta_i = +\delta) = P(\Delta_i = -\delta) = 0.5$

For $q \to \infty \;\Rightarrow\; \delta \to 0$

$\Rightarrow$ Then it holds

$$\boxed{P(\xi \leq \varepsilon < \xi + d\xi) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\xi^2}{2 \cdot \sigma^2}\right) d\xi = f(\xi)\, d\xi}$$

with σ = constant value, to be defined (tbd)

$f(\xi)$ = Density of <u>normal distribution</u>

History of Central Limit Theorem (CLT)

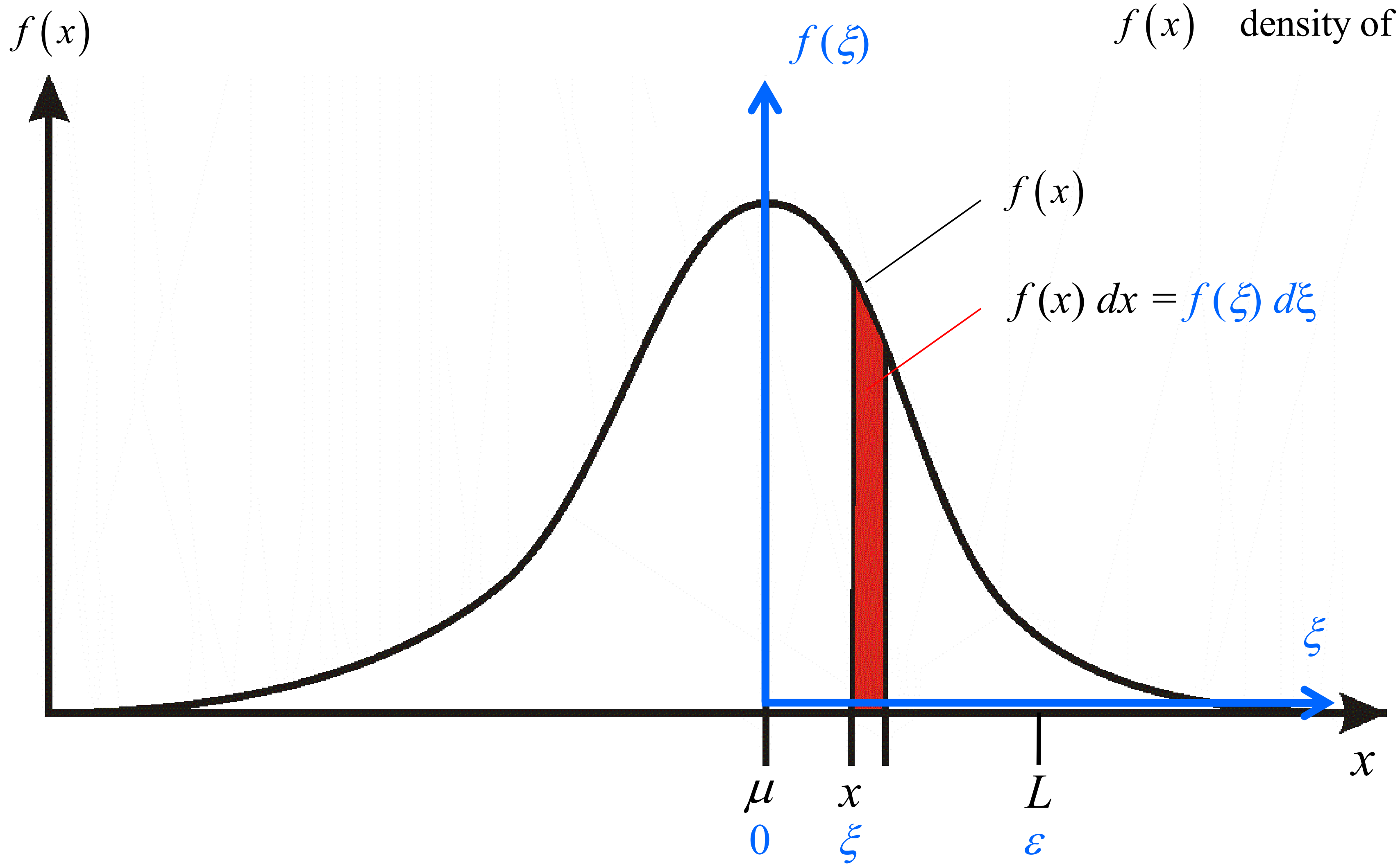
de Moivre 1730 Laplace 1892	First ideas, no proof
Gauss 1820	introduction into error theory (→ “Gauss Distribution”)
Lyapunov 1901	Proof!

In practice: If many “elementary errors” have (more or less) the same magnitude, we can ***assume*** a normal distribution.

Given: random variable $L = \mu + \varepsilon \rightarrow \varepsilon = L - \mu$ with: μ expectation $E(L)$

ε random deviation

$f(x)$ density of L



Shifting the coordinate system:

$$x = \mu + \xi \rightarrow \underbrace{\xi = x - \mu} \rightarrow d\xi = dx \rightarrow f(x) = f(\xi)$$

Substitution yields

$$P(\xi \le \varepsilon < \xi + d\xi) = \underbrace{f(\xi)}_{f(x)}\underbrace{d\xi}_{dx}$$

$$P(x - \cancel{\mu} \le L - \cancel{\mu} < x - \cancel{\mu} + dx)$$

$$P(\quad x \quad \le \quad L \quad < \quad x + dx \quad) = f(x)\,dx$$

Normal distribution from Central Limit Theorem: $f(\xi)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\frac{\xi^2}{\sigma^2}\right)$

Substitution $\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right)=f(x)$

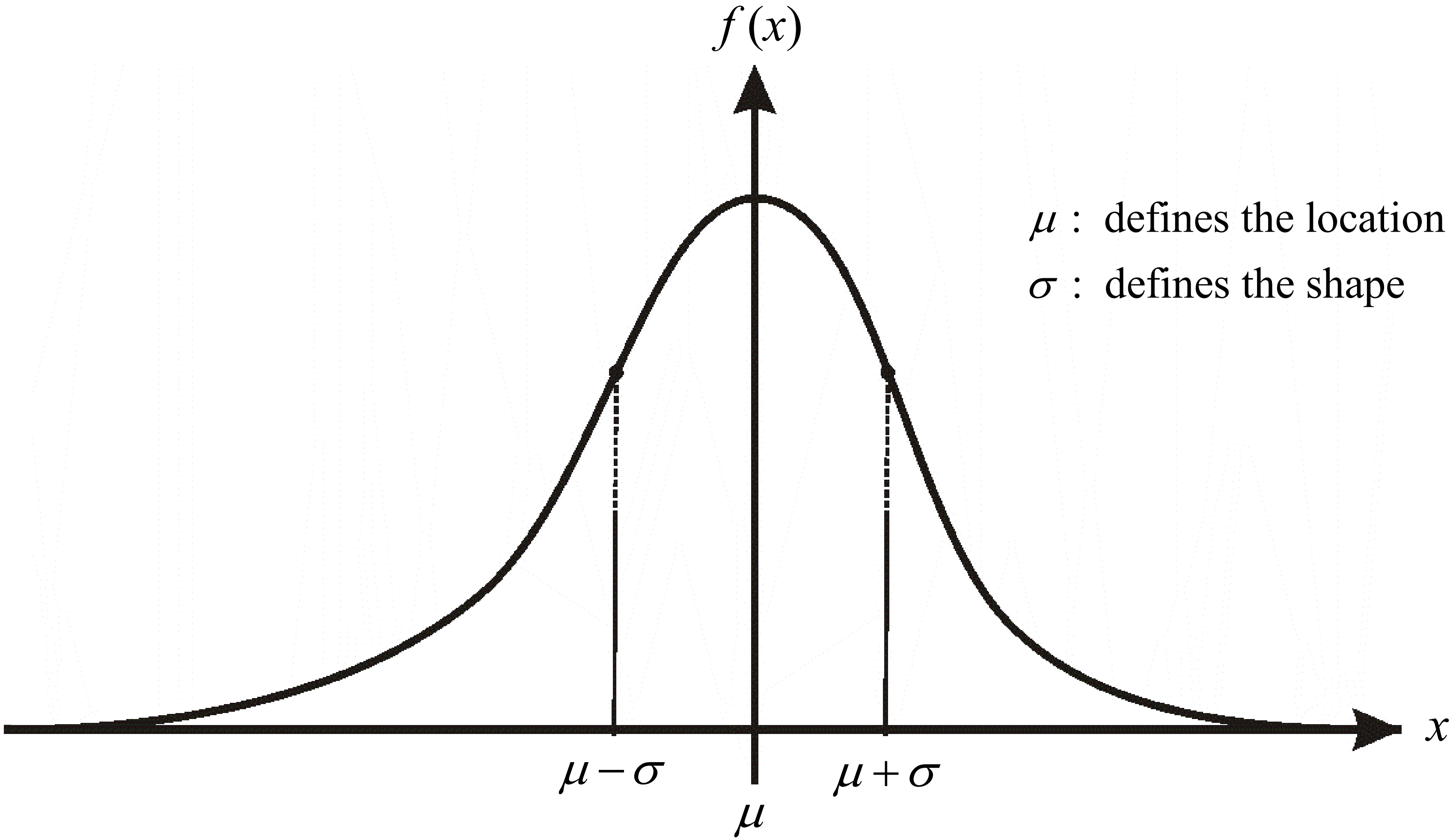
$$\boxed{P(x\le L<x+dx)=\underbrace{\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right)}_{f(x)}dx}$$

Probability density and thus the distribution function of a normally distributed random variable is completely defined by **two** parameters: Expectation μ and variance σ^2 !

$$\boxed{L\sim N\left(\mu,\sigma^2\right)}$$

"L follows the normal distribution with expectation μ and variance σ^2 "

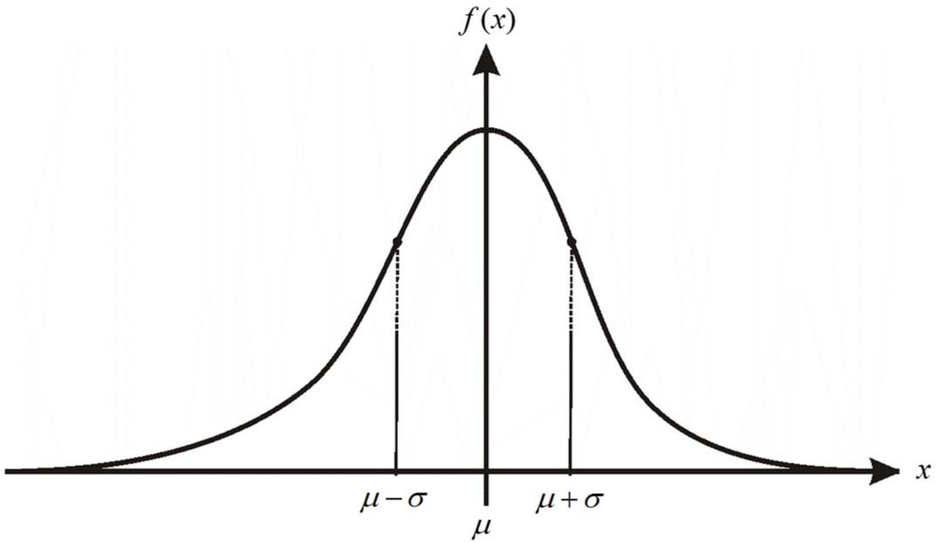
Normal Distribution Density Function



μ : defines the location

σ : defines the shape

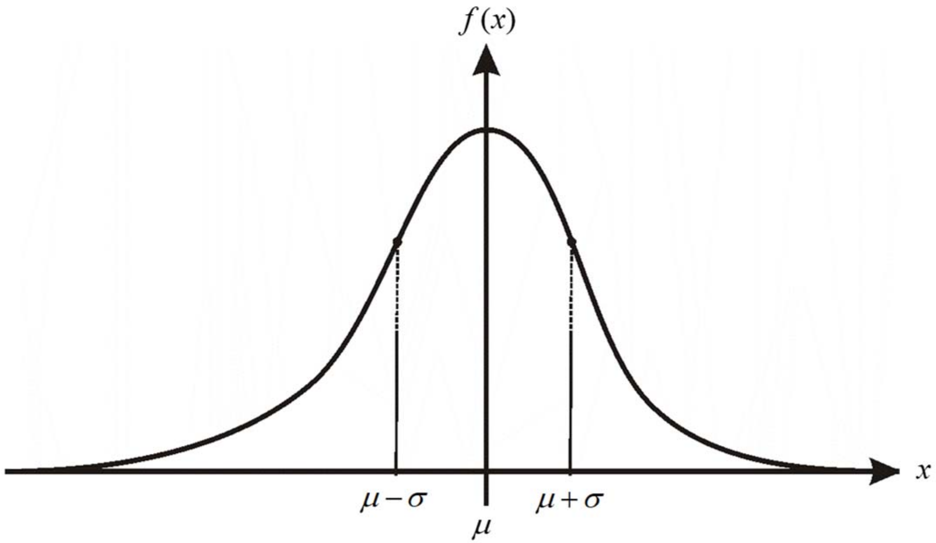
1. The normal density function $f(x)$ is symmetric about the expectation μ. Therefore all odd central moments are zero. Also the median and mode, which are two parameters of location sometimes used in practice, are equal to the expectation μ.
2. The maximum density value for the standardised variable is 0.399.
3. The density function approaches zero asymptotically as x goes to .
4. The density function has two points of inflection at $x = \mu \pm \sigma$



5. The probability from x taking values within x_1 and x_2 is given by the area between the x axis, the density function curve, and the boundaries of the interval $x = x_1, x = x_2$.

   In particular the probabilities for the deviation from the mean within some multiples of σ are as follows:

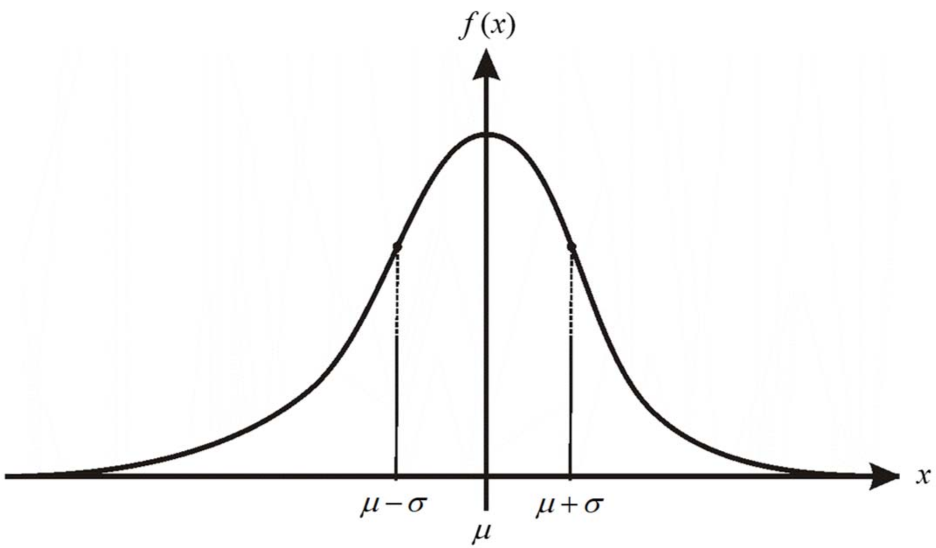
$$P[-\ \sigma < x - \mu < +\sigma] = 0.6827$$
$$P[-2\sigma < x - \mu < +2\sigma] = 0.9545$$
$$P[-3\sigma < x - \mu < +3\sigma] = 0.9973$$



6. The abscissae associated with intervals covering probabilities of 0.90, 0.95 and 0.99 are

$$P[-1.645\sigma < x - \mu < +1.645\sigma] = 0.90$$
$$P[-1.960\sigma < x - \mu < +1.960\sigma] = 0.95$$
$$P[-2.576\sigma < x - \mu < +2.576\sigma] = 0.99$$

7. The probability that x takes on values on either side of μ (that is, either larger or smaller than μ) is equal 0.5.



The theoretical and practical importance of the normal distribution is due to the “central limit theorem” which states that the sum $\sum_{i=1}^{q} \Delta_i$ of q independent elementary errors $\Delta_1, \Delta_2, \ldots, \Delta_q$ will be asymptotically normally distributed as $n \to \infty$.

In practical applications, normal distributions are encountered very often. In particular, random variables that represent measurements in photogrammetry, geodesy or surveying are often nearly normally distributed.

Propositions

1. $f(x)$ fulfills the necessary conditions for a probability density
2. $E(L) = \mu$ Translation allowed
3. $E(\varepsilon^2) = E\{(L-\mu)^2\} = \sigma^2$ Expectation of ε^2 is the variance σ^2 (interpretation of constant factor σ)

Proof: Propositions hold true if:

$$\left.\begin{aligned}
&1)\quad \int_{-\infty}^{+\infty} f(x)dx \overset{!}{=} 1 \\
&2)\quad E(L) = \int_{-\infty}^{+\infty} x \cdot f(x)dx \overset{!}{=} \mu \\
&3)\quad E(\varepsilon^2) = E\{(L-\mu)^2\} = \int_{-\infty}^{+\infty} (x-\mu)^2 \cdot f(x)dx \overset{!}{=} \sigma^2
\end{aligned}\right\} \text{to be poved}$$

Intermediate step: Integrals from formulary (e.g. Bronshtein, Handbook of Mathematics)

$$\int_{-\infty}^{+\infty} \exp\left(-\frac{t^2}{2}\right) dt = \sqrt{2\pi}$$

$$\int_{-\infty}^{+\infty} t \cdot \exp\left(-\frac{t^2}{2}\right) dt = 0$$

$$\int_{-\infty}^{+\infty} t^2 \cdot \exp\left(-\frac{t^2}{2}\right) dt = \sqrt{2\pi}$$

Substitution: $y = \frac{x-\mu}{\sigma} \quad \Rightarrow \quad x = \sigma y + \mu \quad \Rightarrow \quad dx = \sigma dy$

Proof of proposition 1)

$$\int_{-\infty}^{+\infty} f(x)dx = \frac{1}{\cancel{\sigma}\sqrt{2\pi}} \underbrace{\int_{-\infty}^{+\infty} \exp\left(-\frac{y^2}{2}\right) \cdot \cancel{\sigma} dy}_{\sqrt{2\pi}} = 1 \quad \checkmark$$

Proof of proposition 2)

$$E(L) = \int_{-\infty}^{+\infty} x \cdot f(x)dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \underbrace{(\sigma \cdot y + \mu) \cdot \exp\left(-\frac{y^2}{2}\right) \cdot \sigma dy}_{\text{multiplying out} \rightarrow \text{2 integrals}}$$

$$= \frac{\sigma}{\sigma\sqrt{2\pi}} \underbrace{\int_{-\infty}^{+\infty} y \cdot \exp\left(-\frac{y^2}{2}\right) dy}_{0} + \mu \underbrace{\int_{-\infty}^{+\infty} \exp\left(-\frac{y^2}{2}\right) dy}_{\sqrt{2\pi}} = \frac{\sigma \cdot \mu \cdot \sqrt{2\pi}}{\sigma\sqrt{2\pi}} = \mu \quad \checkmark$$

Proof of proposition 3)

$$E\left(\varepsilon^2\right) = E\{(L-\mu)^2\}$$

$$= \int_{-\infty}^{+\infty} (x-\mu)^2 \cdot f(x) dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \sigma^2 \cdot y^2 \cdot \exp\left(-\frac{y^2}{2}\right) \cdot \sigma dy$$

$$= \frac{\sigma^3}{\sigma\sqrt{2\pi}} \underbrace{\int_{-\infty}^{+\infty} y^2 \cdot \exp\left(-\frac{y^2}{2}\right) dy}_{\sqrt{2\pi}} = \frac{\sigma^3\sqrt{2\pi}}{\sigma\sqrt{2\pi}} = \sigma^2 \quad \checkmark$$

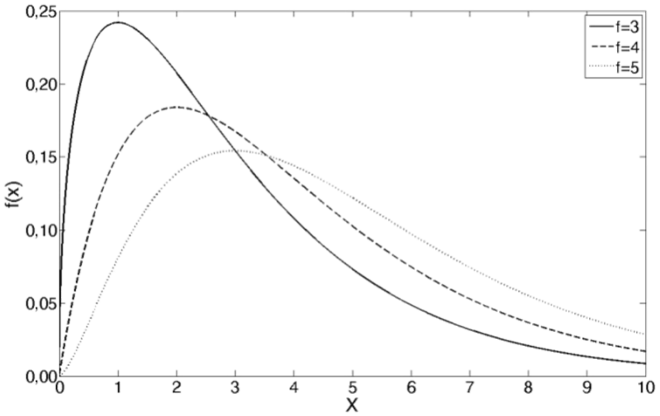
$\Rightarrow$ Expectation μ and variance σ^2 are the only parameters of the normal distribution!

$$\boxed{L \sim N\left(\mu, \sigma^2\right)}$$

Adjustment Theory I

Chapter 5: The Gaussian or Normal Distribution

Prof. Dr.-Ing. Frank Neitzel



Adjustment Theory I

Chapter 6 – Introduction to Least Squares Adjustment

Prof. Dr.-Ing Frank Neitzel | Institute of Geodesy and Geoinformation Science

Version: 18 November 2024

Contents

6. Introduction to Least Squares Adjustment

6.1 Introduction

$$\underset{n}{\boxed{\text{observations (measurements)}}} \longrightarrow \boxed{\text{functional model}} \longrightarrow \underset{u}{\boxed{\text{unknowns}}}$$

Until now we have considered
minimal configurations

$$\boxed{n = u}$$

$\Rightarrow$ Resulting equation system has a **unique solution**,
but
the reliability of the solution is 0%
$\Rightarrow$ no chance to detect blunders

6.1 Introduction

Example: Equation system

Functional model

$$7 = 3x + y$$
$$8 = 4x + 2y$$

obervations, unknowns, fixed values

$$n = 2, \quad u = 2$$

$\Rightarrow$ Minimal configuration,
unique solution

6.1 Introduction

Now: We consider
overdetermined configurations

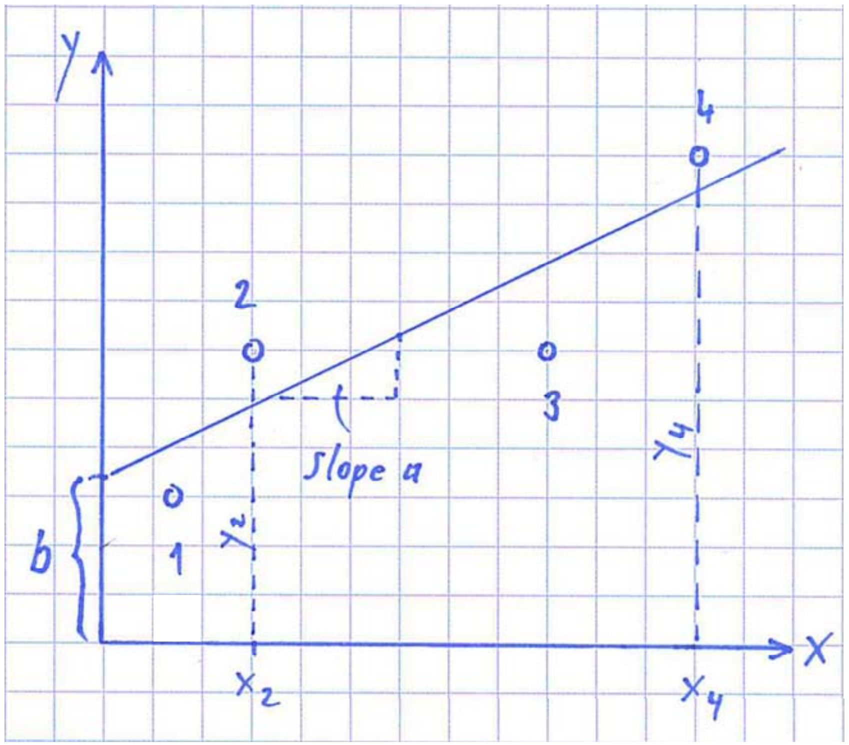
$$\boxed{n > u}$$

Fundamental principle in Geodesy!

$\Rightarrow$ **Reliable** solution
$\Rightarrow$ Increase of **precision** of the solution

6.1 Introduction

Example: Computation of slope a and intercept b of a straight line



Functional model

$$y = a\,x + b$$

$$y_1 = a\,x_1 + b$$
$$y_2 = a\,x_2 + b$$
$$y_3 = a\,x_3 + b$$
$$y_4 = a\,x_4 + b$$

obervations, unknowns, fixed values

$$n = 4, \quad u = 2$$

$\Rightarrow$ Overdetermined configuration

6.1 Introduction

We know:

Our measurements are affected by random errors

$\rightarrow$ Functional model can **not** be satisfied

$\rightarrow$ To overcome this inconsistency, we introduce residuals v_i and obtain …

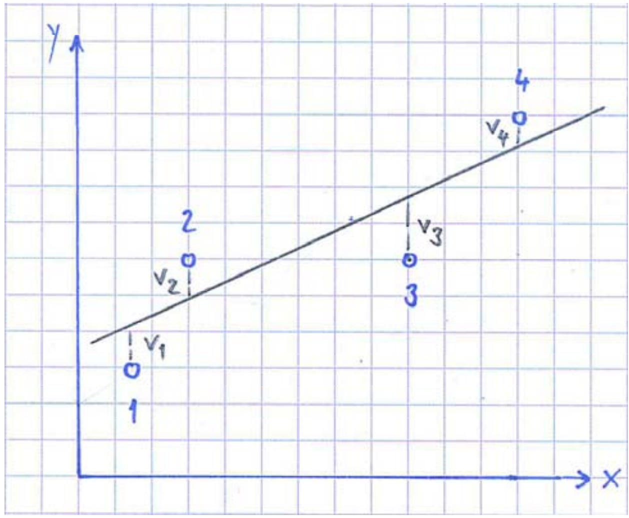
… Observation equations	Functional model
$y + v_i = a\,x + b$	$y = a\,x + b$
$y_1 + v_1 = a\,x_1 + b$	$y_1 = a\,x_1 + b$
$y_2 + v_2 = a\,x_2 + b$	$y_2 = a\,x_2 + b$
$y_3 + v_3 = a\,x_3 + b$	$y_3 = a\,x_3 + b$
$y_4 + v_4 = a\,x_4 + b$	$y_4 = a\,x_4 + b$

Problem:

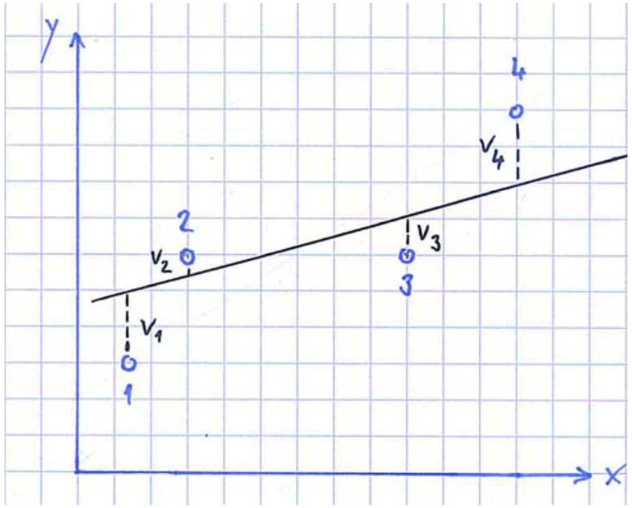
No unique solution, we can chose arbitrary straight lines as solution

6.1 Introduction

Possible solution 1



Possible solution 2



How to solve the problem?

⇒ We have to introduce a suitable constraint (target function) for the residuals to obtain **one specific solution**

⇒ Solution of overdetermined equation systems under consideration of a target function for the residuals is called **Adjustment Calculation**

⇒ Many different target functions for the residuals possible

⇒ We want to introduce specific target function that yields **Least Squares Adjustment**

6.1 Introduction

$$\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v} \rightarrow \min$$

Least Squares Adjustment

or

Method of Least Squares

6.1 Introduction

Technische Universität Berlin

Adjustment Theory deals with the optimal combination of redundant measurements under consideration of their stochastic properties together with the estimation of unknown parameters.

6.2 Historical Development

Historical development concerning the optimal combination of redundant measurements

<u>Method of selected points (before 1750)</u>

- ▶ Select only as many observations (“points”) as there are unknowns
- ▶ Remaining unused observations can be used to validate the estimated result
- ▶ Suppose that we use n observations → we obtain $\binom{n}{u}$ choices

6.2 Historical Development

Technische Universität Berlin

Method of averages (ca. 1750)

- In 1714: Longitude Prize for determination of a ships longitude offered by British government

- Thomas Mayer (1723-1762, German mathematician):
  Determination of longitude (rather time) from motion of the moon.
  He obtained overdetermined equation System:

$$\mathbf{L}_{27\times1} = \mathbf{A}_{27\times3} \cdot \mathbf{x}_{3\times1}$$

27 observations, 3 unknowns

6.2 Historical Development

- Mayer's adjustment strategy:
  - Distribute observations in 3 groups
  - Sum up the equations within each group
  - Solve the 3×3 system

<u>Euler's attempt (1749)</u>

- Leonhard Euler (1707-1783, Swiss mathematician and physicist)
- Orbital motion of Saturn under influence of Jupiter
- Prize (1748) of the Academy of Science, Paris
- 75 observations from the years 1582-1745, 6 unknowns

→ Euler could not solve the problem

6.2 Historical Development

Technische Universität Berlin

Laplace's attempt (ca. 1787)

- Pierre-Simon Laplace (1749-1827, French mathematician and astronomer)
- Motion of Saturn
- Best data: 24 observations, reformulated: 4 unknowns
- Approach: Like Mayer but other combinations

6.2 Historical Development

Technische Universität Berlin

Method of least deviations (1760)

- Roger Boscovich (1711-1787, Croatian Jesuit, mathematician and physicist)
- Ellipticity of the earth
- 5 observations (Quito, Cape Town, Rome, Paris, Lapland), 2 unknowns
- First attempt:
  - All $\binom{5}{2} = \frac{5!}{2!(5-2)!} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1 \cdot 3 \cdot 2 \cdot 1} = 10$ combinations with 2 observations each

6.2 Historical Development

▶ First attempt [cont.]:

- 10 systems of equations (2×2) → 10 solutions, comparison of results
- His result shows gross variations of ellipticity → reject ellipsoidal hypothesis

▶ Second attempt:

- Mean deviation (or sum of deviations) should be zero

$$\sum_{i=1}^{n} v_i = 0$$

and sum of absolute deviations should be minimum

$$\sum_{i=1}^{n} |v_i| \rightarrow \min$$

→ This is an objective adjustment criterion, known (today) as L_1-norm estimation: Detection of outliers (blunders)

6.2 Historical Development

Technische Universität Berlin

Method of least squares (ca. 1805)

- Adrien-Marie Legendre (1752-1833)
- He published his method of least squares (in French “moindres carrés”)
- Application: Determination of orbits of comets
- After Legendre’s publication C.F. Gauss (1777-1855, German mathematician, astronomer, geodesist and physicist) states that he has already developed and used the method of least squares in 1794
- Today: It is acknowledged that Gauss’ claim of priority is very likely valid

Technische Universität Berlin

6.3.1 Gauss' First Argument

▶ We assume that our measurements

- Contain no systematic deviations
- Contain no blunders
- Density function of observations is known → Gaussian or Normal distribution

‖ Under the assumption of normal distribution we want to obtain the most probable solution ‖

▶ If we assume normally distributed measurements
→ random deviations ε_i are also normally distributed with density function:

$$f(\varepsilon_i) = \frac{1}{\sigma_i \cdot \sqrt{2\pi}} \mathrm{e}^{\left(-\frac{1}{2} \cdot \frac{\varepsilon_i^2}{\sigma_i^2}\right)}$$

6.3.1 Gauss‘ First Argument

- The same distribution can be applied for the case that we consider the empirical residuals v_i

$$f(v_i) = \frac{1}{\sigma_i \cdot \sqrt{2\pi}} \mathrm{e}^{\left(-\frac{1}{2} \cdot \frac{v_i^2}{\sigma_i^2}\right)}$$

- Now, if we consider the joint occurrence of all residuals we obtain the overall probability density from:

$$\Omega = \frac{1}{\sigma_1 \cdot \sqrt{2\pi}} \mathrm{e}^{\left(-\frac{1}{2} \cdot \frac{v_1^2}{\sigma_1^2}\right)} \cdot \frac{1}{\sigma_2 \cdot \sqrt{2\pi}} \mathrm{e}^{\left(-\frac{1}{2} \cdot \frac{v_2^2}{\sigma_2^2}\right)} \cdot \ldots \cdot \frac{1}{\sigma_n \cdot \sqrt{2\pi}} \mathrm{e}^{\left(-\frac{1}{2} \cdot \frac{v_n^2}{\sigma_n^2}\right)}$$

$$\Omega = \left(\prod_{i=1}^{n} \frac{1}{\sigma_i \cdot \sqrt{2\pi}}\right) \mathrm{e}^{(-K)} \qquad \text{with } K = \frac{1}{2} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} v_i^2$$

6.3.1 Gauss‘ First Argument

Wanted: Maximum of density function

We search for values v_i that yield maximum Ω → That is the case if K obtains a minimum value

$$\sum_{i=1}^{n} \frac{1}{\sigma_i^2} v_i^2 \rightarrow \min \quad \text{with} \quad p_i = \frac{1}{\sigma_i^2} \quad \text{“weight of observations } l_i\text{”}$$

Weight matrix $\mathbf{P}$, here diagonal matrix $\mathbf{P} = \begin{bmatrix} p_{11} & & 0 \\ & \ddots & \\ 0 & & p_{nn} \end{bmatrix}$

$$\boxed{\sum_{i=1}^{n} p_i\, v_i^2 \rightarrow \min}$$

In matrix notation:

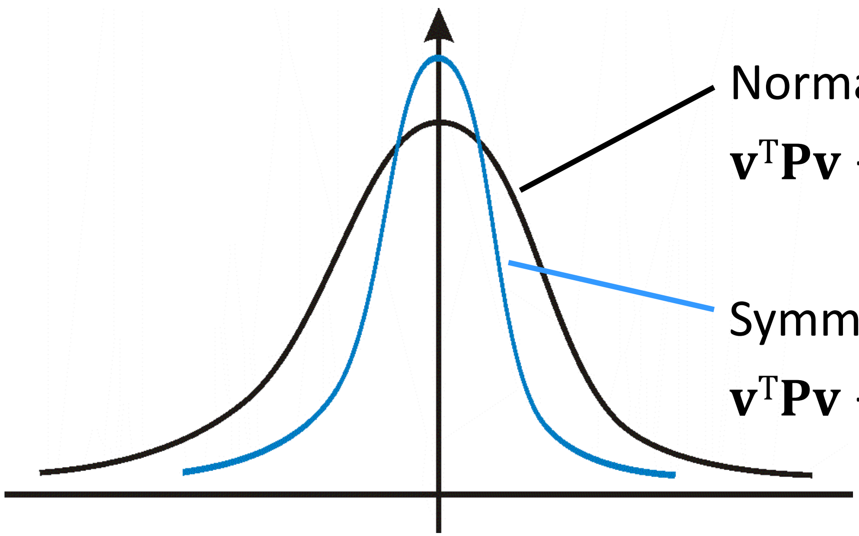
$$\boxed{\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v} \rightarrow \min}$$

6.3.2 Gauss‘ Second Argument

Gauss was not satisfied with his first argument

- In his “Theoria Combinationes” he did <u>not</u> apply the <u>normal distribution</u>
- He has shown that the method of least squares yields the smallest standard deviations if the frequency function $f(\varepsilon)$ is only symmetric to $\varepsilon = 0$.

  But: In this case, the results are no longer the most probable solution!
  Results are referred to instead as the “most plausible” or “most appropriate”



Normal distribution:

$\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v} \rightarrow$ min yields “most probable” solution

Symmetric frequency function $\neq$ normal dist. :

$\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v} \rightarrow$ min yields “most plausible” solution

6.3.3 Extension to correlated observations

Until now we have considered uncorrelated observations, weights represented by a diagonal matrix

$$\mathbf{P}_{n\times n} = \begin{bmatrix} p_{11} & & & 0 \\ & p_{22} & & \\ & & \ddots & \\ 0 & & & p_{nn} \end{bmatrix}$$

Gauss' derivations have been extended by e.g. Helmert, Tienstra and others:

1. Instead of original observations (e.g. horizontal directions) we can introduce derived observations (angles) if we consider the <u>correlations</u>.
2. We want to consider observations with known mathematical or physical correlations

6.3.3 Extension to correlated observations

Generalisation:

$$\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v} = \mathbf{v}^{\mathrm{T}}\mathbf{Q}^{-1}\mathbf{v} \rightarrow \min$$

with $\qquad \mathbf{Q} = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n1} & q_{n2} & \cdots & q_{nn} \end{bmatrix}$

$$\mathbf{\Sigma}_{LL} = \sigma_0^2 \cdot \mathbf{Q} \qquad \Rightarrow \qquad \mathbf{Q} = \frac{1}{\sigma_0^2} \cdot \mathbf{\Sigma}_{LL}$$

6.3.3 Extension to correlated observations

We know: $\mathbf{Q}$ is symmetric and non-singular → $\mathbf{P}$ is symmetric and non-singular

→ we can apply Cholesky decomposition: $\mathbf{P} = \mathbf{C}^{\mathrm{T}}\mathbf{C}$

$\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v}$ with $\mathbf{v} = \hat{\mathbf{l}} - \mathbf{l}$

$$
\begin{aligned}
\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v} &= \left(\hat{\mathbf{l}} - \mathbf{l}\right)^{\mathrm{T}}\mathbf{P}\left(\hat{\mathbf{l}} - \mathbf{l}\right) && \text{with} \quad \mathbf{P} = \mathbf{C}^{\mathrm{T}}\mathbf{C} \\
&= \left(\hat{\mathbf{l}} - \mathbf{l}\right)^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\mathbf{C}\left(\hat{\mathbf{l}} - \mathbf{l}\right) \\
&= \left(\hat{\mathbf{l}}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}} - \mathbf{l}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\right) \cdot \left(\mathbf{C}\hat{\mathbf{l}} - \mathbf{C}\mathbf{l}\right) && \text{with} \quad \mathbf{l}' = \mathbf{C}\mathbf{l} \text{ and } \quad \hat{\mathbf{l}}' = \mathbf{C}\hat{\mathbf{l}} \\
&= \left(\hat{\mathbf{l}}' - \mathbf{l}'\right)^{\mathrm{T}} \cdot \left(\hat{\mathbf{l}}' - \mathbf{l}'\right) && \text{with} \quad \mathbf{v}' = \hat{\mathbf{l}}' - \mathbf{l}' \\
&= (\mathbf{v}')^{\mathrm{T}} \cdot (\mathbf{v}') && \text{equally weighted}
\end{aligned}
$$

$\boxed{\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v} = (\mathbf{v}')^{\mathrm{T}} \cdot (\mathbf{v}')}$ $\quad (\mathbf{v}')^{\mathrm{T}} \cdot (\mathbf{v}') \rightarrow \min$ $\quad$ can be solved via Gauss

6.3.3 Extension to correlated observations

→ Adjustment with $\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v} \rightarrow \min$ yields same result as $(\mathbf{v}')^{\mathrm{T}} \cdot (\mathbf{v}')$

→ We can in general introduce correlated observations by introducing $\mathbf{P} = \mathbf{Q}^{-1}$ as weight matrix

$$\underset{n\times n}{\mathbf{P}} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}$$

6.4 Functional Model and Stochastic Model

6.4.1 Functional model

- Functional model in adjustment computation is a set of equations that represents an adjustment condition
- If the functional model represents the geometrical or physical situation adequately, observation errors can be expected to conform to the normal distribution
- Example: Functional model $\alpha + \beta + \gamma = 200 \text{ gon}$ in a triangle.
  But: This model is only adequate if the survey is limited to a small region.
  Large areas: Spherical excess has to be considered.

→ If the functional model does not fit the geometrical or physical situation, an incorrect adjustment will result!

6.4.1 Functional Model

There are two basic forms for functional models:

▶ Conditional adjustment $\mathbf{\Phi}(\mathbf{L}) = \mathbf{0}$

▶ Parametric adjustment $\mathbf{L} = \mathbf{\Phi}(\mathbf{X})$

Gauss-Markov-Model

▶ Conditional adjustment:

Geometric conditions are enforced on the observations and their residuals, e.g. $\alpha + v_\alpha + \beta + v_\beta + \gamma + v_\gamma = 200 \text{ gon}$

- Advantage: Small equation systems
- Disadvantage: Often difficult and time consuming to find conditions, e.g. in complicated networks

→ Not well suited to computers

6.4.1 Functional Model

▶ Parametric adjustment:

Observations are expressed as functions of unknown parameters, e.g.

$$s_{ij} + v_{ij} = \sqrt{\left(x_j - x_i\right)^2 + \left(y_j - y_i\right)^2}$$

- Disadvantage: large equation systems
- Advantages:
  - "Standard" formulas can be applied
  - Well suited to computers

→ Of course, conditional and parametric adjustment yield same results.

This semester: parametric adjustment

Next semester: conditional adjustment

6.4.2 Stochastic Model

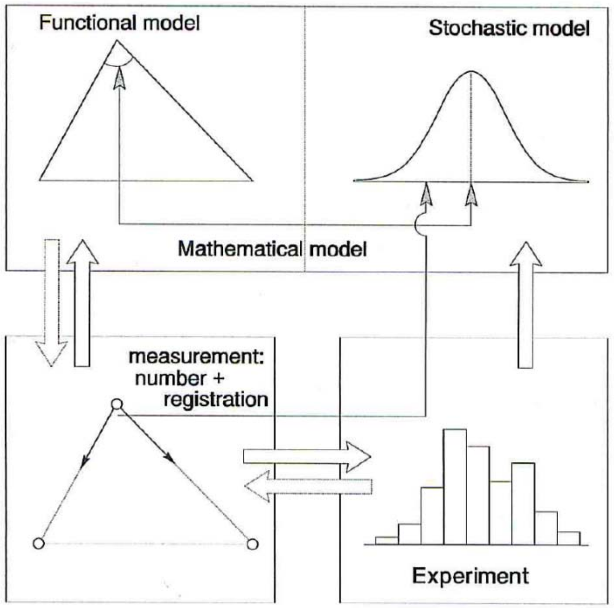
- Determination of variances and, subsequently, the weights of the observations is know as <u>stochastic model</u> in least squares adjustment
- Weight of an observation controls the amount of correction it receives during the adjustment → It is <u>very</u> important to select a proper stochastic (weight) model!
- Attention: When doing an “unweighted” adjustment (= all obs. have the same weight) then the stochastic model is created implicitly
- Failures to select the stochastic model properly will
  - Influence the adjusted parameters
  - Affect the ability to isolate blunders

→ Combination of functional and stochastic models is called the <u>mathematical model</u>

6.4.2 Stochastic Model

Technische Universität Berlin

Diagram of the fundamental relations in adjustment theory



Both, stochastic and functional model must be correct

if the adjustment is to yield the most probable values!

6.5 Observation Equations

- Equations that relate observed quantities (measurements) to <u>both</u> observational residuals and unknown parameters are called <u>observation equations</u>

- Functional model, e.g.

$$L_1 = x + y \qquad (1)$$

$$L_2 = 2x - y \qquad (2)$$

$$\underbrace{L_3}_{n\text{ obs.}} = \underbrace{x - y}_{u\text{ unknowns}} \qquad (3)$$

$n = 3 \quad u = 2 \quad \rightarrow \quad n > u$

6.5 Observation Equations

▶ <u>If $n > u$</u> functional model can only be fulfilled by “true values”

$$"\sim" = \text{true value}$$

$$\tilde{L}_1 = \tilde{x} + \tilde{y}$$

$$\tilde{L}_2 = 2\tilde{x} - \tilde{y}$$

$$\tilde{L}_3 = \tilde{x} - \tilde{y}$$

6.5 Observation Equations

- Problem:

  We don't know the true values

  → Functional model has usually no solution

- Solution:

  We introduce residuals for the observations
  →Resulting set of equations is called

  Observation equations (residual equations)

$$L_1 + v_1 = \hat{x} + \hat{y}$$
$$L_2 + v_2 = 2\hat{x} - \hat{y}$$
$$L_3 + v_3 = \hat{x} - \hat{y}$$

with $\hat{x}, \hat{y} =$ adjusted parameters

6.6 Least squares adjustment without matrix notation

- **Problem:**

  This equation system has no unique solution

- **Solution:**

  We introduce a target function for the residuals

$$\sum_{i=1}^{n} p_i v_i^2$$

  and we search for a solution with

$$\sum_{i=1}^{n} p_i v_i^2 \rightarrow \min$$

  → Least squares adjustment

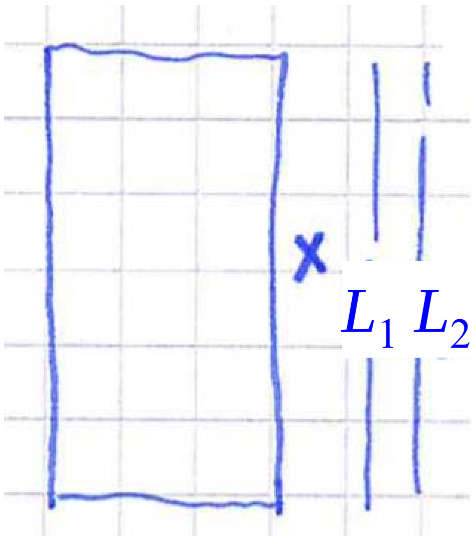
6.6 Least squares adjustment without matrix notation

Example:

Length of the classroom has been measured n-times with same precision

▶ Given: obs. $L_1, L_2, \dots, L_n$, equally weighted

▶ Wanted: Adjusted unknown $\hat{x}$



Functional model:

$$\begin{aligned} L_1 &= x \\ L_2 &= x \\ &\vdots \\ L_n &= x \end{aligned}$$

Stochastic model:

$$p_1 = p_2 = \cdots = p_n = 1$$

6.6 Least squares adjustment without matrix notation

Observation equations:

$$L_1 + v_1 = \hat{x}$$
$$L_2 + v_2 = \hat{x}$$
$$\vdots$$
$$L_n + v_n = \hat{x}$$

Rearranging:

$$v_1 = \hat{x} - L_1$$
$$v_2 = \hat{x} - L_2$$
$$\vdots$$
$$v_n = \hat{x} - L_n$$

$$\sum_{i=1}^{n} p_i v_i^2 \rightarrow \min$$

$$\rightarrow 1 \cdot (\hat{x} - L_1)^2 + 1 \cdot (\hat{x} - L_2)^2 + \cdots + 1 \cdot (\hat{x} - L_n)^2 \rightarrow \min$$

6.6 Least squares adjustment without matrix notation

How to obtain the min?

- ▶ Minimum value of a function can be found by taking its first derivative
- ▶ Equate the resulting function with zero

→ Taking first derivative with respect to x and setting the resulting function equal to zero yields

$$\Sigma p_i v_i^2 = 1 \cdot (\hat{x} - L_1)^2 + 1 \cdot (\hat{x} - L_2)^2 + \cdots + 1 \cdot (\hat{x} - L_n)^2$$

$$\frac{\partial \Sigma p_i v_i^2}{\partial \hat{x}} = \underbrace{2 \cdot (\hat{x} - L_1) + 2 \cdot (\hat{x} - L_2) + \cdots + 2 \cdot (\hat{x} - L_n) = 0}_{\text{normal equation}}$$

6.6 Least squares adjustment without matrix notation

Solution of normal equation:

$$(\hat{x} - L_1) + (\hat{x} - L_2) + \cdots + (\hat{x} - L_n) = 0$$

$$n \cdot \hat{x} = L_1 + L_2 + \cdots + L_n$$

Adjusted unknown:

$$\boxed{\hat{x} = \frac{L_1 + L_2 + \cdots + L_n}{n}}$$

$$\boxed{\hat{x} = \frac{\Sigma L_i}{n}}$$

6.6 Least squares adjustment without matrix notation

Technische Universität Berlin

<u>Example</u>:

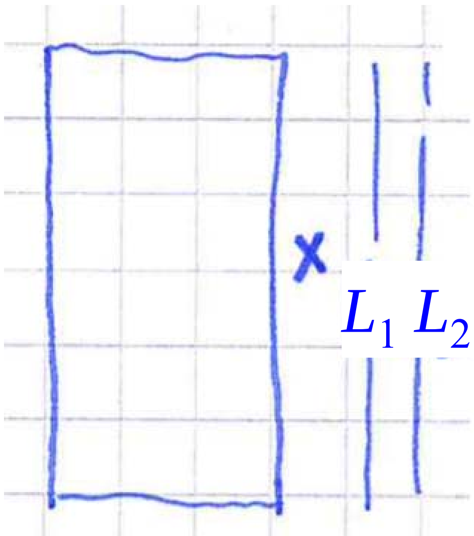
Length of the classroom has been measured n-times with different precision (standard deviation) σ_i

▶ Given: L_1, σ_1

L_2, σ_2

$\vdots$

L_n, σ_n



▶ Wanted: Adjusted unknown $\hat{x}$

6.6 Least squares adjustment without matrix notation

Functional model:

$$
\begin{aligned}
L_1 &= x \\
L_2 &= x \\
&\vdots \\
L_n &= x
\end{aligned}
$$

Stochastic model:

$$p_1 = \left(\frac{1}{\sigma_1}\right)^2, p_2 = \left(\frac{1}{\sigma_2}\right)^2, \ldots, p_n = \left(\frac{1}{\sigma_n}\right)^2$$

6.6 Least squares adjustment without matrix notation

Observation equations:

$$L_1 + v_1 = \hat{x}$$
$$L_2 + v_2 = \hat{x}$$
$$\vdots$$
$$L_n + v_n = \hat{x}$$

Rearranging:

$$v_1 = \hat{x} - L_1$$
$$v_2 = \hat{x} - L_2$$
$$\vdots$$
$$v_n = \hat{x} - L_n$$

$$\sum_{i=1}^{n} p_i v_i^2 \rightarrow \min$$

$$\rightarrow \quad p_1(\hat{x} - L_1)^2 + p_2(\hat{x} - L_2)^2 + \cdots + p_n(\hat{x} - L_n)^2 \rightarrow \min$$

$$\frac{\partial \Sigma p_i v_i^2}{\partial \hat{x}} = \underbrace{2p_1(\hat{x} - L_1) + 2p_2(\hat{x} - L_2) + \cdots + 2p_n(\hat{x} - L_n) = 0}_{\underline{\text{normal equation}}}$$

6.6 Least squares adjustment without matrix notation

Solution of normal equation:

$$p_1(\hat{x} - L_1) + p_2(\hat{x} - L_2) + \cdots + p_n(\hat{x} - L_n) = 0$$

$$p_1\hat{x} - p_1L_1 + p_2\hat{x} - p_2L_2 + \cdots + p_n\hat{x} - p_nL_n = 0$$

$$\hat{x} \cdot \Sigma p_i = \Sigma p_i L_i$$

Adjusted unknown:

$$\boxed{\hat{x} = \frac{\Sigma p_i L_i}{\Sigma p_i}}$$

Example:

Functional model:

$$3.0 = x + y$$
$$1.5 = 2x - y$$
$$0.2 = x - y$$

Values	$3.0, 1.5, 0.2$	are observations
Parameters	x, y	are unknowns

Stochastic model for the observations:

$$p_1 = 1, \qquad p_2 = 1, \qquad p_3 = 1$$

6.6 Least squares adjustment without matrix notation

Observation equations:

$$3.0 + v_1 = \hat{x} + \hat{y}$$
$$1.5 + v_2 = 2\hat{x} - \hat{y}$$
$$0.2 + v_3 = \hat{x} - \hat{y}$$

Rearranging:

$$v_1 = \hat{x} + \hat{y} - 3.0$$
$$v_2 = 2\hat{x} - \hat{y} - 1.5$$
$$v_3 = \hat{x} - \hat{y} - 0.2$$

$$\sum_{i=1}^{n} p_i v_i^2 \rightarrow \min$$

$$\rightarrow \; 1 \cdot (\hat{x} + \hat{y} - 3.0)^2 + 1 \cdot (2\hat{x} - \hat{y} - 1.5)^2 + 1 \cdot (\hat{x} - \hat{y} - 0.2)^2 \rightarrow \min$$

6.6 Least squares adjustment without matrix notation

Normal equations:

$$\Sigma p_i v_i^2 = 1 \cdot (\hat{x} + \hat{y} - 3.0)^2 + 1 \cdot (2\hat{x} - \hat{y} - 1.5)^2 + 1 \cdot (\hat{x} - \hat{y} - 0.2)^2$$

$$\frac{\partial \Sigma p_i v_i^2}{\partial \hat{x}} = \boxed{2(\hat{x} + \hat{y} - 3.0) + 2 \cdot (2\hat{x} - \hat{y} - 1.5) \cdot 2 + 2 \cdot (\hat{x} - \hat{y} - 0.2) = 0}$$

$$\frac{\partial \Sigma p_i v_i^2}{\partial \hat{y}} = \boxed{2(\hat{x} + \hat{y} - 3.0) + 2 \cdot (2\hat{x} - \hat{y} - 1.5) \cdot (-1) + 2 \cdot (\hat{x} - \hat{y} - 0.2) \cdot (-1) = 0}$$

<u>normal equations</u>

$$\hat{x} + \hat{y} - 3.0 + 4\hat{x} - 2\hat{y} - 3.0 + \hat{x} - \hat{y} - 0.2 = 0$$

$$\hat{x} + \hat{y} - 3.0 - 2\hat{x} + \hat{y} + 1.5 - \hat{x} + \hat{y} + 0.2 = 0$$

$$6\hat{x} - 2\hat{y} = 6.2 \qquad (1)$$

$$-2\hat{x} + 3\hat{y} = 1.3 \qquad (2)$$

6.6 Least squares adjustment without matrix notation

Solution of normal equations:

$(1) + 3 \cdot (2)$: $\quad 7\hat{y} = 10.1 \quad \Rightarrow \quad \hat{y} = 1.443$

$\hat{y}$ in (1): $\quad \hat{x} = \frac{6.2 + 2\hat{y}}{6} \quad \Rightarrow \quad \hat{x} = 1.514$

Residuals:

Can be computed from observation equations

$$
\begin{aligned}
v_1 &= -0.044 \\
v_2 &= 0.085 \\
v_3 &= -0.128
\end{aligned}
$$

6.6 Least squares adjustment without matrix notation

Same example, but now we know the precision (standard deviation) s_i of the measured values

$$3.0, \qquad s_1 = 4\text{ cm}$$
$$1.5, \qquad s_2 = 2\text{ cm}$$
$$0.2, \qquad s_3 = 1\text{ cm}$$

Stochastic model:

How to obtain weights p_1, p_2, p_3?

$$p_1 = \frac{1}{(s_1)^2} \qquad p_2 = \frac{1}{(s_2)^2} \qquad p_3 = \frac{1}{(s_3)^2}$$

$$p_1 = \frac{1}{16} \qquad p_2 = \frac{1}{4} \qquad p_3 = 1$$

6.6 Least squares adjustment without matrix notation

$$\sum_{i=1}^{n} p_i v_i^2 \rightarrow \min$$

$$\rightarrow \quad \frac{1}{16}\cdot(\hat{x}+\hat{y}-3.0)^2+\frac{1}{4}\cdot(2\hat{x}-\hat{y}-1.5)^2+1\cdot(\hat{x}-\hat{y}-0.2)^2 \rightarrow \min$$

$$\frac{\Sigma p_i v_i^2}{\partial\hat{x}} = 2\cdot\frac{1}{16}(\hat{x}+\hat{y}-3.0)+2\cdot\frac{1}{4}(2\hat{x}-\hat{y}-1.5)\cdot 2+2\cdot(\hat{x}-\hat{y}-0.2)=0$$

$$\frac{\Sigma p_i v_i^2}{\partial\hat{y}} = 2\cdot\frac{1}{16}(\hat{x}+\hat{y}-3.0)+2\cdot\frac{1}{4}(2\hat{x}-\hat{y}-1.5)(-1)+2\cdot(\hat{x}-\hat{y}-0.2)(-1)=0$$

6.6 Least squares adjustment without matrix notation

$$\hat{x} + \hat{y} - 3.0 \ + 16\hat{x} - 8\hat{y} - 12 + 16\hat{x} - 16\hat{y} - 3.2 = 0$$

$$\hat{x} + \hat{y} - 3.0 \ + 8\hat{x} + 4\hat{y} - 6 - 16\hat{x} + 16\hat{y} + 3.2 = 0$$

$$33\hat{x} - 23\hat{y} = 18.2 \qquad (1)$$

$$-23\hat{x} + 21\hat{y} = -6.2 \qquad (2)$$

Solution of normal equations:

$21 \cdot (1) + 23 \cdot (2)$: $\quad (21 \cdot 33 - 23 \cdot 23)\hat{x} = (21 \cdot 18.2) - (23 \cdot 6.2)$

$$164\,\hat{x} = 239.6$$

$$\hat{x} = 1.4610$$

$\hat{x}$ in (1): $\quad \hat{y} = 1.3049$

Residuals:

Solution $\hat{x}, \hat{y}$ in observation equations → v_i

6.6 Least squares adjustment without matrix notation

$$\hat{x} + \hat{y} - 3.0 \; + 16\hat{x} - 8\hat{y} - 12 + 16\hat{x} - 16\hat{y} - 3.2 = 0$$

$$\hat{x} + \hat{y} - 3.0 \; + 8\hat{x} + 4\hat{y} - 6 - 16\hat{x} + 16\hat{y} + 3.2 = 0$$

$$33\hat{x} - 23\hat{y} = 18.2 \qquad (1)$$

$$-23\hat{x} + 21\hat{y} = -6.2 \qquad (2)$$

Solution of normal equations:

$$\underbrace{\begin{bmatrix} 33 & -23 \\ -23 & 21 \end{bmatrix}}_{\mathbf{N}} \underbrace{\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix}}_{\hat{\mathbf{X}}} = \underbrace{\begin{bmatrix} 18.2 \\ -6.2 \end{bmatrix}}_{\mathbf{n}}$$

$$\mathbf{N}\hat{\mathbf{X}} = \mathbf{n}$$

$$\hat{\mathbf{X}} = \mathbf{N}^{-1}\mathbf{n}$$

6.7 Least squares adjustment in matrix notation

Now: Application of matrices to build normal equations

6.7.1 Linear functional models

▶ Observations:

$$
\begin{aligned}
&3.0, \qquad s_1 = 4\ \text{cm} \\
&1.5, \qquad s_2 = 2\ \text{cm} \\
&0.2, \qquad s_3 = 1\ \text{cm}
\end{aligned}
$$

Observation vector:

$$\mathbf{L}_{n\times 1} = \begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{bmatrix} \qquad \text{here} \qquad \mathbf{L}_{3\times 1} = \begin{bmatrix} 3.0 \\ 1.5 \\ 0.2 \end{bmatrix}$$

6.7.1 Linear functional models

Stochastic model of $\mathbf{L}$:

- For theoretical standard deviations σ_i: ($\boldsymbol{\Sigma}_{LL}$: VCM of $\mathbf{L}$)

$$\boldsymbol{\Sigma}_{LL_{n\times n}} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix}$$

- For empirical standard deviations s_i:

$$\mathbf{S}_{LL_{n\times n}} = \begin{bmatrix} s_1^2 & s_{12} & \cdots & s_{1n} \\ s_{21} & s_2^2 & \cdots & s_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n1} & s_{n2} & \cdots & s_n^2 \end{bmatrix}$$

6.7.1 Linear functional models

▶ We choose arbitrary value σ_0

$$\mathbf{\Sigma}_{LL} = \sigma_0^2 \cdot \mathbf{Q}_{LL_{n\times n}} \quad \Rightarrow \quad \mathbf{Q}_{LL_{n\times n}} = \frac{1}{\sigma_0^2} \cdot \mathbf{\Sigma}_{LL_{n\times n}}$$

- σ_0: theoretical reference standard deviation or
  reference standard deviation à priori

- $\mathbf{Q}_{LL}$: Cofactor matrix of $\mathbf{L}$

- $\mathbf{P}$: Weight matrix of $\mathbf{L}$

$$\boxed{\mathbf{P}_{n\times n} = \mathbf{Q}_{LL}^{-1}{}_{n\times n}}$$

$\mathbf{Q}_{LL}$ regular

6.7.1 Linear functional models

- In our example:

$$\sigma_0 = 1 \qquad \text{(usually used in practice)}$$

$$\mathbf{S}_{LL} = \mathbf{Q}_{LL} = \begin{bmatrix} (4\ \mathrm{cm})^2 & 0 & 0 \\ 0 & (2\ \mathrm{cm})^2 & 0 \\ 0 & 0 & (1\ \mathrm{cm})^2 \end{bmatrix}$$

$$\Rightarrow \mathbf{P} = \mathbf{Q}_{LL}^{-1} = \begin{bmatrix} \frac{1}{16\ \mathrm{cm}^2} & 0 & 0 \\ 0 & \frac{1}{4\ \mathrm{cm}^2} & 0 \\ 0 & 0 & \frac{1}{1\ \mathrm{cm}^2} \end{bmatrix}$$

6.7.1 Linear functional models

Functional model:

$$3.0 = 1x + 1y \qquad \underbrace{\begin{bmatrix} 3.0 \\ 1.5 \\ 0.2 \end{bmatrix}}_{\mathbf{L}_{n\times 1}} = \underbrace{\begin{bmatrix} 1 & 1 \\ 2 & -1 \\ 1 & -1 \end{bmatrix}}_{\mathbf{A}_{n\times u}} \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{\mathbf{X}_{u\times 1}}$$

$$1.5 = 2x - 1y$$

$$0.2 = 1x - 1y$$

- Observation vector $\mathbf{L}_{n\times 1}$
- Matrix with coefficients of the linear functional model
  → coefficient matrix or design matrix $\mathbf{A}_{n\times u}$
- Vector of unknowns $\mathbf{X}_{u\times 1}$

$$\boxed{\mathbf{L}_{n\times 1} = \mathbf{A}_{n\times u}\mathbf{X}_{u\times 1}}$$

6.7.1 Linear functional models

Observation equations:

$$\begin{aligned} 3.0 + v_1 &= \hat{x} + \hat{y} \\ 1.5 + v_2 &= 2\hat{x} - \hat{y} \\ 0.2 + v_3 &= \hat{x} - \hat{y} \end{aligned}$$

Vector of residuals:

$$\mathbf{v}_{n\times 1} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \qquad \text{here} \qquad \mathbf{v}_{3\times 1} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

Vector of adjusted unknowns:

$$\hat{\mathbf{X}}_{u\times 1} = \begin{bmatrix} \hat{X}_1 \\ \hat{X}_2 \\ \vdots \\ \hat{X}_u \end{bmatrix} \qquad \text{here} \qquad \hat{\mathbf{X}}_{2\times 1} = \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix}$$

6.7.1 Linear functional models

$$\boxed{\mathbf{L}_{n\times 1} + \mathbf{v}_{n\times 1} = \mathbf{A}_{n\times m}\widehat{\mathbf{X}}_{u\times 1}}$$

$\widehat{\ }$: adjusted value

$$\mathbf{v} = \mathbf{A}\widehat{\mathbf{X}} - \mathbf{L}$$

$$\Omega = \sum_{i=1}^{n} p_i v_i^2 \rightarrow \min$$

in matrix notation $\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v} \rightarrow \min$

with $\mathbf{v} = \mathbf{A}\widehat{\mathbf{X}} - \mathbf{L}$

6.7.1 Linear functional models

$$\begin{aligned}
\Omega &= \left(\mathbf{A}\widehat{\mathbf{X}} - \mathbf{L}\right)^{\mathrm{T}}\mathbf{P}\left(\mathbf{A}\widehat{\mathbf{X}} - \mathbf{L}\right) \\
&= \left(\widehat{\mathbf{X}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} - \mathbf{L}^{\mathrm{T}}\right)\mathbf{P}\left(\mathbf{A}\widehat{\mathbf{X}} - \mathbf{L}\right) \\
&= (\widehat{\mathbf{X}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P} - \mathbf{L}^{\mathrm{T}}\mathbf{P})(\mathbf{A}\widehat{\mathbf{X}} - \mathbf{L}) \\
&= \widehat{\mathbf{X}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\widehat{\mathbf{X}} - \widehat{\mathbf{X}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{L} - \mathbf{L}^{\mathrm{T}}\mathbf{P}\mathbf{A}\widehat{\mathbf{X}} + \mathbf{L}^{\mathrm{T}}\mathbf{P}\mathbf{L} \\
&= \widehat{\mathbf{X}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\widehat{\mathbf{X}} - \underbrace{\widehat{\mathbf{X}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{L}}_{\text{scalar}} - \underbrace{\left(\widehat{\mathbf{X}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{L}\right)^{\mathrm{T}}}_{\text{scalar}} + \mathbf{L}^{\mathrm{T}}\mathbf{P}\mathbf{L} \\
&= \widehat{\mathbf{X}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\widehat{\mathbf{X}} - 2\cdot\widehat{\mathbf{X}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{L} + \mathbf{L}^{\mathrm{T}}\mathbf{P}\mathbf{L}
\end{aligned}$$

Minimum:

$$\frac{\partial\Omega}{\partial\widehat{\mathbf{X}}^{\mathrm{T}}} = 2\cdot\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\widehat{\mathbf{X}} - 2\cdot\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{L} = 0$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\widehat{\mathbf{X}} = \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{L}$$

6.7.1 Linear functional models

Technische Universität Berlin

Normal Equations:

$$\underbrace{\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}}_{\substack{\mathbf{N}\\ u\times u}}\hat{\mathbf{X}}=\underbrace{\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{L}}_{\substack{\mathbf{n}\\ u\times 1}}$$

- $\mathbf{N}_{u\times u} = \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}$: Normal matrix
- $\mathbf{n}_{u\times 1} = \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{L}$: Right hand side of normal equations

If $\mathbf{N}$ regular → we can compute $\mathbf{N}^{-1}$

Solution of normal equations:

$$\widehat{\mathbf{X}} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{L}$$

$$\widehat{\mathbf{X}} = \mathbf{N}^{-1}\mathbf{n}$$

Residuals: $\mathbf{v} = \mathbf{A}\widehat{\mathbf{X}} - \mathbf{L}$

6.7.1 Linear functional models

Adjusted observations:

$$\boxed{\hat{\mathbf{L}}_{n\times 1} = \mathbf{L}_{n\times 1} + \mathbf{v}_{n\times 1}}$$

Final check:

$$\boxed{\hat{\mathbf{L}} = \mathbf{\Phi}(\hat{\mathbf{X}})}$$

Original functional model

$$\hat{\mathbf{L}} - \mathbf{\Phi}(\hat{\mathbf{X}}) \overset{!}{=} \mathbf{0}$$

zero within computing precision

6.7.1 Linear functional models

▶ Linear functional model

→ we obtain linear normal equations

→ very easy to solve (linear algebra)

▶ Empirical reference standard deviation (or empirical reference variance s_0^2)

- $n =$ number of observations
- $u =$ number of unknowns

$$s_0 = \sqrt{\frac{\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v}}{n-u}}$$

6.7.2 Variance-Covariance Matrices (VCM) for the results

VCM for the vector of unknowns $\widehat{\mathbf{X}}$:

- We know

$$\widehat{\mathbf{X}} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{L} \qquad (1)$$

→ For $\mathbf{L}$ we have given $\boldsymbol{\Sigma}_{LL}$ resp. $\mathbf{S}_{LL}$ → $\mathbf{Q}_{LL}$

- Question: What is the VCM ($\boldsymbol{\Sigma}_{\hat{X}\hat{X}}$) of $\widehat{\mathbf{X}}$

- From Variance-Covariance Propagation (VCP), see Section 4.4, we know

$$\mathbf{x} = \mathbf{F}\mathbf{L}, \qquad \mathbf{Q}_{LL}$$

$$\mathbf{Q}_{xx} = \mathbf{F}\mathbf{Q}_{LL}\mathbf{F}^{\mathrm{T}}$$

Technische Universität Berlin

- Now we apply VCP to (1) $\hat{\mathbf{X}} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{L}$:

$$\mathbf{Q}_{\hat{X}\hat{X}} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\underbrace{\mathbf{P}\mathbf{Q}_{LL}}_{\mathbf{I}}\mathbf{P}\mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1}$$

$$= \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1}\underbrace{\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1}}_{\mathbf{I}}$$

$$\boxed{\mathbf{Q}_{\hat{X}\hat{X}} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1}}$$ Cofactor Matrix of the unknowns

$$\mathbf{Q}_{\hat{X}\hat{X}} = \mathbf{N}^{-1}$$ Inverse of normal matrix

$$\boldsymbol{\Sigma}_{\hat{X}\hat{X}} = s_0^2 \cdot \mathbf{Q}_{\hat{X}\hat{X}}$$ VCM of the unknowns

VCM for functions of the parameters $\widehat{\mathbf{X}}$:

For any linear function of the parameters

$$\mathbf{f} = \mathbf{F}\widehat{\mathbf{X}}$$

we can apply VCP to obtain the cofactor matrix of $\mathbf{f}$ as

$$\mathbf{Q}_{ff} = \mathbf{F}\mathbf{Q}_{XX}\mathbf{F}^{\mathrm{T}} \qquad (2)$$

Technische Universität Berlin

1. Cofactor matrix $\mathbf{Q}_{\hat{L}\hat{L}}$ for the adjusted observations:

We know

$$\hat{\mathbf{L}} = \mathbf{L} + \mathbf{v} = \mathbf{A}\hat{\mathbf{X}}$$

Application of (2) yields

- Cofactor matrix of the adjusted observations:

$$\boxed{\mathbf{Q}_{\hat{L}\hat{L}} = \mathbf{A}\mathbf{Q}_{\hat{X}\hat{X}}\mathbf{A}^{\mathrm{T}}}$$

- VCM of the adjusted observations:

$$\boldsymbol{\Sigma}_{\hat{L}\hat{L}} = s_0^2 \cdot \mathbf{Q}_{\hat{L}\hat{L}}$$

Technische Universität Berlin

1. Cofactor matrix $\mathbf{Q}_{vv}$ for the residuals:

Residuals are obtained from

$$\mathbf{v} = \mathbf{A}\hat{\mathbf{X}} - \mathbf{L}$$

with

$$\hat{\mathbf{X}} = \underbrace{\mathbf{N}^{-1}}_{\mathbf{Q}_{\hat{X}\hat{X}}}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{L}$$

$$\mathbf{v} = \mathbf{A}\mathbf{Q}_{\hat{X}\hat{X}}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{L} - \mathbf{L}$$
$$\mathbf{v} = \left(\mathbf{A}\mathbf{Q}_{\hat{X}\hat{X}}\mathbf{A}^{\mathrm{T}}\mathbf{P} - \mathbf{I}\right)\mathbf{L}$$

→ Residuals as function of the observations

→ Application of VCP

6.7.2 Variance-Covariance Matrices (VCM) for the results

$$\mathbf{Q}_{vv} = \left(\mathbf{A}\mathbf{Q}_{\hat{X}\hat{X}}\mathbf{A}^{\mathrm{T}}\mathbf{P} - \mathbf{I}\right)\mathbf{Q}_{LL}\left(\mathbf{P}\mathbf{A}\mathbf{Q}_{\hat{X}\hat{X}}\mathbf{A}^{\mathrm{T}} - \mathbf{I}\right)$$

$$= \mathbf{A}\mathbf{Q}_{\hat{X}\hat{X}}\underbrace{\mathbf{A}^{\mathrm{T}}\underbrace{\mathbf{P}\mathbf{Q}_{LL}}_{\mathbf{I}}\mathbf{P}\mathbf{A}}_{\mathbf{Q}_{\hat{X}\hat{X}}^{-1}}\mathbf{Q}_{\hat{X}\hat{X}}\mathbf{A}^{\mathrm{T}} - \underbrace{\mathbf{Q}_{LL}\mathbf{P}}_{\mathbf{I}}\mathbf{A}\mathbf{Q}_{\hat{X}\hat{X}}\mathbf{A}^{\mathrm{T}} - \mathbf{A}\mathbf{Q}_{\hat{X}\hat{X}}\mathbf{A}^{\mathrm{T}}\underbrace{\mathbf{P}\mathbf{Q}_{LL}}_{\mathbf{I}} + \mathbf{Q}_{LL}$$

$$= \mathbf{A}\mathbf{Q}_{\hat{X}\hat{X}}\mathbf{A}^{\mathrm{T}} - \mathbf{A}\mathbf{Q}_{\hat{X}\hat{X}}\mathbf{A}^{\mathrm{T}} - \mathbf{A}\mathbf{Q}_{\hat{X}\hat{X}}\mathbf{A}^{\mathrm{T}} + \mathbf{Q}_{LL}$$

$$= \mathbf{Q}_{LL} - \underbrace{\mathbf{A}\mathbf{Q}_{\hat{X}\hat{X}}\mathbf{A}^{\mathrm{T}}}_{\mathbf{Q}_{\hat{L}\hat{L}}}$$

- Cofactor matrix of the residuals:

$$\boxed{\mathbf{Q}_{vv} = \mathbf{Q}_{LL} - \mathbf{Q}_{\hat{L}\hat{L}}}$$

- VCM of the residuals:

$$\mathbf{\Sigma}_{vv} = s_0^2 \cdot \mathbf{Q}_{vv}$$

6.7.3 Linear functional models, Summary

Technische Universität Berlin

Least-squares Adjustment for Linear Adjustment Problems

Linear functional model for the unknowns:

$$\begin{aligned}
L_1 &= a_{11}X_1 + a_{12}X_2 + \cdots + a_{1u}X_u \\
L_2 &= a_{21}X_1 + a_{22}X_2 + \cdots + a_{2u}X_u \\
&\vdots \qquad\qquad \vdots \\
L_n &= a_{n1}X_1 + a_{n2}X_2 + \cdots + a_{nu}X_u
\end{aligned}$$

Vector of observations:

$$\underset{n,1}{\mathbf{L}} = \begin{bmatrix} L_1 & L_2 & \cdots & L_n \end{bmatrix}^{\mathrm{T}}$$

Variance covariance matrix of the observations:

$$\underset{n,n}{\mathbf{\Sigma_{LL}}} = \begin{bmatrix} \sigma_{L_1}^2 & \sigma_{L_1L_2} & \cdots & \sigma_{L_1L_n} \\ \sigma_{L_2L_1} & \sigma_{L_2}^2 & \cdots & \sigma_{L_2L_n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{L_nL_1} & \sigma_{L_nL_2} & \cdots & \sigma_{L_n}^2 \end{bmatrix}$$ with theoretical values σ_i

$$\underset{n,n}{\mathbf{S_{LL}}} = \begin{bmatrix} s_{L_1}^2 & s_{L_1L_2} & \cdots & s_{L_1L_n} \\ s_{L_2L_1} & s_{L_2}^2 & \cdots & s_{L_2L_n} \\ \vdots & \vdots & \ddots & \vdots \\ s_{L_nL_1} & s_{L_nL_2} & \cdots & s_{L_n}^2 \end{bmatrix}$$ with empirical values s_i

6.7.3 Linear functional models, Summary

Theoretical reference standard deviation:	σ_0 (or theoretical reference variance σ_0^2)
Cofactor matrix of the observations:	$\underset{n,n}{\mathbf{Q}_{\mathrm{LL}}} = \frac{1}{\sigma_0^2}\underset{n,n}{\mathbf{\Sigma}_{\mathrm{LL}}}$ respectively $\underset{n,n}{\mathbf{Q}_{\mathrm{LL}}} = \frac{1}{\sigma_0^2}\underset{n,n}{\mathbf{S}_{\mathrm{LL}}}$
Weight matrix of the observations:	$\underset{n,n}{\mathbf{P}} = \underset{n,n}{\mathbf{Q}_{\mathrm{LL}}^{-1}}$
Vector of adjusted unknowns:	$\underset{u,1}{\hat{\mathbf{X}}} = \begin{bmatrix} \hat{X}_1 & \hat{X}_2 & \cdots & \hat{X}_u \end{bmatrix}^{\mathrm{T}}$
Matrix of coefficients of the linear functional model:	$\underset{n,u}{\mathbf{A}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1u} \\ a_{21} & a_{22} & \cdots & a_{2u} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nu} \end{bmatrix}$ "Design Matrix"
Vector of residuals:	$\underset{n,1}{\mathbf{v}} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}^{\mathrm{T}}$
Observation equations:	$\underset{n,1}{\mathbf{L}} + \underset{n,1}{\mathbf{v}} = \underset{n,u}{\mathbf{A}}\,\underset{u,1}{\hat{\mathbf{X}}}$
Normal equations:	$\underset{u,n}{\mathbf{A}^{\mathrm{T}}}\,\underset{n,n}{\mathbf{P}}\,\underset{n,u}{\mathbf{A}}\,\underset{u,1}{\hat{\mathbf{X}}} = \underset{u,n}{\mathbf{A}^{\mathrm{T}}}\,\underset{n,n}{\mathbf{P}}\,\underset{n,1}{\mathbf{L}}$
Normal matrix:	$\underset{u,u}{\mathbf{N}} = \underset{u,n}{\mathbf{A}^{\mathrm{T}}}\,\underset{n,n}{\mathbf{P}}\,\underset{n,u}{\mathbf{A}}$
Right hand side of normal equations:	$\underset{u,1}{\mathbf{n}} = \underset{u,n}{\mathbf{A}^{\mathrm{T}}}\,\underset{n,n}{\mathbf{P}}\,\underset{n,1}{\mathbf{L}}$
Normal equations:	$\underset{u,u}{\mathbf{N}}\,\underset{u,1}{\hat{\mathbf{X}}} = \underset{u,1}{\mathbf{n}}$

6.7.3 Linear functional models, Summary

Technische Universität Berlin

Inversion of normal matrix:	$\underset{u,u}{\mathbf{Q}_{\hat{X}\hat{X}}} = \underset{u,u}{\mathbf{N}^{-1}}$
Solution for the unknowns:	$\underset{u,1}{\hat{\mathbf{X}}} = \underset{u,u}{\mathbf{Q}_{\hat{X}\hat{X}}}\,\underset{u,1}{\mathbf{n}}$
Vector of residuals:	$\underset{n,1}{\mathbf{v}} = \underset{n,u}{\mathbf{A}}\,\underset{u,1}{\hat{\mathbf{X}}} - \underset{n,1}{\mathbf{L}}$
Vector of adjusted observations:	$\underset{n,1}{\hat{\mathbf{L}}} = \underset{n,1}{\mathbf{L}} + \underset{n,1}{\mathbf{v}}$
Final check:	$\underset{n,1}{\hat{\mathbf{L}}} \overset{!}{=} \underset{n,1}{\mathbf{\Phi}}\left(\underset{u,1}{\hat{\mathbf{X}}}\right)$
Empirical reference standard deviation:	$s_0 = \sqrt{\dfrac{\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v}}{n-u}}$ (or empirical reference variance s_0^2)

6.7.3 Linear functional models, Summary

Cofactor matrix of adjusted unknowns:	$\underset{u,u}{\mathbf{Q}_{\hat{\mathbf{X}}\hat{\mathbf{X}}}}$
VCM of adjusted unknowns:	$\underset{u,u}{\mathbf{\Sigma}_{\hat{\mathbf{X}}\hat{\mathbf{X}}}} = s_0^2 \underset{u,u}{\mathbf{Q}_{\hat{\mathbf{X}}\hat{\mathbf{X}}}}$
Cofactor matrix of adjusted observations:	$\underset{n,n}{\mathbf{Q}_{\hat{\mathbf{L}}\hat{\mathbf{L}}}} = \underset{n,u}{\mathbf{A}} \underset{u,u}{\mathbf{Q}_{\hat{\mathbf{X}}\hat{\mathbf{X}}}} \underset{u,n}{\mathbf{A}^{\mathrm{T}}}$
VCM of adjusted observations:	$\underset{n,n}{\mathbf{\Sigma}_{\hat{\mathbf{L}}\hat{\mathbf{L}}}} = s_0^2 \underset{n,n}{\mathbf{Q}_{\hat{\mathbf{L}}\hat{\mathbf{L}}}}$
Cofactor matrix of the residuals:	$\underset{n,n}{\mathbf{Q}_{\mathbf{vv}}} = \underset{n,n}{\mathbf{Q}_{\mathbf{LL}}} - \underset{n,n}{\mathbf{Q}_{\hat{\mathbf{L}}\hat{\mathbf{L}}}}$
VCM of the residuals:	$\underset{n,n}{\mathbf{\Sigma}_{\mathbf{vv}}} = s_0^2 \underset{n,n}{\mathbf{Q}_{\mathbf{vv}}}$

Adjustment_Theory_I_LSA_Linear.pdf

6.7.4 Nonlinear functional models

Example:

Find least squares solution for the following system of nonlinear equations

Functional model:

$$
\begin{aligned}
x + y - 2y^2 &= -4 \\
x^2 + y^2 &= 8 \\
3x^2 - y^2 &= 7.7
\end{aligned}
$$

Stochastic model for the observation:

$$p_1 = 1, \qquad p_2 = 1, \qquad p_3 = 1$$

6.7.4 Nonlinear functional models

Observation equations:

$$-4 + v_1 = \hat{x} + \hat{y} - 2\hat{y}^2$$

$$8 + v_2 = \hat{x}^2 + \hat{y}^2$$

$$7.7 + v_3 = 3\hat{x}^2 - \hat{y}^2$$

Rearranging:

$$v_1 = \hat{x} + \hat{y} - 2\hat{y}^2 + 4$$

$$v_2 = \hat{x}^2 + \hat{y}^2 - 8$$

$$v_3 = 3\hat{x}^2 - \hat{y}^2 - 7.7$$

$$\sum_{i=1}^{n} p_i v_i^2 \rightarrow \min$$

$$\rightarrow 1 \cdot (\hat{x} + \hat{y} - 2\hat{y}^2 + 4)^2 + 1 \cdot (\hat{x}^2 + \hat{y}^2 - 8)^2 + 1 \cdot (3\hat{x}^2 - \hat{y}^2 - 7.7)^2 \rightarrow \min$$

6.7.4 Nonlinear functional models

$$\frac{\partial \Sigma p_i v_i^2}{\partial x} =$$
$$2(\hat{x} + \hat{y} - 2\hat{y}^2 + 4) + 2(\hat{x}^2 + \hat{y}^2 - 8)(2\hat{x}) + 2(3\hat{x}^2 - \hat{y}^2 - 7.7)(6\hat{x}) = 0$$

$$\frac{\partial \Sigma p_i v_i^2}{\partial y} =$$
$$2(\hat{x} + \hat{y} - 2\hat{y}^2 + 4)(1 - 4\hat{y}) + 2(\hat{x}^2 + \hat{y}^2 - 8)(2\hat{y}) + 2(3\hat{x}^2 - \hat{y}^2 - 7.7)(-2\hat{y}) = 0$$

Nonlinear normal equations:

$$\boxed{\begin{array}{c} \hat{x} + \hat{y} - 2\hat{y}^2 + 4 + 2\hat{x}^3 + 2\hat{x}\hat{y}^2 - 16\hat{x} + 18\hat{x}^3 - 6\hat{x}\hat{y}^2 - 46.2\hat{x} = 0 \\ \hat{x} - 6\hat{y}^2 + 4 - 4\hat{x}\hat{y} + 8\hat{y}^3 - 15\hat{y} + 2\hat{x}^2\hat{y} + 2\hat{y}^3 - 16\hat{y} - 6\hat{x}^2\hat{y} + 2\hat{y}^3 + 15.4\hat{y} = 0 \end{array}}$$

<u>nonlinear normal equations</u>

6.7.4 Nonlinear functional models

▶ Nonlinear functional model → nonlinear normal equations

→ Question: How to solve nonlinear normal equations?

a) Direct solution (if possible, e.g. for quadratic equations), or

b) Linearisation (Taylor-Approximation) of normal equations and iterative solution, or

c) Heuristic Approaches, or

d) Replace the original functional model by a linearised functional model and iterative solution

6.7.4 Nonlinear functional models

Solution from linearisation of the original functional model

→ We replace our nonlinear functional model by an approximation (linear functional model)

Attention: We have to compute an iterative solution!

6.7.4 Nonlinear functional models

Nonlinear functional model:

$$\begin{aligned} -4.0 &= x + y - 2y^2 \\ 8.0 &= x^2 + y^2 \\ 7.7 &= 3x^2 - y^2 \end{aligned}$$

$$\mathbf{L}_{n\times 1} = \mathbf{\Phi}_{n\times 1}(\mathbf{X}_{u\times 1}) = \begin{bmatrix} \varphi_1(\mathbf{X}_{u\times 1}) \\ \varphi_2(\mathbf{X}_{u\times 1}) \\ \vdots \\ \varphi_n(\mathbf{X}_{u\times 1}) \end{bmatrix}$$

Question: How can we find a linear approximation?

→ Taylor series approximation for each equation

6.7.4 Nonlinear functional models

Taylor series approximation:

$$\mathbf{L}_{n\times 1} = \mathbf{\Phi}_{n\times 1}\left(\mathbf{X}_{u\times 1}\right) = \underbrace{\mathbf{\Phi}_{n\times 1}(\mathbf{X}^0_{u\times 1}) + \left(\frac{\partial \mathbf{\Phi}(\mathbf{X})}{\partial \mathbf{X}}\right)_{0_{n\times u}} (\mathbf{X} - \mathbf{X}^0)_{u\times 1}}_{\text{First-order Taylor series approximation at the place of the initial values } \mathbf{X}^0} + \cdots$$

Start of iterative computing

$\mathbf{X}^0$: Initial values (starting vales, approximations) for the unknowns

Initial values have to be close to the final solution, otherwise

– computation will not converge, or
– computation can converge to a local min for $\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v} \rightarrow \min$ → worst case!

6.7.4 Nonlinear functional models

We introduce:

$\mathbf{L}^0_{n\times 1} = \mathbf{\Phi}_{n\times 1}(\mathbf{X}^0_{u\times 1})$ observations as functions of the unknowns

$\mathbf{l}_{n\times 1} = \mathbf{L}_{n\times 1} - \mathbf{L}^0_{n\times 1}$ vector of reduced observations

$\mathbf{\Sigma}_{ll} = \mathbf{\Sigma}_{LL}$ from variance covariance propagation
with functional model $\mathbf{l} = \mathbf{L} - \mathbf{L}^0$

$\mathbf{x}_{u\times 1} = \mathbf{X} - \mathbf{X}^0$ vector of reduced unknowns
vector of corrections for the unknowns

$\mathbf{J}_{n\times u} = \left(\frac{\partial \mathbf{\Phi}(\mathbf{X})}{\partial \mathbf{X}}\right)_{0_{n\times u}}$ Jacobian matrix, contains the partial derivatives

of each equation of the functional model

6.7.4 Nonlinear functional models

Linearised functional model:

$$\underbrace{\mathbf{L} - \underbrace{\boldsymbol{\Phi}(\mathbf{X}^0)}_{\mathbf{L}^0}}_{\mathbf{l}} = \underbrace{\left(\frac{\partial \boldsymbol{\Phi}(\mathbf{X})}{\partial \mathbf{X}}\right)_0}_{\mathbf{J}} \underbrace{(\mathbf{X} - \mathbf{X}^0)}_{\mathbf{x}}$$

$$\boxed{\mathbf{l}_{n\times 1} = \mathbf{J}_{n\times u}\mathbf{x}_{u\times 1}}$$

with

$$\mathbf{l}_{n\times 1} = \begin{bmatrix} L_1 - \varphi_1(\mathbf{X}^0) \\ L_2 - \varphi_2(\mathbf{X}^0) \\ \vdots \\ L_n - \varphi_n(\mathbf{X}^0) \end{bmatrix}$$

and with

$$\mathbf{J}_{n\times u} = \begin{bmatrix} \dfrac{\partial\varphi_1(\mathbf{X}^0)}{\partial X_1^0} & \dfrac{\partial\varphi_1(\mathbf{X}^0)}{\partial X_2^0} & \cdots & \dfrac{\partial\varphi_1(\mathbf{X}^0)}{\partial X_u^0} \\ \dfrac{\partial\varphi_2(\mathbf{X}^0)}{\partial X_1^0} & \dfrac{\partial\varphi_2(\mathbf{X}^0)}{\partial X_2^0} & \cdots & \dfrac{\partial\varphi_2(\mathbf{X}^0)}{\partial X_u^0} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial\varphi_n(\mathbf{X}^0)}{\partial X_1^0} & \dfrac{\partial\varphi_n(\mathbf{X}^0)}{\partial X_2^0} & \cdots & \dfrac{\partial\varphi_n(\mathbf{X}^0)}{\partial X_u^0} \end{bmatrix} = \begin{bmatrix} j_{11} & j_{12} & \cdots & j_{1u} \\ j_{21} & j_{22} & \cdots & j_{2u} \\ \vdots & \vdots & \ddots & \vdots \\ j_{n1} & j_{n2} & \cdots & j_{nu} \end{bmatrix}$$

6.7.4 Nonlinear functional models

Now: We take elements of matrix $\mathbf{J}$ and insert them into matrix $\mathbf{A}$

$$\mathbf{A}_{n\times u} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1u} \\ a_{21} & a_{22} & \cdots & a_{2u} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nu} \end{bmatrix}$$

with

$$\begin{array}{cccc} a_{11} = j_{11}, & a_{12} = j_{12}, & \dots, & a_{1u} = j_{1u} \\ & \vdots & & \\ a_{n1} = j_{n1}, & a_{n2} = j_{n2}, & \dots, & a_{nu} = j_{nu} \end{array}$$

6.7.4 Nonlinear functional models

We obtain:

$$\boxed{\mathbf{l}_{n\times 1} = \mathbf{A}_{n\times u}\,\mathbf{x}_{u\times 1}}$$

Advantage of linearisation of the functional model:

We can apply the simple algorithm for the linear case!

Observation equations:

$$\boxed{\mathbf{l}_{n\times 1} + \mathbf{v}_{n\times 1} = \mathbf{A}_{n\times u}\,\hat{\mathbf{x}}_{u\times 1}}$$

with

6.7.4 Nonlinear functional models

Vector of residuals:

$$\mathbf{v}_{n\times 1} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \quad \text{here} \quad \mathbf{v}_{3\times 1} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

Vector of adjusted reduced unknowns:

$$\hat{\mathbf{x}}_{u\times 1} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \vdots \\ \hat{x}_u \end{bmatrix} \quad \text{here} \quad \hat{\mathbf{x}}_{2\times 1} = \begin{bmatrix} d\hat{x} \\ d\hat{y} \end{bmatrix}$$

6.7.4 Nonlinear functional models

$$\Omega = \sum_{i=1}^{n} p_i v_i^2 \rightarrow \min$$

In matrix notation: $\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v} \rightarrow \min$ with $\mathbf{v} = \mathbf{A}\hat{\mathbf{x}} - \mathbf{l}$

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

$$\vdots$$

$$= \hat{\mathbf{x}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\hat{\mathbf{x}} - 2 \cdot \hat{\mathbf{x}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l} + \mathbf{l}^{\mathrm{T}}\mathbf{P}\mathbf{l}$$

Minimum:

$$\frac{\partial \Omega}{\partial \hat{\mathbf{x}}^{\mathrm{T}}} = 2 \cdot \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\hat{\mathbf{x}} - 2 \cdot \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l} = 0$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l}$$

6.7.4 Nonlinear functional models

Normal Equations:

$$\underbrace{\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}}_{\substack{\mathbf{N} \\ u \times u}} \hat{\mathbf{x}} = \underbrace{\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l}}_{\substack{\mathbf{n} \\ u \times 1}}$$

- $\mathbf{N}_{u \times u} = \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}$: Normal matrix
- $\mathbf{n}_{u \times 1} = \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l}$: Right hand side of normal equations

If $\mathbf{N}$ regular → we can compute $\mathbf{N}^{-1}$

Solution of normal equations:

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

$$\hat{\mathbf{x}} = \mathbf{N}^{-1}\mathbf{n}$$

What is $\hat{\mathbf{x}}$? $\rightarrow$ $\hat{\mathbf{x}}$ is a correction for the initial values $\mathbf{X}^0$

6.7.4 Nonlinear functional models

Adjusted unknowns:

$$\boxed{\hat{\mathbf{X}} = \mathbf{X}^0 + \hat{\mathbf{x}}}$$

Residuals:

$$\boxed{\mathbf{v} = \mathbf{A}\hat{\mathbf{x}} - \mathbf{l}}$$

Adjusted observations:

$$\boxed{\hat{\mathbf{L}} = \mathbf{L} + \mathbf{v}}$$

6.7.4 Nonlinear functional models

Attention: $\widehat{\mathbf{X}}$ is solution for the linearised problem (approximation)

→ We have to perform iterative computation to reach final solution

→ We have to introduce $\widehat{\mathbf{X}}$ as new approximation $\mathbf{X}^0$ and perform the adjustment again and again!

Question: How often do we have to repeat the computation?

Answer: Until

a) the corrections for the unknowns $\hat{\mathbf{x}}$ become <u>sufficiently small</u>

b) final check $\hat{\mathbf{L}} - \mathbf{\Phi}(\widehat{\mathbf{X}}) \overset{!}{\approx} \mathbf{0}$

6.7.4 Nonlinear functional models

a) „sufficiently small“

If our unknowns are coordinates and we want to obtain results with $\mathrm{mm}\ (10^{-3}\mathrm{m}) \rightarrow \max|\hat{x}_i|$ for $i = 1, \ldots, u \leq 10^{-5}$

Rule: $\max|\hat{x}_i|$ should be two orders of magnitude smaller than the resolution of the result

b) $\hat{\mathbf{L}} - \boldsymbol{\Phi}(\hat{\mathbf{X}}) \overset{!}{\approx} \mathbf{0}$ $\boldsymbol{\Phi}(\hat{\mathbf{X}})$: Original nonlinear functional model

Result contains:

- Computation error
- Linearisation error

→ "How far away is my solution for the linearized problem from the solution for the nonlinear problem"

6.7.4 Nonlinear functional models

<u>Rule</u>:

$\max\left|\hat{L}_i - \Phi_i(\hat{\mathbf{X}})\right|$ should be smaller than 10^{-8}

→ If these break-off conditions are not met,
<u>continue</u> with iterative computation

→ **Attention:** Introduce $\hat{\mathbf{X}}$ as new starting values $\mathbf{X}^0$
and start iterative computing at step " "

6.7.5 Variance-Covariance Matrices (VCM) for the results

Precision measures (see also Section 6.7.2):

- Empirical reference standard deviation: $s_0 = \sqrt{\frac{\mathbf{v}^\mathrm{T}\mathbf{P}\mathbf{v}}{n-u}}$
- $\mathbf{Q}_{\hat{X}\hat{X}} = \mathbf{Q}_{\hat{x}\hat{x}}$ from variance covariance propagation
  with functional model $\widehat{\mathbf{X}} = \mathbf{X}^0 + \hat{\mathbf{x}}$
- VCM of adjusted unknowns: $\mathbf{\Sigma}_{\hat{X}\hat{X}} = s_0^2\mathbf{Q}_{\hat{X}\hat{X}}$ with $\mathbf{Q}_{\hat{X}\hat{X}} = \mathbf{N}^{-1}$
- Cofactor matrix of adjusted observations: $\mathbf{Q}_{\hat{L}\hat{L}} = \mathbf{A}\mathbf{Q}_{\hat{X}\hat{X}}\mathbf{A}^\mathrm{T}$
- VCM of adjusted observations: $\mathbf{\Sigma}_{\hat{L}\hat{L}} = s_0^2\mathbf{Q}_{\hat{L}\hat{L}}$
- Cofactor matrix of residuals: $\mathbf{Q}_{vv} = \mathbf{Q}_{LL} - \mathbf{Q}_{\hat{L}\hat{L}}$
- VCM of residuals: $\mathbf{\Sigma}_{vv} = s_0^2\mathbf{Q}_{vv}$

6.7.6 Nonlinear functional models, Example

<u>Example</u>: Find least squares solution for the unknown parameters and compute their standard deviation

Functional model:

$$
\begin{aligned}
\varphi_1: &\quad 4.0 = x + y - 2y^2 && (1)\\
\varphi_2: &\quad 8.0 = x^2 + y^2 && (2)\\
\varphi_3: &\quad 7.7 = 3x^2 - y^2 && (3)
\end{aligned}
$$

Values $4.0, 8.0, 7.7$ are equally weighted and uncorrelated observations

Parameters x, y are unknowns

Observation vector:

$$\mathbf{L} = \begin{bmatrix} 4.0 \\ 8.0 \\ 7.7 \end{bmatrix}$$

6.7.6 Nonlinear functional models, Example

Stochastic model of the observations:

$$p_1 = 1, \quad p_2 = 1, \quad p_3 = 1$$

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{I}$$

Vector of adjusted unknowns:

$$\hat{\mathbf{X}} = \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix}$$

Nonlinear functional model → solution from iterative computing with linearised functional model → introduction of approximate values x^0, y^0

6.7.6 Nonlinear functional models, Example

Vector of starting values:

$$\mathbf{X}^0 = \begin{bmatrix} x^0 \\ y^0 \end{bmatrix}$$

Vector of adjusted reduced unknowns:

$$\hat{\mathbf{x}} = \begin{bmatrix} d\hat{x} \\ d\hat{y} \end{bmatrix} = \hat{\mathbf{X}} - \mathbf{X}^0 = \begin{bmatrix} \hat{x} - x^0 \\ \hat{y} - y^0 \end{bmatrix}$$

Vector of reduced observations:

$$\mathbf{l} = \begin{bmatrix} 4.0 - (x^0 + y^0 - 2(y^0)^2) \\ 8.0 - ((x^0)^2 + (y^0)^2) \\ 7.7 - (3(x^0)^2 - (y^0)^2) \end{bmatrix}$$

6.7.6 Nonlinear functional models, Example

Technische Universität Berlin

Jacobian matrix:

$$\mathbf{J} = \begin{array}{c} \\ \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{array} \begin{array}{c} \begin{array}{cc} x^0 & y^0 \end{array} \\ \begin{bmatrix} \dfrac{\partial \varphi_1}{\partial x^0} & \dfrac{\partial \varphi_1}{\partial y^0} \\ \dfrac{\partial \varphi_2}{\partial x^0} & \dfrac{\partial \varphi_2}{\partial y^0} \\ \dfrac{\partial \varphi_3}{\partial x^0} & \dfrac{\partial \varphi_3}{\partial y^0} \end{bmatrix} \end{array} = \begin{array}{c} \\ \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{array} \begin{array}{c} \begin{array}{cc} x^0 & y^0 \end{array} \\ \begin{bmatrix} 1 & (1-4y^0) \\ 2x^0 & 2y^0 \\ 6x^0 & -2y^0 \end{bmatrix} \end{array}$$

Design matrix:

$$\mathbf{A} = \begin{bmatrix} 1 & (1-4y^0) \\ 2x^0 & 2y^0 \\ 6x^0 & -2y^0 \end{bmatrix}$$

6.7.6 Nonlinear functional models, Example

Normal equations:

$$\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l}$$

Solutions of normal equations:

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

Adjusted unknowns:

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

Residuals:

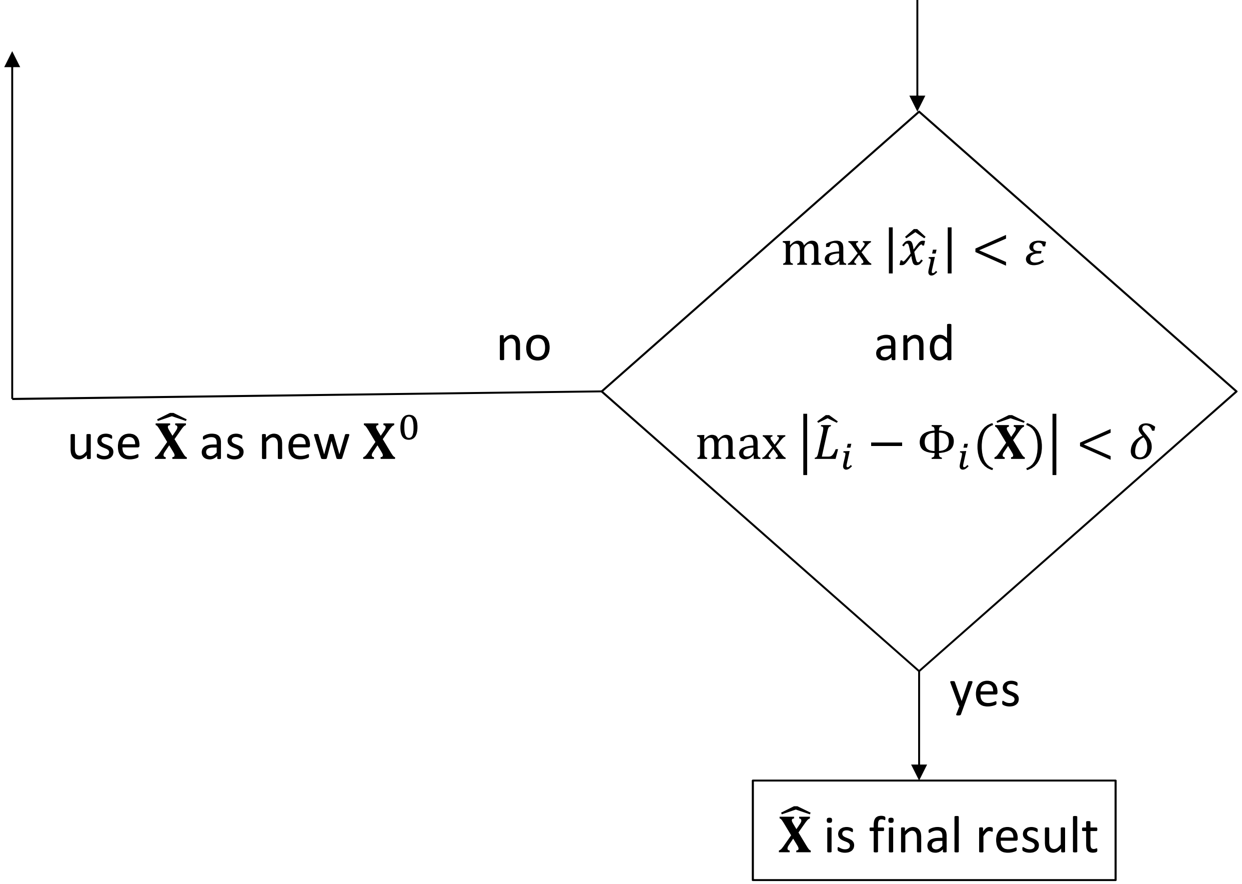
$$\mathbf{v} = \mathbf{A}\hat{\mathbf{x}} - \mathbf{l}$$

Adjusted observations:

$$\hat{\mathbf{L}} = \mathbf{L} + \mathbf{v}$$

6.7.6 Nonlinear functional models, Example

Check:



6.7.6 Nonlinear functional models, Example

Empirical reference standard deviation:

$$s_0 = \sqrt{\frac{\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v}}{n-u}}$$

VCM of adjusted unknowns:

$$\mathbf{\Sigma}_{\hat{X}\hat{X}} = s_0^2 \cdot \mathbf{Q}_{\hat{X}\hat{X}} \qquad \text{with} \qquad \mathbf{Q}_{\hat{X}\hat{X}} = \mathbf{N}^{-1}$$

6.7.6 Nonlinear functional models, Example

Standard deviation of unknowns:

$$\mathbf{\Sigma}_{\hat{X}\hat{X}} = s_0^2 \cdot \begin{bmatrix} q_{\hat{x}\hat{x}} & q_{\hat{x}\hat{y}} \\ q_{\hat{y}\hat{x}} & q_{\hat{y}\hat{y}} \end{bmatrix}$$

- Cofactor for unknown value x: $q_{\hat{x}\hat{x}}$
- Cofactor for unknown value y: $q_{\hat{y}\hat{y}}$

$$s_{\hat{x}} = \sqrt{s_0^2 \cdot q_{\hat{x}\hat{x}}} = s_0 \cdot \sqrt{q_{\hat{x}\hat{x}}}$$

$$s_{\hat{y}} = \sqrt{s_0^2 \cdot q_{\hat{y}\hat{y}}} = s_0 \cdot \sqrt{q_{\hat{y}\hat{y}}}$$

6.7.7 Nonlinear functional models, Summary

Least-squares Adjustment for Nonlinear Adjustment Problems

- Iterative solution with linearized functional model -

Nonlinear functional model for the unknowns:

$$\begin{aligned} L_1 &= \varphi_1(X_1, X_2, \cdots, X_u) \\ L_2 &= \varphi_2(X_1, X_2, \cdots, X_u) \\ \vdots & \qquad \vdots \\ L_n &= \varphi_n(X_1, X_2, \cdots, X_u) \end{aligned}$$

Nonlinear functional model in matrix notation:

$$\underset{n,1}{\mathbf{L}} = \underset{n,1}{\mathbf{\Phi}}(\underset{u,1}{\mathbf{X}}) = \begin{bmatrix} \varphi_1(\underset{u,1}{\mathbf{X}}) \\ \varphi_2(\underset{u,1}{\mathbf{X}}) \\ \vdots \\ \varphi_n(\underset{u,1}{\mathbf{X}}) \end{bmatrix}$$

Vector of observations:

$$\underset{n,1}{\mathbf{L}} = [L_1 \quad L_2 \quad \cdots \quad L_n]^{\mathrm{T}}$$

Variance covariance matrix of the observations:

$$\underset{n,n}{\mathbf{\Sigma}_{\mathbf{LL}}} = \begin{bmatrix} \sigma_{L_1}^2 & \sigma_{L_1L_2} & \cdots & \sigma_{L_1L_n} \\ \sigma_{L_2L_1} & \sigma_{L_2}^2 & \cdots & \sigma_{L_2L_n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{L_nL_1} & \sigma_{L_nL_2} & \cdots & \sigma_{L_n}^2 \end{bmatrix}$$ with theoretical values σ_i

$$\underset{n,n}{\mathbf{S}_{\mathbf{LL}}} = \begin{bmatrix} s_{L_1}^2 & s_{L_1L_2} & \cdots & s_{L_1L_n} \\ s_{L_2L_1} & s_{L_2}^2 & \cdots & s_{L_2L_n} \\ \vdots & \vdots & \ddots & \vdots \\ s_{L_nL_1} & s_{L_nL_2} & \cdots & s_{L_n}^2 \end{bmatrix}$$ with empirical values s_i

6.7.7 Nonlinear functional models, Summary

VCM of the reduced observations from variance covariance propagation with the functional model $\mathbf{l} = \mathbf{L} - \mathbf{L}^0$	$\underset{n,n}{\boldsymbol{\Sigma}_{\mathbf{ll}}} = \underset{n,n}{\boldsymbol{\Sigma}_{\mathbf{LL}}}$	with theoretical values σ_i
	$\underset{n,n}{\mathbf{S}_{\mathbf{ll}}} = \underset{n,n}{\mathbf{S}_{\mathbf{LL}}}$	with empirical values s_i
Theoretical reference standard deviation:	σ_0 (or theoretical reference variance σ_0^2)	
Cofactor matrix of the observations and reduced observations:	$\underset{n,n}{\mathbf{Q}_{\mathbf{LL}}} = \frac{1}{\sigma_0^2} \underset{n,n}{\boldsymbol{\Sigma}_{\mathbf{LL}}}$ respectively	$\underset{n,n}{\mathbf{Q}_{\mathbf{LL}}} = \frac{1}{\sigma_0^2} \underset{n,n}{\mathbf{S}_{\mathbf{LL}}}$
Weight matrix of the observations and reduced observations:	$\underset{n,n}{\mathbf{P}} = \underset{n,n}{\mathbf{Q}_{\mathbf{LL}}^{-1}}$	
Vector of adjusted unknowns:	$\underset{u,1}{\hat{\mathbf{X}}} = \begin{bmatrix} \hat{X}_1 & \hat{X}_2 & \cdots & \hat{X}_u \end{bmatrix}^{\mathrm{T}}$	
Vector of $\begin{cases} \text{initial values} \\ \text{starting values} \\ \text{approximations} \end{cases}$ for the unknowns:	$\underset{u,1}{\mathbf{X}^0} = \begin{bmatrix} X_1^0 & X_2^0 & \cdots & X_u^0 \end{bmatrix}^{\mathrm{T}}$	
Vector of adjusted reduced unknowns:	$\underset{u,1}{\hat{\mathbf{x}}} = \underset{u,1}{\hat{\mathbf{X}}} - \underset{u,1}{\mathbf{X}^0}$	

6.7.7 Nonlinear functional models, Summary

→ Observations as functions of the approximations for the unknowns: $\underset{n,1}{\mathbf{L}^0} = \underset{n,1}{\mathbf{\Phi}}\left(\underset{u,1}{\mathbf{X}^0}\right)$

Vector of reduced observations: $\underset{n,1}{\mathbf{l}} = \underset{n,1}{\mathbf{L}} - \underset{n,1}{\mathbf{L}^0}$

Jacobian matrix:

$$\underset{n,u}{\mathbf{J}} = \left(\frac{\partial \mathbf{\Phi}(\mathbf{X})}{\partial \mathbf{X}}\right)_{\mathbf{X}=\mathbf{X}^0} = \begin{bmatrix} \dfrac{\partial \varphi_1(\mathbf{X})}{\partial X_1} & \dfrac{\partial \varphi_1(\mathbf{X})}{\partial X_2} & \cdots & \dfrac{\partial \varphi_1(\mathbf{X})}{\partial X_u} \\ \dfrac{\partial \varphi_2(\mathbf{X})}{\partial X_1} & \dfrac{\partial \varphi_2(\mathbf{X})}{\partial X_2} & \cdots & \dfrac{\partial \varphi_2(\mathbf{X})}{\partial X_u} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial \varphi_n(\mathbf{X})}{\partial X_1} & \dfrac{\partial \varphi_n(\mathbf{X})}{\partial X_2} & \cdots & \dfrac{\partial \varphi_n(\mathbf{X})}{\partial X_u} \end{bmatrix}_{\mathbf{X}=\mathbf{X}^0}$$

Coefficient matrix of the linearized functional model: $\underset{n,u}{\mathbf{A}} = \underset{n,u}{\mathbf{J}}$ "Design Matrix"

Vector of residuals: $\underset{n,1}{\mathbf{v}} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}^{\mathrm{T}}$

Observation equations: $\underset{n,1}{\mathbf{l}} + \underset{n,1}{\mathbf{v}} = \underset{n,u}{\mathbf{A}}\,\underset{u,1}{\hat{\mathbf{x}}}$

Normal equations: $\underset{u,n}{\mathbf{A}^{\mathrm{T}}}\,\underset{n,n}{\mathbf{P}}\,\underset{n,u}{\mathbf{A}}\,\underset{u,1}{\hat{\mathbf{x}}} = \underset{u,n}{\mathbf{A}^{\mathrm{T}}}\,\underset{n,n}{\mathbf{P}}\,\underset{n,1}{\mathbf{l}}$

Normal matrix: $\underset{u,u}{\mathbf{N}} = \underset{u,n}{\mathbf{A}^{\mathrm{T}}}\,\underset{n,n}{\mathbf{P}}\,\underset{n,u}{\mathbf{A}}$

6.7.7 Nonlinear functional models, Summary

Right hand side of normal equations: $\underset{u,1}{\mathbf{n}} = \underset{u,n}{\mathbf{A}^{\mathrm{T}}} \underset{n,n}{\mathbf{P}} \underset{n,1}{\mathbf{l}}$

Normal equations: $\underset{u,u}{\mathbf{N}} \underset{u,1}{\hat{\mathbf{x}}} = \underset{u,1}{\mathbf{n}}$

Inversion of normal matrix: $\underset{u,u}{\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}} = \underset{u,u}{\mathbf{N}^{-1}}$

Solution of normal equations: $\underset{u,1}{\hat{\mathbf{x}}} = \underset{u,u}{\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}} \underset{u,1}{\mathbf{n}}$

Solution for the unknowns: $\underset{u,1}{\hat{\mathbf{X}}} = \underset{u,1}{\hat{\mathbf{X}}^0} + \underset{u,1}{\hat{\mathbf{x}}}$

Vector of residuals: $\underset{n,1}{\mathbf{v}} = \underset{n,u}{\mathbf{A}} \underset{u,1}{\hat{\mathbf{x}}} - \underset{n,1}{\mathbf{l}}$

Vector of adjusted observations: $\underset{n,1}{\hat{\mathbf{L}}} = \underset{n,1}{\mathbf{L}} + \underset{n,1}{\mathbf{v}}$

Check 1: $\max|\hat{x}_i| \le \varepsilon \quad \forall\, i = 1, \ldots, u$ MatLab: `max(abs(x_hat))`

Check 2: $\max\left|\hat{L}_i - \varphi_i\left(\hat{\mathbf{X}}\right)\right| \le \delta \quad \forall\, i = 1, \ldots, n$

If $\left\{ \left(\max|\hat{x}_i| \le \varepsilon \quad \forall\, i = 1, \ldots, u\right) \wedge \left(\max\left|\hat{L}_i - \varphi_i\left(\hat{\mathbf{X}}\right)\right| \le \delta \quad \forall\, i = 1, \ldots, n\right) \right\}$

$\hat{\mathbf{X}}$ is the solution for the nonlinear adjustment problem

Else

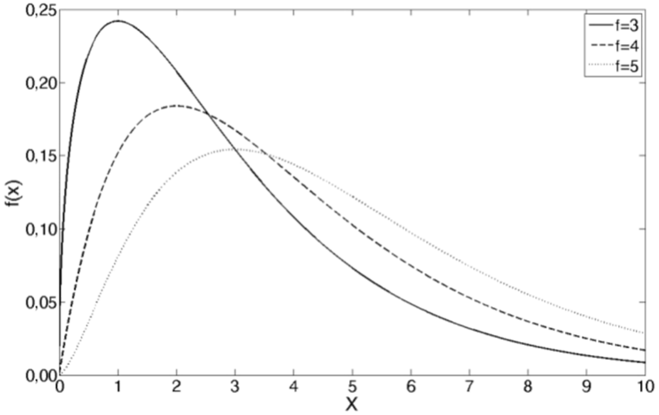
Use $\hat{\mathbf{X}}$ as new approximation for the unknowns $\mathbf{X}^0$ and continue with step "→"

6.7.7 Nonlinear functional models, Summary

Technische Universität Berlin

Empirical reference standard deviation:	$s_0 = \sqrt{\dfrac{\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v}}{n-u}}$ (or empirical reference variance s_0^2)
Cofactor matrix of adjusted unknowns from variance covariance propagation with the functional model $\hat{\mathbf{X}} = \mathbf{X}^0 + \hat{\mathbf{x}}$	$\underset{n,n}{\mathbf{Q}_{\hat{\mathbf{X}}\hat{\mathbf{X}}}} = \underset{n,n}{\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}}$
VCM of adjusted unknowns:	$\underset{u,u}{\mathbf{\Sigma}_{\hat{\mathbf{X}}\hat{\mathbf{X}}}} = s_0^2 \underset{u,u}{\mathbf{Q}_{\hat{\mathbf{X}}\hat{\mathbf{X}}}}$
Cofactor matrix of adjusted observations:	$\underset{n,n}{\mathbf{Q}_{\hat{\mathbf{L}}\hat{\mathbf{L}}}} = \underset{n,u}{\mathbf{A}} \underset{u,u}{\mathbf{Q}_{\hat{\mathbf{X}}\hat{\mathbf{X}}}} \underset{u,n}{\mathbf{A}^{\mathrm{T}}}$
VCM of adjusted observations:	$\underset{n,n}{\mathbf{\Sigma}_{\hat{\mathbf{L}}\hat{\mathbf{L}}}} = s_0^2 \underset{n,n}{\mathbf{Q}_{\hat{\mathbf{L}}\hat{\mathbf{L}}}}$
Cofactor matrix of the residuals:	$\underset{n,n}{\mathbf{Q}_{\mathbf{vv}}} = \underset{n,n}{\mathbf{Q}_{\mathbf{LL}}} - \underset{n,n}{\mathbf{Q}_{\hat{\mathbf{L}}\hat{\mathbf{L}}}}$
VCM of the residuals:	$\underset{n,n}{\mathbf{\Sigma}_{\mathbf{vv}}} = s_0^2 \underset{n,n}{\mathbf{Q}_{\mathbf{vv}}}$

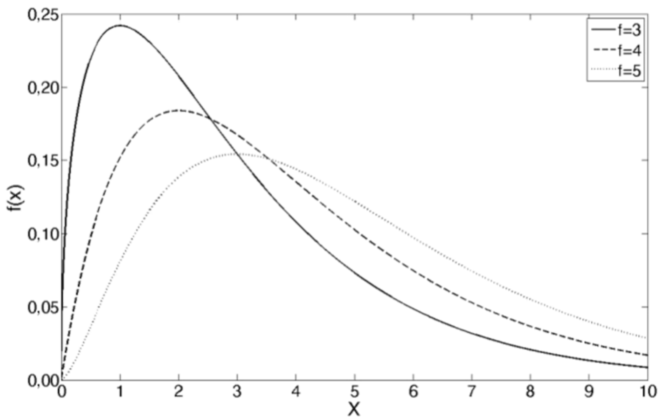
Adjustment_Theory_I_LSA_Nonlinear.pdf



Adjustment Theory I

Chapter 6 – Introduction to Least Squares Adjustment

Prof. Dr.-Ing Frank Neitzel | Institute of Geodesy and Geoinformation Science



Technische Universität Berlin

Adjustment Theory I

Chapter 7 – Applications of Least Squares Adjustment

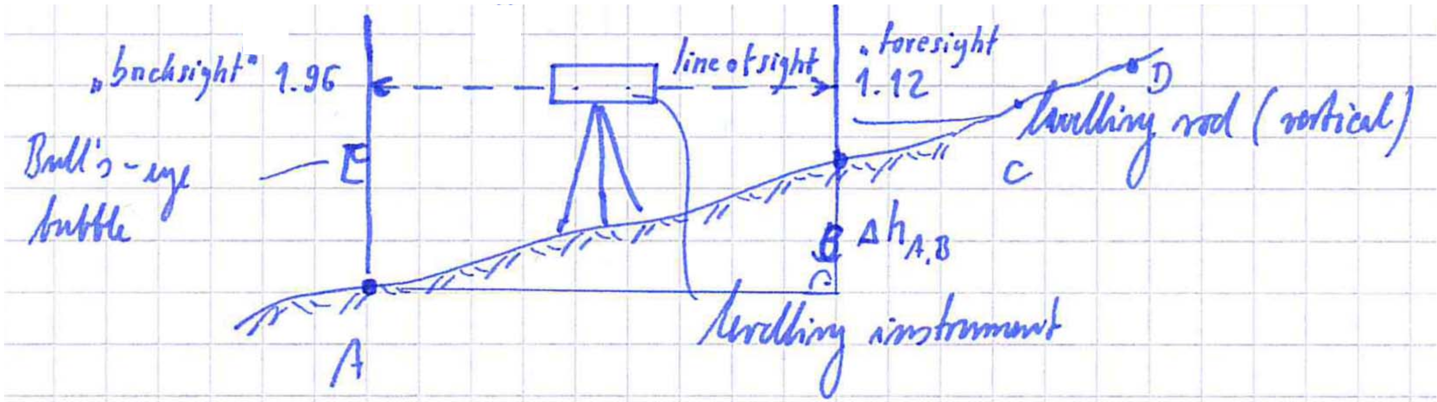
Prof. Dr.-Ing Frank Neitzel | Institute of Geodesy and Geoinformation Science

Version: 09.12.2024

Contents

7.1 Adjustment of Levelling Networks

Basic idea of differential levelling:
Determination of the elevation difference between points A and B by measuring the vertical distance from a horizontal line of sight



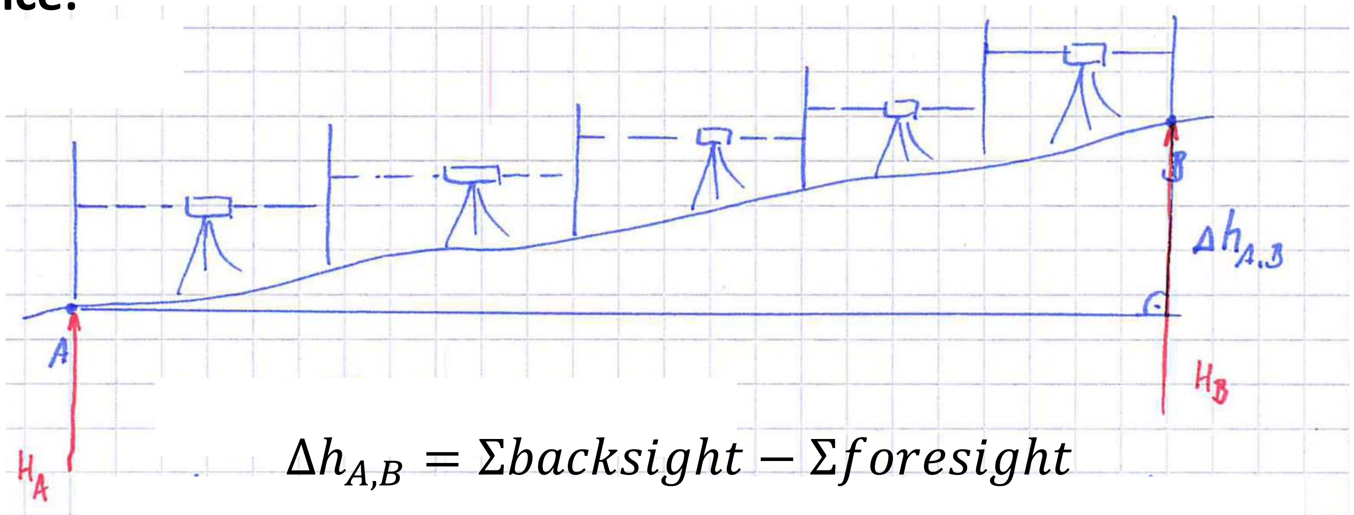
Levelling instrument ≙ Instrument capable of generating a horizontal line of sight!

Elevation difference between points A and B: $1.96 - 1.12 = 0.84$

$$\text{"backsight"} - \text{"foresight"} = \Delta h$$

7.1 Adjustment of Levelling Networks

In practice:



$$\Delta h_{A,B} = \Sigma backsight - \Sigma foresight$$

→ Δh is regarded as observation with corresponding precision / weight

For an adjustment within the Gauss-Markov Model (parametric adjustment)

$$\mathbf{L} = \boldsymbol{\Phi}(\mathbf{X})$$

we have to introduce appropriate unknowns to express our observations as functions of unknowns

7.1 Adjustment of Levelling Networks

Question: What are appropriate unknowns in a levelling network (height network)?

Answer: Heights! (1D-coordinates)

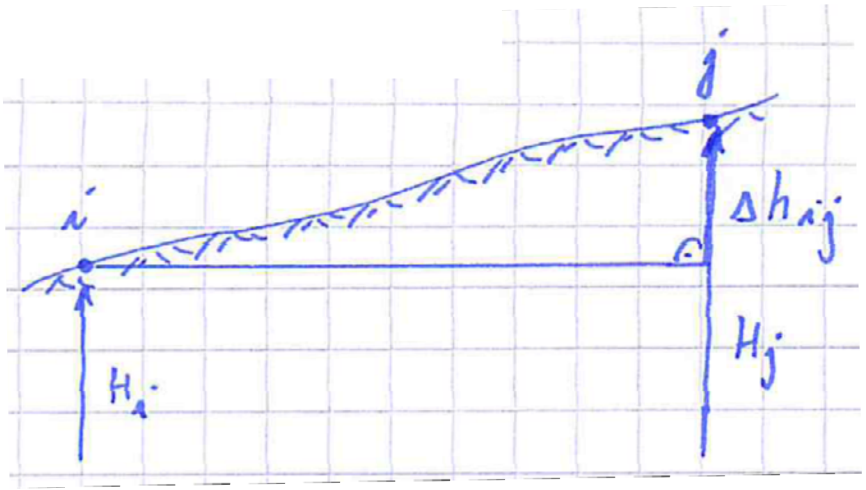
$$\underbrace{\Delta h_{A,B}}_{\text{observation}} = \underbrace{H_B - H_A}_{\text{unknowns}}$$

Please note:

In many cases the introduction of coordinates as unknowns is recommended in the Gauss-Markov Model

7.1 Adjustment of Levelling Networks

Functional model:



Note:

$\Delta h_{i,j}$ → relative measurements

H_j, H_i → absolute values

$$\boxed{\Delta h_{ij} = H_j - H_i}$$

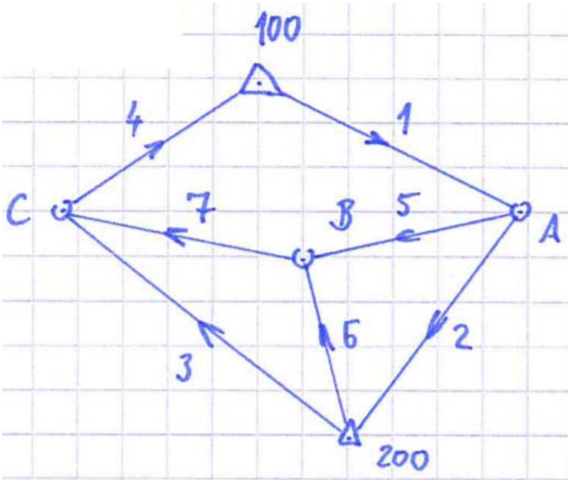
Observation equations:

Height differences are introduced as observations. As with all observations these are subject to random deviations → We have to introduce residuals $v_{\Delta h_{ij}}$

$$\boxed{\Delta h_{ij} + v_{\Delta h_{ij}} = H_j - H_i}$$

7.1 Adjustment of Levelling Networks

Example 1:



Given:

- Heights of the benchmark points

Point No.	Height H [m]
100	100.000
200	107.500

- Observed elevation differences

Line	Observed elevation diff. [m]
1	5.100
2	2.340
3	−1.250
4	−6.130
5	−0.680
6	−3.000
7	1.700

The heights of the benchmark points can be regarded as fixed (error-free) values, the observations are equally weighted and uncorrelated

Determine the adjusted heights of points A, B, C and their standard deviation

7.1 Adjustment of Levelling Networks

General considerations:

- What are our unknowns?

  → Heights of points A, B, C

  → We introduce H_A, H_B, H_C

- What are our observations?

  → Elevation differences

  → $\Delta h_{100,A}, \Delta h_{A,200}, \Delta h_{200,C}, \Delta h_{C,100}, \Delta h_{A,B}, \Delta h_{200,B}, \Delta h_{B,C}$

- What are our fixed values?

  → H_{100}, H_{200}

- Redundancy?

  → $r = n - u \ \rightarrow r = 7 - 3 \ \rightarrow r = 4$

7.1 Adjustment of Levelling Networks

Functional model:

$$\Delta h_{100,A} = H_A - H_{100}$$
$$\Delta h_{A,200} = H_{200} - H_A$$
$$\Delta h_{200,C} = H_C - H_{200}$$
$$\Delta h_{C,100} = H_{100} - H_C$$
$$\Delta h_{A,B} = H_B - H_A$$
$$\Delta h_{200,B} = H_B - H_{200}$$
$$\Delta h_{B,C} = H_C - H_B$$

We insert the fixed values for H_{100} and H_{200} and bring them to the left-hand side of the equations …

… Functional model

$$\begin{aligned}
&\varphi_1\text{: } && \Delta h_{100,A} + 100.000 = H_A \\
&\varphi_2\text{: } && \Delta h_{A,200} - 107.500 = -H_A \\
&\varphi_3\text{: } && \Delta h_{200,C} + 107.500 = H_C \\
&\varphi_4\text{: } && \Delta h_{C,100} - 100.000 = -H_C \\
&\varphi_5\text{: } && \Delta h_{A,B} \qquad\qquad\qquad = H_B - H_A \\
&\varphi_6\text{: } && \Delta h_{200,B} + 107.500 = H_B \\
&\varphi_7\text{: } && \underbrace{\Delta h_{B,C} \qquad\qquad\quad}_{\text{observation vector } \mathbf{L}'} = H_C - H_B
\end{aligned}$$

Linear or nonlinear? → Linear!

Observation equations:

$$
\begin{aligned}
\Delta h_{100,A} + 100.000 + v_1 &= \widehat{H}_A \\
\Delta h_{A,200} - 107.500 + v_2 &= -\widehat{H}_A \\
\Delta h_{200,C} + 107.500 + v_3 &= \widehat{H}_C \\
\Delta h_{C,100} - 100.000 + v_4 &= -\widehat{H}_C \\
\Delta h_{A,B} \qquad\qquad\quad + v_5 &= \widehat{H}_B - \widehat{H}_A \\
\Delta h_{200,B} + 107.500 + v_6 &= \widehat{H}_B \\
\Delta h_{B,C} \qquad\qquad\quad + v_7 &= \widehat{H}_C - \widehat{H}_B
\end{aligned}
$$

Observation vector:

$$\mathbf{L}' = \begin{bmatrix} \Delta h_{100,A} + 100.000 \\ \Delta h_{A,200} - 107.500 \\ \Delta h_{200,C} + 107.500 \\ \Delta h_{C,100} - 100.000 \\ \Delta h_{A,B} \\ \Delta h_{200,B} + 107.500 \\ \Delta h_{B,C} \end{bmatrix}$$

Vector of residuals:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{bmatrix}$$

7.1 Adjustment of Levelling Networks

Stochastic Model of the observations:

$$p_1 = 1, p_2 = 1, \dots, p_7 = 1 \;\rightarrow \mathbf{P} = \mathbf{I}$$

Vector of unknowns:

$$\hat{\mathbf{X}} = \begin{bmatrix} \hat{H}_A \\ \hat{H}_B \\ \hat{H}_C \end{bmatrix}$$

Design Matrix (Matrix with coefficients of the linear functional model):

$$\mathbf{A} = \begin{array}{c} \\ \varphi_1 \\ \varphi_2 \\ \varphi_3 \\ \varphi_4 \\ \varphi_5 \\ \varphi_6 \\ \varphi_7 \end{array} \begin{array}{c} \begin{array}{ccc} H_A & H_B & H_C \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \end{array}$$

7.1 Adjustment of Levelling Networks

Normal equations:

$$\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\widehat{\mathbf{X}} = \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{L}' \qquad \text{with} \qquad \mathbf{P} = \mathbf{I}$$

$$\widehat{\mathbf{X}} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{L}' \qquad \text{with} \qquad \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1} = \mathbf{Q}_{\hat{X}\hat{X}}$$

Residuals:

$$\mathbf{v} = \mathbf{A}\widehat{\mathbf{X}} - \mathbf{L}'$$

Adjusted observations:

$$\hat{\mathbf{L}}' = \mathbf{L}' + \mathbf{v}$$

Final check:

$$\hat{\mathbf{L}}' - \boldsymbol{\Phi}(\hat{\mathbf{X}}) \overset{!}{=} \mathbf{0}$$ → zero within computing precision

Computer: $\hat{\mathbf{L}}' - \boldsymbol{\Phi}(\hat{\mathbf{X}}) \leq \delta$ → e.g. 10^{-12}

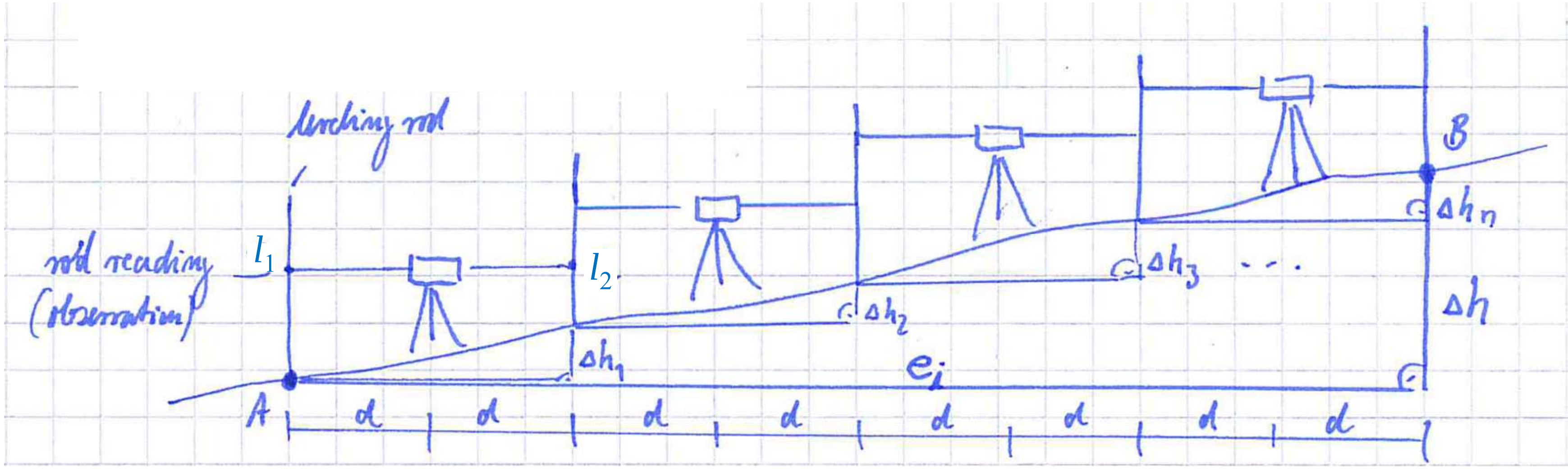
Empirical reference standard deviation:

$$s_0 = \sqrt{\frac{\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v}}{n-u}}$$

VCM of adjusted unknowns:

$$\boldsymbol{\Sigma}_{\hat{X}\hat{X}} = \mathrm{s}_0^2 \cdot \mathbf{Q}_{\hat{X}\hat{X}}$$

Weights in Differential Levelling



d: Length of sight distance

e_i: Length of the course between surveyed points, here points A and B

7.1 Adjustment of Levelling Networks

- Standard deviation in rod reading is usually expressed as ratio of the estimated standard error in rod reading per unit sight distance length
  → Standard deviation for rod reading

$$\sigma_{l_i} = d_i \cdot \sigma_{l/d}$$

d_i: length of sight distance

$$\Delta h_1 = l_1 - l_2$$

with equal length of sight distance

$$\sigma_{l_1} = \sigma_{l_2} = d \cdot \sigma_{l/d}$$

$$\rightarrow \sigma_{\Delta h_1}^2 = \left(d \cdot \sigma_{l/d}\right)^2 + \left(d \cdot \sigma_{l/d}\right)^2$$

$$= 2 \cdot d^2 \cdot \sigma_{l/d}^2$$

7.1 Adjustment of Levelling Networks

- Height difference Δh:

$$\Delta h = \Delta h_1 + \Delta h_2 + \cdots + \Delta h_n$$

$$\rightarrow \quad \sigma_{\Delta h}^2 = 2 \cdot n \cdot d^2 \cdot \sigma_{l/d}^2$$

- n as a function of length of sight distance d and length of the course e_i

$$n = \frac{e_i}{2d}$$

Example: $e_i = 200\ \mathrm{m},\ \ d = 25\ \mathrm{m} \quad \rightarrow \quad n = \frac{200\ \mathrm{m}}{2{\cdot}25\ \mathrm{m}} = 4$

$$\rightarrow \quad \sigma_{\Delta h}^2 = 2 \cdot \frac{e_i}{2d} \cdot d^2 \cdot \sigma_{l/d}^2 = e_i \cdot d \cdot \sigma_{l/d}^2$$

with $d, \sigma_{l/d}$ constant values, we introduce $k = d \cdot \sigma_{l/d}^2$

$$\boxed{\sigma_{\Delta h}^2 = e_i \cdot k}$$

7.1 Adjustment of Levelling Networks

- We know: Weights are the inverse values of the variances

$$p_i = \frac{1}{e_i \cdot k}$$

Since k is a constant and weights are relative, equation can be simplified to

$$\boxed{p_i = \frac{1}{e_i}}$$

Weights of different levelling lines are inversely proportional to their length. And since any course length is proportional to its number of instrument setups, weights are also inversely proportional to the number of instrument setups.

7.1 Adjustment of Levelling Networks

Example 1:

Line	Length e [km]	Weight p	Rel. weights p
1	4	0.25	3
2	3	$0.\overline{3}$	4
3	2	0.5	6
4	3	$\vdots$	$\vdots$
5	2		
6	2		
7	2		

Remember: Weights are relative

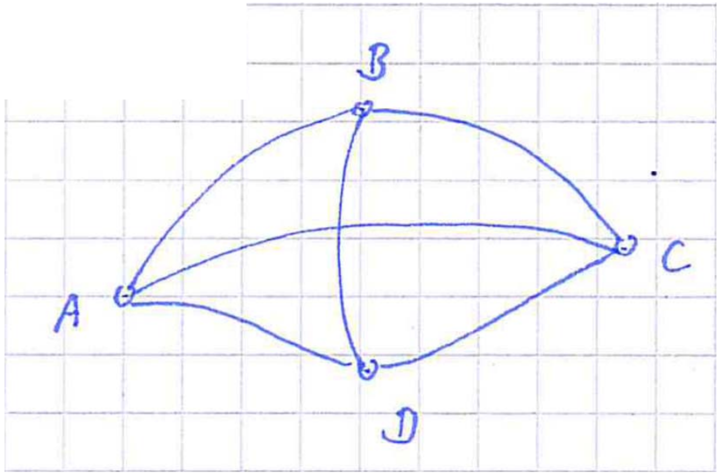
→ We can introduce relative weights

7.1 Adjustment of Levelling Networks

Example 2:

Given: Observed elevation differences and their standard deviations

From	To	Δh [m]	σ [m]
A	B	10.509	0.006
B	C	5.360	0.004
C	D	−8.235	0.005
D	A	−7.348	0.003
B	D	−3.167	0.004
A	C	15.881	0.012



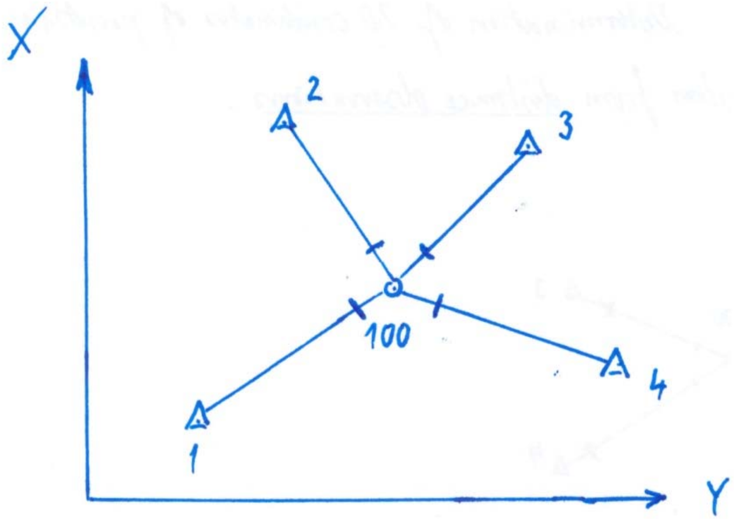
Determine the adjusted observations and their standard deviation

Problem: No height(s) of benchmark point(s) given $\rightarrow$ Solution?

7.2 Adjustment of Horizontal Surveys: Trilateration

Basic idea of trilateration:

Determination of 2D coordinates of points in a plane Cartesian coordinate system from <u>distance observations</u>

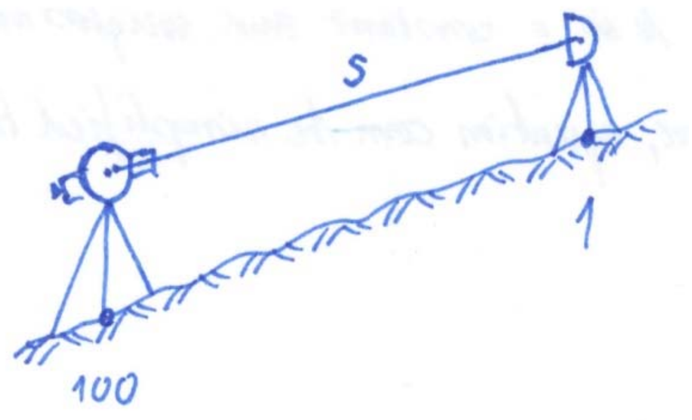


Cartesian coordinates in 2D

- Gauss-Krueger coordinates
- UTM coordinates

→ Projected coordinates into a plane

Distance measurement between point 100 and 1



Question: What does our distance measurement between point 100 and point 1 have to do with distance between points in Gauss-Krueger coordinates?

Answer: Nothing!

Measurement with EDM: s

↓

Atmospheric correction: s_a

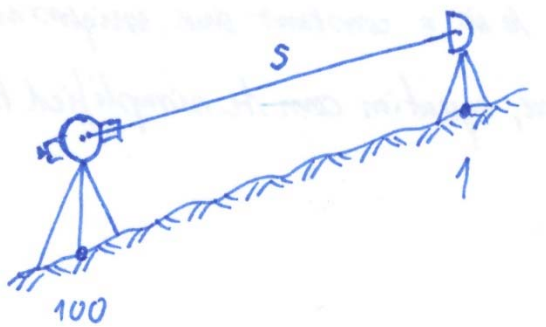
↓

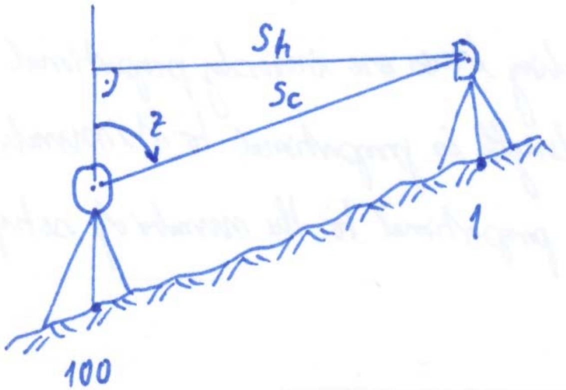
Correction for calibration parameters (zero correction, scale factor): s_c

↓

Reduction into the local horizontal plane: s_h

↓



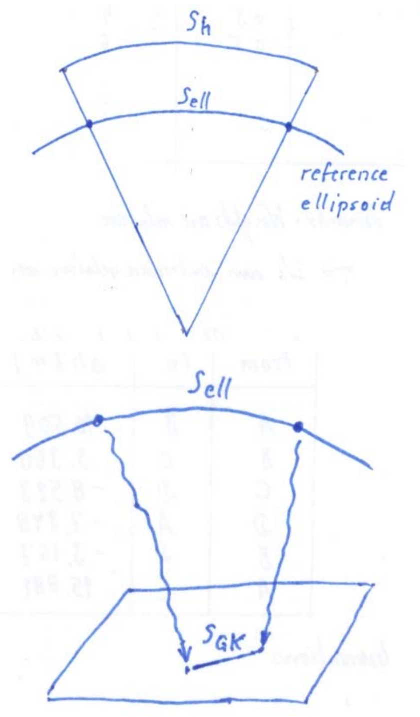


7.2 Adjustment of Horizontal Surveys: Trilateration

Height reduction
(projection onto the surface
of the reference ellipsoid): s_{ell}

↓

Projection into the GK- or
UTM-plane (by formulas): s_{GK}



$\rightarrow$ Finally s_{GK} corresponds with the given Gauss-Krueger coordinates

Attention: Pre-processing of distance measurements <u>must</u> be performed

7.2 Adjustment of Horizontal Surveys: Trilateration

In practice:

- s_{GK} or s_{UTM} is regarded as observation with corresponding precision/weight

  or

- s_h is regarded as observation and reduction and projection is performed within the application of the adjustment software

  $\rightarrow$ Check the pre-settings!

For an adjustment within the Gauss-Markov Model (parametric adjustment)

$$\mathbf{L} = \boldsymbol{\Phi}(\mathbf{X})$$

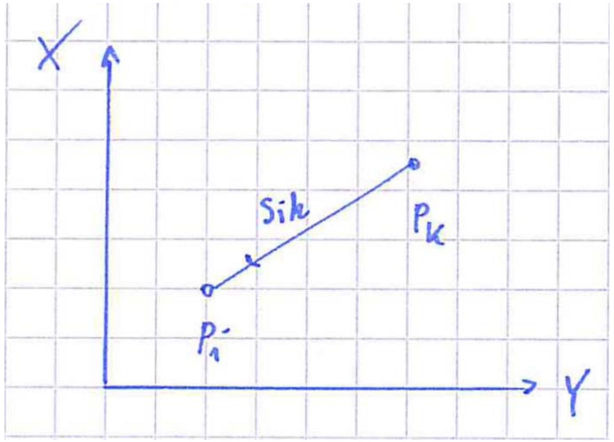
we have to introduce appropriate unknowns to express our observations as functions of the unknowns

$\rightarrow$ We introduce 2D coordinates as unknowns

Functional model:

Measured from “i” to “k”

$$s_{ik} = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2}$$



Observation equations:

$$s_{ik} + v_{s_{ik}} = \sqrt{(\hat{x}_k - \hat{x}_i)^2 + (\hat{y}_k - \hat{y}_i)^2}$$

Nonlinear functional model

$\rightarrow$ for least squares adjustment we need a linearised functional model

$\rightarrow$ Jacobian matrix with partial derivatives

Partial derivatives:

$$\frac{\partial s_{ik}}{\partial x_k} = \frac{1}{2\sqrt{}} \cdot 2(x_k - x_i) = \frac{x_k - x_i}{s_{ik}} = \frac{\Delta x_{ik}}{s_{ik}}$$

$$\frac{\partial s_{ik}}{\partial x_i} = \frac{1}{2\sqrt{}} \cdot 2(x_k - x_i) \cdot (-1) \qquad = \frac{-\Delta x_{ik}}{s_{ik}}$$

$$\frac{\partial s_{ik}}{\partial y_k} = \frac{\Delta y_{ik}}{s_{ik}}$$

$$\frac{\partial s_{ik}}{\partial y_i} = \frac{-\Delta y_{ik}}{s_{ik}}$$

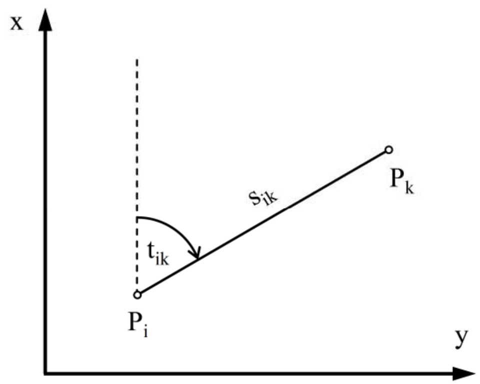
→ See handout (Partial derivatives of geodetic observation equations)

Technische Universität Berlin

Partial derivatives of geodetic observation equations

Given: Observation Equations
Searched: Partial derivatives with respect to the unknowns for the linearization of the observation equations

1. Distances $s_{ik} = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2}$



Notation information:

Index i: Station
Index k: Target

measured from ‚i' to ‚k'

Partial derivatives can be further simplified by using the equations for the grid bearing t_{ik} and distance s_{ik} .

Partial Derivatives:

$$\frac{\partial s_{ik}}{\partial x_k} = \frac{x_k - x_i}{s_{ik}} = \frac{\Delta x_{ik}}{s_{ik}} = \cos t_{ik} \qquad \frac{\partial s_{ik}}{\partial x_i} = -\frac{x_k - x_i}{s_{ik}} = \frac{-\Delta x}{s_{ik}} = -\cos t_{ik}$$

$$\frac{\partial s_{ik}}{\partial y_k} = \frac{y_k - y_i}{s_{ik}} = \frac{\Delta y_{ik}}{s_{ik}} = \sin t_{ik} \qquad \frac{\partial s_{ik}}{\partial y_i} = -\frac{y_k - y_i}{s_{ik}} = -\frac{\Delta y_{ik}}{s_{ik}} = -\sin t_{ik}$$

Adjustment_Theory_I_Derivatives.pdf

Weights in trilateration networks

- Precision for distances from electronic distance measurement given in

$$\sigma_{s_i} = a_1 + a_2 \cdot d_i$$

a_1: constant part of precision
a_2: standard error per unit sight distance length
d_i: length of sight distance

7.2 Adjustment of Horizontal Surveys: Trilateration

- Typical values for the precision of an EDM:

$$3 \text{ mm} + 2 \text{ ppm}$$

ppm: parts per million $\widehat{=}$ mm per km

$\rightarrow$ Standard deviation for a distance of

$$\begin{aligned} 500 \text{ m}&: 4 \text{ mm} \\ 1000 \text{ m}&: 5 \text{ mm} \\ 2000 \text{ m}&: 7 \text{ mm} \end{aligned}$$

7.2 Adjustment of Horizontal Surveys: Trilateration

- Variance matrix of the observations

$$\mathbf{\Sigma}_{LL} = \begin{bmatrix} \sigma_{s_1}^2 & & & 0 \\ & \sigma_{s_2}^2 & & \\ & & \ddots & \\ 0 & & & \sigma_{s_n}^2 \end{bmatrix}$$

- With reference variance σ_0^2
- Cofactor matrix of observations: $\mathbf{Q}_{LL} = \frac{1}{\sigma_0^2}\mathbf{\Sigma}_{LL}$
- Weight matrix of observations: $\mathbf{P} = \mathbf{Q}_{LL}^{-1}$

7.2 Adjustment of Horizontal Surveys: Trilateration

Technische Universität Berlin

Example

The measurements of the trilateration network depicted in Figure 1 are listed in Table 2. The points 1, 2 and 3 are control points (error-free) and their Gauss-Krueger coordinates are given in Table 1. Calculate the adjusted Gauss-Krueger coordinates of point 100 using least squares adjustment.

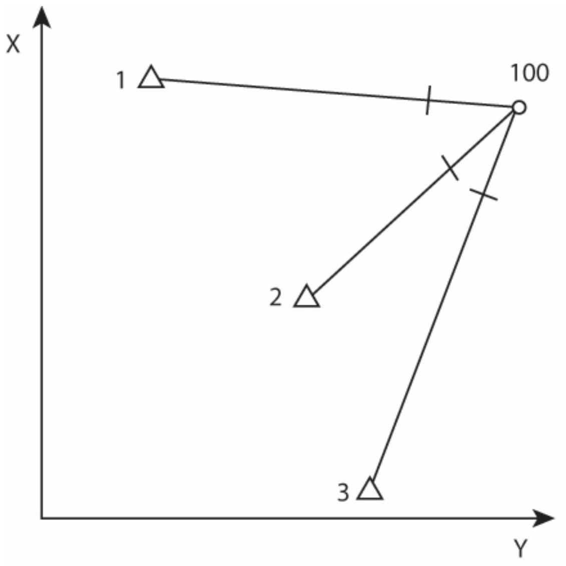


Figure 1: Trilateration network

Table 1: Gauss-Krueger coordinates of control points

Point No.	y [m]	x [m]
1	865.400	4527.150
2	2432.550	2047.250
3	2865.220	27.150

Approximate values for the coordinates of point 100:

$$y_{100}^{0} : 6861.3 \; ; \; x_{100}^{0} : 3727.8$$

(graphical coordinates from a map)

7.2 Adjustment of Horizontal Surveys: Trilateration

Technische Universität Berlin

Table 2: Observed reduced distances

From	To	s [m]
100	1	6049.000
100	2	4736.830
100	3	5446.490

- The distance measurements have been performed with a precision of 1 mm + 2 ppm
- The distances are uncorrelated and already reduced into the Gauss-Krueger projection
- Set up an appropriate functional model as well as the observation equations
- Set up the stochastic model
- Choose appropriate values for the break-off conditions ε and δ and justify your decision
- Solve the normal equation system and determine the Gauss-Krueger coordinates of point 100 as well as their standard deviations
- Calculate the residuals and the adjusted observations as well as their standard deviations

7.2 Adjustment of Horizontal Surveys: Trilateration

General considerations:

- What are our unknowns?

  → Coordinates of point 100

  → We introduce y_{100}, x_{100}

- What are our observations?

  → Distances

  → $s_{100,1}, s_{100,2}, s_{100,3}$

- Observations reduced into projection?

  → Yes!

- What are our fixed values?

  → $y_1, x_1; y_2, x_2; y_3, x_3$

- Redundancy? → $r = n - u \;\rightarrow r = 3 - 2 \;\rightarrow r = 1$

Functional model:

$$s_{100,1} = \sqrt{(x_1 - x_{100})^2 + (y_1 - y_{100})^2}$$
$$s_{100,2} = \sqrt{(x_2 - x_{100})^2 + (y_2 - y_{100})^2}$$
$$s_{100,3} = \sqrt{(x_3 - x_{100})^2 + (y_3 - y_{100})^2}$$

Observation equations:

$$s_{100,1} + v_1 = \sqrt{(x_1 - \hat{x}_{100})^2 + (y_1 - \hat{y}_{100})^2}$$
$$s_{100,2} + v_2 = \sqrt{(x_2 - \hat{x}_{100})^2 + (y_2 - \hat{y}_{100})^2}$$
$$s_{100,3} + v_3 = \sqrt{(x_3 - \hat{x}_{100})^2 + (y_3 - \hat{y}_{100})^2}$$

7.2 Adjustment of Horizontal Surveys: Trilateration

Observation vector:

$$\mathbf{L} = \begin{bmatrix} 6049.000 \\ 4736.830 \\ 5446.490 \end{bmatrix}$$

Stochastic model of the observations:

$$\sigma_1 = 1\ \text{mm} + 2\,\frac{\text{mm}}{\text{km}} \cdot 6.049\ \text{km}$$

$$\sigma_2 = 1\ \text{mm} + 2\,\frac{\text{mm}}{\text{km}} \cdot 4.73683\ \text{km}$$

$$\sigma_3 = 1\ \text{mm} + 2\,\frac{\text{mm}}{\text{km}} \cdot 5.44649\ \text{km}$$

7.2 Adjustment of Horizontal Surveys: Trilateration

$$\mathbf{\Sigma}_{LL} = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix}$$

$$\mathbf{Q}_{LL} = \frac{1}{\sigma_0^2}\mathbf{\Sigma}_{LL}$$

with

$$\sigma_0^2 = 1$$

$$\mathbf{Q}_{LL} = \mathbf{\Sigma}_{LL}$$

$$\rightarrow \qquad \mathbf{P} = \mathbf{Q}_{LL}^{-1}$$

7.2 Adjustment of Horizontal Surveys: Trilateration

Vector of adjusted unknowns:

$$\hat{\mathbf{X}} = \begin{bmatrix} \hat{x}_{100} \\ \hat{y}_{100} \end{bmatrix}$$

Nonlinear functional model

→ Solution from iterative computing with linearised functional model

→ Introduction of approximate values x_{100}^0, y_{100}^0

7.2 Adjustment of Horizontal Surveys: Trilateration

Vector of starting values:

$$\mathbf{X}^0 = \begin{bmatrix} x_{100}^0 \\ y_{100}^0 \end{bmatrix}$$

Vector of adjusted reduced unknowns:

$$\hat{\mathbf{x}} = \hat{\mathbf{X}} - \mathbf{X}^0 = \begin{bmatrix} d\hat{x}_{100} \\ d\hat{y}_{100} \end{bmatrix} = \begin{bmatrix} \hat{x}_{100} - x_{100}^0 \\ \hat{y}_{100} - y_{100}^0 \end{bmatrix}$$

Vector of reduced observations:

$$\mathbf{l} = \begin{bmatrix} 6049.000 - \sqrt{\left(x_1 - x_{100}^0\right)^2 + \left(y_1 - y_{100}^0\right)^2} \\ 4736.830 - \sqrt{\left(x_2 - x_{100}^0\right)^2 + \left(y_2 - y_{100}^0\right)^2} \\ 5446.490 - \sqrt{\left(x_3 - x_{100}^0\right)^2 + \left(y_3 - y_{100}^0\right)^2} \end{bmatrix}$$

7.2 Adjustment of Horizontal Surveys: Trilateration

Jacobian matrix:

$$\mathbf{J} = \begin{array}{c c} & \begin{array}{cc} x_{100}^0 & y_{100}^0 \end{array} \\ \begin{array}{c} s_{100,1} \\[3ex] s_{100,2} \\[3ex] s_{100,3} \end{array} & \begin{bmatrix} \dfrac{\partial s_{100,1}^0}{\partial x_{100}^0} & \dfrac{\partial s_{100,1}^0}{\partial y_{100}^0} \\ \dfrac{\partial s_{100,2}^0}{\partial x_{100}^0} & \dfrac{\partial s_{100,2}^0}{\partial y_{100}^0} \\ \dfrac{\partial s_{100,3}^0}{\partial x_{100}^0} & \dfrac{\partial s_{100,3}^0}{\partial y_{100}^0} \end{bmatrix} \end{array}$$

Technische Universität Berlin

with

$$\frac{\partial s_{100,1}^0}{\partial x_{100}^0} = \frac{1}{2\sqrt{(x_1 - x_{100}^0)^2 + (y_1 - y_{100}^0)^2}} \cdot 2\left(x_1 - x_{100}^0\right) \cdot (-1)$$

$$= \frac{-(x_1 - x_{100}^0)}{s_{100,1}^0}$$

$$\vdots$$

$$\frac{\partial s_{100,3}^0}{\partial y_{100}^0} = \frac{1}{2\sqrt{(x_3 - x_{100}^0)^2 + (y_3 - y_{100}^0)^2}} \cdot 2\left(y_3 - y_{100}^0\right) \cdot (-1)$$

$$= \frac{-(y_3 - y_{100}^0)}{s_{100,3}^0}$$

Technische Universität Berlin

Design matrix:

$$\mathbf{A} = \mathbf{J}$$

Normal equations:

$$\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l}$$

Solution of normal equations:

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

Adjusted unknowns:

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

Residuals:

$$\mathbf{v} = \mathbf{A}\hat{\mathbf{x}} - \mathbf{l}$$

Adjusted observations:

$$\hat{\mathbf{L}} = \mathbf{L} + \mathbf{v}$$

Check:

$$\max |\hat{x}_i| < \varepsilon$$

and

$$\max \left|\hat{L}_i - \Phi_i(\hat{\mathbf{X}})\right| < \delta$$

no

Use $\hat{\mathbf{X}}$ as new $\mathbf{X}^0$

yes

$\hat{\mathbf{X}}$ is final result

7.2 Adjustment of Horizontal Surveys: Trilateration

Empirical reference standard deviation:

$$s_0 = \sqrt{\frac{\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v}}{n-u}}$$

VCM of adjusted unknowns:

$$\mathbf{\Sigma}_{\hat{X}\hat{X}} = \mathrm{s}_0^2 \cdot \mathbf{Q}_{\hat{X}\hat{X}} \qquad \text{with} \qquad \mathbf{Q}_{\hat{X}\hat{X}} = \mathbf{N}^{-1}$$

Standard deviation of unknowns:

$$\mathbf{\Sigma}_{\hat{X}\hat{X}} = s_0^2 \cdot \begin{bmatrix} q_{\hat{x}\hat{x}} & q_{\hat{x}\hat{y}} \\ q_{\hat{y}\hat{x}} & q_{\hat{y}\hat{y}} \end{bmatrix}$$

$q_{\hat{x}\hat{x}}$: Cofactor of unknown value x_{100}

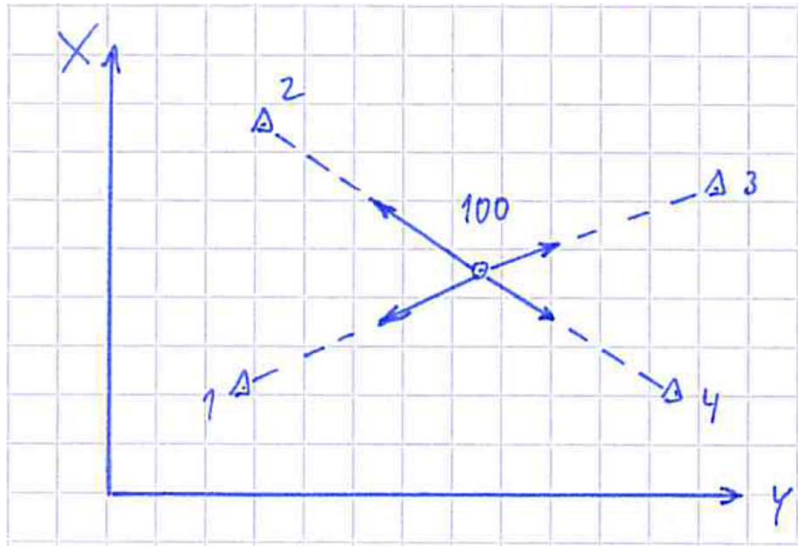
$q_{\hat{y}\hat{y}}$: Cofactor of unknown value y_{100}

$$s_{\hat{x}_{100}} = s_0 \cdot \sqrt{q_{\hat{x}\hat{x}}}$$
$$s_{\hat{y}_{100}} = s_0 \cdot \sqrt{q_{\hat{y}\hat{y}}}$$

Technische Universität Berlin

Basic idea of triangulation:

Determination of 2D coordinates of points in a plane Cartesian coordinate system from observed directions



Cartesian coordinates in 2D

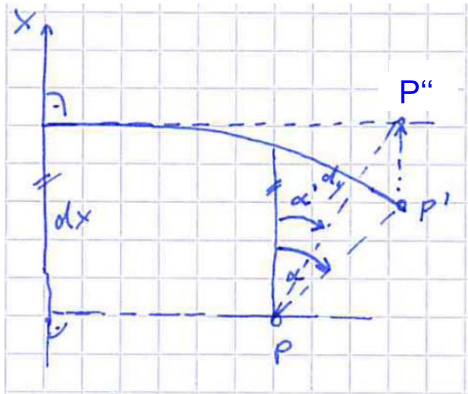
- Gauss-Krueger coordinates
- UTM coordinates

→ Projected coordinates into a plane

→ Conformal mapping

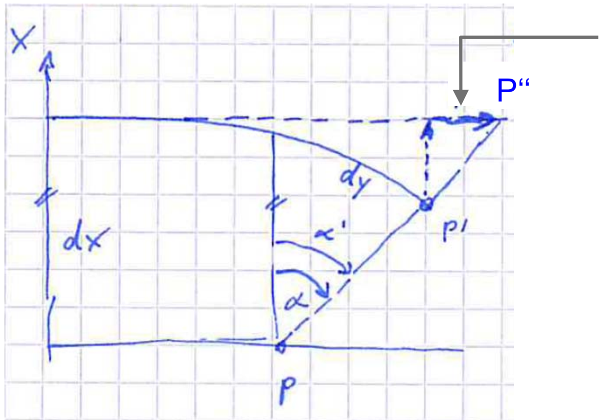
7.3 Adjustment of Horizontal Surveys: Triangulation

Non-conformal and conformal mapping



$$\alpha' \neq \alpha$$

e.g. Soldner coordinates



Ordinate difference is elongated
→ elimination of angular distortion

$$\alpha' = \alpha$$

e.g. Gauss-Krueger coordinates, UTM coordinates

7.3 Adjustment of Horizontal Surveys: Triangulation

Technische Universität Berlin

- A conformal mapping keeps a differential similarity between the original elliptic situation and the maps image

  $\rightarrow$ We can use our measured directions without corrections and reductions and combine them with GK or UTM coordinates

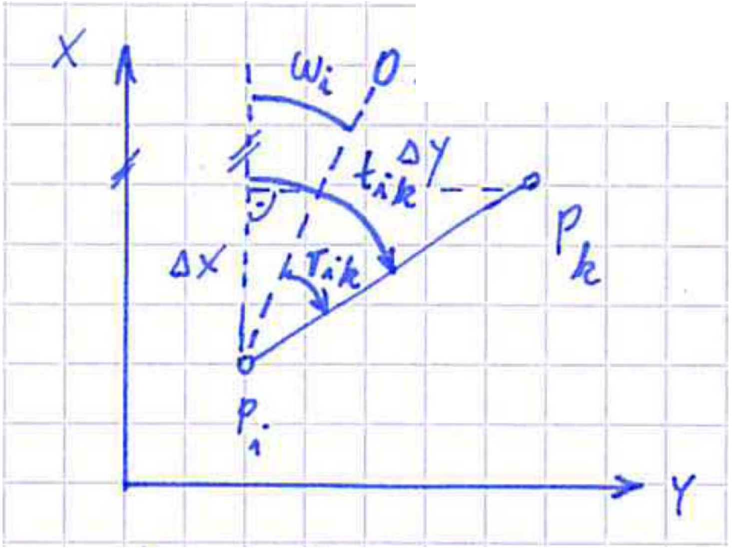
- Directions r_{ik} are our observations with corresponding precision / weight

- For an adjustment within the Gauss-Markov Model (parametric adjustment)

$$\mathbf{L} = \mathbf{\Phi}(\mathbf{X})$$

we have to introduce appropriate unknowns to express our observations as functions of the unknowns

$\rightarrow$ We introduce 2D coordinates as unknowns



0: zero direction of our instrument (tacheometer, theodolite)

i: instrument station

k: foresight station

r_{ik}: measured direction from i to k

t_{ik}: azimuth from i to k

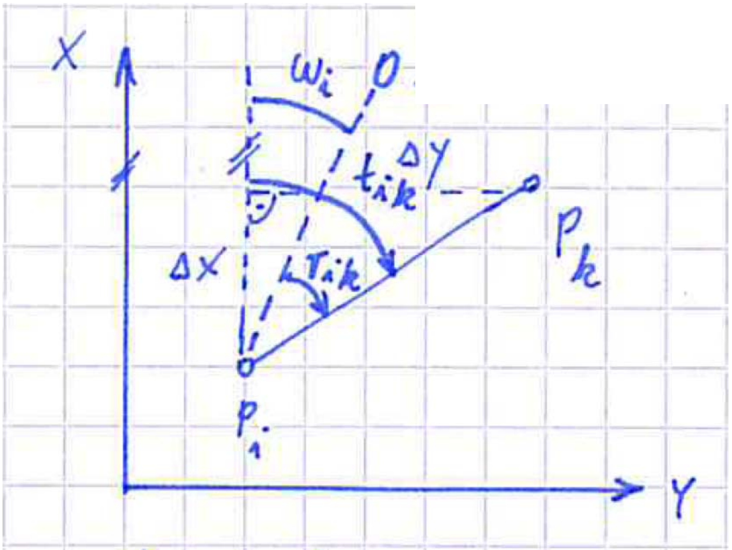
ω_i: orientation unknown

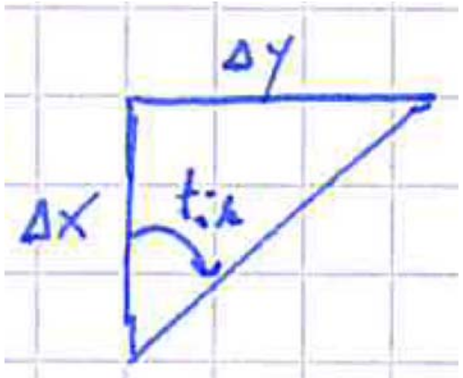
Functional model:

$$r_{ik} = t_{ik} - \omega_i$$

7.3 Adjustment of Horizontal Surveys: Triangulation

How to obtain t_{ik}?





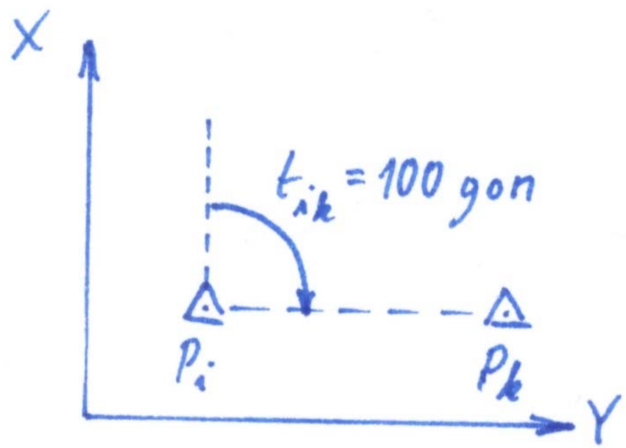
$$\tan t_{ik} = \frac{\Delta y}{\Delta x} \quad \rightarrow \quad t_{ik} = \arctan\frac{\Delta y}{\Delta x}$$

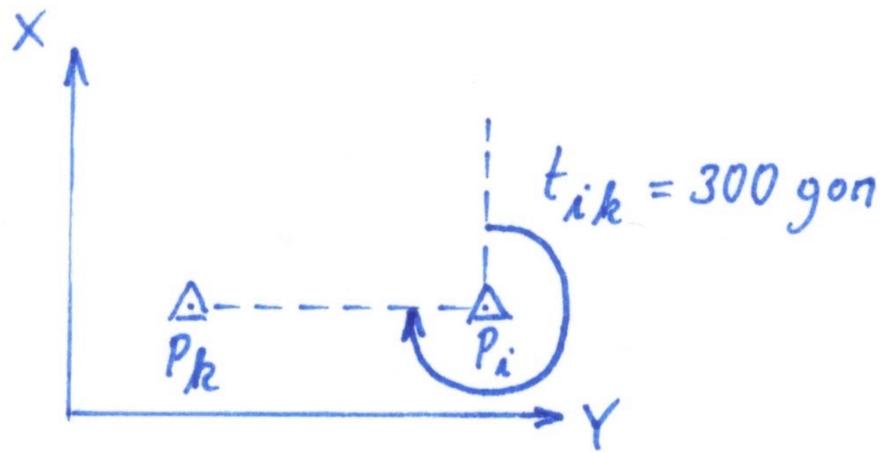
$$\rightarrow \quad t_{ik} = \arctan\frac{y_k - y_i}{x_k - x_i}$$

▶ Attention 1: What happens if $\Delta x = 0$?

→ Cannot use formula

→ two cases possible



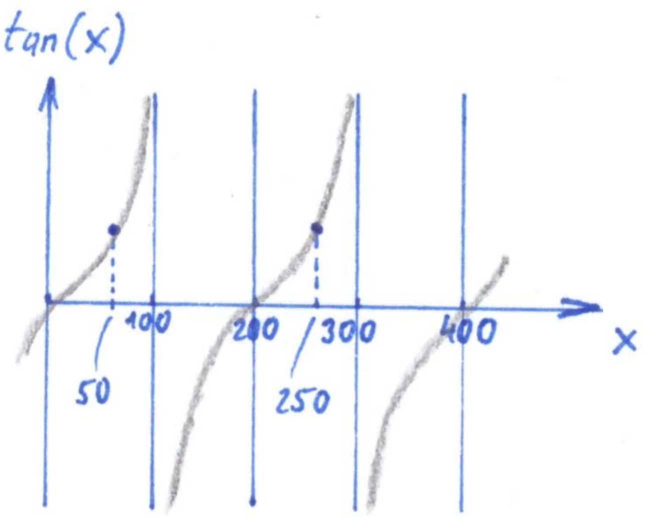


7.3 Adjustment of Horizontal Surveys: Triangulation

► Attention 2:

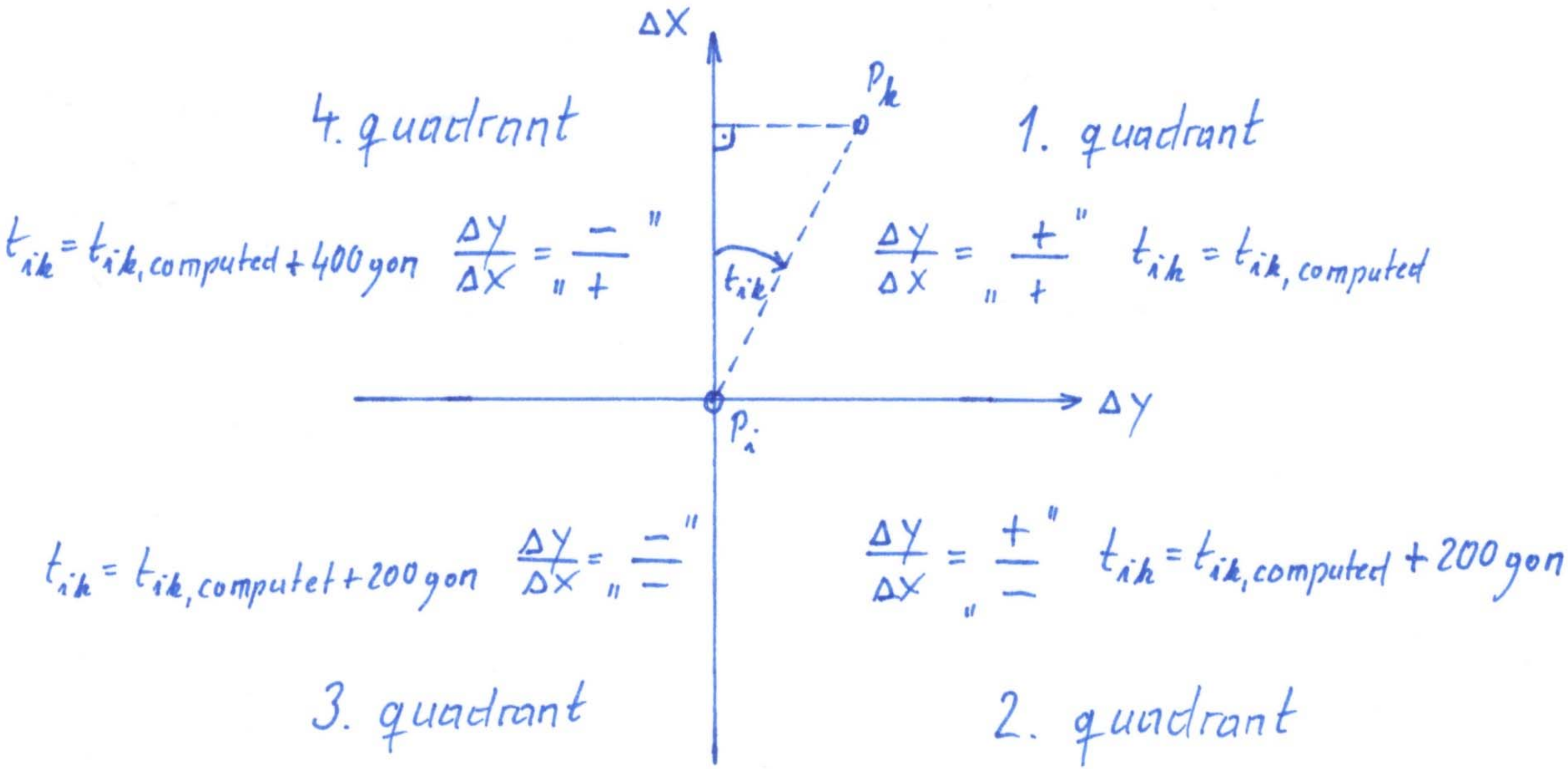
e.g. $\tan 50 \text{ gon} = 1$

$\tan 250 \text{ gon} = 1$



Problem: Which value is the desired one?

Solution: Analysis of the quadrants, $t_{ik,\text{computed}} = \arctan\frac{\Delta y}{\Delta x}$



Remark: “atan2”

7.3 Adjustment of Horizontal Surveys: Triangulation

Functional model:

Measured from “i” to “k”

$$\boxed{r_{ik} = \arctan\frac{y_k - y_i}{x_k - x_i} - \omega_i}$$

Attention:
Quadrants!

Observation equations:

$$\boxed{r_{ik} + v_{ik} = \arctan\frac{\hat{y}_k - \hat{y}_i}{\hat{x}_k - \hat{x}_i} - \hat{\omega}_i}$$

Nonlinear functional model

$\rightarrow$ for least squares adjustment we need a linearised functional model

$\rightarrow$ Jacobian matrix with partial derivatives

Technische Universität Berlin

Partial derivatives:

we know $(\arctan x)' = \frac{1}{1+x^2}$

$$\frac{\partial r_{ik}}{\partial y_k} = \frac{\partial\left(\arctan\frac{y_k-y_i}{x_k-x_i} - \omega_i\right)}{\partial y_k} = \frac{1}{1+\left(\frac{y_k-y_i}{x_k-x_i}\right)^2} \cdot \frac{1}{x_k-x_i}$$

$$= \frac{1}{1+\frac{(y_k-y_i)^2}{(x_k-x_i)^2}} \cdot \frac{1}{x_k-x_i} = \frac{1}{\frac{(x_k-x_i)^2}{(x_k-x_i)^2}+\frac{(y_k-y_i)^2}{(x_k-x_i)^2}} \cdot \frac{1}{x_k-x_i}$$

$$= \frac{(x_k-x_i)^2}{(x_k-x_i)^2+(y_k-y_i)^2} \cdot \frac{1}{x_k-x_i} = \frac{x_k-x_i}{s_{ik}^2} = \frac{\Delta x_{ik}}{s_{ik}^2}$$

7.3 Adjustment of Horizontal Surveys: Triangulation

$$\left.\frac{\partial r_{ik}}{\partial y_i}, \frac{\partial r_{ik}}{\partial x_k}, \frac{\partial r_{ik}}{\partial x_i}\right\}$$ See handout! $$\frac{\partial r_{ik}}{\partial \omega_i} = -1$$

2. Directions $$r_{ik} = \arctan\left(\frac{y_k - y_i}{x_k - x_i}\right) - \omega_i$$

Partial Derivatives:

$$\frac{\partial r_{ik}}{\partial y_k} = \frac{x_k - x_i}{s_{ik}^{\;2}} = \frac{\Delta x_{ik}}{s_{ik}^{\;2}} = \frac{\cos t_{ik}}{s_{ik}} \qquad \frac{\partial r_{ik}}{\partial y_i} = -\frac{x_k - x_i}{s_{ik}^{\;2}} = -\frac{\Delta x_{ik}}{s_{ik}^{\;2}} = -\frac{\cos t_{ik}}{s_{ik}}$$

$$\frac{\partial r_{ik}}{\partial x_k} = -\frac{y_k - y_i}{s_{ik}^{\;2}} = -\frac{\Delta y_{ik}}{s_{ik}^{\;2}} = -\frac{\sin t_{ik}}{s_{ik}} \qquad \frac{\partial r_{ik}}{\partial x_i} = \frac{y_k - y_i}{s_{ik}^{\;2}} = \frac{\Delta y_{ik}}{s_{ik}^{\;2}} = \frac{\sin t_{ik}}{s_{ik}} \qquad \frac{\partial r_{ik}}{\partial \omega_i} = -1$$

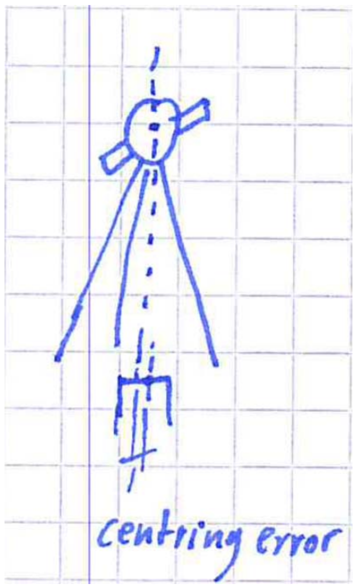
Adjustment_Theory_I_Derivatives.pdf

Technische Universität Berlin

Weights in triangulation networks

- Precision for horizontal directions given in

$$\sigma_{r_i} = \sqrt{b_0^2 + \left(\frac{b_1}{d_i} \cdot \rho\right)^2}$$



$\rho = \frac{200}{\pi}$	if computation in gon
b_0:	constant part of precision → usually given by manufacturer or from experience
b_1:	centring error → often chosen as $1\ \mathrm{mm}$
d_i:	length of sight distance

7.3 Adjustment of Horizontal Surveys: Triangulation

- Typical values for standard deviation σ_{r_i}:

  with $b_0 = 0.3 \text{ mgon}, b_1 = 1 \text{ mm}$

and $d_i = 100 \text{ m}$	→	$\sigma_{r_i} = 0.70 \text{ mgon}$
50 m	→	$\sigma_{r_i} = 1.31 \text{ mgon}$
10 m	→	$\sigma_{r_i} = 6.37 \text{ mgon}$
3 m	→	$\sigma_{r_i} = 21.22 \text{ mgon}$

- Variance matrix of the observations

$$\boldsymbol{\Sigma}_{LL} = \begin{bmatrix} \sigma_{r_1}^2 & & & 0 \\ & \sigma_{r_2}^2 & & \\ & 0 & \ddots & \\ & & & \sigma_{r_n}^2 \end{bmatrix}$$

- With reference variance σ_0^2
- Cofactor matrix of observations: $\mathbf{Q}_{LL} = \frac{1}{\sigma_0^2}\boldsymbol{\Sigma}_{LL}$
- Weight matrix of observations: $\mathbf{P} = \mathbf{Q}_{LL}^{-1}$

7.3 Adjustment of Horizontal Surveys: Triangulation

Technische Universität Berlin

Example

The observed directions of the triangulation network depicted in Figure 1 are listed in Table 2. The points 1, 2, 4, 5 and 6 are control points (error free) and their 2D coordinates are given in Table 1. Calculate the adjusted coordinates of point 3 using least squares adjustment.

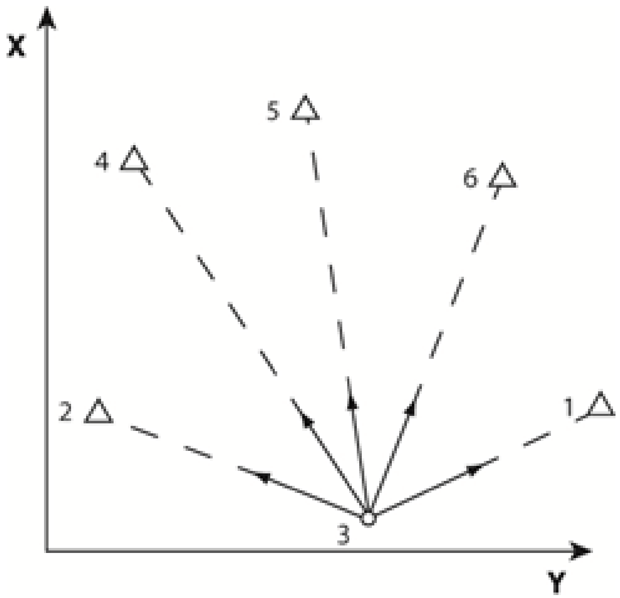


Figure 1: Triangulation network

Table 1: 2D coordinates of control points

Point No.	y [m]	x [m]
1	682.415	321.052
2	203.526	310.527
4	251.992	506.222
5	420.028	522.646
6	594.553	501.494

Approximate values for the coordinates of point 3:

y_3^0 : 493.7 ; x_3^0: 242.9

(graphical coordinates from a map)

7.3 Adjustment of Horizontal Surveys: Triangulation

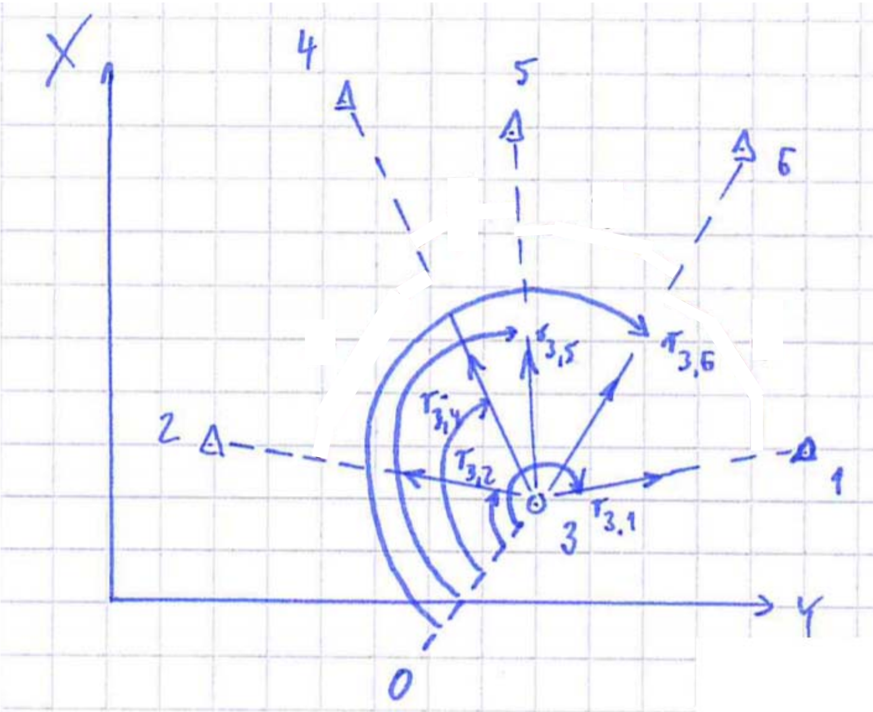
Technische Universität Berlin

Table 2: Observed directions

Instrument station	Foresight station	Direction [gon]
3	1	206.9094
	2	46.5027
	4	84.6449
	5	115.5251
	6	155.5891

- The observed directions are uncorr. and have been obtained with a precision of 1 mgon
- Set up an appropriate functional model as well as the observation equations
- Set up the stochastic model
- Choose appropriate values for the break-off conditions ε and δ and justify your decision
- Solve the normal equation system and determine the 2D coordinates of point 3 as well as their standard deviations
- Calculate the residuals and the adjusted observations as well as their standard deviations

7.3 Adjustment of Horizontal Surveys: Triangulation



Instrument station	Foresight station	Direction [gon]
3	1	206.9094
	2	46.5027
	4	84.6449
	5	115.5251
	6	155.5891

7.3 Adjustment of Horizontal Surveys: Triangulation

General considerations:

- What are our unknowns?

  → Coordinates for point 3, orientation

  → We introduce y_3, x_3, ω_3
- What are our observations?

  → Directions

  → $r_{31}, r_{32}, r_{34}, r_{35}, r_{36}$
- Observations reduced into projection?

  → not necessary
- What are our fixed values?

  → $y_1, x_1; y_2, x_2; y_4, x_4; y_5, x_5; y_6, x_6$
- Redundancy?

  → $r = n - u$ → $r = 5 - 3$ → $r = 2$

Functional model:

$$r_{31} = \arctan\frac{y_1 - y_3}{x_1 - x_3} - \omega_3$$

$$r_{32} = \arctan\frac{y_2 - y_3}{x_2 - x_3} - \omega_3$$

$$r_{34} = \arctan\frac{y_4 - y_3}{x_4 - x_3} - \omega_3$$

$$r_{35} = \arctan\frac{y_5 - y_3}{x_5 - x_3} - \omega_3$$

$$r_{36} = \arctan\frac{y_6 - y_3}{x_6 - x_3} - \omega_3$$

7.3 Adjustment of Horizontal Surveys: Triangulation

Observation equations:

$$r_{31} + v_{r_{31}} = \arctan \ldots$$

$$r_{32} + v_{r_{32}} = \arctan \ldots$$

$$\vdots$$

$$r_{36} + v_{r_{36}} = \arctan \ldots$$

Please note: Perform computation in $[\text{rad}]$!

Observation vector:

$$\mathbf{L} = \begin{bmatrix} 206.9094 \\ 46.5027 \\ 84.6449 \\ 115.5251 \\ 155.5891 \end{bmatrix} \cdot \frac{1}{\rho}$$

with $\rho = \frac{200 \text{ gon}}{\pi}$

Technische Universität Berlin

Stochastic model of the observations:

$$\sigma_i = 1 \text{ mgon} = 0.001 \text{ gon} = \frac{0.001}{\rho} \text{rad}$$

for $i = 1, \ldots, n$

$$\rightarrow \mathbf{\Sigma}_{LL} = \begin{bmatrix} \left(\frac{0.001}{\rho}\right)^2 & & & \\ & \left(\frac{0.001}{\rho}\right)^2 & & \\ & & \ddots & \\ & & & \left(\frac{0.001}{\rho}\right)^2 \end{bmatrix}$$

with $\sigma_0 = 1 \text{ mgon} = \frac{0.001}{\rho} \text{ rad}$

$$\mathbf{Q}_{LL} = \frac{1}{\sigma_0^2} \cdot \mathbf{\Sigma}_{LL} \rightarrow \mathbf{Q}_{LL} = \mathbf{I} \rightarrow \mathbf{P} = \mathbf{I}$$

Vector of adjusted unknowns:

$$\hat{\mathbf{X}} = \begin{bmatrix} \hat{x}_3 \\ \hat{y}_3 \\ \widehat{\omega}_3 \end{bmatrix}$$

Nonlinear functional model

→ Solution from iterative computation with linearised functional model

→ Introduction of appropriate values x_3^0, y_3^0, ω_3^0

→ ω_3^0 is a linear term, we can choose arbitrary values, e.g. $\omega_3^0 = 0$

Technische Universität Berlin

Vector of starting values:

$$\mathbf{X}^0 = \begin{bmatrix} x_3^0 \\ y_3^0 \\ \omega_3^0 \end{bmatrix}$$

Vector of adjusted reduced unknowns:

$$\hat{\mathbf{x}} = \hat{\mathbf{X}} - \mathbf{X}^0 = \begin{bmatrix} d\hat{x}_3 \\ d\hat{y}_3 \\ d\hat{\omega}_3 \end{bmatrix} = \begin{bmatrix} \hat{x}_3 - x_3^0 \\ \hat{y}_3 - y_3^0 \\ \hat{\omega}_3 - \omega_3^0 \end{bmatrix}$$

Technische Universität Berlin

Vector of reduced observations:

$$\mathbf{l} = \begin{bmatrix} 206.9094 \cdot \frac{1}{\rho} - \left(\arctan \frac{y_1 - y_3^0}{x_1 - x_3^0} - \omega_3^0 \right) \\ 46.5027 \cdot \frac{1}{\rho} - \left(\arctan \frac{y_2 - y_3^0}{x_2 - x_3^0} - \omega_3^0 \right) \\ 84.6449 \cdot \frac{1}{\rho} - \left(\arctan \frac{y_4 - y_3^0}{x_4 - x_3^0} - \omega_3^0 \right) \\ 115.5251 \cdot \frac{1}{\rho} - \left(\arctan \frac{y_5 - y_3^0}{x_5 - x_3^0} - \omega_3^0 \right) \\ 155.5891 \cdot \frac{1}{\rho} - \left(\arctan \frac{y_6 - y_3^0}{x_6 - x_3^0} - \omega_3^0 \right) \end{bmatrix}$$

Quadrants!

Technische Universität Berlin

Jacobian matrix:

$$
\mathbf{J} = \begin{array}{c} \\ r_{31} \\ r_{32} \\ r_{34} \\ r_{35} \\ r_{36} \end{array}
\begin{array}{c}
\begin{array}{ccc} x_3^0 & y_3^0 & \omega_3^0 \end{array} \\
\begin{bmatrix}
\dfrac{\partial r_{31}}{\partial x_3^0} & \dfrac{\partial r_{31}}{\partial y_3^0} & \dfrac{\partial r_{31}}{\partial \omega_3^0} \\
\dfrac{\partial r_{32}}{\partial x_3^0} & \vdots & -1 \\
\vdots & \vdots & -1 \\
\vdots & \vdots & -1 \\
\dfrac{\partial r_{36}}{\partial x_3^0} & \dfrac{\partial r_{36}}{\partial y_3^0} & -1
\end{bmatrix}
\end{array}
$$

Partial derivatives → see handout!

7.3 Adjustment of Horizontal Surveys: Triangulation

Technische Universität Berlin

Design matrix:

$$\mathbf{A} = \mathbf{J}$$

Normal equations:

$$\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l}$$

Solution of normal equations:

$$\hat{\mathbf{x}} = \left(\underbrace{\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}}_{\mathbf{N}}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l}$$

Adjusted unknowns:

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

Residuals:

$$\mathbf{v} = \mathbf{A}\hat{\mathbf{x}} - \mathbf{l}$$

Adjusted observations:

$$\hat{\mathbf{L}} = \mathbf{L} + \mathbf{v}$$

Check:

$$\max|\hat{x}_i| < \varepsilon$$

and

$$\max\left|\hat{L}_i - \Phi_i(\widehat{\mathbf{X}})\right| < \delta$$

no

Use $\widehat{\mathbf{X}}$ as new $\mathbf{X}^0$

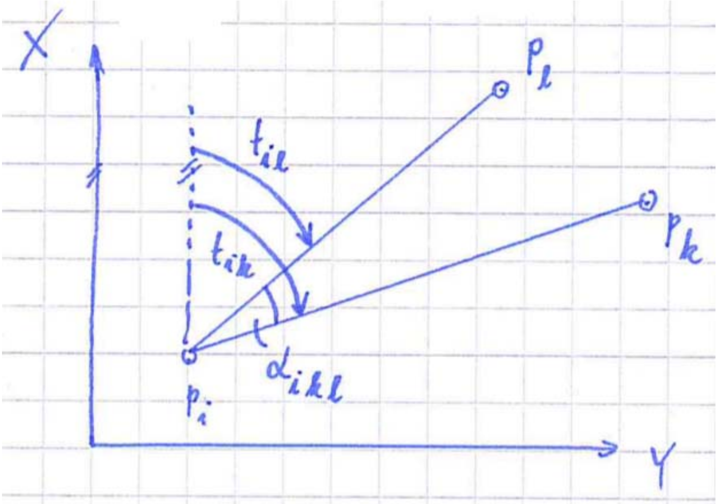
yes

$\boxed{\widehat{\mathbf{X}} \text{ is final result}}$

$$s_0, \boldsymbol{\Sigma}_{\hat{X}\hat{X}}, \dots$$

Conversion of final results from [rad] into [gon]

Triangulation with angles as observations



Functional model:

$$\alpha_{ikl} = t_{ik} - t_{il}$$

$$\boxed{\alpha_{ikl} = \arctan\frac{y_k - y_i}{x_k - x_i} - \arctan\frac{y_l - y_i}{x_l - x_i}}$$

Partial derivatives → See handout!

3. Angles

$$\alpha_{ikl} = \arctan\left(\frac{y_k - y_i}{x_k - x_i}\right) - \arctan\left(\frac{y_l - y_i}{x_l - x_i}\right)$$

Partial Derivatives:

$$\frac{\partial \alpha_{ikl}}{\partial y_k} = \frac{x_k - x_i}{s_{ik}{}^2} = \frac{\Delta x_{ik}}{s_{ik}{}^2} = \frac{\cos t_{ik}}{s_{ik}} \qquad \frac{\partial \alpha_{ikl}}{\partial y_l} = -\frac{x_l - x_i}{s_{il}{}^2} = \frac{-\Delta x_{il}}{s_{il}{}^2} = -\frac{\cos t_{il}}{s_{il}}$$

$$\frac{\partial \alpha_{ikl}}{\partial x_k} = -\frac{y_k - y_i}{s_{ik}{}^2} = -\frac{\Delta y_{ik}}{s_{ik}{}^2} = -\frac{\sin t_{ik}}{s_{ik}} \qquad \frac{\partial \alpha_{ikl}}{\partial x_l} = \frac{y_l - y_i}{s_{il}{}^2} = \frac{\Delta y_{il}}{s_{il}{}^2} = \frac{\sin t_{il}}{s_{il}}$$

$$\frac{\partial \alpha_{ikl}}{\partial y_i} = -\frac{\Delta x_{ik}}{s_{ik}{}^2} + \frac{\Delta x_{il}}{s_{il}{}^2} = -\frac{\cos t_{ik}}{s_{ik}} + \frac{\cos t_{il}}{s_{il}} \qquad \frac{\partial \alpha_{ikl}}{\partial x_i} = \frac{\Delta y_{ik}}{s_{ik}{}^2} - \frac{\Delta y_{il}}{s_{il}{}^2} = \frac{\sin t_{ik}}{s_{ik}} - \frac{\sin t_{il}}{s_{il}}$$

Adjustment_Theory_I_Derivatives.pdf

7.3 Adjustment of Horizontal Surveys: Triangulation

Technische Universität Berlin

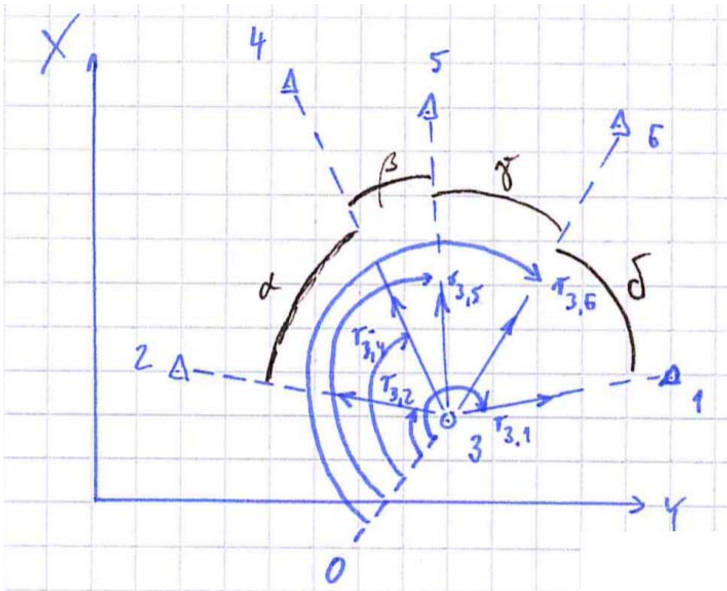
Attention: We measure directions, not angles!

How can we compute angles?

→ Differences of directions!

In our example:

$$\left.\begin{aligned} \alpha &= r_{34} - r_{32} \\ \beta &= r_{35} - r_{34} \\ \gamma &= r_{36} - r_{35} \\ \delta &= r_{31} - r_{36} \end{aligned}\right\} \text{Reduction to interval } [0, \ldots, 400[\text{ gon}$$



→ Derived observations!

→ VCM for these derived observations? From VC propagation!

Technische Universität Berlin

Example

The Gauss-Krueger coordinates of the control points, which can be regarded as fixed (error free) values are listed in Table 1. The measurements of the combined horizontal network depicted in Figure 1 are listed in Table 2. Calculate the adjusted Gauss-Krueger coordinates of point 1 and 15 using least squares adjustment.

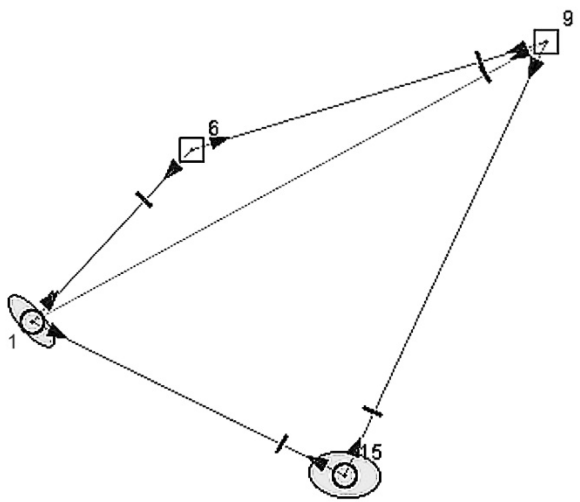


Figure 1: Combined horizontal network

Table 1: Gauss-Krueger coordinates for control points

Point No.	Easting [m]	Northing [m]	Remarks
6	53 17 651.428	49 68 940.373	Fixed point
9	53 24 162.853	49 70 922.160	Fixed point
1	to be computed, see blackboard		Initial values
15			Initial values

7.4 Adjustment of Horizontal Surveys: Combined Network

Table 2: Observed distances and directions

From	To	Horizontal directions [gon]	Horizontal distances [m]
1	6	148.0875	
	15	228.9044	
6	1	248.0883	4307.851
	9	81.1917	
9	15	207.9027	
	1	248.4428	10759.852
	6	261.1921	6806.332
15	1	358.9060	6399.069
	9	57.9014	8751.757

- The distances measurements have been performed with a precision of 10 cm and are already reduced into the Gauss-Krueger projection
- The observation of directions has been performed with a precision of 1 mgon
- All measurements (distances and directions) are uncorrelated

7.4 Adjustment of Horizontal Surveys: Combined Network

- Compute appropriate initial values for the coordinates of points 1 and 15
- Set up an appropriate functional model as well as the observation equations
- Set up the stochastic model
- Choose appropriate values for the break-off conditions ε and δ and justify your decision
- Solve the normal equation system and determine the Gauss-Krueger coordinates of point 1 and 15 as well as their standard deviations
- Calculate the residuals and the adjusted observations as well as their standard deviations

Technische Universität Berlin

General considerations:

- What are our unknowns?

→ Coordinates of points 1 and 15, orientation unknowns

→ We introduce $y_1, x_1;\; y_{15}, x_{15}; \omega_1, \omega_6, \omega_9, \omega_{15}$

- What are our observations?

→ Distances and directions

→ $s_{6,1}, s_{9,1}, s_{9,6}, s_{15,1}, s_{15,9}$

→ $r_{1,6}, r_{1,15}, r_{6,1}, r_{6,9}, r_{9,15}, r_{9,1}, r_{9,6}, r_{15,1}, r_{15,9}$

Technische Universität Berlin

- Observations reduced into projection?

  → Distances: Yes!

  → Directions: Not necessary, conformal coordinates

- What are our fixed values?

  → $y_6, x_6; y_9, x_9$

- Redundancy?

  → $r = n - u$ → $r = 14 - 8$ → $r = 6$

Functional model:

$$s_{6,1} = \sqrt{(y_1 - y_6)^2 + (x_1 - x_6)^2}$$

$$\vdots$$

$$s_{15,9} = \sqrt{(y_9 - y_{15})^2 + (x_9 - x_{15})^2}$$

$$r_{1,6} = \arctan\frac{y_6 - y_1}{x_6 - x_1} - \omega_1$$

$$\vdots$$

$$r_{15,9} = \arctan\frac{y_9 - y_{15}}{x_9 - x_{15}} - \omega_{15}$$

7.4 Adjustment of Horizontal Surveys: Combined Network

Observation equations:

$$s_{6,1} + v_{s_{6,1}} = \sqrt{\quad}$$

$$\vdots$$

$$s_{15,9} + v_{s_{15,9}} = \sqrt{\quad}$$

$$r_{1,6} + v_{r_{1,6}} = \arctan \frac{}{}$$

$$\vdots$$

$$r_{15,9} + v_{r_{15,9}} = \arctan \frac{}{}$$

Please note: Perform computations in [rad]!

Technische Universität Berlin

Observation vector:

$$\mathbf{L} = \begin{bmatrix} s_{6,1} \\ s_{9,1} \\ s_{9,6} \\ s_{15,1} \\ s_{15,9} \\ r_{1,6} \\ r_{1,15} \\ r_{6,1} \\ r_{6,9} \\ r_{9,15} \\ r_{9,1} \\ r_{9,6} \\ r_{15,1} \\ r_{15,9} \end{bmatrix} \quad \text{Values from Table 2} \qquad \left.\begin{matrix} r_{1,6} \\ \vdots \\ r_{15,9} \end{matrix}\right\} \cdot \frac{1}{\rho} \text{ with } \rho = \frac{200\,\text{gon}}{\pi}$$

Technische Universität Berlin

Stochastic model of the observations:

- Distances: $\sigma_{s_i} = 10 \text{ cm} = 0.10 \text{ m}$
- Directions: $\sigma_{r_i} = 1 \text{ mgon} = 0.001 \text{ gon} = \frac{0.001}{\rho} \text{ rad}$

$$\boldsymbol{\Sigma}_{LL} = \begin{array}{c} \\ s_{6,1} \\ s_{9,1} \\ \vdots \\ s_{15,9} \\ \vdots \\ r_{15,9} \end{array} \begin{array}{c} \begin{array}{ccccccc} s_{6,1} & s_{9,1} & \cdots & s_{15,9} & r_{1,6} & \cdots & r_{15,9} \end{array} \\ \left[\begin{array}{cccc:ccc} 0.10^2 & & & & & & \\ & 0.10^2 & & & & 0 & \\ & & \ddots & & & & \\ & & & 0.10^2 & & & \\ \hdashline & & & & \left(\frac{0.001}{\rho}\right)^2 & & \\ & 0 & & & & \ddots & \\ & & & & & & \left(\frac{0.001}{\rho}\right)^2 \end{array}\right] \end{array}$$

7.4 Adjustment of Horizontal Surveys: Combined Network

- Attention:

  Sequence of variances in $\mathbf{\Sigma}_{LL}$ must coincide with the sequence of the observations in vector $\mathbf{L}$

- With reference variance σ_0^2

- Cofactor matrix of observations: $\mathbf{Q}_{LL} = \frac{1}{\sigma_0^2}\mathbf{\Sigma}_{LL}$

- Weight matrix of observations: $\mathbf{P} = \mathbf{Q}_{LL}^{-1}$

Vector of adjusted unknowns:

$$\hat{\mathbf{X}} = [\hat{x}_1 \quad \hat{y}_1 \quad \hat{x}_{15} \quad \hat{y}_{15} \quad \hat{\omega}_1 \quad \hat{\omega}_6 \quad \hat{\omega}_9 \quad \hat{\omega}_{15}]^{\mathrm{T}}$$

→ Nonlinear functional model

→ Solution from iterative computing with linearised functional model

→ Introduction of approximate values $x_1^0, y_1^0; x_{15}^0, y_{15}^0; \omega_1^0, \omega_6^0, \omega_9^0, \omega_{15}^0$

→ $\omega_{1,\dots,15}$: linear terms

Technische Universität Berlin

Vector of starting values:

$$\mathbf{X}^0 = [x_1^0 \quad y_1^0 \quad x_{15}^0 \quad y_{15}^0 \quad \omega_1^0 \quad \omega_6^0 \quad \omega_9^0 \quad \omega_{15}^0]^\mathrm{T}$$

Vector of adjusted reduced unknowns:

$$\hat{\mathbf{x}} = \begin{bmatrix} \hat{x}_1 - x_1^0 \\ \hat{y}_1 - y_1^0 \\ \hat{x}_{15} - x_{15}^0 \\ \hat{y}_{15} - y_{15}^0 \\ \widehat{\omega}_1 - \omega_1^0 \\ \widehat{\omega}_6 - \omega_6^0 \\ \widehat{\omega}_9 - \omega_9^0 \\ \widehat{\omega}_{15} - \omega_{15}^0 \end{bmatrix}$$

Technische Universität Berlin

Vector of reduced observations:

$$\mathbf{l} = \begin{bmatrix} s_{6,1} - s_{6,1}^0 \\ \vdots \\ s_{15,9} - s_{15,9}^0 \\ r_{1,6} - r_{1,6}^0 \\ \vdots \\ r_{15,9} - r_{15,9}^0 \end{bmatrix} = \begin{bmatrix} 4307.851 - \sqrt{(x_1^0 - x_6)^2 + (y_1^0 - y_6)^2} \\ \vdots \\ 8751.757 - \sqrt{(x_9 - x_{15}^0)^2 + (y_9 - y_{15}^0)^2} \\ 148.0875 \cdot \frac{1}{\rho} - \left(\arctan\frac{y_6 - y_1^0}{x_6 - x_1^0} - \omega_1^0\right) \\ \vdots \\ 57.9014 \cdot \frac{1}{\rho} - \left(\arctan\frac{y_9 - y_{15}^0}{x_9 - x_{15}^0} - \omega_{15}^0\right) \end{bmatrix} \left.\vphantom{\begin{matrix} a \\ b \\ c \end{matrix}}\right\} [\text{rad}]$$

Technische Universität Berlin

Jacobian matrix:

Attention: Sequence must coincide with sequence of unknowns in vector $\hat{\mathbf{x}}$

Attention: Sequence must coincide with sequence of observations in vector $\mathbf{L}$

$$\mathbf{J} = \begin{array}{c} \\ s_{6,1} \\ s_{9,1} \\ \vdots \\ r_{1,6} \\ \vdots \\ r_{15,9} \end{array} \begin{array}{c} \begin{array}{cccccccc} x_1 & y_1 & & x_{15} & y_{15} & \omega_1 & \omega_6 & \omega_9 & \omega_{15} \end{array} \\ \begin{bmatrix} \frac{\partial s_{6,1}}{\partial x_1} & \frac{\partial s_{6,1}}{\partial y_1} & \cdots & & & \\ \frac{\partial s_{9,1}}{\partial x_1} & \frac{\partial s_{9,1}}{\partial y_1} & \cdots & & & \\ & & & & & \\ \frac{\partial r_{1,6}}{\partial x_1} & \frac{\partial r_{1,6}}{\partial y_1} & & & \ddots & \\ \vdots & \vdots & & & & \\ \frac{\partial r_{15,9}}{\partial x_1} & \frac{\partial r_{15,9}}{\partial y_1} & \cdots & & & \frac{\partial r_{15,9}}{\partial \omega_{15}} \end{bmatrix} \end{array}$$

Partial derivatives → see handout!

Design matrix:

$$\mathbf{A} = \mathbf{J}$$

Normal equations:

$$\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l}$$

Solution of normal equations:

$$\hat{\mathbf{x}} = \big(\underbrace{\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}}_{\mathbf{N}}\big)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l}$$

Adjusted unknowns:

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

Residuals:

$$\mathbf{v} = \mathbf{A}\hat{\mathbf{x}} - \mathbf{l}$$

Adjusted observations:

$$\hat{\mathbf{L}} = \mathbf{L} + \mathbf{v}$$

Check:

$$\max|\hat{x}_i| < \varepsilon$$

and

$$\max\left|\hat{L}_i - \Phi_i(\widehat{\mathbf{X}})\right| < \delta$$

no

Use $\widehat{\mathbf{X}}$ as new $\mathbf{X}^0$

yes

$\widehat{\mathbf{X}}$ is final result

Conversion of final results from [rad] into [gon]

Empirical reference standard deviation:

$$s_0 = \sqrt{\frac{\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v}}{n-u}}$$

VCM of adjusted unknowns:

$$\mathbf{\Sigma}_{\hat{X}\hat{X}} = s_0^2 \cdot \mathbf{Q}_{\hat{X}\hat{X}} \qquad \text{with} \qquad \mathbf{Q}_{\hat{X}\hat{X}} = \mathbf{N}^{-1}$$

Standard deviation of unknowns:

Computed from diagonal elements (square root) of $\mathbf{\Sigma}_{\hat{X}\hat{X}}$

7.4 Adjustment of Horizontal Surveys: Combined Network

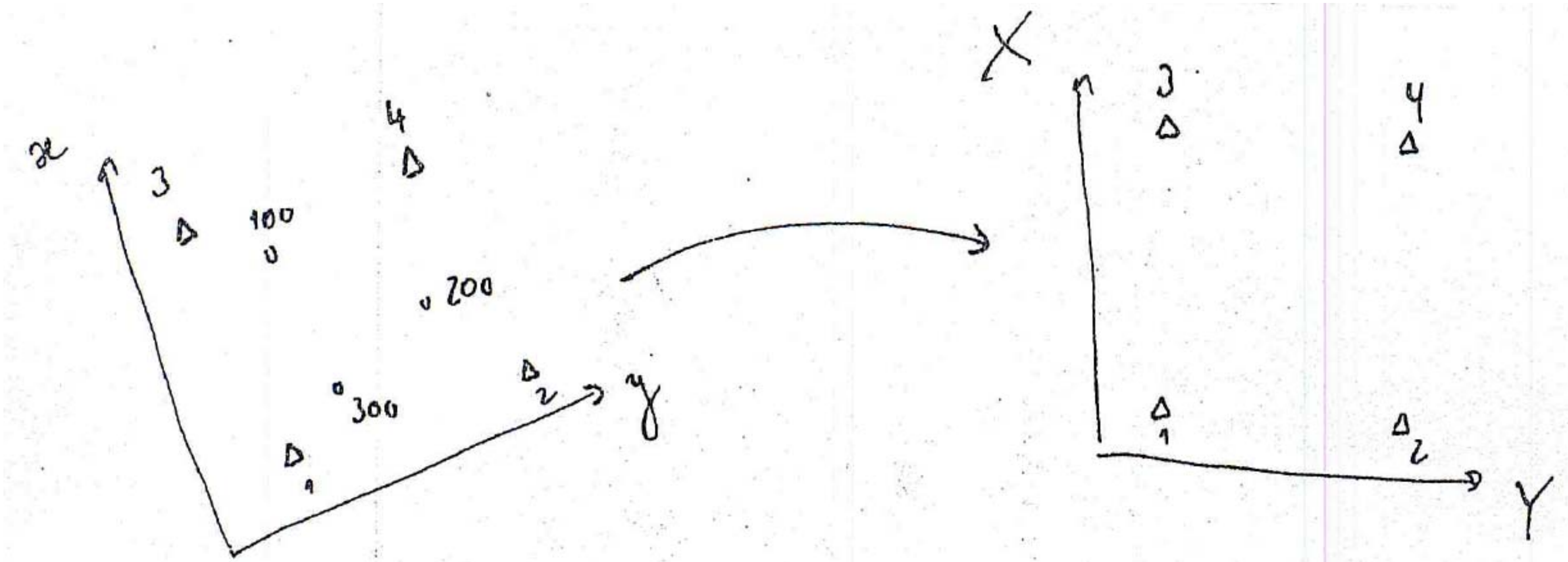
Remark:

Nowadays geodetic networks with measurement of

- Distances
- Directions
- GNSS baselines

Course „Transformation of Geodetic Networks“
3rd semester

Introduction



Start System
(System A, System 1,
local system)

Target System
(System B, System 2,
global system)

7.5 2D Coordinate Transformation

- Coordinates of p identical points (here 1, 2, 3, 4) known in both systems
  → Transformation parameters can be computed via adjustment computation
- Application of the transformation parameters to transform “new points” (here 100, 200, 300) from the start system into the target system

For arbitrary points in the start system (local system) we denote

$$\mathbf{x}_i = \begin{bmatrix} x_i \\ y_i \end{bmatrix}$$

For arbitrary points in the target system (global system)

$$\mathbf{X}_i = \begin{bmatrix} X_i \\ Y_i \end{bmatrix}$$

Technische Universität Berlin

Transformation parameters

- Translation vector $\mathbf{t} = \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix}$

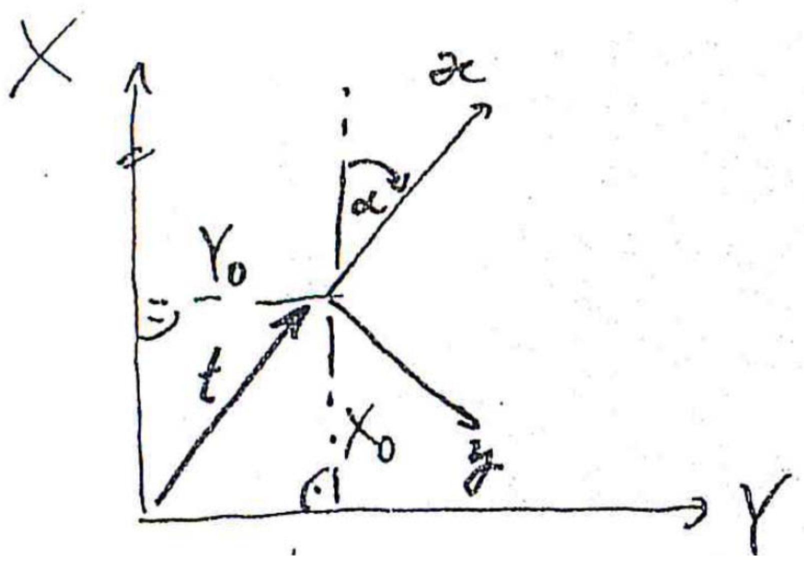


Figure: Translation

7.5 2D Coordinate Transformation

- Rotation α
  - Axes of coordinate systems as unit vectors (length = 1)
  - Rotation around $\vec{e}_Z = \vec{e}_z$-axis with angle α

  → Base vectors $\vec{e}_X$, $\vec{e}_Y$ as functions of $\vec{e}_x$ and $\vec{e}_y$

  → $\vec{e}_X = \vec{e}_x \cdot \cos\alpha - \vec{e}_y \cdot \sin\alpha$

  $\vec{e}_Y = \vec{e}_x \cdot \sin\alpha + \vec{e}_y \cdot \cos\alpha$

  In matrix notation:

$$\begin{bmatrix} \vec{e}_X \\ \vec{e}_Y \end{bmatrix} = \underbrace{\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}}_{\substack{\text{Rotation Matrix} \\ \mathbf{R}_z(\alpha)}} \begin{bmatrix} \vec{e}_x \\ \vec{e}_y \end{bmatrix}$$

- Scale factor m

  → Scaling of both axes $\vec{e}_x$ and $\vec{e}_y$

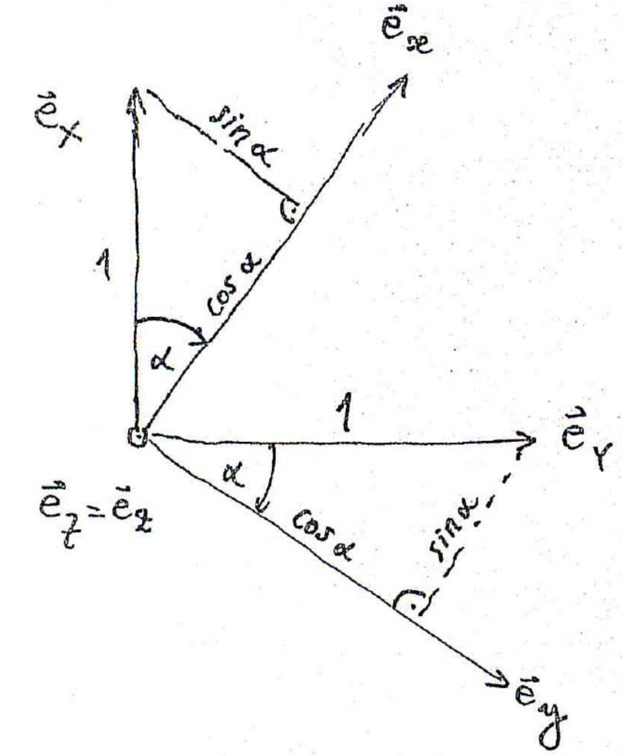


Figure: Rotation

7.5 2D Coordinate Transformation

- How many points for a minimal configuration?
- What are the dangers of using a minimal configuration?
- How can we avoid these problems?

7.5 2D Coordinate Transformation

Transformation with coordinates in the start system as fixed values and stochastic coordinates in the target system (Overdetermined 4-parameter transformation)

With parameters $\mathbf{t}$, $\mathbf{R}_z(\alpha)$ and m we can set up the transformation equation:

$$\begin{bmatrix} X_i \\ Y_i \end{bmatrix} = m \cdot \mathbf{R}_z(\alpha) \cdot \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \mathbf{t}$$

Multiplying out yields:

$$X_i = m \cdot \cos\alpha \cdot x_i - m \cdot \sin\alpha \cdot y_i + X_0$$
$$Y_i = m \cdot \sin\alpha \cdot x_i + m \cdot \cos\alpha \cdot y_i + Y_0$$

Nonlinear equations

Substitutions:

$$a = m \cdot \cos\alpha$$
$$o = m \cdot \sin\alpha$$

→

$$\boxed{\begin{aligned} X_i &= a \cdot x_i - o \cdot y_i + X_0 \\ Y_i &= o \cdot x_i + a \cdot y_i + Y_0 \end{aligned}}$$

Linear equations

7.5 2D Coordinate Transformation

Question: What are the observations, unknowns?

$$X_i = a \cdot x_i - \mathrm{o} \cdot y_i + X_0$$
$$Y_i = o \cdot x_i + a \cdot y_i + Y_0$$

unknowns
observations

x_i, y_i are considered as error-free (fixed) values

→ Classical form of "Helmert-Transformation"

Observation equations for the adjustment within GM-model:

$$X_i + v_{X_i} = \hat{a} \cdot x_i - \hat{\mathrm{o}} \cdot y_i + \hat{X}_0$$
$$Y_i + v_{Y_i} = \hat{\mathrm{o}} \cdot x_i + \hat{a} \cdot y_i + \hat{Y}_0$$

7.5 2D Coordinate Transformation

Observation vector: $\mathbf{L} = \begin{bmatrix} X_1 \\ Y_1 \\ X_2 \\ Y_2 \\ \vdots \\ X_P \\ Y_P \end{bmatrix}$

Vector of unknowns: $\mathbf{X} = \begin{bmatrix} a \\ o \\ X_0 \\ Y_0 \end{bmatrix}$

Design Matrix:
$$\mathbf{A} = \begin{array}{c} \\ X_1 \\ Y_1 \\ X_2 \\ Y_2 \\ \vdots \\ X_P \\ Y_P \end{array} \begin{array}{c} \begin{array}{cccc} a & o & X_0 & Y_0 \end{array} \\ \begin{bmatrix} x_1 & -y_1 & 1 & 0 \\ y_1 & x_1 & 0 & 1 \\ x_2 & -y_2 & 1 & 0 \\ y_2 & x_2 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_P & -y_P & 1 & 0 \\ y_P & x_P & 0 & 1 \end{bmatrix} \end{array}$$

7.5 2D Coordinate Transformation

Stochastic Model: In practice oftentimes $\mathbf{P} = \mathbf{I}$

→ Observations (coordinates in the target system) are regarded as equally weighted and uncorrelated quantities

Solution from: $\underbrace{\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}}_{\mathbf{N}}\,\widehat{\mathbf{X}} = \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{L}$ with $\mathbf{Q}_{xx} = \mathbf{N}^{-1}$

$$\widehat{\mathbf{X}} = \mathbf{Q}_{xx}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{L}$$

Unknowns: $\hat{a}, \hat{o}, \hat{X}_0, \hat{Y}_0$ from $\widehat{\mathbf{X}}$

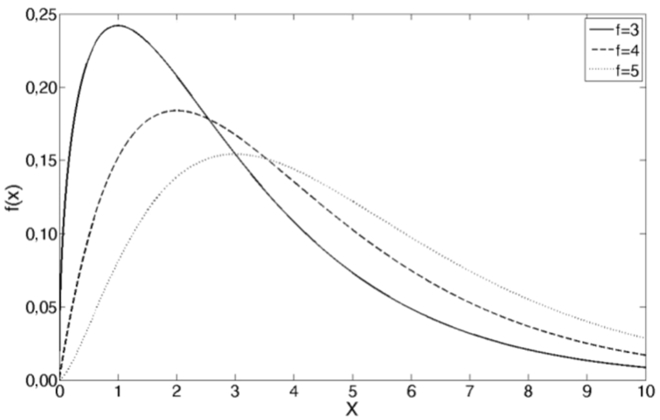
$$\hat{\alpha} = \arctan\frac{\hat{o}}{\hat{a}} \quad , \quad \widehat{m} = \sqrt{\hat{a}^2 + \hat{o}^2}$$

Standard deviations s_α and s_m from variance-covariance propagation

Afterwards:

- Computation of residuals for the identical points
- Computation of s_0

⋮



Adjustment Theory I

Chapter 8 – Least Squares Adjustment with constraints for the unknown parameters

Prof. Dr.-Ing Frank Neitzel | Institute of Geodesy and Geoinformation Science

Version: 03.02.2025

Contents

8. Least-squares adjustment with constraints for the unknown parameters

8.1 Introduction

When performing an adjustment, it is sometimes necessary

a) to fix an unknown parameter to a specific value, e.g.
   - Height of a specific point in differential levelling
   - 2D coordinates of a control station in a horizontal survey

b) to fix a function of unknown parameters to obtain a specific value, e.g.
   - Elevation difference between two stations has to be fixed in differential leveling
   - Direction or length of a line has to be held to a specific value in a horizontal survey

$\Rightarrow$ We have to consider <u>constraints for the unknown parameters</u>

$\Rightarrow$ “Gauss-Markov Model with Constraints” (for the unknown parameters)

8.1 Introduction

Example: (from Section 6.7)

$$3.0 = x + y$$
$$1.5 = 2x - y$$
$$0.2 = x - y$$

Values $3.0,\ 1.5,\ 0.2$ are observations with standard deviations $s_1 = 4\text{ cm}, s_2 = 2\text{ cm}, s_3 = 1\text{ cm}$

We know: Determine the adjusted unknowns $\hat{x}$ and $\hat{y}$

Now: Determine the adjusted unknowns $\hat{x}$ and $\hat{y}$ in such a way that $\hat{y} = 0.9\,\hat{x}$ holds true for the adjusted unknowns

There are two possibilities for a rigorous solution:

1. Elimination of unknowns by inserting the constraint into the equation system
2. Determine the constraint minimum of $\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v}$ under consideration of the constraints (Helmert's Method, see Helmert 1872)

And one approximate solution

Introduce constraints as an additional observation with a very high weight
$\rightarrow$ This "observation" will have residual close to 0 $\rightarrow$ value is "fixed"

8.1 Introduction

Inserting the constraint into the equation system

Redundancy of the original problem (without constraint): $r = n - u$

$$r = 3 - 2$$
$$r = 1$$

Now we consider the constraint: $\hat{y} = 0.9\,\hat{x}$

$\rightarrow$ inserting yields

$$\begin{aligned} 3.0 &= x + 0.9x \\ 1.5 &= 2x - 0.9x \\ 0.2 &= x - 0.9x \end{aligned} \quad \Rightarrow \quad \begin{aligned} 3.0 &= 1.9x \\ 1.5 &= 1.1x \\ 0.2 &= 0.1x \end{aligned}$$

Redundancy under consideration of constraint: $r = n - u$

$$r = 3 - 1$$
$$r = 2$$

General formula : $\boxed{r = n - u + n_{\mathrm{b}}}$

redundancy of the original problem $(n - u)$

\+ number of constraints n_{b}

8. Least-squares adjustment with constraints for the unknown parameters

8.2 Linear Case

Constraint minimum of $\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v}$

Linear functional model (between $\mathbf{L}$ and $\mathbf{X}$)

$$\mathbf{L} = \mathbf{A}\mathbf{X}$$

and additional linear constraints for the unknowns

$$\mathbf{C}\mathbf{X} = \mathbf{c}$$

We obtain

linear observation equations $\mathbf{L} + \mathbf{v} = \mathbf{A}\widehat{\mathbf{X}}$ (1)

linear constraints $\mathbf{c} = \mathbf{C}\widehat{\mathbf{X}}$ (2)

8.2 Linear Case

We can write (1) and (2) as

$$\mathbf{v} = \mathbf{A}\hat{\mathbf{X}} - \mathbf{L}$$
$$\mathbf{0} = \mathbf{C}\hat{\mathbf{X}} - \mathbf{c}$$

$\Omega = \mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v} \rightarrow$ min with constraints $\mathbf{C}\hat{\mathbf{X}} - \mathbf{c} = \mathbf{0}$

We apply Lagrange method for constraint minima:

$$\Omega = (\mathbf{A}\hat{\mathbf{X}} - \mathbf{L})^{\mathrm{T}}\mathbf{P}(\mathbf{A}\hat{\mathbf{X}} - \mathbf{L}) + 2\mathbf{k}^{\mathrm{T}}(\mathbf{C}\hat{\mathbf{X}} - \mathbf{c})$$

$\vdots$

$\mathbf{k} \rightarrow$ vector of Lagrange multipliers (LM)
(vector of correlates)

$$\Omega = \hat{\mathbf{X}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\hat{\mathbf{X}} - 2\hat{\mathbf{X}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}\,\mathbf{L} + \mathbf{L}^{\mathrm{T}}\mathbf{P}\,\mathbf{L} + 2\mathbf{k}^{\mathrm{T}}\mathbf{C}\hat{\mathbf{X}} - 2\mathbf{k}^{\mathrm{T}}\mathbf{c}$$

8.2 Linear Case

Minimum

$$\frac{\partial\Omega}{\partial\hat{\mathbf{X}}^{\mathrm{T}}} = 2\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\hat{\mathbf{X}} - 2\,\mathbf{A}^{\mathrm{T}}\mathbf{P}\,\mathbf{L} \;+ 2\,\mathbf{C}^{\mathrm{T}}\mathbf{k} \qquad = 0$$

$$\frac{\partial\Omega}{\partial\mathbf{k}^{\mathrm{T}}} = \qquad 2\,\mathbf{C}\hat{\mathbf{X}} - 2\mathbf{c} \qquad\qquad = 0$$

Linear equation system

$$\boxed{\left[\begin{array}{c:c} \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A} & \mathbf{C}^{\mathrm{T}} \\ \hdashline \mathbf{C} & \mathbf{0} \end{array}\right] \cdot \left[\begin{array}{c} \hat{\mathbf{X}} \\ \hdashline \mathbf{k} \end{array}\right] = \left[\begin{array}{c} \mathbf{A}^{\mathrm{T}}\mathbf{P}\,\mathbf{L} \\ \hdashline \mathbf{c} \end{array}\right]}$$

"extended normal matrix", if matrix regular …

… Solution

$$\boxed{\left[\begin{array}{c} \hat{\mathbf{X}} \\ \hdashline \mathbf{k} \end{array}\right] = \left[\begin{array}{c:c} \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A} & \mathbf{C}^{\mathrm{T}} \\ \hdashline \mathbf{C} & \mathbf{0} \end{array}\right]^{-1} \left[\begin{array}{c} \mathbf{A}^{\mathrm{T}}\mathbf{P}\,\mathbf{L} \\ \hdashline \mathbf{c} \end{array}\right]}$$

Residuals

$$\boxed{\mathbf{v} = \mathbf{A}\hat{\mathbf{X}} - \mathbf{L}}$$

Adjusted obs.

$$\boxed{\hat{\mathbf{L}} = \mathbf{L} + \mathbf{v}}$$

8.2 Linear Case

Empirical reference standard deviation:

$$s_0 = \sqrt{\frac{\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v}}{n - u + n_{\mathrm{b}}}}$$

VCM of adjusted unknowns:

$$\mathbf{\Sigma}_{\hat{X}\hat{X}} = \mathrm{s}_0^2 \cdot \mathbf{Q}_{\hat{X}\hat{X}} \qquad \text{with} \qquad \mathbf{Q}_{\hat{X}\hat{X}} = \mathbf{Q}_{11}$$

$$\begin{bmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A} & \mathbf{C}^{\mathrm{T}} \\ \mathbf{C} & \mathbf{0} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix}$$

Standard deviation of unknowns:

Computed from diagonal elements (square root) of $\mathbf{\Sigma}_{\hat{X}\hat{X}}$

8. Least-squares adjustment with constraints for the unknown parameters

8.3 Nonlinear Case

Constraint minimum of $\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v}$

Nonlinear, differentiable functional model (between $\mathbf{L}$ and $\mathbf{X}$)

$$\mathbf{L} = \boldsymbol{\phi}(\mathbf{X})$$

and additional nonlinear, differentiable constraints for the unknowns

$$\boldsymbol{\phi}_{\mathrm{B}}(\mathbf{X}) = \mathbf{c}$$

Now we introduce approximate values $\mathbf{X}^0$ and perform first-order Taylor Series approximation (see section 6.8.4)

We obtain

linearized observation equations $\mathbf{L} + \mathbf{v} = \mathbf{A}\left(\hat{\mathbf{X}} - \mathbf{X}^0\right) + \boldsymbol{\phi}(\mathbf{X}^0)$ (1)

linearized constraints $\mathbf{c} = \mathbf{C}\left(\hat{\mathbf{X}} - \mathbf{X}^0\right) + \boldsymbol{\phi}_{\mathrm{B}}(\mathbf{X}^0)$ (2)

with the Jacobian matrices

$$\mathbf{J}_1 = \frac{\partial \boldsymbol{\phi}(\mathbf{X})}{\partial \mathbf{X}}\Big|_{\mathbf{X}^0} \quad \text{and} \quad \mathbf{J}_2 = \frac{\partial \boldsymbol{\phi}_{\mathrm{B}}(\mathbf{X})}{\partial \mathbf{X}}\Big|_{\mathbf{X}^0}$$

we obtain the matrices $\mathbf{A} = \mathbf{J}_1$ and $\mathbf{C} = \mathbf{J}_2$

8.3 Nonlinear Case

We introduce

$\mathbf{l} = \mathbf{L} - \boldsymbol{\phi}(\mathbf{X}^0)$ vector of reduced observations

$\mathbf{w} = \boldsymbol{\phi}_\mathrm{B}(\mathbf{X}^0) - \mathbf{c}$ vector of misclosures

$\hat{\mathbf{x}} = \hat{\mathbf{X}} - \mathbf{X}^0$ vector of reduced unknowns

We can write (1) and (2) as

$$\mathbf{v} = \mathbf{A}\hat{\mathbf{x}} - \mathbf{l}$$

$$\mathbf{0} = \mathbf{C}\hat{\mathbf{x}} + \mathbf{w}$$

$\Omega = \mathbf{v}^\mathrm{T}\mathbf{P}\mathbf{v} \rightarrow$ min with constraints $\mathbf{C}\hat{\mathbf{x}} + \mathbf{w} = \mathbf{0}$

We apply Lagrange method for constraint minima:

$$\Omega = (\mathbf{A}\hat{\mathbf{x}} - \mathbf{l})^\mathrm{T}\mathbf{P}(\mathbf{A}\hat{\mathbf{x}} - \mathbf{l}) + 2\mathbf{k}^\mathrm{T}(\mathbf{C}\hat{\mathbf{x}} + \mathbf{w})$$

$\vdots$

$\mathbf{k}$ → vector of Lagrange multipliers (LM) (vector of correlates)

$$\Omega = \hat{\mathbf{x}}^\mathrm{T}\mathbf{A}^\mathrm{T}\mathbf{P}\mathbf{A}\hat{\mathbf{x}} - 2\hat{\mathbf{x}}^\mathrm{T}\mathbf{A}^\mathrm{T}\mathbf{P}\,\mathbf{l} + \mathbf{l}^\mathrm{T}\mathbf{P}\,\mathbf{l} + 2\mathbf{k}^\mathrm{T}\mathbf{C}\hat{\mathbf{x}} + 2\mathbf{k}^\mathrm{T}\mathbf{w}$$

8.3 Nonlinear Case

Minimum

$$\frac{\partial\Omega}{\partial\hat{\mathbf{x}}^{\mathrm{T}}} = 2\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\hat{\mathbf{x}} - 2\,\mathbf{A}^{\mathrm{T}}\mathbf{P}\,\mathbf{l} \qquad + 2\,\mathbf{C}^{\mathrm{T}}\mathbf{k} \qquad = 0$$

$$\frac{\partial\Omega}{\partial\mathbf{k}^{\mathrm{T}}} = \qquad 2\,\mathbf{C}\hat{\mathbf{x}} + 2\mathbf{w} \quad = 0$$

Linear equation system

$$\boxed{\left[\begin{array}{c:c} \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A} & \mathbf{C}^{\mathrm{T}} \\ \hdashline \mathbf{C} & \mathbf{0} \end{array}\right] \cdot \left[\begin{array}{c} \hat{\mathbf{x}} \\ \hdashline \mathbf{k} \end{array}\right] = \left[\begin{array}{c} \mathbf{A}^{\mathrm{T}}\mathbf{P}\,\mathbf{l} \\ \hdashline -\mathbf{w} \end{array}\right]}$$

"extended normal matrix", if matrix regular …

… **Solution**

$$\boxed{\left[\begin{array}{c} \hat{\mathbf{x}} \\ \hdashline \mathbf{k} \end{array}\right] = \left[\begin{array}{c:c} \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A} & \mathbf{C}^{\mathrm{T}} \\ \hdashline \mathbf{C} & \mathbf{0} \end{array}\right]^{-1} \left[\begin{array}{c} \mathbf{A}^{\mathrm{T}}\mathbf{P}\,\mathbf{l} \\ \hdashline -\mathbf{w} \end{array}\right]}$$

Residuals

$$\boxed{\mathbf{v} = \mathbf{A}\hat{\mathbf{x}} - \mathbf{l}}$$

Adjusted obs.

$$\boxed{\hat{\mathbf{L}} = \mathbf{L} + \mathbf{v}}$$

What is $\hat{\mathbf{x}}$?

$\rightarrow$ $\hat{\mathbf{x}}$ is a correction for initial values $\mathbf{X}^0$

$\rightarrow$ **Adjusted unknowns**

$$\boxed{\hat{\mathbf{X}} = \mathbf{X}^0 + \hat{\mathbf{x}}}$$

8.3 Nonlinear Case

Check:

$$\max |\hat{x}_i| < \varepsilon$$

and

$$\max \left|\hat{L}_i - \Phi_i(\widehat{\mathbf{X}})\right| < \delta$$

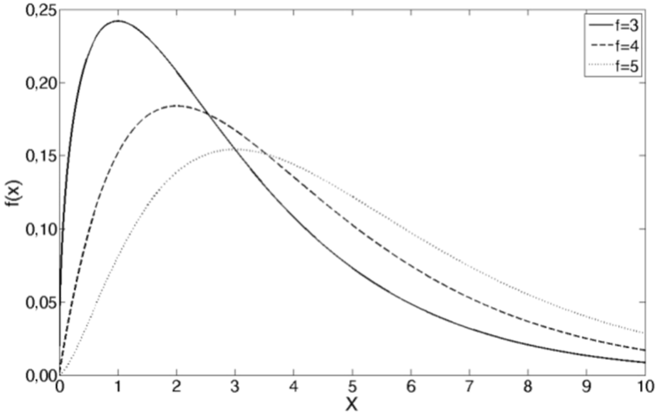
no

Use $\widehat{\mathbf{X}}$ as new $\mathbf{X}^0$

yes

$\widehat{\mathbf{X}}$ is final result

Computation of s_0, $\mathbf{Q}_{\hat{X}\hat{X}}$, $\mathbf{\Sigma}_{\hat{X}\hat{X}}$ same as for the linear case

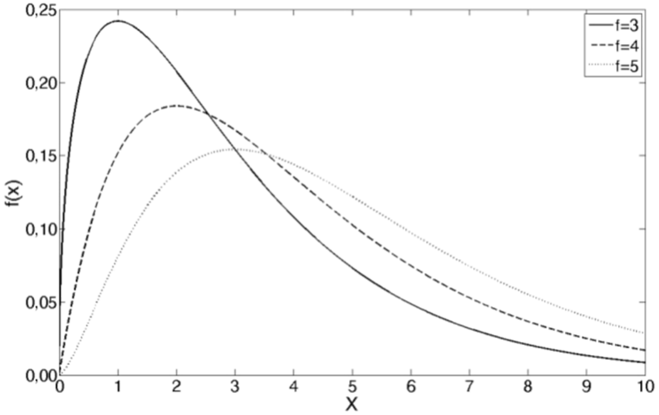


Technische Universität Berlin

Adjustment Theory I

Chapter 8 – Least Squares Adjustment with constraints for the unknown parameters

Prof. Dr.-Ing Frank Neitzel | Institute of Geodesy and Geoinformation Science



Adjustment Theory I

Chapter 9 – Least Squares Adjustment with constant values in the functional model

Prof. Dr.-Ing Frank Neitzel | Institute of Geodesy and Geoinformation Science

Version: 09.02.2025

Contents

9 Adjustment with constant values

Until now:

Functional model

$$\mathbf{L} = \mathbf{AX}$$

Now:

$$\mathbf{L} = \mathbf{AX} + \mathbf{Bc}$$

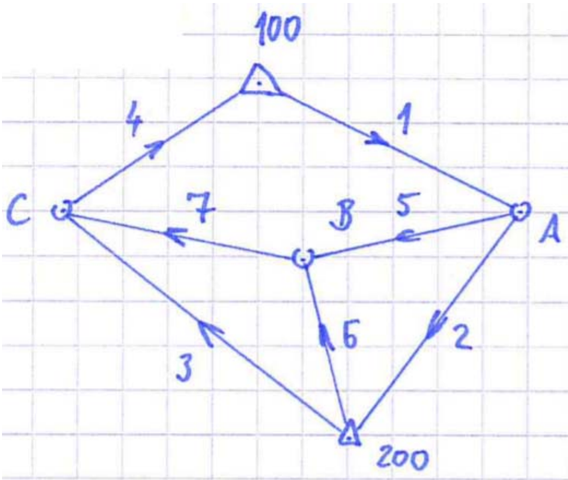
with

$\mathbf{c}$ … Vector of constant (fixed / error free) values
$\mathbf{B}$ … Coefficient matrix of constant (fixed / error free) values

Consideration of additive or subtractive constant values
in adjustment problems with a linear functional model

9 Adjustment with constant values

Example from Section 7.1:



Given:

- Heights of the benchmark points

Point No.	Height H [m]
100	100.000
200	107.500

- Observed elevation differences

Line	Observed elevation diff. [m]
1	5.100
2	2.340
3	−1.250
4	−6.130
5	−0.680
6	−3.000
7	1.700

The heights of the benchmark points can be regarded as fixed (error-free) values, the observations are equally weighted and uncorrelated

Determine the adjusted heights of points A, B, C and their standard deviation

9 Adjustment with constant values

General considerations:

- What are our unknowns?

  → Heights of points A, B, C

  → We introduce H_A, H_B, H_C

- What are our observations?

  → Elevation differences

  → $\Delta h_{100,A}, \Delta h_{A,200}, \Delta h_{200,C}, \Delta h_{C,100}, \Delta h_{A,B}, \Delta h_{200,B}, \Delta h_{B,C}$

- What are our fixed values?

  → H_{100}, H_{200}

- Redundancy?

  → $r = n - u \;\rightarrow r = 7 - 3 \;\rightarrow r = 4$

9 Adjustment with constant values

Functional model:

$$\Delta h_{100,A} = H_A - H_{100}$$
$$\Delta h_{A,200} = H_{200} - H_A$$
$$\Delta h_{200,C} = H_C - H_{200}$$
$$\Delta h_{C,100} = H_{100} - H_C$$
$$\Delta h_{A,B} = H_B - H_A$$
$$\Delta h_{200,B} = H_B - H_{200}$$
$$\Delta h_{B,C} = H_C - H_B$$

We insert the constant (fixed) values for H_{100} and H_{200} and bring them to the left-hand side of the equations …

… Functional model

$$\varphi_1: \Delta h_{100,A} + 100.000 = H_A$$
$$\varphi_2: \Delta h_{A,200} - 107.500 = -H_A$$
$$\varphi_3: \Delta h_{200,C} + 107.500 = H_C$$
$$\varphi_4: \Delta h_{C,100} - 100.000 = -H_C$$
$$\varphi_5: \Delta h_{A,B} = H_B - H_A$$
$$\varphi_6: \Delta h_{200,B} + 107.500 = H_B$$
$$\varphi_7: \Delta h_{B,C} = H_C - H_B$$

observation vector $\mathbf{L}'$

Linear or nonlinear? → Linear!

9 Adjustment with constant values

Observation equations:

$$
\begin{aligned}
\Delta h_{100,A} + 100.000 + v_1 &= \hat{H}_A \\
\Delta h_{A,200} - 107.500 + v_2 &= -\hat{H}_A \\
\Delta h_{200,C} + 107.500 + v_3 &= \hat{H}_C \\
\Delta h_{C,100} - 100.000 + v_4 &= -\hat{H}_C \\
\Delta h_{A,B} \qquad\qquad\quad + v_5 &= \hat{H}_B - \hat{H}_A \\
\Delta h_{200,B} + 107.500 + v_6 &= \hat{H}_B \\
\Delta h_{B,C} \qquad\qquad\quad + v_7 &= \hat{H}_C - \hat{H}_B
\end{aligned}
$$

9 Adjustment with constant values

Observation vector:

$$\mathbf{L}' = \begin{bmatrix} \Delta h_{100,A} + 100.000 \\ \Delta h_{A,200} - 107.500 \\ \Delta h_{200,C} + 107.500 \\ \Delta h_{C,100} - 100.000 \\ \Delta h_{A,B} \\ \Delta h_{200,B} + 107.500 \\ \Delta h_{B,C} \end{bmatrix}$$

Vector of residuals:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{bmatrix}$$

9 Adjustment with constant values

Stochastic Model of the observations:

$$p_1 = 1, p_2 = 1, \dots, p_7 = 1 \;\rightarrow \mathbf{P} = \mathbf{I}$$

Vector of unknowns:

$$\hat{\mathbf{X}} = \begin{bmatrix} \hat{H}_A \\ \hat{H}_B \\ \hat{H}_C \end{bmatrix}$$

Design Matrix (Matrix with coefficients of the linear functional model):

$$\mathbf{A} = \begin{array}{c} \\ \varphi_1 \\ \varphi_2 \\ \varphi_3 \\ \varphi_4 \\ \varphi_5 \\ \varphi_6 \\ \varphi_7 \end{array} \begin{array}{c} \begin{array}{ccc} H_A & H_B & H_C \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \end{array}$$

9 Adjustment with constant values

Normal equations:

$$\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\widehat{\mathbf{X}} = \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{L}' \qquad \text{with} \qquad \mathbf{P} = \mathbf{I}$$

$$\widehat{\mathbf{X}} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{L}' \qquad \text{with} \qquad \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1} = \mathbf{Q}_{\hat{X}\hat{X}}$$

Residuals:

$$\mathbf{v} = \mathbf{A}\widehat{\mathbf{X}} - \mathbf{L}'$$

Adjusted observations:

$$\hat{\mathbf{L}}' = \mathbf{L}' + \mathbf{v}$$

9 Adjustment with constant values

Final check:

$$\hat{\mathbf{L}}' - \boldsymbol{\Phi}(\hat{\mathbf{X}}) \overset{!}{=} \mathbf{0}$$ → zero within computing precision

Computer: $\hat{\mathbf{L}}' - \boldsymbol{\Phi}(\hat{\mathbf{X}}) \leq \delta$ → e.g. 10^{-12}

Empirical reference standard deviation:

$$s_0 = \sqrt{\frac{\mathbf{v}^{\mathrm{T}}\mathbf{P}\mathbf{v}}{n-u}}$$

VCM of adjusted unknowns:

$$\boldsymbol{\Sigma}_{\hat{X}\hat{X}} = \mathrm{s}_0^2 \cdot \mathbf{Q}_{\hat{X}\hat{X}}$$

9 Adjustment with constant values

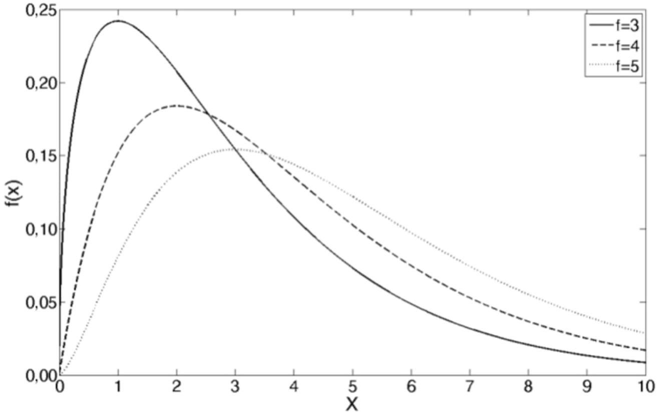
Derivation of a generalized approach (see blackboard)

$$\mathbf{L} + \mathbf{v} = \mathbf{A}\hat{\mathbf{X}} + \mathbf{Bc}$$

$$\vdots$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\hat{\mathbf{X}} = \mathbf{A}^{\mathrm{T}}\mathbf{P}(\mathbf{L} - \mathbf{Bc})$$

$$\hat{\mathbf{X}} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}(\mathbf{L} - \mathbf{Bc})$$



Adjustment Theory I

Chapter 9 – Least Squares Adjustment with constant values in the functional model

Prof. Dr.-Ing Frank Neitzel | Institute of Geodesy and Geoinformation Science