LEAST SQUARES ADJUSTMENT
A MODERN APPROACH

by
PETER MEISSL

Part B: Stochastic Approach
Toward Least Squares Adjustment

MITTEILUNGEN
der geodätischen Institute der Technischen Universität Graz
Folge 43

Graz, 1982

Herausgeber:
Geodätische Institute der Technischen Universität Graz

Redaktion für diese Folge:
Abteilung für Mathematische Geodäsie und Geoinformatik
des Institutes für Theoretische Geodäsie

Mit freundlicher Genehmigung der Geodätischen Institute der Technischen Universität Graz wurde diese Folge am Institut für Geodäsie und Geoinformation der Universität Bonn eingescannt.

Druck und Herstellung:
Druck- und Kopierzentrum der Technischen Universität Graz

Adresse:
Technische Universität Graz
Rechbauerstraße 12
A-8010 Graz, Österreich.

Preface

For his lectures at the Tongji University in Shanghai and at other institutions in China in November - December 1981, Peter Meissl prepared a set of lecture notes on contemporary least-squares adjustment and applications. Subsequently he worked on correcting and expanding them, but this was interrupted by his tragic death on May 22, 1982. (For Peter Meissl's life and work, the reader is referred to his biography by Franz Allmer, Mitteilungen der geodätischen Institute der Technischen Universität Graz, Folge 44, 1983.)

In view of the unique importance of this work, the Institute of Theoretical Geodesy decided to edit the manuscript posthumously and to publish the book in the series of the Geodetic Institutes of the Technical University, Graz, although Peter Meissl himself would certainly have included additional topics such as inner adjustment theory, expanded others such as the theory of large networks, and polished the manuscript much more before being satisfied with its publication.

The finishing of the book is due to Peter Meissl's closest associates: Dr. Norbert Bartelme, Dr. Helmut Fuchs, Dr. Bernhard Hofmann-Wellenhof, Dipl.-Ing. Wolf-Dieter Schuh and Dipl.-Ing. Manfred Wieser. In addition to being responsible for the careful editing of the manuscript, they also prepared the printing text using the word processing facilities of the computer WANG 2200 MVP.

A glance at the table of contents shows that this book is a thoroughly modern text on least-squares adjustment. In the contemporary spirit, the usual linear algebra is treated in the context of general linear spaces, which makes possible an easy transition to Hilbert space important for advanced topics. Also modern is the division into an algebraic and geometric approach (without statistics) and a stochastic approach, including statistical tests. Applications to Doppler observations, large networks, geodetic data bases, and splines essentially increase the practical usefulness. Although the book develops adjustment theory in a systematic and self-contained way, it will be best appreciated by readers who already have some elementary previous knowledge of adjustment computations.

The book needs no recommendation. Both students and research workers will find it indispensable. It is a fitting memorial of a great scientist.

Helmut Moritz

TABLE OF CONTENTS

A.5. Projectors

1. Decomposition of a vector space into a direct sum of subspaces
2. Orthocomplementary subspaces
3. Theorem by Pythagoras
4. Matrix representation of orthogonal projectors
5. Projections of functionals

A.6. Least squares adjustment

1. Projecting the vector of observations
2. Inhomogeneous form of least squares adjustment
3. Fundamental rectangular triangle of least squares adjustment
4. Least squares adjustment by projecting functionals

A.7. Partitioned matrices

1. Definitions
2. Computational rules
3. Block diagonality
4. Block Gauss elimination
5. Theoretical background of partitioned matrices

A.8. Isometric mappings between inner product spaces

1. Definitions
2. Preservation of inner products
3. Matrix representation
4. Examples of isometric mappings
5. Canonical transformation of an adjustment problem

A.9. Partial reduction

1. Partitioning the set of parameters
2. Partial reduction of the normal equations
3. Orthogonal decomposition of the parameter space
4. Partially reduced observation equations
5. Alternative derivation of the partially reduced observation equations

A.10. Adjustment phased with respect to observations

1. Formulation of the problem
2. Addition of normal equations
3. Updating the solution of the previous phase
4. Geometrical insight
5. Pre-elimination of group-internal unknowns
6. Helmert blocking

A.11. Complementary extremum principles in least squares adjustment

1. Basic geometric principle
2. Reformulation for linear manifolds
3. Adjustment by minimizing the norm of the residuals
4. Adjustment by minimizing variances

A.12. Generalized inverses

1. Range space and null space of a linear operator
2. g-inverse
3. Reflexive generalized inverse
4. Generalized inverse with least squares property
5. Generalized inverse with minimum norm property
6. Minimum norm least squares inverse
7. Pseudo inverse

A.13. Adjustment of rank-deficient systems

1. Formulation of the problem
2. Solution via generalized inverse of A
3. A minimum norm property of the covariance matrix of adjusted parameters
4. Solution via singular normal equations
5. Calculation of the ℓm-inverse
6. Application to free network adjustment

B. STOCHASTIC APPROACH TOWARD LEAST SQUARES ADJUSTMENT

B.1. Probabilities

1. Relative frequencies
2. Probability space
3. Examples
4. Calculus of probabilities

B.2. Random variables

1. One-dimensional random variables
2. Probability density function
3. n-dimensional random variables
4. Functions of random variables
5. Marginal distribution
6. Stochastic independence

B.3. Expectation, variances and covariances

1. Expectation of a one-dimensional random variable
2. Variance of a one-dimensional random variable
3. Various kinds of observation errors
4. Simple computational rules for expectation and variance
5. The case of higher dimensional random variables
6. Covariance matrix
7. Propagation of expectations and covariances
8. Important special cases
9. Zero correlation and stochastic independence

B.4. Gauss-Markoff model of least squares adjustment

1. Stochastic model
2. Unbiased estimates
3. Best linear unbiased estimation
4. Error calculus

B.5. Applications of the error propagation law

1. Triangle with three measured sides
2. First fundamental problem in the plane
3. Error ellipses
4. Polar survey with redundancy
5. Area calculated from polar survey
6. Conventionally adjusted regular traverse
7. Rigorously adjusted regular traverse
8. Systematic errors in a regular traverse

C. CONFIDENCE REGIONS AND TESTS OF LINEAR HYPOTHESES

C.1. Probability distributions used in statistical tests

1. One-dimensional Gauss distribution (normal distribution)
2. Multi-dimensional Gauss distribution (normal distribution)
3. Chi-squared distribution (χ^2-distribution)
4. Student's distribution (t-distribution)
5. Fisher's distribution (F-distribution)

C.2. Canonical transformation

1. Preliminaries
2. Making the functionals a part of the parameters
3. Orthogonal decomposition of the space L_A
4. Orthogonal decomposition of L into L_A and L_B
5. Orthonormalizing the bases of the subspaces

C.3. Distribution of various quantities resulting from least squares adjustment

1. Joint distribution of BLUE's and residuals
2. Distribution of the "weighted sum of residuals"
3. Expressions in $\Phi\tilde{x}$ and v having χ^2- or t-distribution
4. Expressions in $\tilde{x}$ and v having t-distribution

C.4. Confidence regions

1. Confidence intervals for one-dimensional Gauss variables
2. Application to the Gauss-Markoff model with known unit weight error
3. Studentization
4. Confidence regions for σ^2
5. Ellipsoidal confidence regions for sets of linear estimates

C.5. Tests of linear hypotheses

1. Linear hypotheses
2. Tests of variances
3. A simple example
4. A sophisticated example

D. SPECIAL TOPICS

D.1. Adjustment of Doppler observations

1. Transit system
2. Observing a difference in light travel time
3. Frequency shift
4. Technique of cycle counting
5. Parameters accounting for receiver imperfections
6. Transformations into an earth-fixed frame
7. Parameters accounting for orbit corrections
8. Linearization of the observation equations
9. Single station adjustment
10. Multi-station adjustment

D.2. Geodetic data bases

1. Storage media
2. Requirements for geodetic data bases
3. The data base of NGS

D.3. Cholesky's algorithm applied to normal equations of geodetic networks

1. Cholesky's algorithm for a general symmetric positive definite system
2. Partial reduction by Cholesky's algorithm
3. Geodetic normal equations
4. Geodetic interpretation of the partially Cholesky-reduced system
5. Problem of station ordering

D.4. One-dimensional cubic spline interpolation

1. Introduction
2. Parameterizing a cubic polynomial
3. Conditions at the inner nodes
4. Boundary conditions
5. Tridiagonal linear system
6. Modification of the periodic case
7. Interpolation of curves in the plane
8. Splines viewed as a vector space
9. The locality of splines

D.5. Two-dimensional spline interpolation

1. Introduction
2. Bicubic polynomials
3. Hermite bicubic interpolation
4. Bicubic splines

D.6. Geometry of exact spline interpolation

1. Formulation of the problem
2. Definition of splines
3. Existence and uniqueness of splines
4. Minimum properties of splines
5. Other examples
6. Prediction as a special case of spline interpolation
7. Noise-free collocation with trend parameters

D.7. Approximation with splines

1. Introduction
2. Approximation in one dimension
3. Basis splines with local support
4. Two dimensions

B. THE STOCHASTIC APPROACH TOWARD LEAST SQUARES ADJUSTMENT

1. Probabilities.

1.1 Relative frequencies.

Imagine a box filled with N identically shaped cards. Each of the individual cards carries one out of the letters A,B,C. Suppose that the corresponding absolute frequencies are N_A, N_B, N_C. Of course

$$N_A + N_B + N_C = N$$

We also introduce the relative frequencies

$$f_A = \frac{N_A}{N}, \quad f_B = \frac{N_B}{N}, \quad f_C = \frac{N_C}{N}$$

It follows that

$$f_A + f_B + f_C = 1$$

Suppose that the contents of the box are thoroughly mixed and that one card is drawn at random. We call this experiment an elementary event.

An event is defined as a set of elementary events. All events related to the present experiment are quickly listed as follows:

Φ	the empty set ... the impossible event
A	the writing on the drawn card shows the letter A
B	similar
C	similar
$A \cup B$	the letter A or the letter B is written on the card
$A \cup C$	similar
$B \cup C$	similar
$\Omega = A \cup B \cup C$	any letter is written on the card ... the certain event

We do not hesitate to assign probabilities to these events as follows

$$p(\Phi) = 0$$

$$p(A) = f_A, \quad p(B) = f_B, \quad p(C) = f_C$$

$$p(A \cup B) = f_A + f_B, \quad p(A \cup C) = f_A + f_C, \quad p(B \cup C) = f_B + f_C,$$

$$p(\Omega) = 1$$

There are alternative ways to identify an event. The symbol $\bar{A}$ denotes the event "Not A", i.e. the letter A does not appear on the drawn card. Obviously $\bar{A} = B \cup C$. Consequently $p(\bar{A}) = p(B \cup C) = 1 - f_A = f_B + f_C$.

1.2. Probability space.

Let Ω be a set of elements. The elements are called elementary events. Let Σ be a collection of subsets of Ω. Σ is a set of sets. The subsets of Ω which belong to Σ are called events. Not all possible subsets of Ω may be in Σ. We require the following properties of Σ.

(1) $\Omega \in \Sigma$

(2) Σ is closed with respect to complementation (if $A \in \Sigma$ then $\bar{A} \in \Sigma$), and with respect to the formation of countable unions (if $A_1, A_2, \ldots \in \Sigma$, then $A_1 \cup A_2 \cup \ldots \in \Sigma$).

It follows that $\Phi \in \Sigma$. (Ω is in Σ, hence $\Phi = \bar{\Omega}$ must be in Σ). Furthermore, if $A_1, A_2, \ldots$ is a countable sequence of events, $\bigcap_{i=1}^{\infty} A_i = A_1 \cap A_2 \cap \ldots$ must be in Σ. This holds because

$$\overline{\bigcap_{i=1}^{\infty} A_i} = \bigcup_{i=1}^{\infty} \bar{A}_i$$

A collection of subsets having the indicated properties is called a sigma-field of subsets (events). Frequently Σ is also called a Borel field of sets.

Let p be a functional defined on the subsets in Σ. Thus p assigns a number to any $A \in \Sigma$. This number is denoted p(A). It is called the probability of the event A. (There is no point asking whether the functional p is linear. The domain Σ is not necessarily a vector space!).

We require the following properties of the functional p.

(1) $0 \leq p(A) \leq 1$

(2) $p(\Omega) = 1$

(3) Let A_1, A_2... be a sequence of mutually nonintersecting events, i.e.

$A_i \cap A_j = 0$, if $i \neq j$.

Then

$$p(A_1 \cup A_2 \cup \ldots) = p(A_1) + p(A_2) + \ldots$$

i.e.

$$p\left\{\bigcup_{n=1}^{\infty} A_n\right\} = \sum_{n=1}^{\infty} p(A_n), \qquad \text{if} \quad A_i \cap A_j = 0 \text{ for } i \neq j$$

1.3. Examples.

1.3.1. Drawing cards out of a box.

Events and probabilities are defined in section 1.1.

1.3.2. Shooting against a target butt fixed to a wall.

Events are Borel sets of points in the plane. Borel sets in the plane are defined as follows. Suppose that a Cartesian coordinate system is chosen in the plane. The Borel field of sets in the plane is then defined as the smallest Σ field containing all rectangles of the form

$$a \leq x \leq b$$

$$c \leq y \leq d$$

It can be shown that most two dimensional sets which "can be imagined" are in Σ. It is an amusing and nontrivial task to show that for example the set of all points x, y fulfilling $ax + by < c$ is in Σ. The interior of a region bounded by a nonintersecting smooth curve is also a Borel set. Consider now a function $f(x,y)$, called a probability density function, having the following properties

$$f(x,y) \geq 0$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\, dxdy = 1$$

The density assigns a probability to any rectangle $a \leq x \leq b$, $c \leq y \leq d$ by means of

$$p\{a \leq x \leq b,\ c \leq y \leq d\} = \int_a^b \int_c^d f(x,y)\, dxdy$$

The properties of Borel sets and probabilities propagate the functional $p(A)$ from rectangles to any Borel set in the plane.

1.4. Calculus of probabilities.

We restrict ourselves to the following simple rules. Some of them have been anticipated in section 1.1.

$$p(\bar{A}) = 1 - p(A), \quad \bar{A} = \Omega - A$$

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

$$A \subset B \text{ implies } p(A) \leq p(B)$$

The calculus of probabilities is modelled after the calculus of relative frequencies.

2. Random variables.

2.1. One dimensional random variables.

Let Ω, Σ, p be the 3 ingredients of a probability space. Let X be a function mapping Ω into the real line R. X is called measureable, if for any real x the set

$$\{\omega \in \Omega \mid X(\omega) \leq x\}$$

is a member of Σ.

Remark: The notation $\xi(\omega)$ would be more appropriate, because $\xi(\omega)$ is a real number. However, we use the conventional notation $X(\omega)$.

To the mathematician a random variable is nothing but a measureable function. To the practician any number resulting from a random experiment is a random variable. Random variables quantify the outcome of an experiment. Instead of results like "A", "B", "C", "head", "tail", "male", "female", "win", "loss", "wet", "dry", "low", "medium", "big" etc. we get numbers. Any number resulting from a geodetic measurement will be considered as a random variable.

Since the above set $\{\omega \in \Omega \mid X(\omega) \leq x\}$ is in Σ, a probability may be assigned to it. We introduce the distribution function

$$F(x) = p\{\omega \in \Omega \mid X(\omega) \leq x\}$$

For the sake of brevity we write this as

$$F(x) = p\{X \leq x\}$$

The following properties of F(x) follow

$$0 \leq F(x) \leq 1$$

$$F(x_1) \leq F(x_2) \ \ldots \text{ if } x_1 \leq x_2 \quad \text{(monotonicity)}$$

$$\lim_{x \to -\infty} F(x) = 0$$

$$\lim_{x \to \infty} F(x) = 1$$

The distribution function makes it easy to specify probabilities for events $a \leq X \leq b$

$$p\{a \leq X \leq b\} = F(b) - F(a)$$

It follows that probabilities can be defined for any event of the form

$$\{X \in A\}$$

where A is a one-dimensional Borel set.

2.2. Probability density function.

Assume that F(x) is differentiable. Then

$$f(x) = F'(x)$$

$$F(x) = \int_{-\infty}^{x} f(y)dy$$

f(x) is called the probability density function of X. It follows that

$$p\{a \leq X \leq b\} = \int_a^b f(x)dx$$

The following properties hold

$$f(x) \geq 0$$

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

2.3. n-dimensional random variables.

Let $X_1, X_2, \ldots, X_n$ be n one-dimensional random variables. Any X_i maps Ω into the real line. We may introduce the n-dimensional random variable or random vector

$$X = \begin{bmatrix} X_1 \\ \cdot \\ \cdot \\ \cdot \\ X_n \end{bmatrix}$$

It is a mapping (function) from Ω into R^n. Because any X_i is measureable, any set of the form

$$\{\omega \in \Omega \mid X_1(\omega) \leq x_1, \ldots, X_n(\omega) \leq x_n\}$$

is an element of Σ. We may introduce the joint distribution function of $X_1, \ldots, X_n$ by defining

$$F(x_1, \ldots, x_n) = p\{\omega \in \Omega \mid X_1(\omega) \leq x_1, \ldots, X_n(\omega) \leq x_n\}$$

shorter

$$F(x_1,\ldots,x_n) = p\{X_1 \leq x_1,\ldots,X_n \leq x_n\}$$

$F(x_1,\ldots,x_n)$ is a scalar function of n independent variables having the following properties

$$0 \leq F(x_1,\ldots,x_n) \leq 1$$

$$F(x_1,\ldots,x_i,\ldots,x_n) \leq F(x_1,\ldots,\bar{x}_i,\ldots,x_n) \quad \text{if} \quad x_i \leq \bar{x}_i$$

(monotonicity in each variable separately)

$$\lim_{x_i \to -\infty} F(x_1,\ldots,x_n) = 0, \quad i=1,\ldots,n$$

$$\lim_{\substack{x_1 \to \infty \\ \vdots \\ x_n \to \infty}} F(x_1,\ldots,x_n) = 1$$

The probability of an event of the form

$$p\{a_i \leq X_i \leq b_i,\ i=1,\ldots,n\}$$

is obtained as

$$p\{a_i \leq X_i \leq b_i,\ i=1,\ldots,n\} = \Delta_1\Delta_2 \ldots \Delta_n F(x_1,\ldots,x_n)$$

Here Δ_i is the difference operator

$$\Delta_i \phi(x_1,\ldots,x_i,\ldots,x_n) = \phi(x_1,\ldots,b_i,\ldots,x_n) - \phi(x_1,\ldots,a_i,\ldots,x_n)$$

For example, in case of n=2 we have

$$p\{a_1 \leq x_1 \leq b_1,\ a_2 \leq x_2 \leq b_2\} =$$
$$F(b_1,b_2) - F(a_1,b_2) - F(b_1,a_2) + F(a_1,a_2)$$

The mechanism of Σ-fields and Borel sets propagates the assignment of probabilities to any set of the form

$$\{X \in A\}$$

where A is a Borel set in R^n. (Borel sets in R^n are the smallest Σ-field including all rectangular boxes with faces parallel to the coordinate planes. It turns out that the choice of a coordinate system does not affect the field of Borel sets.)

If $F(x_1,\ldots,x_n)$ is differentiable, we may introduce the density function $f(x_1,\ldots,x_n)$ by

$$f(x_1,\ldots,x_n) = \frac{\partial}{\partial x_1}\ldots\frac{\partial}{\partial x_n} F(x_1,\ldots,x_n)$$

It follows that

$$F(x_1,\ldots,x_n) = \int_{-\infty}^{x_1}\ldots\int_{-\infty}^{x_n} f(y_1,\ldots,y_n)\, dy_1\ldots dy_n$$

The probability

$$p\{a_i \leq X_i \leq b_i,\ i=1,\ldots,n\}$$

is given by the integral

$$\int_{a_1}^{b_1} \ldots \int_{a_n}^{b_n} f(x_1,\ldots,x_n)\, dx_1 \ldots dx_n$$

Moreover, we have for any Borel set

$$p\{X \in A\} = \int\int\ldots\int_A f(x_1,\ldots,x_n)\, dx_1 \ldots dx_n$$

(The integral is always defined in the sense of Lebesgue. In most cases it is identical to the familiar Riemann-integral.)

2.4. Functions of random variables.

Let $\varphi(x_1,\ldots,x_m)$ be a vector valued function mapping R^m into R^n.

$$y = \varphi(x)$$

is a shorthand notation for

$$\begin{aligned} y_1 &= \varphi(x_1,\ldots,x_m) \\ y_2 &= \varphi(x_1,\ldots,x_m) \\ &\ldots\ldots\ldots\ldots \\ y_n &= \varphi(x_1,\ldots,x_m) \end{aligned}$$

The function is called measureable, provided that the sets

$$\{x \in R^m \mid \varphi_1(x) \leq y_1, \ldots, \varphi_n(x) \leq y_n\}$$

are Borel sets in R^m for any choice of $y_1, \ldots, y_n$.

If X is an m-dimensional random variable, then

$$Y = \varphi(X)$$

is an n-dimensional random variable which is the image of X under the mapping φ.

If $F(x_1, \ldots, x_m)$ is the distribution function of X, then the distribution function $G(y_1, \ldots, y_n)$ of Y may in principle be deduced from

$$G(y_1, \ldots, y_n) = p\{x \in R^m \mid \varphi_1(x) \leq y_1, \ldots, \varphi_n(x) \leq y_n\}$$

The probability of the Borel set on the right hand side may be deduced from $F(x_1, \ldots, x_m)$.

Remark: If n=m and if the mapping $y = \varphi(x)$ is one to one and differentiable, and if X has a probability density $f(x_1, \ldots, x_n)$, then also Y has a probability density $g(y_1, \ldots, y_n)$ which is given by

$$g(y_1, \ldots, y_n) = f(x_1, \ldots, x_n) \left| \frac{\partial \varphi}{\partial x} \right|^{-1}$$

Here

$$\left|\frac{\partial \varphi}{\partial x}\right| = \begin{vmatrix} \frac{\partial \varphi_1}{\partial x_1} & \cdots\cdots & \frac{\partial \varphi_1}{\partial x_n} \\ \cdots\cdots & & \cdots\cdots \\ \cdots\cdots & & \cdots\cdots \\ \cdots\cdots & & \cdots\cdots \\ \frac{\partial \varphi_n}{\partial x_1} & \cdots\cdots & \frac{\partial \varphi_n}{\partial x_n} \end{vmatrix}$$

is the Jacobian determinant of the mapping. The proof follows from the familiar rule of substituting variables in an n-dimensional integral.

2.5. Marginal distribution.

Let

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} .$$

be a 2-dimensional random variable. Let $F(x_1, x_2)$ be the distribution function and $f(x_1, x_2)$ the density. Suppose that we are interested in the distribution of X_1 alone. We denote by $F_{(1)}(x_1)$ the distribution of X_1 and call it the marginal distribution. Similarly, we call the corresponding density $f_{(1)}(x_1)$ the marginal density. Obviously we have

$$F_{(1)}(x_1) = p\{X_1 \leq x_1\} = p\{X_1 \leq x_1,\ X_2 \leq \infty\} = F(x_1, \infty)$$

The corresponding density is obtained as

$$f_{(1)}(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_2$$

The procedure generalizes to higher dimensions in an obvious way. Let X be n-dimensional and partitioned as

$$X = \begin{bmatrix} X_1 \\ \cdot \\ \cdot \\ X_m \\ \hline X_{m+1} \\ \cdot \\ \cdot \\ X_n \end{bmatrix} = \begin{bmatrix} X_{(1)} \\ X_{(2)} \end{bmatrix} \text{, say}$$

The marginal distribution function of the m-dimensional random variable $X_{(1)}$ is obtained as

$$F_{(1)}(x_1, \ldots, x_m) = F(x_1, \ldots, x_m, \infty, \ldots, \infty)$$

The density follows from

$$f_{(1)}(x_1, \ldots, x_m) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(x_1, \ldots, x_m, x_{m+1}, \ldots, x_n)\, dx_{m+1} \ldots dx_n$$

2.6. Stochastic independence.

Assume again a 2-dimensional random variable

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

Suppose that the density decomposes as

$$f(x_1,x_2) = f_{(1)}(x_1)\, f_{(2)}(x_2)$$

The two 1-dimensional random variables X_1, X_2 are then called stochastically independent. The probability of the joint event $\{X_1 \in A_1,\ X_2 \in A_2\}$ is

$$p\{X_1 \in A_1,\ X_2 \in A_2\} = \int_A \int_A f(x_1,x_2)\, dx_1 dx_2 =$$
$$= \int_A \int_A f_{(1)}(x_1) f_{(2)}(x_2)\, dx_1 dx_2 = \int_A f_{(1)}(x_1)\, dx_1 * \int_A f_{(2)}(x_2)\, dx_2 =$$
$$= p\{X_1 \in A_1\}\, p\{X_2 \in A_2\}$$

This allows one to view X_1 and X_2 as the outcomes of two completely independent random experiments. There is no coupling between these two experiments. Knowing the outcome of experiment 1 tells us nothing about the outcome of experiment 2.

The concept of stochastic independence carries over to more than two dimensions. Represent an n-dimensional random variable as

$$X = \left[\begin{array}{c} X_1 \\ \cdot \\ \cdot \\ X_m \\ \hline X_{m+1} \\ \cdot \\ \cdot \\ X_n \end{array}\right] = \begin{bmatrix} X_{(1)} \\ X_{(2)} \end{bmatrix}$$

If the density decomposes as

$$f(x_1,\ldots,x_m,x_{m+1},\ldots,x_n) = f_{(1)}(x_1,\ldots,x_m)\, f_{(2)}(x_{m+1},\ldots,x_n)$$

Then the two subvectors $X_{(1)}, X_{(2)}$ are called stochastically independent. Again it holds that

$$p\{X_{(1)} \in A_1,\ X_{(2)} \in A_2\} = p\{X_{(1)} \in A_1\}\, p\{X_{(2)} \in A_2\}$$

for any Borel sets A_1, A_2 of appropriate dimensions.

An interesting special case arises if

$$f(x_1,\ldots,x_n) = f_{(1)}(x_1)\, f_{(2)}(x_2) \ldots f_{(n)}(x_n)$$

The components $X_1, \ldots, X_n$ are then mutually independent. We have

$$p\{X_1 \in A_1, \ldots, X_n \in A_n\} = p\{X_1 \in A_1\}\, p\{X_2 \in A_2\} \ldots p\{X_n \in A_n\}$$

Another point of interest is the following one. Suppose that the two subvectors $X_{(1)}$, $X_{(2)}$ of X are stochastically independent. Let

$$Y_{(1)} = \varphi_1(X_{(1)})$$
$$Y_{(2)} = \varphi_2(X_{(2)})$$

be two functions, where $Y_{(i)}$ depends on $X_{(i)}$ only. Then $Y_{(1)}$, $Y_{(2)}$ are also stochastically independent.

3. Expectation, variances and covariances.

3.1. Expectation of a one-dimensional random variable.

Let X be a 1-dimensional random variable having a probability density f(x). The expectation E(X) is defined as

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

Obviously E(X) represents some mean value of the random variable X. The various possible outcomes of X are averaged in agreement with f(x) as weight function. If a random variable is observed many times, and if the arithmetic mean of all outcomes is taken, one expects that the arithmetic mean is very close to E(X).

The above integral may not exist. However, in our applications existence is always tacitly assumed.

Let the one-dimensional random variable Y be a function of X

$$Y = \varphi(X)$$

Y has an expectation of its own. It may be defined in two different ways.

(1) The density g(y) of Y may be calculated as indicated in section 2.4. One then defines

$$E(Y) = \int_{-\infty}^{\infty} y\, g(y)\, dy$$

(2) One may directly define

$$E(Y) = \int_{-\infty}^{\infty} \varphi(x) f(x)\, dx$$

It may be shown that both definitions coincide. The equivalence proof is easy if X and Y have densities and if the mapping $Y = \varphi(X)$ is one to one and smooth.

3.2. The variance of a 1-dimensional random variable.

We denote for the sake of brevity:

$$\mu = E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

We put $\varphi(X) = (X-\mu)^2$ and calculate the expectation of $\varphi(X)$. The result is the socalled variance of X:

$$\sigma^2(X) = E((X-\mu)^2) = \int_{-\infty}^{\infty} (x-\mu)^2\, f(x)\, dx$$

Obviously $\sigma^2(X)$ measures in some way, how strongly X varies around its expectation. If the density f(x) is very much concentrated around $\mu = E(X)$, we anticipate a small variance. If f(x) has very wide and strong tails, the variance will be large.

3.3. Various kinds of observation errors.

In geodesy an observation ℓ is imagined as the superposition of a true value λ plus an observation error ε.

$$\ell = \lambda + \varepsilon$$

The observation error is a random variable, so is the observation ℓ. The true value λ is an (unknown) constant.

The expectation of the observation error

$$E(\varepsilon)$$

is called the systematic error of the observation. It systematically falsifies the observation, because

$$E(\ell) = E(\lambda+\varepsilon) = \int_{-\infty}^{\infty}(\lambda+\varepsilon)f(\varepsilon)d\varepsilon =$$
$$\lambda\underbrace{\int_{-\infty}^{\infty}f(\varepsilon)d\varepsilon}_{1} + \int_{-\infty}^{\infty}\varepsilon f(\varepsilon)d\varepsilon = \lambda + E(\varepsilon)$$

The underlying assumption in least squares adjustment is

$$E(\varepsilon) = 0$$

This assumption is frequently violated. Hence least squares adjustment is not always optimal.

The variance of ε

$$\sigma^2(\varepsilon) = \int_{-\infty}^{\infty} (\varepsilon - E(\varepsilon))^2 \, f(\varepsilon) \, d\varepsilon$$

is called mean square error. The square root

$$\sigma(\varepsilon) = \sqrt{\int_{-\infty}^{\infty} (\varepsilon - E(\varepsilon))^2 \, f(\varepsilon) \, d\varepsilon}$$

is called root mean square error. The variance of ℓ is also $\sigma^2(\varepsilon)$, because

$$\begin{aligned} \sigma^2(\ell) &= \int_{-\infty}^{\infty} [\lambda+\varepsilon-(\lambda+E(\varepsilon))]^2 \, f(\varepsilon) \, d\varepsilon \\ &= \int_{-\infty}^{\infty} (\varepsilon-E(\varepsilon))^2 \, f(\varepsilon) \, d\varepsilon = \sigma^2(\varepsilon) \end{aligned}$$

3.4. Simple computational rules for $E(X)$, $\sigma^2(X)$.

Some of these rules have been used in the previous subsection. Let c, c_1, c_2 denote any constants. Then:

$$\begin{aligned} E(cX) &= c\,E(X) \\ E(c+X) &= c + E(X) \\ E(c_1+c_2X) &= c_1 + c_2E(X) \\ \sigma^2(cX) &= c^2\sigma^2(X) \\ \sigma(cX) &= c\,\sigma(X) \\ \sigma^2(c+X) &= \sigma^2(X) \\ \sigma^2(c_1+c_2X) &= c_2{}^2\sigma^2(X) \end{aligned}$$

3.5. The case of higher dimensional random variables.

Let $X = (X_1,\ldots,X_m)^T$ be an m-dimensional random variable. Let $f(x_1,\ldots,x_m)$ be the joint density function. The expectation of X is defined as a vector

$$E(X) = \begin{bmatrix} E(X_1) \\ E(X_2) \\ . \\ . \\ E(X_m) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ . \\ . \\ \mu_m \end{bmatrix}$$

With

$$E(X_i) = \mu_i = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} x_i f(x_1,\ldots,x_m)\, dx_1 \ldots dx_m$$
$$= \int_{-\infty}^{\infty} x_i f_{(i)}(x_i)\, dx_i$$

Here

$$f_{(i)}(x_i) = \int_{-\infty}^{\infty} f(x_1,\ldots,x_{i-1},x_i,x_{i+1},\ldots,x_m)\, dx_1 \ldots dx_{i-1}, dx_{i+1} \ldots dx_m$$

is the socalled marginal density of x_i. It is the density of X_i considered as a one-dimensional random variable. Confer section 2.5.

Let

$$y = \varphi(x)$$

denote a mapping from R_m into R_n. In component notation we have

$$y_1 = \varphi_1(x_1,\ldots,x_m)$$
$$y_2 = \varphi_2(x_1,\ldots,x_m)$$
$$\ldots\ldots\ldots\ldots$$
$$y_n = \varphi_n(x_1,\ldots,x_m)$$

The mapping also maps the random variable X onto the random variable Y.

$$Y = \varphi(X)$$

The expectation of Y may be computed in two different ways:

(1) The joint density $g(y_1,\ldots,y_n)$ of Y may be calculated. One then defines

$$E(Y) = \begin{bmatrix} E(Y_1) \\ \vdots \\ E(Y_n) \end{bmatrix} = \begin{bmatrix} \nu_1 \\ \vdots \\ \nu_n \end{bmatrix}$$

with

$$E(Y_i) = \int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} y_i g(y_1,\ldots,y_n)\, dy_1\ldots dy_n$$

$$= \int_{-\infty}^{\infty} y_i g_{(i)}(y_i)\, dy_i$$

Here $g_{(i)}$ is again the marginal density of the component Y_i.

(2) One defines directly

$$E(Y_i) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \varphi_i(x_1, \ldots, x_m)\, dx_1 \ldots dx_m$$

Both definitions are consistent.

3.6. Covariance matrix.

The covariance matrix of the random vector X is defined as

$$\Sigma = \Sigma(X) = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1m} \\ \vdots & & \vdots \\ \sigma_{m1} & \cdots & \sigma_{mm} \end{bmatrix}$$

with

$$\sigma_{ij} = E\{(X_i - \mu_i)(X_j - \mu_j)\} =$$

$$= \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} (x_i - \mu_i)(x_j - \mu_j)\, f(x_1 \ldots x_m)\, dx_1 \ldots dx_m$$

The diagonals

$$\sigma_{ii} = \sigma^2(X_i) = E((X_i - \mu_i)^2)$$

are just the variances of the individual components of X considered as onedimensional random variables. The off-diagonals

$$\sigma_{ij} = E\{(X_i - \mu_i)(X_j - \mu_j)\}$$

are something new and deserve discussion. Clearly, σ_{ij} measures in some way a coupling between the deviations of X_i from its expectation μ_i and the deviations of X_j from its expectation μ_j. If X_i and X_j have a tendency to deviate either both positively or both negatively from their expectations then σ_{ij} will be positive. This does not mean that a positive $X_i - \mu_i$ cannot occur together with a negative $X_j - \mu_j$. However, in the majority of cases the signs will be coupled as indicated. Similarly, σ_{ij} will be negative, if a positive $X_i - \mu_i$ prefers to be coupled to a negative $X_j - \mu_j$ and vice versa.

3.7. Propagation of expectations and covariances.

It suffices to consider linear inhomogeneous mappings

$$Y = AX + b$$

Here X is m-dimensional and Y is n-dimensional. A is a known n×m matrix and b is a known n-vector. From the linearity of integrals we derive at once the following important relations.

$$E(Y) = E(AX+b) = A\,E(X) + b$$
$$\Sigma(Y) = \Sigma(AX+b) = A\,\Sigma(X)\,A^T$$

Proof of the first law.

$$E(Y_i) = E\left(\sum_{j=1}^{m} a_{ij}X_j + b_i\right) =$$

$$= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left(\sum_j a_{ij}x_j + b_i\right) f(x_1,\ldots,x_m)\, dx_1\ldots dx_m$$

$$= \sum_{j=1}^{m} a_{ij} \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} x_j f(x_1,\ldots,x_m)\, dx_1\ldots dx_m +$$

$$+ b_i \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(x_1,\ldots,x_m)\, dx_1\ldots dx_m = \sum_{j=1}^{m} a_{ij}E(X_j) + b_i$$

as was to be shown.

Proof of the second law.

Denote

$$\Sigma' = \Sigma(Y) = \begin{bmatrix} \sigma'_{11} & \cdots & \sigma'_{1n} \\ \vdots & & \vdots \\ \sigma'_{n1} & \cdots & \sigma'_{nn} \end{bmatrix}$$

Then

$$\sigma'_{ij} = E\left\{(Y_i - \nu_i)(Y_j - \nu_j)\right\} =$$

$$= E\left\{\left(\sum_{k=1}^{n} a_{ik}X_k + b_i - \sum_{k=1}^{m} a_{ik}\mu_k - b_i\right)\left(\sum_{l=1}^{m} a_{jl}X_l + b_j - \sum_{l=1}^{m} a_{jl}\mu_l - b_j\right)\right\}$$

$$= E\left\{\sum_k a_{ik}(X_k - \mu_k)\sum_l a_{jl}(X_l - \mu_l)\right\}$$

$$= \sum_{kl} a_{ik}a_{jl}E\left\{(X_k - \mu_k)\cdot(X_l - \mu_l)\right\} = \sum_{kl} a_{ik}a_{jl}\sigma_{kl}$$

This is equivalent to

$$\Sigma' = A\Sigma A^T$$

which was to be shown.

Remark: The second law is called the law of propagation of covariances. In geodesy it is usually and simply called "the" error propagation law.

3.8. Important special cases.

If X_1 and X_2 are two random variables having a joint distribution then

$$E(X_1 + X_2) = E(X_1) + E(X_2)$$

Also

$$E\{\lambda_1 X_1 + \lambda_2 X_2\} = \lambda_1 E(X_1) + \lambda_2 E(X_2)$$

This is the linearity property of the expectation.

If X is a random vector, and if $\Sigma(X)$ is of diagonal form

$$\Sigma(X) = \begin{bmatrix} \sigma_{11} & 0 & \dots & 0 \\ \cdot & \sigma_{22} & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ 0 & \dots & \dots & \sigma_{nn} \end{bmatrix}$$

Then the components of X are called mutually uncorrelated. If Y is a onedimensional function

$$Y = a_1 x_1 + \ldots + a_n x_n = a^T X$$

Then

$$\sigma^2(Y) = a^T \Sigma(X) a = \sum_{i=1}^{n} a_i^2 \sigma^2(X_i)$$

3.9. Zero correlation and stochastic independence.

Suppose that X is 2-dimensional

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

Let the covariance matrix be

$$\Sigma(X) = \begin{bmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{bmatrix}$$

The two components are then uncorrelated.

Assume now that X_1, X_2 are stochastically independent. As we know from section 2.6, this is equivalent to

$$f(x_1, x_2) = f_{(1)}(x_1)\, f_{(2)}(x_2)$$

We show that stochastic independence implies zero correlation:

$$\sigma_{12} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1-\mu_1)(x_2-\mu_2)\, f(x_1,x_2)\, dx_1 dx_2, \quad \mu_i = E(x_i), \quad i=1,2$$

$$\sigma_{12} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1-\mu_1)(x_2-\mu_2)\, f_{(1)}(x_1)\, f_{(2)}(x_2)\, dx_1 dx_2$$

$$= \int_{-\infty}^{\infty} (x_1-\mu_1)\, f_{(1)}(x_1)\, dx_1 \int_{-\infty}^{\infty} (x_2-\mu_2)\, f_{(2)}(x_2)\, dx_2$$

$$= E(X_1-\mu_1)\, E(X_2-\mu_2) = 0.0 = 0$$

The converse is not true. Zero correlation does not necessarily imply stochastic independence.

The concept of zero correlation generalizes to more than two dimensions. Represent an n-dimensional random vector as

$$X = \left[\begin{array}{c} X_1 \\ \cdot \\ \cdot \\ X_m \\ \hline X_{m+1} \\ \cdot \\ \cdot \\ X_n \end{array}\right] = \left[\begin{array}{c} X_{(1)} \\ X_{(2)} \end{array}\right]$$

Suppose that the covariance matrix looks as

$$\Sigma(X) = \left[\begin{array}{ccc|ccc} \sigma_{11} & \cdots & \sigma_{1m} & & & \\ \vdots & & \vdots & & 0 & \\ \sigma_{m1} & \cdots & \sigma_{mm} & & & \\ \hline & & & \sigma_{m+1,m+1} & \cdots & \sigma_{m+1,n} \\ & 0 & & \vdots & & \vdots \\ & & & \sigma_{n,m+1} & \cdots & \sigma_{nn} \end{array}\right] = \left|\begin{array}{cc} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{array}\right|$$

The two subvectors $X_{(1)}$, $X_{(2)}$ are then called uncorrelated.

In section 2.6 we called $X_{(1)}$, $X_{(2)}$ stochastically independent if

$$f(x_1,\ldots,x_n) = f_{(1)}(x_1,\ldots,x_m)\, f_{(2)}(x_{m+1},\ldots,x_n)$$

One shows again that stochastic independence implies zero correlation. The converse is, however, generally not true.

4. The Gauss-Markoff model of least squares adjustment.

4.1. The stochastic model.

Remark on notation: From now on it will be completely impossible to adhere to the convention of using Latin letters for vectors and Greek letters for coordinates.

Let $\mathbf{l}$ be the vector of observations. Its components are denoted l_i, as in section A.6.1.

$$\mathbf{l} = \begin{bmatrix} l_1 \\ l_2 \\ . \\ . \\ l_n \end{bmatrix}$$

$\mathbf{l}$ is viewed as an n-dimensional random variable. In the mathematical sense, $\mathbf{l}$ comprises n measurable functions $l_i(\omega)$ mapping the set Ω introduced in section 1.2 into R^n. The image space R^n will be denoted L in the following. It is a vector space. Sometimes L is called sample space or space of realizations of $\mathbf{l}$. Once the observations are taken, the result are n numbers, the coordinates of a vector in L. Although it is logically unsatisfactory, this vector will occasionally be denoted by the same letter $\mathbf{l}$.

We introduce the vector ε of observation errors and the vector λ of "true" observables by the equation

$$\mathbf{l} = \lambda + \varepsilon$$

The components of λ are constants which are generally unknown. They are true angles, distances, height differences, i.e. observable quantities unaffected by observation errors. As a vector of constants, λ can be viewed as a vector in the n-dimensional space L. Although λ is generally unknown, some a priori information on it is available. It is known that the n unknown quantities λ_i are linearly expressible in terms of m unknown parameters $x_1, \ldots, x_m$, whereby $m < n$.

$$\lambda_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m$$
$$\lambda_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\lambda_n = a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m$$

Shortly

$$\lambda = Ax$$

with

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{bmatrix}$$

The n×m matrix A is assumed of rank m. Note that we use Latin letters now for the elements of A and the coordinates of x.

Example: Consider a leveling network involving the stations P_0, P_1, P_2, P_3. The height $H_0 = 0$ of P_0 is known. The heights H_1, H_2, H_3 of stations P_1, P_2. P_3 are unknown. Assume height difference measurements ℓ_{01}, ℓ_{03}, ℓ_{12}, ℓ_{13}, ℓ_{23}.

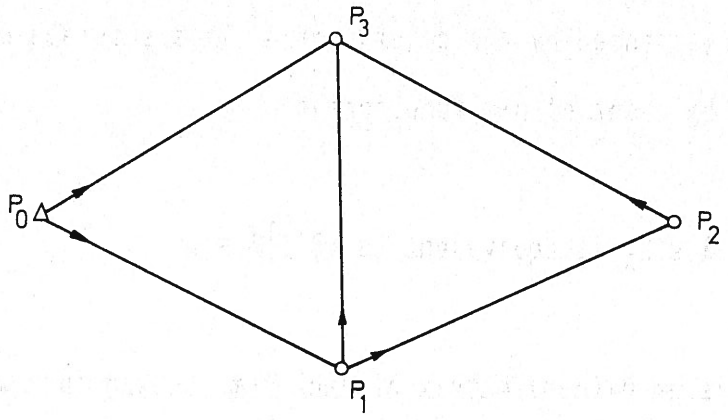

Fig. 4.1.

The true (unknown) height differences are

$$h_{01} = H_1 - H_0 = H_1$$
$$h_{03} = H_3 - H_0 = H_3$$
$$h_{12} = H_2 - H_1$$
$$h_{13} = H_3 - H_1$$
$$h_{23} = H_3 - H_2$$

Thus we have

$$\begin{bmatrix} h_{01} \\ h_{03} \\ h_{12} \\ h_{13} \\ h_{23} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} H_1 \\ H_2 \\ H_3 \end{bmatrix}$$

corresponding to $\lambda = Ax$.

Remark: The requirement $\lambda = Ax$ is equivalent to restricting the vector λ to the

subspace L_A spanned by the columns of A. This subspace could equally well be described by a set of n-m functionals:

$$\lambda \in L_A \text{ is equivalent to } (B')^T\lambda = 0$$

where B' is an $n\times(n-m)$ matrix of rank n-m, chosen in a way that

$$(B')^T A = 0$$

This would lead to the concept of adjustment by conditions.

Example: Refering to the above introduced leveling network, we have two condition equations of the form

$$h_{01} + h_{13} - h_{03} = 0$$
$$h_{12} + h_{23} - h_{13} = 0$$

i.e.

$$\begin{bmatrix} 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} h_{01} \\ h_{03} \\ h_{12} \\ h_{13} \\ h_{23} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This corresponds to

$$B^T\lambda = 0$$

We adhere to the model of adjustment by parameters x in

$$\lambda = Ax$$

Recall the representation of the observations ℓ in terms of true observables λ and observation errors ε.

$$\ell = \lambda + \varepsilon$$

We now introduce the important requirement

$$E(\varepsilon) = 0$$

It implies

$$E(\ell) = \lambda$$

or

$$E(\ell) = Ax$$

The requirement $E(\varepsilon) = 0$ means that the systematic error of the observations is assumed as zero. The observations are "unbiased". The postulate of unbiased observations is far reaching in theory. Unfortunately it is rarely fulfilled in practice. Many difficulties encountered during the practical application of least squares adjustment are caused by the failure of the observations to be truly unbiased.

We assume the covariance matrix of ε in the form

$$\Sigma(\varepsilon) = Q\sigma^2$$

Here Q is a known symmetric and positive definite n×n matrix. In most applications Q will be a diagonal matrix. The scalar factor σ^2 is called mean square unit weight error. It may be assumed either as known or as unknown.

Since the random vectors $\mathbf{l}$ and ε differ only by the constant vector λ, i.e. $\mathbf{l} = \lambda+\varepsilon$, the covariance matrix of $\mathbf{l}$ is identical to that of ε:

$$\Sigma(\mathbf{l}) = \Sigma(\varepsilon) = Q\sigma^2$$

We summarize the basic assumptions of the Gauss-Markoff model as follows

$$\boxed{\begin{aligned} E(\mathbf{l}) &= Ax \\ \Sigma(\mathbf{l}) &= Q\sigma^2 \end{aligned}}$$

The vector $\mathbf{l}$ is the vector of observations. The known matrix A is sometimes called design matrix. The unknown vector x is the vector of parameters. The known matrix Q is positive definite. Its inverse

$$P = Q^{-1}$$

is called the matrix of observational weights. The scalar σ^2 is either known or unknown. It is called the mean square unit weight error.

4.2. Unbiased estimates.

Let φ denote a linear functional on L_A. Because any vector in L_A is identified by its coordinates $x = (x_1,\ldots,x_m)^T$ with respect to the bases represented by the columns of A, the functional φ may be represented as

$$\varphi = \varphi^T x = \varphi_1 x_1 + \varphi_2 x_2 + \ldots + \varphi_m x_m$$

Thus a functional on L_A is a linear homogeneous function of the unknown parameters.

Example: Refering to the above introduced leveling network, we have an example for a functional by

$$\varphi^T H = \left(\frac{1}{4} \frac{1}{4} \frac{1}{4} \right) \begin{bmatrix} H_1 \\ H_2 \\ H_3 \end{bmatrix} = = \frac{1}{4} (H_1 + H_2 + H_3)$$

It is the mean value of the heights $H_0 = 0$ and H_1, H_2, H_3. Another example is simply

$$\varphi^T H = \left(0\ 1\ 0 \right) \begin{bmatrix} H_1 \\ H_2 \\ H_3 \end{bmatrix} = H_2$$

the height of the station H_2.

A third example is:

$$\varphi^T H = (0 \; -1 \; 1) \begin{bmatrix} H_1 \\ H_2 \\ H_3 \end{bmatrix} = H_3 - H_2 = h_{23}$$

the height difference $H_3 - H_2$.

Example: Consider a network adjustment. The observations are angles, distances, azimuths, Doppler positions etc. The vector ℓ represents the observation increments after linearization of the observation equations. The parameters x represent coordinate increments. A functional φ may be the (linearized) distance between two remote stations. Alternatively it may refer to an azimuth between any pair of stations, or to an angle between a triplet of stations. Also the area defined by a polygon whose corners are a subset of all stations is (after linearization) a functional of the considered type.

Remark: It is important to note that also any component x_i of the parameter vector x may be viewed as a functional on L_A. Recall section 2.8 where we pointed out that coordinates may be viewed as functionals. The functional x_i is represented by

$$\varphi^T = (0,\ldots,0,1,0,\ldots,0)$$

i.e. by the i-th row of the m×m unit matrix.

Remark: It is further important to note that any component λ_i of the vector $\lambda = E(\ell)$ may be viewed as a functional on L_A. Indeed we have

$$\lambda_i = a_{i1}x_1 + \ldots + a_{im}x_m$$

This means that λ_i is represented by

$$\varphi^T = (a_{i1},\ldots,a_{im})$$

i.e. by the i-th row of the matrix A.

The parameters x are unknown. Hence the value of the functional $\varphi(x) = \varphi^T x$ is unknown. The purpose of statistical estimation is to give estimates of $\varphi = \varphi(x)$ in terms of the observed values of ℓ. Such an estimate is denoted $\hat{\varphi}$. It is a function of the random vector ℓ. Thus $\hat{\varphi}$ is a random variable. It has an expectation $E(\hat{\varphi})$ and a variance $\sigma^2(\hat{\varphi})$.

We restrict ourselves to linear estimates

$$\hat{\varphi} = \beta_1\ell_1 + \beta_2\ell_2 + \ldots + \beta_n\ell_n$$

Shortly

$$\hat{\varphi} = \beta^T\ell, \quad \beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}$$

Remark: The vector β^T can be viewed as the 1×n matrix representation of a linear

functional defined on L.

The expectation of $\hat{\varphi}$ is

$$E(\hat{\varphi}) = E(\beta^T \ell) = \beta^T E(\ell) = \beta^T \lambda = \beta^T A x$$

It is seen that the expectation $E(\hat{\varphi})$ is a linear form in the unknowns x. The requirement

$$E(\hat{\varphi}) = \beta^T A x = \varphi^T x \quad \text{for any } x$$

results in the fundamental concept of a linear unbiased estimate (abbreviated LUE).

The requirement

$$\beta^T A x = \varphi^T x \quad \text{for any } x$$

is equivalent to

$$\beta^T A = \varphi^T$$

4.3. Best linear unbiased estimation.

The variance of the linear estimate

$$\hat{\varphi} = \beta^T \ell$$

is given by

$$\sigma^2(\hat{\varphi}) = \beta^T \Sigma(\ell) \beta = \beta^T Q \beta \, \sigma^2$$

We are looking for an unbiased estimate having a minimal variance. It will be denoted $\tilde{\varphi}$ and called best linear unbiased estimate (BLUE).

The BLUE is the solution of the following extremum problem. Find β such that

$$\beta^T Q \beta = \text{Minimum}$$

subject to

$$\beta^T A = \varphi^T$$

The minimum problem is purely algebraic and can be attacked by means of Lagrange-multipliers. However, we may also make use of the fact that the minimum problem was essentially solved in an earlier section 5.5, using a different notation.

As indicated above, nothing can prevent us from viewing β^T as the matrix representer of a linear functional β on L. The space L was called the sample space. It is the space of all possible realizations of the random vector ℓ of observations. Recall that λ was assumed in L and that L_A was assumed as a subspace of L.

We assign a norm to our functional β by putting

$$\|\beta\|^2 = (\beta,\beta) = \beta^T Q \beta$$

Up to the factor σ^2 the norm of the functional b is nothing but the variance of $\beta^T \ell$ viewed as a linear function of the random vector ℓ.

The matrix Q is interpreted as representing the inner product of functionals on L:

$$(\beta,\gamma) = \beta^T Q \gamma$$

In agreement with section A.4.9, Q is called the reproducing kernel of L. It follows that the inner product in L is represented by the matrix

$$P = Q^{-1}$$

The functional $\varphi = \varphi^T x$ is only defined on L_A. The requirement

$$\beta^T A = \varphi^T$$

is equivalent to requiring that the functional β shall coincide with the functional φ if applied to vectors out of L_A. The functional β is an extension of φ. Confer section A.5.5. Among all such functionals we search for one having minimal norm. According to section A.5.5. on projection of functionals the solution is given by a functional $\tilde{\varphi}$ defined on the whole of L, coinciding with φ

on L_A and vanishing on L_B, the orthocomplement of L_A in L:

$$\tilde{\varphi}(\ell) = \varphi(P_A\ell), \qquad \ell \in L$$

Given a vector $\ell \in L$, we form $P_A\ell$. We find the coordinates $\tilde{x}$ of $P_A\ell$:

$$P_A\ell = A\tilde{x}$$

and we form

$$\tilde{\varphi}(\ell) = \varphi^T\tilde{x}$$

The projection $P_A\ell$ is given by

$$P_A\ell = A(A^TPA)^{-1}A^TP\ell$$

Comparing with

$$P_A\ell = A\tilde{x}$$

we recognize

$$\tilde{x} = (A^TPA)^{-1}A^TP\ell$$

Hence

$$\tilde{\varphi}(\ell) = \varphi^T(A^TPA)^{-1}A^TP\ell$$

This is also obtained by the following rule. Solve the normals

$$(A^T P A)\tilde{x} = A^T P \ell$$

and apply the functional φ to the adjusted parameters $\tilde{x}$:

$$\tilde{\varphi}(\ell) = \varphi^T \tilde{x}$$

The minimal variance $\sigma^2(\tilde{\varphi}(\ell))$ of the best estimater for $\varphi = \varphi^T x$ is given by

$$\sigma^2(\tilde{\varphi}(\ell)) = \beta^T Q \beta \, \sigma^2, \quad \beta = \varphi^T (A^T P A)^{-1} A^T P$$

Because $PQ = I$, one obtains

$$\sigma^2(\tilde{\varphi})) = \varphi^T (A^T P A)^{-1} \varphi \, \sigma^2$$

4.4. Error calculus.

The adjusted parameters $\tilde{x}$ are linear functions of the observations

$$\tilde{x} = (A^T P A)^{-1} A^T P \ell = B\ell$$

The expectation of $\tilde{x}$ is x. The estimators $\tilde{x}$ are unbiased

$$\begin{aligned} E(\tilde{x}) &= E(B\ell) = BE(\ell) = BAx \\ &= (A^T P A)^{-1} A^T P A x = x \end{aligned}$$

The $\tilde{x}$ are BLUE for x (i.e. $\tilde{x}_i$ is the BLUE of x_i). This is also seen by putting $\varphi^T = (0,\ldots,0,1,0,\ldots,0)$ where the 1 appears at the i-th position.

The covariance matrix of $\tilde{x}$ is obtained as

$$\begin{aligned}\Sigma(\tilde{x}) &= B\Sigma(\ell)B^T = BQB^T\sigma^2 = \\ &= (A^TPA)^{-1}A^TPQPA(A^TPA)^{-1}\sigma^2 = (A^TPA)^{-1}(A^TPA)(A^TPA)^{-1}\sigma^2 \\ &= (A^TPA)^{-1}\sigma^2\end{aligned}$$

The covariance matrix of the adjusted parameters is the inverse of the normal equation matrix multiplied by σ^2.

Remark: Note the validity of the error propagation law

$$\tilde{\varphi}(\ell) = \varphi^T\tilde{x}$$

Hence

$$\sigma^2(\tilde{\varphi}(\ell)) = \sigma^2(\varphi^T\tilde{x}) = \varphi^T\Sigma(\tilde{x})\varphi = \varphi^T(A^TPA)^{-1}\varphi\,\sigma^2$$

as before.

Instead of a linear functional $\varphi = \varphi^T x$, we now consider a set of p functionals

$$\Phi = \Phi x$$

Here Φ is a p×m matrix

$$\Phi = \begin{bmatrix} \varphi_{11} & \cdots\cdots & \varphi_{1m} \\ \vdots & & \vdots \\ \varphi_{p1} & \cdots\cdots & \varphi_{pm} \end{bmatrix}$$

Any row of Φ represents one linear functional. Hence we immediately obtain that the best unbiased estimators for all functionals comprised by Φ are given by

$$\tilde{\Phi} = \Phi\tilde{x}$$

Their covariance matrix is

$$\Sigma(\tilde{\Phi}) = \Phi\Sigma(\tilde{x})\Phi^T = \Phi(A^TPA)^{-1}\Phi^T\sigma^2$$

One example for Φ is given by

$$\Phi = Ax$$

The best estimators are

$$\tilde{\Phi} = A\tilde{x} = \tilde{l}$$

It is usual to call them "adjusted observations". Their covariance matrix is

$$\Sigma(\tilde{l}) = A(A^TPA)^{-1}A^T\sigma^2$$

This is sometimes called the "a posteriori covariance matrix" of the

observations. The "a priori covariance matrix" is, of course, given by $\Sigma(\underline{l}) = Q\sigma^2$.

The residuals v are defined as the difference between adjusted and observed values.

$$v = \tilde{\underline{l}} - \underline{l} = A\tilde{x} - \underline{l}$$

Inserting for $\tilde{x}$, we obtain

$$v = -(I-A(A^TPA)^{-1}A^TP)\underline{l}$$
$$= -(I-P_A)\underline{l} = -P_B\underline{l}$$

The covariance matrix of v is obtained as

$$\Sigma(v) = P_B Q P_B^T \sigma^2 =$$
$$= (Q-A(A^TPA)^{-1}A^T)\sigma^2$$

One notices:

$$\Sigma(\underline{l}) = \Sigma(\tilde{\underline{l}}) + \Sigma(v)$$

This is the familiar theorem of Pythagoras in a new disguise.

It is also interesting to calculate the common covariance matrix of $\tilde{\ell}$ and v.

$$\tilde{\ell} = A\tilde{x} = A(A^T PA)^{-1}A^T P\ell = P_A\ell$$

$$v = -(I-A(A^T PA)^{-1}A^T P)\ell = -(I-P_A)\ell$$

One finds

$$\Sigma\begin{bmatrix}\tilde{\ell}\\ v\end{bmatrix} = \begin{bmatrix}\Sigma(\tilde{\ell}) & 0\\ 0 & \Sigma(v)\end{bmatrix}$$

where $\Sigma(\tilde{\ell})$ and $\Sigma(v)$ are the expressions derived above. The remarkable thing is the zero correlation between $\tilde{\ell}$ and v. Zero correlation is the stochastic counterpart to the geometric concept of orthogonality.

5. Applications of the error propagation law.

5.1. Triangle with three measured sides.

Consider the triangle of fig. 5.1.

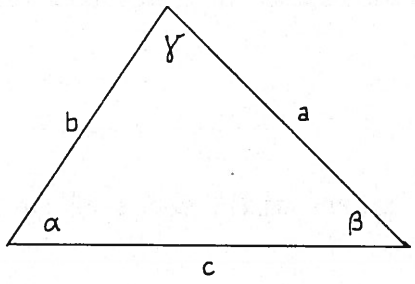

Fig. 5.1.

Assume that a,b,c are measured. Let the covariance matrix of the measurements be

$$\sum \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} m_a^2 & 0 & 0 \\ 0 & m_b^2 & 0 \\ 0 & 0 & m_c^2 \end{bmatrix}$$

What is the variance of the angle α?

By the law of cosine we have

$$a^2 = b^2 + c^2 - 2bc \cos \alpha$$

or

$$\cos \alpha = \frac{b^2 + c^2 - a^2}{2bc}, \quad \alpha = \arccos \frac{b^2 + c^2 - a^2}{2bc}$$

We have expressed α as a function of the observations a,b,c. However, the function is not linear, and does not allow an application of the error propagation law in its standard form. We must linearize the dependence of α on a,b,c.

We assume that m_a, m_b, m_c are small compared to a,b,c. We represent

$$a = a_0 + \Delta a, \quad b = b_0 + \Delta b, \quad c = c_0 + \Delta c$$

Here a_0, b_0, c_0 are fixed values, Δa, Δb, Δc are random variables. Because they differ from a,b,c only by constants a_0, b_0, c_0, the covariance matrix of Δa, Δb, Δc is the same as that one of a,b,c:

$$\sum \begin{bmatrix} \Delta a \\ \Delta b \\ \Delta c \end{bmatrix} = \sum \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} m_a^2 & 0 & 0 \\ 0 & m_b^2 & 0 \\ 0 & 0 & m_c^2 \end{bmatrix}$$

Call

$$\alpha_0 = \arccos \frac{b_0^2 + c_0^2 - a_0^2}{2b_0c_0}$$

and let

$$\alpha = \alpha_0 + \Delta\alpha$$

then

$$a^2 = b^2 + c^2 - 2bc \cos \alpha$$

goes over into

$$(a_0 + \Delta a)^2 = (b_0 + \Delta b)^2 + (c_0 + \Delta c)^2 - 2(b_0 + \Delta b)(c_0 + \Delta c) \cos(\alpha_0 + \Delta \alpha)$$

Applying Taylor's formula, and keeping only the linear terms, gives

$$2a_0 \Delta a = 2b_0 \Delta b + 2c_0 \Delta c - 2(c_0 \Delta b + b_0 \Delta c) \cos \alpha_0 + 2b_0 c_0 \sin \alpha_0 \Delta \alpha$$

The constant terms cancel due to the consistency of a_0, b_0, c_0 and α_0. We solve for $\Delta \alpha$ obtaining

$$\Delta \alpha = \frac{1}{b_0 c_0 \sin \alpha_0} \left\{ a_0 \Delta a - (b_0 - c_0 \cos \alpha_0) \Delta b - (c_0 - b_0 \cos \alpha_0) \Delta c \right\}$$

We abbreviate this as

$$\Delta \alpha = C_{\alpha a} \Delta a + C_{\alpha b} \Delta b + C_{\alpha c} \Delta c$$

This is the desired linearized relationship. We immediately get the variance of $\Delta \alpha$ by

$$m_\alpha^2 = C_{\alpha a}^2\, m_a^2 + C_{\alpha b}^2\, m_b^2 + C_{\alpha c}^2\, m_c^2$$

The geometric meaning of $C_{\alpha a}$, $C_{\alpha b}$, $C_{\alpha c}$ is seen from

$$C_{\alpha a} = \frac{2a_o}{F}, \quad C_{\alpha b} = \frac{2p_{ba}}{F}, \quad C_{\alpha c} = \frac{2p_{ca}}{F}$$

with F being the area of the triangle, and p_{ba} and p_{ca} being the projections of b and c onto a as shown in fig. 5.2.

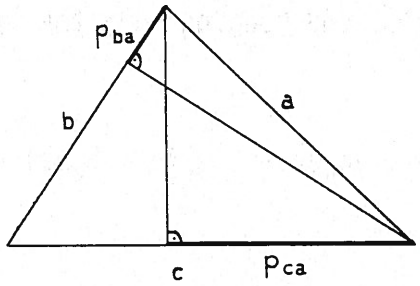

Fig. 5.2.

The covariance matrix of all angles α,β,γ (which is the same as that one of $\Delta\alpha$, $\Delta\beta$, $\Delta\gamma$) is obtained

$$\sum \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} C_{\alpha a} & C_{\alpha b} & C_{\alpha c} \\ C_{\beta a} & C_{\beta b} & C_{\beta c} \\ C_{\gamma a} & C_{\gamma b} & C_{\gamma c} \end{bmatrix} \cdot \begin{bmatrix} m_a^2 & 0 & 0 \\ 0 & m_b^2 & 0 \\ 0 & 0 & m_c^2 \end{bmatrix} \cdot \begin{bmatrix} C_{\alpha a} & C_{\beta a} & C_{\gamma a} \\ C_{\alpha b} & C_{\beta b} & C_{\gamma b} \\ C_{\alpha c} & C_{\beta c} & C_{\gamma c} \end{bmatrix}$$

It is interesting to calculate the variance of

$$\alpha + \beta + \gamma$$

Because this sum equals π, which is a constant, we must have

$$\sigma^2(\alpha+\beta+\gamma) = (1\ 1\ 1)\ \Sigma(\alpha,\beta,\gamma)\ (1\ 1\ 1)^T = 0$$

This is easily verified to be correct. From the geometric meaning of the C's one easily recognizes that e.g.

$$C_{\alpha a} + C_{\beta a} + C_{\gamma a} = 0$$

Remark. If b+c = a, the area F of the triangle becomes zero. The quantities $C_{\alpha a}$, $C_{\alpha b}$, $C_{\alpha c}$ then are infinite. The mean square error m_α^2 becomes infinitely large. This is not entirely meaningful since one could argue that an angle of a triangle is bounded within the interval $[0,\pi]$. The reason for m_α^2 tending to infinity for $b+c \rightarrow a$ is the linearization of the dependency of α on a,b,c. For $b+c \rightarrow a$ the higher order terms are no longer negligible. The degeneracy of a triangle into a line segment is a critical configuration. The angle α becomes very poorly determined. The linearized theory signals this degeneracy. However, it exaggerates somewhat by letting the error of α tend to infinity.

5.2. The first fundamental problem in the plane.

Let a point P with coordinates x_0, y_0 be given. Assume that its coordinates are known and free of any error. Let the distance s and the azimuth α to another point P with unknown coordinates x,y be measured. Let the covariance matrix of s,α be

$$\sum \begin{bmatrix} s \\ \alpha \end{bmatrix} = \begin{bmatrix} m_s^2 & 0 \\ 0 & m_\alpha^2 \end{bmatrix}$$

The coordinates x,y are computed by

$$x = x_0 + s \cos\alpha$$
$$y = y_0 + s \sin\alpha$$

What is the covariance matrix of x,y ? Linearizing the relationship we get

$$\Delta x = \cos\alpha \, \Delta s - s \sin\alpha \, \Delta\alpha$$
$$\Delta y = \sin\alpha \, \Delta s + s \cos\alpha \, \Delta\alpha$$

For simplicity we have refrained from distinguishing between s,α and their approximate values. Applying the error propagation law we find

$$\sum \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\alpha & -s \sin\alpha \\ \sin\alpha & -s \cos\alpha \end{bmatrix} \cdot \begin{bmatrix} m_s^2 & 0 \\ 0 & m_\alpha^2 \end{bmatrix} \cdot \begin{bmatrix} \cos\alpha & \sin\alpha \\ -s \sin\alpha & s \cos\alpha \end{bmatrix}$$

Calling

$$\sum \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} m_{xx} & m_{xy} \\ m_{yx} & m_{yy} \end{bmatrix}$$

we find

$$m_{xx} = m_s^2 \cos^2\alpha + (s m_\alpha)^2 \sin^2\alpha$$
$$m_{xy} = [m_s^2 - (s m_\alpha)^2] \cos\alpha \sin\alpha$$

$$m_{yy} = m_s^2 \sin^2\alpha + (s m_\alpha)^2 \cos^2\alpha$$

Discussion. We see that x and y are correlated unless

$$m_s = s\, m_\alpha$$

i.e. unless the error of the distance s equals the error of the lateral deviation in P due to the azimuth error. If this equality holds, it also implies

$$m_{xx} = m_{yy}$$

The accuracy of P is equally good in all directions. We shall make this statement precise in the next subsection.

5.3. Error ellipses.

In this subsection we prefer to call the coordinates in the plane x_1 and x_2 instead of x and y. Suppose then that

$$X = (X_1\ X_2)^T$$

are the random coordinates of a point in the plane. Let the covariance matrix of X be

$$\Sigma(X) = M = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}$$

M must be positive semidefinite. We assume that M^{-1} exists, i.e. we assume that M is positive definite.

Definition. Let $x = (x_1, x_2)^T$ be points on the curve

$$x^T M^{-1} x = 1$$

This curve is an ellipse. It is called the error ellipse of the point.

The curve is an ellipse because M^{-1} is positive definite. (The inverse of a positive definite matrix is also positive definite. The proof runs as follows: $x^TM^{-1}x = x^TM^{-1}MM^{-1}x = (M^{-1}x)^TM(M^{-1}x) = y^TMy$, with $y = M^{-1}x$. Now $y^TMy > 0$ because M is positive definite.)

Let

$$\xi = (\xi_1\ \xi_2)^T = (\cos\varphi\ \sin\varphi)^T$$

be a unit vector in the plane. ξ defines a direction which is also given by the direction angle φ. We call

$$\xi^TX = \xi_1X_1 + \xi_2X_2 = X_1\cos\varphi + X_2\sin\varphi$$

the error of the point in the direction ξ. Note: ξ^TX is the projection of X

onto the line having direction ξ.

By the error propagation law we get

$$\sigma^2(\xi^T X) = \xi^T M \xi = m_{11}\xi_1^2 + 2m_{12}\xi_1\xi_2 + m_{22}\xi_2^2 =$$
$$m_{11}\cos^2\varphi + 2m_{12}\cos\varphi\,\sin\varphi + m_{22}\sin^2\varphi = S^2(\varphi), \text{ say}$$

An ellipse is a closed convex curve. Any convex curve has a support function $p(\varphi)$ with respect to a chosen point located in its interior. See fig. 5.3 for the definition of the support function.

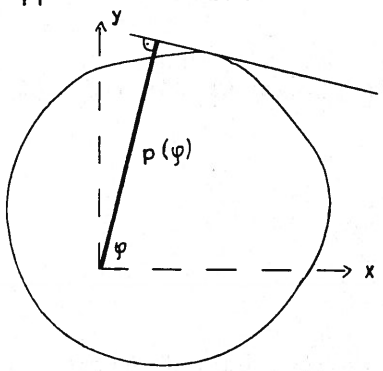

Fig. 5.3.

<u>Theorem.</u> $S(\varphi)$ is the support function of the error ellipse with respect to the center of the ellipse (see fig. 5.4).

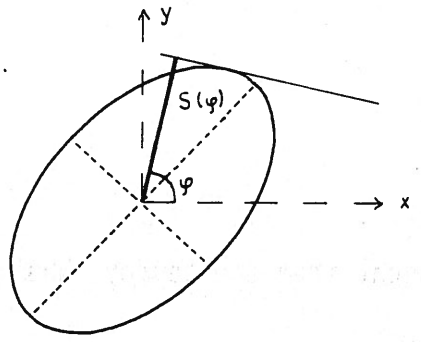

Fig. 5.4.

Proof. Let x_0 be a point on the error ellipse. The tangent line through this point is given by

$$x_0^T M^{-1} x = 1$$

[For a proof note that the gradient of $(x^T M^{-1} x - 1)$ taken at $x = x_0$ is a vector orthogonal to the ellipse. This vector is $2\, M^{-1} x_0$. We may cut its length by 1/2 obtaining $M^{-1} x_0$. Thus the tangent line is $(M^{-1} x_0)^T (x - x_0) = 0$, or $x^T M^{-1} x \;\; x^T M^{-1} x_0 = 0$ or $x_0^T M^{-1} x = 1$, because $x_0^T M^{-1} x_0 = 1$.]

Introducing the normal unit vector of the tangent

$$\xi = \| M^{-1} x_0 \|^{-1} \, M^{-1} x_0$$

we write the tangent line in its "Hesse" form

$$\xi^T x = \| M^{-1} x_0 \|^{-1}$$

we see that

$$p(\varphi) = \| M^{-1} x_0 \|^{-1}$$

is the distance of the tangent from the center. What we want to prove is $p(\varphi) = S(\varphi)$. Now

$$S(\varphi)^2 = \xi^T M \xi = \|M^{-1}x_0\|^{-2}\,(M^{-1}x_0)^T M(M^{-1}x_0) = \|M^{-1}x_0\|^{-2}\,x_0^T M^{-1}x_0 = \|M^{-1}x_0\|^{-2} = p(\varphi)^2$$

because $x_0^T M^{-1} x_0 = 1$. This proves the theorem.

Remark. The theorem and its proof carry over from two to n dimensions with hardly any change. (Of course it takes more than one angle φ to define a direction in n dimensions. A direction in R^n is in the best way defined by a unit vector ξ.)

Remark. The question for the maximal and minimal values of $S(\varphi)$ is equivalent to the question for the direction of the principal axes of the error ellipse. In two dimensions one finds the directions as solutions to the equation

$$\tan 2\varphi = \frac{2\,m_{12}}{m_{11} - m_{22}}$$

If $m_{11} = m_{22}$ then φ becomes indeterminate. The error ellipse becomes a circle.

5.4. Polar survey with redundancy.

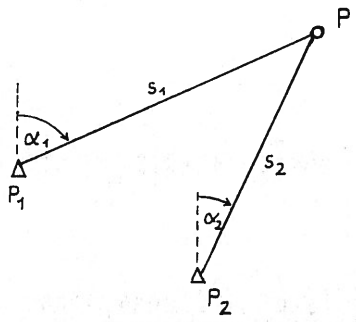

Fig. 5.5.

We return to the problem of section 5.2. However, this time there are two reference points P_1, P_2 whose coordinates are assumed free of any error. The measurements are s_1, α_1, s_2, α_2. Their common covariance matrix be diagonal having the values

$$\Sigma(s_1,\alpha_1,s_2,\alpha_2) = \mathrm{diag}(m_{s_1}^2,\ m_{\alpha_1}^2,\ m_{s_2}^2,\ m_{\alpha_2}^2)$$

5.4.1. The practician's solution.

The practician calculates the coordinates of point P twice, once from P_1 and once from P_2. He then takes the arithmetic mean of the two solutions. Thus he calculates

$$x = \tfrac{1}{2}\{x_1 + s_1\cos\alpha_1 + x_2 + s_2\cos\alpha_2\}$$
$$y = \tfrac{1}{2}\{y_1 + s_1\sin\alpha_1 + y_2 + s_2\sin\alpha_2\}$$

In order to propagate errors we linearize:

$$\Delta x = \tfrac{1}{2}\{\cos\alpha_1\,\Delta s_1 - s_1\sin\alpha_1\,\Delta\alpha_1 + \cos\alpha_2\,\Delta s_2 - s_2\sin\alpha_2\,\Delta\alpha_2\}$$
$$\Delta y = \tfrac{1}{2}\{\sin\alpha_1\,\Delta s_1 + s_1\cos\alpha_1\,\Delta\alpha_1 + \sin\alpha_2\,\Delta s_2 + s_2\cos\alpha_2\,\Delta\alpha_2\}$$

This gives

$$m_{xx} = \tfrac{1}{4}\left\{m_{s_1}^2\cos^2\alpha_1 + (s_1 m_{\alpha_1})^2\sin^2\alpha_1 + m_{s_2}^2\cos^2\alpha_2 + (s_2 m_{\alpha_2})^2\sin^2\alpha_2\right\}$$

$$m_{xy} = \tfrac{1}{4}\left\{[m_{s_1}^2-(s_1 m_{\alpha_1})^2]\cos\alpha_1\sin\alpha_1 + [m_{s_2}^2-(s_2 m_{\alpha_2})^2]\cos\alpha_2\sin\alpha_2\right\}$$

$$m_{yy} = \frac{1}{4}\left\{m_{s_1}^2 \sin^2\alpha_1 + (s_1 m_{\alpha_1})^2 \cos^2\alpha_1 + m_{s_2}^2 \sin^2\alpha_2 + (s_2 m_{\alpha_2})^2 \cos^2\alpha_2\right\}$$

One also recognizes that

$$\Sigma(x,y) = \frac{1}{4}\left[\Sigma_1(x,y) + \Sigma_2(x,y)\right]$$

where $\Sigma_1(x,y)$, $\Sigma_2(x,y)$ are the covariance matrices obtained for P in section 5.2, if the role of P_o is taken either by P_1 or by P_2.

5.4.2. The adjustment expert's solution.

The observation equations (in nonlinear and implicit form) are

$$x = x_1 + (s_1 + v_{s_1}) \cos(\alpha_1 + v_{\alpha_1})$$
$$y = y_1 + (s_1 + v_{s_1}) \sin(\alpha_1 + v_{\alpha_1})$$
$$x = x_2 + (s_2 + v_{s_2}) \cos(\alpha_2 + v_{\alpha_2})$$
$$y = y_2 + (s_2 + v_{s_2}) \sin(\alpha_2 + v_{\alpha_2})$$

We assume approximate values $x^{(o)}$, $y^{(o)}$ for x and y. We introduce coordinate increments Δx, Δy by

$$x = x^{(o)} + \Delta x$$
$$y = y^{(o)} + \Delta y$$

The linearized observation equations are

$$\Delta x = s_1 \cos\alpha_1 - (x^{(o)} - x_1) + \cos\alpha_1 \, v_{s_1} - \sin\alpha_1 s_1 v_{\alpha_1}$$
$$\Delta y = s_1 \sin\alpha_1 - (y^{(o)} - y_1) + \sin\alpha_1 \, v_{s_1} + \cos\alpha_1 s_1 v_{\alpha_1}$$
$$\Delta x = s_2 \cos\alpha_2 - (x^{(o)} - x_2) + \cos\alpha_2 \, v_{s_2} - \sin\alpha_2 s_2 v_{\alpha_2}$$
$$\Delta y = s_2 \sin\alpha_2 - (y^{(o)} - y_2) + \sin\alpha_2 \, v_{s_2} + \cos\alpha_2 s_2 v_{\alpha_2}$$

Although there are adjustment schemes which could handle observation equations in this form, we prefer to solve for v_{s_1}, v_{α_1}, v_{s_2}, v_{α_2}.

Multiplying the first equation by $\cos\alpha_1$, the second one by $\sin\alpha_1$ and adding, we obtain

$$s_1 - (x^{(o)} - x_1) \cos\alpha_1 - (y^{(o)} - y_1) \sin\alpha_1 + v_{s_1} = \cos\alpha_1 \, \Delta x + \sin\alpha_1 \, \Delta y$$

By similar manoeuvres we obtain three more equations

$$\frac{1}{s_1} (x^{(o)} - x_1) \sin\alpha_1 - \frac{1}{s_1} (y^{(o)} - y_1) \cos\alpha_1 + v_{\alpha_1} = \frac{-\sin\alpha_1}{s_1} \Delta x + \frac{\cos\alpha_1}{s_1} \Delta y$$
$$s_2 - (x^{(o)} - x_2) \cos\alpha_2 - (y^{(o)} - y_2) \sin\alpha_2 + v_{s_2} = \cos\alpha_2 \, \Delta x + \sin\alpha_2 \, \Delta y$$
$$\frac{1}{s_2} (x^{(o)} - x_2) \sin\alpha_2 - \frac{1}{s_2} (y^{(o)} - y_2) \cos\alpha_2 + v_{\alpha_2} = \frac{-\sin\alpha_2}{s_2} \Delta x + \frac{\cos\alpha_2}{s_2} \Delta y$$

The weights must be proportional to the reciprocal mean square errors. Thus we use

$$p_{s_1} = m_{s_1}^{-2}, \quad p_{\alpha_1} = m_{\alpha_1}^{-2}$$
$$p_{s_2} = m_{s_2}^{-2}, \quad p_{\alpha_2} = m_{\alpha_2}^{-2}$$

We exhibit the normal equation matrix. Its elements are

$$g_{11} = m_{s_1}^{-2} \cos^2\alpha_1 + (s_1 m_{\alpha_1})^{-2} \sin^2\alpha_1 + m_{s_2}^{-2} \cos^2\alpha_2 + (s_2 m_{\alpha_2})^{-2} \sin^2\alpha_2$$

$$g_{12} = \left\{ m_{s_1}^{-2} - (s_1 m_{\alpha_1})^{-2} \right\} \cos\alpha_1 \sin\alpha_1 + \left\{ m_{s_2}^{-2} - (s_2 m_{\alpha_2})^{-2} \right\} \cos\alpha_2 \sin\alpha_2$$

$$g_{22} = m_{s_1}^{-2} \sin^2\alpha_1 + (s_1 m_{\alpha_1})^{-2} \cos^2\alpha_1 + m_{s_2}^{-2} \sin^2\alpha_2 + (s_2 m_{\alpha_2})^{-2} \cos^2\alpha_2$$

The inverse matrix

$$G^{-1} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}^{-1}$$

is the covariance matrix of the rigorously adjusted coordinates x,y.

5.4.3. Comparing the two solutions.

Call Σ_p the covariance matrix from the practician's solution and Σ_a that one from the adjustment, i.e. $\Sigma_a = G^{-1}$. Because Σ_a is the covariance of the best linear unbiased estimates, any linear function of the rigorously adjusted coordinates must have a variance less than or equal to the variance of the same linear function applied to the practician's coordinates. This must in particular be true for linear functions

$$\cos\varphi\, x + \sin\varphi\, y$$

which led to the concept of error ellipses. From this one may infer that the

error ellipse of the adjusted coordinates must be situated in the interior of the practician's error ellipse. See fig. 5.6.

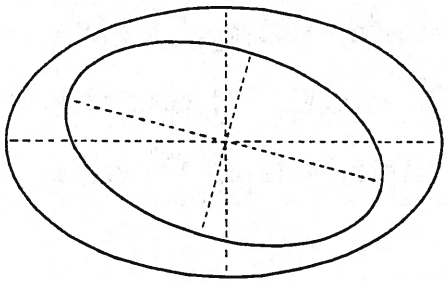

Fig. 5.6.

5.5. An area calculated from polar survey.

Consider the problem illustrated by fig. 5.7.

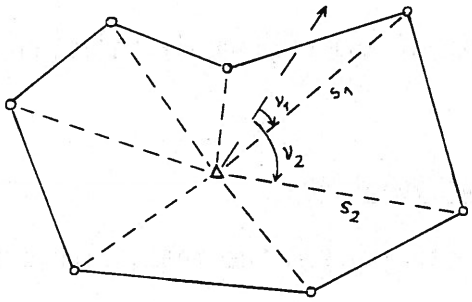

Fig. 5.7.

From a station P_0 whose coordinates in a local system we take as $x_0 = y_0 = 0$, we measure distances s_i and direction angles ν_i to the points on the circumference of a polygon. The direction angles are taken with respect to a local axis as shown in fig. 5.7.

The area is calculated by the following polygon

$$F = \frac{1}{2} \sum_{i=1}^{n} s_i \, s_{i+1} \sin(\nu_{i+1} - \nu_i) \quad \text{(node n+1 equals node 1)}$$

This formula is nonlinear. The linearization is

$$\Delta F = \frac{1}{2} \sum_{i=1}^{n} \{ \Delta s_i \, s_{i+1} \sin(\nu_{i+1} - \nu_i) + \Delta s_{i+1} \, s_i \sin(\nu_{i+1} - \nu_i) - \Delta \nu_i \, s_i \, s_{i+1} \cos(\nu_{i+1} - \nu_i) + \Delta \nu_{i+1} \, s_i \, s_{i+1} \cos(\nu_{i+1} - \nu_i) \}$$

or

$$\Delta F = \frac{1}{2} \sum_{i=1}^{n} \{ \Delta s_i \, [s_{i+1} \sin(\nu_{i+1} - \nu_i) + s_{i-1} \sin(\nu_i - \nu_{i-1})] + s_i \Delta \nu_i \, [-s_{i+1} \cos(\nu_{i+1} \nu_i) + s_{i-1} \cos(\nu_i - \nu_{i-1})] \}$$

$$= \frac{1}{2} \sum_{i=1}^{n} \{ a_i \, \Delta s_i + b_i \, (s_i \Delta \nu_i) \} , \quad \text{say.}$$

Assuming uncorrelated measurement errors with standard deviations m_{s_i}, m_{ν_i}, we find

$$m_F^2 = \frac{1}{4} \sum_{i=1}^{n} [a_i^2 \, m_{s_i}^2 + b_i^2 \, s_i^2 \, m_{\nu_i}^2]$$

The geometric meaning of a_i, b_i is explained in fig. 5.8.

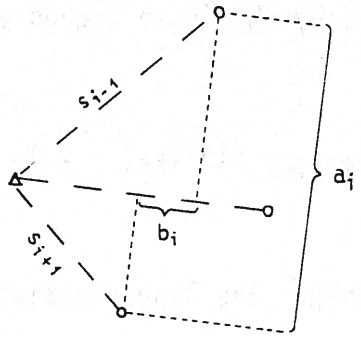

Fig. 5.8.

Thus, while a_i rarely becomes zero, the b_i may vanish under certain symmetry conditions. For example, if the polygon is regular, and if the station P_0 is located at the center, then m_F will depend only on the accuracy of the distances s_i. This again must be seen in the light of the first order Taylor expansion which underlies the error analysis.

5.6. Conventionally adjusted regular traverse.

In fig. 5.9, stations marked by triangles are fixed, those marked by circles are unknown.

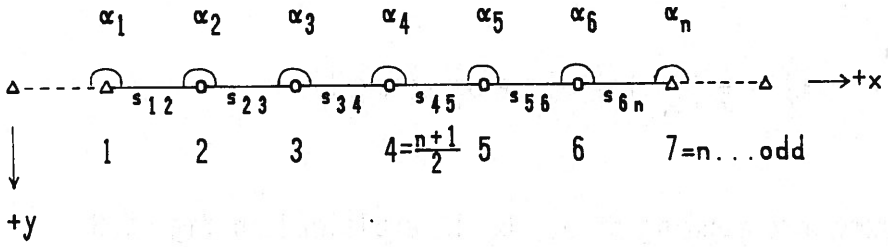

Fig. 5.9.

We assume that the distances $s_{i,i+1}$, i=1,...,n-1 are measured, as well as the angles α_i, i=1,...,n. Let m_s, m_α be the corresponding r.m.s. errors. We are

interested in the accuracy of coordinates ξ,μ of the central station (n+1)/2.

We must express those coordinates ξ,μ in terms of measurements, Because the measurements are redundant, we must know how the traverse is adjusted.

We assume that, as usual, the angles are adjusted first. They are constrained by

$$\alpha_1 + \alpha_2 + \ldots + \alpha_n = n\,\varkappa$$

Due to measurement errors, there will be a discrepancy

$$\alpha_1 + \alpha_2 + \ldots + \alpha_n = n\,\varkappa + w_\alpha$$

It is divided by n, and each angle α_i is replaced by

$$\bar{\alpha}_i = \alpha_i - \frac{w_\alpha}{n}$$

Now the coordinates of the last station are calculated by considering a free traverse leading from station 1 to station n, and having measurements $s_{i,i+1}$ and $\bar{\alpha}_i$.

One gets:

$$x_n = \sum_{i=1}^{n-1} s_{i,i+1} \cos \nu_{i,i+1}, \quad y_n = \sum_{i=1}^{n-1} s_{i,i+1} \sin \nu_{i,i+1}$$

with

$$\nu_{i,i+1} = \sum_{j=1}^{i} \bar{\alpha}_j - i\pi$$

We must linearize this, using approximate values $s_{i,i+1} = 1$, $\nu_{i,i+1} = 0$:

$$\Delta x_n = \sum_{i=1}^{n-1} \Delta s_{i,i+1}, \quad \Delta y_n = \sum_{i=1}^{n-1} \Delta \nu_{i,i+1}$$

For $\Delta \nu_{i,i+1}$, we substitute

$$\Delta \nu_{i,i+1} = \sum_{j=1}^{i} \Delta \alpha_j$$

For $\Delta \bar{\alpha}_j$, we substitute

$$\Delta \bar{\alpha}_j = \Delta \alpha_j - \frac{1}{n} \Delta w_\alpha$$

For Δw_α, we get

$$\Delta w_\alpha = \sum_{k=1}^{n} \Delta \alpha_k$$

Carrying out all these substitutions, we obtain:

$$\Delta y_n = \sum_{i=1}^{n-1} \sum_{j=1}^{i} \left\{ \Delta \alpha_j - \frac{1}{n} \sum_{k=1}^{n} \Delta \alpha_k \right\} =$$

$$= \sum_{j=1}^{n-1} \left\{ \sum_{i=j}^{n-1} 1 \right\} \Delta \alpha_j - \frac{1}{n} \sum_{i=1}^{n-1} \left\{ \sum_{j=1}^{i} 1 \right\} \sum_{k=1}^{n} \Delta \alpha_k$$

$$= \sum_{j=1}^{n-1} (n-j)\,\Delta\alpha_j - \left\{ \frac{1}{n} \sum_{i=1}^{n-1} i \right\} \sum_{k=1}^{n} \Delta\alpha_k$$

$$= \sum_{j=1}^{n} (n-j)\,\Delta\alpha_j - \frac{1}{n}\,\frac{n(n-1)}{2} \sum_{k=1}^{n} \Delta\alpha_k$$

$$= \sum_{j=1}^{n} \left\{ \frac{n+1}{2} - j \right\} \Delta\alpha_j$$

Summarizing:

$$\Delta x_n = \sum_{i=1}^{n-1} \Delta s_{i,i+1}, \quad \Delta y_n = \sum_{i=1}^{n} \left\{ \frac{n+1}{2} - i \right\} \Delta\alpha_i$$

A similar calculation for the midpoint $\frac{n+1}{2}$ gives:

$$\Delta\xi = \Delta x_{(n+1)/2} = \sum_{i=1}^{(n-1)/2} \Delta s_{i,i+1}$$

$$\Delta\eta = \Delta y_{(n+1)/2} = \sum_{i=1}^{(n-1)/2} \sum_{j=1}^{i} \left\{ \Delta\alpha_j - \frac{1}{n} \sum_{k=1}^{n} \Delta\alpha_k \right\}$$

$$= \sum_{j=1}^{(n-1)/2} \left\{ \sum_{i=j}^{(n-1)/2} 1 \right\} \Delta\alpha_j - \frac{1}{n} \left\{ \sum_{i=1}^{(n-1)/2} \sum_{j=1}^{i} 1 \right\} \sum_{k=1}^{n} \Delta\alpha_k$$

$$= \sum_{j=1}^{(n-1)/2} \left\{ \frac{n+1}{2} - j \right\} \Delta\alpha_j - \frac{1}{n}\,\frac{n-1}{2}\,\frac{n+1}{4} \sum_{k=1}^{n} \Delta\alpha_k$$

$$= \sum_{j=1}^{(n-1)/2} \left\{ \frac{3n^2+4n+1}{8n} - j \right\} \Delta\alpha_j - \frac{n^2-1}{8n} \sum_{k=(n+1)/2}^{n} \Delta\alpha_k$$

Summarizing:

$$\Delta\xi = \sum_{i=1}^{(n-1)/2} \Delta s_{i,i+1}$$

$$\Delta\eta = \sum_{i=1}^{(n-1)/2} \left\{ \frac{3n^2+4n+1}{8n} - i \right\} \Delta\alpha_i - \frac{n^2-1}{8n} \sum_{i=(n+1)/2}^{n} \Delta\alpha_i$$

The coordinate discrepancies at station n are now redistributed. This results in adjusted coordinates at the midpoint given by

$$\Delta\tilde{\xi} = \Delta\xi - \frac{1}{2}\Delta x_n = \frac{1}{2} \sum_{i=1}^{(n-1)/2} \Delta s_{i,i+1} - \frac{1}{2} \sum_{i=(n+1)/2}^{n-1} \Delta s_{i,i+1}$$

$$\Delta\tilde{\eta} = \Delta\eta - \frac{1}{2}\Delta y_n = \sum_{i=1}^{(n-1)/2} \left\{ \frac{3n^2+4n+1}{8n} - i - \frac{1}{2}\left(\frac{n+1}{2} - i \right) \right\} \Delta\alpha_i$$

$$+ \sum_{i=(n+1)/2}^{n} \left\{ -\frac{n^2-1}{8n} - \frac{1}{2}\left(\frac{n+1}{2} - i \right) \right\} \Delta\alpha_i$$

$$= \sum_{i=1}^{(n-1)/2} \left\{ \frac{(n+1)^2}{8n} - \frac{i}{2} \right\} \Delta\alpha_i + \sum_{i=(n+1)/2}^{n} \left\{ \frac{(n+1)^2}{8n} - \frac{n-i+1}{2} \right\} \Delta\alpha_i$$

One verifies that the coefficients of $\Delta\alpha_i$ and $\Delta\alpha_{n-i+1}$ are equal. This is desirable from reasons of symmetry.

From the above expressions follow the mean square errors of $\tilde{\xi},\tilde{\eta}$ as

$$m_{\tilde{\xi}}^2 = \frac{n-1}{4} m_s^2$$

$$m_{\tilde{\eta}}^2 = 2 \sum_{i=1}^{(n-1)/2} \left\{ \frac{(n+1)^2}{8n} - \frac{i}{2} \right\}^2 + \left\{ \frac{(n+1)^2}{8n} - \frac{n+1}{4} \right\}^2$$

$$= \frac{1}{2} \sum_{i=1}^{(n-1)/2} \left\{ \frac{(n+1)^2}{4n} - i \right\}^2 + \left\{ \frac{(n+1)(n-1)}{8n} \right\}^2$$

$$= \frac{1}{2} \left\{ \frac{(n+1)^4(n-1)}{32\,n^2} - \frac{(n+1)^2}{2n} \cdot \frac{(n+1)(n-1)}{8} + \frac{n-1}{2} \frac{n+1}{2} \frac{n}{6} \right\} +$$

$$\left\{ \frac{(n+1)(n-1)}{8n} \right\}^2$$

We obtain:

$$m_{\tilde{\eta}}^2 = \frac{(n^2-1)(n^2+3)}{192n}$$

The leading term is implied by

$$m_{\tilde{\eta}}^2 = \frac{n^3}{192}$$

Summarizing:

$$m_{\tilde{\xi}}^2 = \frac{n-1}{4} m_s^2, \qquad m_{\tilde{\xi}} \doteq \frac{\sqrt{n}}{2} m_s$$

$$m_{\tilde{\eta}}^2 = \frac{(n^2-1)(n^2+3)}{192n}, \qquad m_{\tilde{\eta}} \doteq \frac{n}{8} \frac{\sqrt{n}}{\sqrt{3}}$$

These results confirm our intuition: in a stretched traverse, the transversal accuracy is far inferior to the longitudinal accuracy.

5.7. Rigorously adjusted regular traverse.

It will turn out that, in case of a regular traverse, conventional adjustment is equivalent to rigorous adjustment.

Rigorous adjustment is most easily done by conditions. The three linearized conditions are

$$\sum_{i=1}^{n-1} \Delta s_{i,i+1} = 0 \qquad (\text{from } \Delta x_n = 0)$$
$$\sum_{i=1}^{n} \Delta\alpha_i = 0 \qquad (\text{from } \sum_{i=1}^{n} \alpha_i = n\pi)$$
$$\sum_{i=1}^{n-1} \sum_{j=1}^{i} \Delta\alpha_j = 0 \qquad (\text{from } \Delta y_n = 0)$$

The adjustment problem decomposes into one for distances and one for angles. We deal only with the problem for angles which is the more difficult one. The conditions are rewritten as

$$\Delta\alpha_1 + \Delta\alpha_2 + \ldots + \Delta\alpha_n = 0$$
$$(n-1)\Delta\alpha_1 + (n-2)\Delta\alpha_2 + \ldots + 0\cdot\Delta\alpha_n = 0$$

We orthogonalize the conditions by subtracting a proper multiple of the first one from the second, obtaining:

$$\Delta\alpha_1 + \Delta\alpha_2 + \ldots + \Delta\alpha_n = 0$$
$$\frac{n-1}{2}\Delta\alpha_1 + \frac{n-3}{2}\Delta\alpha_2 + \ldots - \frac{n-1}{2}\Delta\alpha_n = 0$$

The normal equations are

$$\begin{bmatrix} n & 0 \\ 0 & \sum_{i=-(n-1)/2}^{+(n-1)/2} i^2 \end{bmatrix} \cdot \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

or

$$\begin{bmatrix} n & 0 \\ 0 & \frac{n(n^2-1)}{12} \end{bmatrix} \cdot \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

The covariance matrix of the adjusted angles is

$$\Sigma(\Delta\tilde{\alpha}) = I - \begin{bmatrix} 1 & \frac{n-1}{2} \\ 1 & \frac{n-3}{2} \\ \vdots & \vdots \\ 1 & -\frac{n-1}{2} \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{n} & 0 \\ 0 & \frac{12}{n(n^2-1)} \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & \dots & 1 \\ \frac{n-1}{2} & \frac{n-3}{2} & \dots & -\frac{n-1}{2} \end{bmatrix}$$

The best estimate for $\Delta\eta$ is

$$\Delta\tilde{\eta} = \frac{n-1}{2}\,\Delta\tilde{\alpha}_1 + \frac{n-3}{2}\,\Delta\tilde{\alpha}_2 + \dots + \Delta\tilde{\alpha}_{(n-1)/2}$$

In view of the condition

$$\frac{n-1}{2}\,\Delta\tilde{\alpha}_1 + \frac{n-3}{2}\,\Delta\tilde{\alpha}_2 + \dots - \frac{n-1}{2}\,\Delta\tilde{\alpha}_n = 0$$

we may rewrite the best estimate $\Delta\tilde{\eta}$ as:

$$\Delta\tilde{\eta} = \frac{n-1}{4}\,\Delta\tilde{\alpha}_1 + \frac{n-3}{4}\,\Delta\tilde{\alpha}_2 + \dots \quad 0\,\Delta\tilde{\alpha}_{(n+1)/2} + \dots + \dots + \frac{n-3}{4}\,\Delta\tilde{\alpha}_{n-1} + \frac{n-1}{4}\,\Delta\tilde{\alpha}_n$$

The mean square error of $\Delta\tilde{\eta}$ is obtained as

$$m_{\tilde{\eta}}^2 = \frac{1}{4}\, 2 \sum_{i=1}^{(n-1)/2} i^2 - \frac{1}{4} \left\{ 2 \sum_{i=1}^{(n-1)/2} i \right\}^2 . \frac{1}{n} =$$

$$= \frac{1}{2}\, \frac{n-1}{2}\, \frac{n+1}{2}\, \frac{n}{6} - \frac{1}{4}\, 4 \left\{ \frac{n-1}{2}\, \frac{n+1}{4} \right\}^2 \frac{1}{n} =$$

$$= \frac{n(n^2-1)}{48} - \frac{(n^2-1)^2}{64n} = \frac{(n^2-1)(n^2+3)}{192n} \doteq \frac{n^3}{192}.$$

This is exactly the same result as it was obtained in the previous section.

5.8. Systematic errors in a regular traverse.

We shall answer the following question. Suppose that the actual error of any of the angles α_i is bounded by a small quantity ε. What is the maximal error in $\tilde{\eta}$ that can arise under this assumption ?

From the previous sections we know the best estimator for $\Delta\eta$ as

$$\Delta\tilde{\eta} = \sum_{i=1}^{(n-1)/2} \left\{ \frac{(n+1)^2}{8n} - \frac{i}{2} \right\} \Delta\alpha_i + \sum_{i=(n+1)/2}^{n} \left\{ \frac{(n+1)^2}{8n} - \frac{n-i+1}{2} \right\} \Delta\alpha_i$$

The maximal error obviously is

$$\Delta\tilde{\eta}_{MAX} = \left\{ \sum_{i=1}^{(n-1)/2} \left| \frac{(n+1)^2}{8n} - \frac{i}{2} \right| + \sum_{i=(n+1)/2}^{n} \left| \frac{(n+1)^2}{8n} - \frac{n-i+1}{2} \right| \right\} \varepsilon$$

For large n, this sum is approximated by

$$\Delta\tilde{\eta}_{MAX} \doteq \left\{ \sum_{i=1}^{(n-1)/2} \left| \frac{n}{8} - \frac{i}{2} \right| + \sum_{i=(n+1)/2}^{n} \left| \frac{n}{8} - \frac{n-i+1}{2} \right| \right\} \varepsilon$$

$$\doteq 4 \sum_{i=1}^{n/4} \left(\frac{n}{8} - \frac{i}{2} \right) \varepsilon = 2 \sum_{i=1}^{n/4} \left(\frac{n}{4} - i \right) \varepsilon \doteq \frac{n^2}{16} \varepsilon$$

Thus

$$\Delta\tilde{\eta}_{MAX} \doteq \frac{n^2}{16} \varepsilon$$

Comparing this with the r.m.s. error derived earlier

$$m_{\tilde{\eta}} \doteq \frac{n \sqrt{n}}{8 \sqrt{3}}$$

we arrive at the following conclusion:

For n tending to infinity, the ratio

$$\Delta\tilde{\eta}_{MAX} / m_{\tilde{\eta}} \doteq \frac{\sqrt{3}}{2} \sqrt{n} \varepsilon$$

tends to infinity. In a large traverse, the maximal effect due to systematic errors eventually outgrows the effect of random errors. This statement, at least in a quantitative way, carries over to other large adjustment schemes, e.g. those of large continental networks.

MITTEILUNGEN DER GEODÄTISCHEN INSTITUTE
DER TECHNISCHEN UNIVERSITÄT GRAZ

(früher: Institutsmitteilungen des I. Geodätischen Institutes
der Technischen Hochschule in Graz)

Bisher erschienene Folgen:

Nr., Autor/Herausgeber, Titel, Seiten, Preis (ö.S.)

1 HUBENY K. (1959): Formeln und Tafeln zur Berechnung der 2. Hauptaufgabe auf dem Bessel'schen Ellipsoid für Strecken bis 300 km im Bereich zwischen den geographischen Breiten 45° und 57°30'. 33 Seiten, öS 45.

2 HUBENY K. (1959): Formeln und Tafeln zur Berechnung der geodätischen Hauptaufgaben über Normalschnitte für beliebige Ellipsoide und beliebige Entfernungen. 29 Seiten, öS 45.

3 MORITZ H. (1959): Untersuchungen über eine direkte Lösung der 2. Hauptaufgabe auf dem Rotationsellipsoid für beliebige Entfernungen. 16 Seiten, öS 35.

4 HUBENY K. (1960): Die Lösung der geodätischen Hauptaufgaben nach Bessel-Jordan. Erweiterte und neue Formeln sowie Tafeln für die Ellipsoide von Bessel und Hayford im Bereich zwischen den geographischen Breiten 45° und 58°. 65 Seiten, öS 50.

5 HUBENY K. (1960): Formeln und Tafeln zur Berechnung der geodätischen Hauptaufgaben über beliebige Entfernungen (Internationales Ellipsoid). 44 Seiten, öS 50.

6 HUBENY K., K. RINNER (1966): Vorlesungen am II. Fortbildungskurs für Praktiker an der Technischen Hochschule in Graz vom 5. bis 7. Oktober 1964. 156 Seiten, öS 50.

7 RINNER K. (1967): Geodätische Programme im Rechenzentrum Graz (Stand 9. Oktober 1967). 191 Seiten, öS 50.

8 RINNER K. (1968): Vorlesungen am III. Fortbildungskurs für Praktiker an der Technischen Hochschule in Graz vom 9. bis 12. Oktober 1967. 282 Seiten, öS 50.

9 RINNER K., G. BRANDSTÄTTER (1971): Forschungsberichte über Erdgezeiten und Satellitengeodäsie. 121 Seiten, öS 50.

10 FELDBACHER F., K. HUBENY, K. RINNER (1971): Beiträge zur ellipsoidischen Geometrie und zu Mikrowellen- und Lasermessungen für große Entfernungen. 78 Seiten, öS 50.

11 RINNER K. (1972): Proceedings of the International Symposium "Satellite and Terrestrial Triangulation"; 2 volumes: 1) Sessions of the West European Sub-Commission of the International Commission for Artificial Satellites, I.A.G.; 2) Sessions of the Special Study Group 1.26 of the I.A.G. 612 Seiten, öS 300.

12 BITTMANN O., G. KRAJICEK, P. MEISSL (1973): Microcomputer Compucorp 320 G und 322 G, die Benützung und Anwendungsbeispiele für die Vermessungstechnik. 53 Seiten, öS 30.

13 RINNER K. (1973): Berichte über Forschungsarbeiten. 57 Seiten. öS 50.

14 FRIEDL J., G. KRAJICEK, P. MEISSL (1974): Taschenrechner Hewlett-Packard HP-45, die Benützung und Anwendungsbeispiele für die Vermessungstechnik. 60 Seiten, öS 50.

15 BARTELME N., P. MEISSL (1974): Strength Analysis of Distance Networks. 57 Seiten, öS 50.

16 CHESI G., K. RINNER (1974): Tabellen zur meteorologischen Reduktion von Entfernungsmessungen mit dem Geodimeter 8. 100 Seiten, öS 100.

17 BITTMANN O., P. MEISSL (1974): Empfohlene Algorithmen zur Programmierung geodätischer Rechenaufgaben. I. Einfache Koordinatenrechnungen in der Ebene, 35 Seiten, öS 50.

18 MEISSL P., K. RINNER (1975): Vorträge am IV. Fortbildungskurs für Praktiker des Vermessungswesens an der Technischen Universität in Graz vom 25. bis 27. November 1974. 290 Seiten, öS 130.

19 LACHAPELLE G. (1975): Determination of the Geoid using Heteregeneous Data. 121 Seiten, öS 150.

20 MEISSL P., H. MORITZ, K. RINNER (1975): Contributions of the Graz Group to the XVI. General Assembly of IUGG/IAG in Grenoble. 308 Seiten, öS 150.

21 BENZ F., K. RINNER (1976): Verfahren zur Verminderung des Einflusses der Bodenreflexion bei der Entfernungsmessung mit Mikrowellen. 97 Seiten, öS 100.

22 RINNER K. (1976): Bericht über Laser- und Mikrowellenmessungen im Testnetz Steiermark. 109 Seiten, öS 100.

23 RINNER K. (1976): Bericht zur Meeresgeodäsie und Satellitengeodäsie. 111 Seiten, öS 100.

24 MEISSL P. (1976): Empfohlene Algorithmen zur Programmierung geodätischer Rechenaufgaben. II. Punktverwaltung mittels Massenspeicher. 69 Seiten, öS 50.

25 MEISSL P., K. STUBENVOLL (1977): Ein Computer-Programmsystem zur Verdichtung trigonometrischer Netze. 129 Seiten, öS 100.

26 KRYNSKI J., H. NOE, K.P. SCHWARZ, H. SONKEL (1977): Numerical Studies and Programs for Interpolation and Collocation. 67 Seiten, öS 50.

27 HUBENY K., A. REITHOFER (1977): Isotherme Koordinatensysteme und konforme Abbildungen des Rotationsellipsoides mit Tafeln und Programmen zur konformen Abbildung für die Ellipsoide von Bessel, Hayford, Krasowsky und für das Referenzellipsoid 1967. 222 Seiten, öS 200.

28 SONKEL H. (1977): Die Darstellung geodätischer Integralformeln durch bikubische Spline-Funktionen. 161 Seiten, öS 150.

29 LEBERL F. (1977): Proceedings of the International Symposium on Image Processing - Interactions with Photogrammetry and Remote Sensing, 3-5 October, 1977, Graz. 250 Seiten, öS 160.

30 ALLMER F. (1977): Dr. Ing. h.c. Eduard Ritter von Orel, dem Erfinder des Stereo-Autographen zum 100. Geburtstag. 41 Seiten, öS 50.

31 KRYNSKY J. (1978): Possibilities of Low-low Satellite Tracking for Local Geoid Improvement. 67 Seiten, öS 70.

32 GERONTOPOULOS P. (1978): Molodensky's Problem in the Plane. 160 Seiten, öS 200.

33 LEBERL F. (1980): Beiträge zur Radargrammetrie und digitalen Bildverarbeitung. 230 Seiten, öS 200.

34 HUBENY K. (1980): Die Klothoide (Formeln, Tafeln, Beispiele). 122 Seiten, öS 120.

35 RINNER K. et al. (1980): Festschrift zur Emeritierung von Prof. Dr. K. Hubeny. 200 Seiten, öS 200.

36 NOE H. (1980): Numerical Investigations on the Problem of Molodensky. 80 Seiten, öS 90.

37 BARTELME N., B. HOFMANN-WELLENHOF, P. MEISSL (1980): Empfohlene Algorithmen zur Programmierung geodätischer Rechenaufgaben. III. Zugriff auf Meßdatendatei. 200 Seiten, öS 180.

38 LICHTENEGGER H., K. RINNER (1982): Verzeichnis der Habilitationen, Dissertationen, Diplomarbeiten sowie von Seminar- und Proseminararbeiten 1960-1981. 81 Seiten, öS 100.

39 CHEN C.-Y. (1982): Geodetic Datum and Doppler Positioning (Dissertation). 255 Seiten, öS S 200.

40 MORITZ H. et al. (1982): Geodaesia Universalis. Festschrift Karl Rinner zum 70. Geburtstag. 382 Seiten, öS 250.

41 MORITZ H., H. SONKEL (1982): Geodesy and Global Geodynamics. 689 Seiten, öS 400.

42 RINNER K., H. LICHTENEGGER (1982): Proceedings of the International Symposium "Education in Geodesy", Sept. 27-29, 1982, Graz, Austria.

43 MEISSL P. (1982): Least Squares Adjustment. A Modern Approach. 440 Seiten, öS 450.

44 ALLMER F. (1983): Biographie Peter Meissl (in Vorbereitung -in preparation).