The Bayesian Approach in Two-Step Modeling of Deformations[1]

K. R. Koch, Bonn and
H. B. Papo, Haifa

Mit der Bayes-Statistik werden Deformationen durch ein Verfahren bestimmt, das aus zwei Schritten besteht. Im ersten Schritt werden die Koordinaten von Punkten eines Netzes in bezug auf ein gewähltes Datum aus Messungen unterschiedlicher Natur und verschiedener Epochen geschätzt. In einem zweiten Schritt werden die Koordinatenänderungen durch weitere unbekannte Parameter ausgedrückt. Drei unterschiedliche Verfahren werden vorgeschlagen.

1 Introduction

We assume that a series of measurements has been conducted in a given network over a period of time. The object of the measurements is monitoring variations in point positions relative to a given fixed reference frame. As such frames are generally not available, a stable subset of points of the network is selected to define through them the datum. Different kinds of measurements are considered at each epoch so that the variance compopments for the observations at each epoch are estimated together with the point positions. This establishes the first step of the analysis.
The time series of estimated coordinates are subjected in a second step to an additional analysis where the variations of the coordinates with respect to the established datum are modeled by means of a new set of parameters. Three models are being considered in this second step: the first one is the general linear model, the second one is the model of the free net adjustment (PAPO 1985; PAPO 1986; KOCH and PAPO 1985) and the third one the well known model of prediction and filtering, the so-called model of collocation. Prior information is introduced for the unknown parameters of the first model. This leads to an analysis which corresponds to the regularization of ill-posed problems. The second model leaves the shape of the network adjusted in the first step unchanged, while the third model allows the introduction of system noise.
Separation between step one and step two is done mostly for convenience. There is a clear advantage if after completing step one the bulk of the measurements can be left behind. This is particularly relevant, as a number of different models needs to be tested in step two. Of course, both steps can be combined in one single step. It is shown that under certain conditions both approaches give identical results.
Bayesian statistics is applied for the analysis. Bayesian statistics has in comparison to traditional statistics the following advantage. By means of the posterior density function for the unknown parameters obtained from Bayes' theorem the unknown parameters are not only estimated, but also confidence regions for the unknown parameters may be established or hypotheses for the parameters be tested. Posterior density functions will therefore be given for the variance components, for their ratios and for the signals of collocation, see KOCH (2000, p. 132 and 174) and KOCH and KUSCHE (2002).
The basic ideas of a two-step analysis of deformations have been presented by PAPO and PERELMUTER (1993). Here we augment their ideas by estimating variance components in the first step and the regularization of the matrix of normal equations of the linear model in the second step. In addition the posterior density functions mentioned above are given.
This paper is organized in six chapters. The next one gives the analysis of the first step, while Chapter 3 presents the linear model for the second step. Chapter 4 deals with the model of the extended free net adjustment and Chapter 5 finally contains the model of prediction and filtering. For each model the analysis in a one-step solution is derived and the conditions for the equivalance of the one-step and two-step solutions are given. The paper finishes with the conclusions.

2 Analysis for the First Step

A network of points has been established to detect and monitor motion and deformation of the earth's crust in a certain region. Measurements have been taken at o different time epochs to detect the movements of the points of the network. The measurements at time epoch t_i are collected in the $n_i \times 1$ vector $\boldsymbol{y}_i$ with $i \in \{1, \ldots, o\}$. The vector $\boldsymbol{y}_i$ is a random vector with covariance matrix $D(\boldsymbol{y}_i|\sigma_i^2) = \sigma_i^2 \boldsymbol{P}_i^{-1}$ given by the known positive definite

1 Parts of the paper have been presented at the Weikko A. Heiskanen Symposium in Columbus, Ohio, Oct. 1–4, 2002

weight matrix $\boldsymbol{P}_i$ and the variance factor σ_i^2. Each observation vector $\boldsymbol{y}_i$ may contain different kinds of observations, like GPS-measurements, distances of leveling so that σ_i^2 is considered an unknown random parameter. The observation vectors $\boldsymbol{y}_i$ and $\boldsymbol{y}_j$ are assumed to be independent for $i \neq j$.

On the basis of prior knowledge or by iterative hypothesis testing, as shown for instance in Koch (1985) and Papo and Perelmuter (1989), the points of the network which moved can be separated from the points which retained their positions with respect to each other. The coordinates of the moving points at epoch t_i are collected in the $u_i \times 1$ vector $\boldsymbol{\beta}_i$ with $i \in \{1, \ldots, o\}$. The vectors $\boldsymbol{\beta}_i$ thus establish a time series of coordinates. The coordinates of the remaining points, which did not move, are identical over all epochs and are assembled in the $u_f \times 1$ vectors $\boldsymbol{\beta}_f$. The $u \times 1$ vector $\boldsymbol{\beta}$ of all unknown coordinates, which in Bayesian analysis is a random vector, is thus given with $u = \boldsymbol{\Sigma}_{i=1}^{0}\, u_i + u_f$ by

$$\boldsymbol{\beta} = |\boldsymbol{\beta}_1', \boldsymbol{\beta}_2', \ldots, \boldsymbol{\beta}_o', \boldsymbol{\beta}_f', |' \tag{2.1}$$

Let the $n_i \times u_i$ matrix $\boldsymbol{X}_i$ with $i \in \{1, \ldots, o\}$ be the coefficient matrix for the observations $\boldsymbol{y}_i$ related to the vector $\boldsymbol{\beta}_i$ and the $n_i \times u_f$ matrix $\boldsymbol{X}_{fi}$ with $i \in \{1, \ldots, o\}$ be the coefficient matrix for the observations $\boldsymbol{y}_i$ connected with the vector $\boldsymbol{\beta}_f$. These matrices are obtained if necessary by a linearization.

Because of the unknown variance factors σ_i^2, the so-called variance components, we obtain with $\boldsymbol{e}_i$ being the errors of $\boldsymbol{y}_i$ the linear model with unknown variance components given in the formulation of Bayesian statistics, see for instance Koch (2000, p. 145).

$$\begin{vmatrix} \boldsymbol{X}_1 & \boldsymbol{0} & \ldots & \boldsymbol{0} & \boldsymbol{X}_{f1} \\ \boldsymbol{0} & \boldsymbol{X}_2 & \ldots & \boldsymbol{0} & \boldsymbol{X}_{f2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \boldsymbol{0} & \boldsymbol{0} & \ldots & \boldsymbol{X}_0 & \boldsymbol{X}_{f0} \end{vmatrix} \begin{vmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \ldots \\ \boldsymbol{\beta}_0 \\ \boldsymbol{\beta}_f \end{vmatrix} = \begin{vmatrix} \boldsymbol{y}_1 + \boldsymbol{e}_1 \\ \boldsymbol{y}_2 + \boldsymbol{e}_2 \\ \ldots \\ \boldsymbol{y}_0 + \boldsymbol{e}_0 \end{vmatrix} \tag{2.2}$$

with

$$D\left(\begin{vmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \ldots \\ \boldsymbol{y}_0 \end{vmatrix} |\boldsymbol{\sigma} \right) = \pi \boldsymbol{P}^{-1} = \sigma_1^2 \boldsymbol{V}_1 + \sigma_2^2 \boldsymbol{V}_2 + \ldots + \sigma_0^2 \boldsymbol{V}_0 \tag{2.3}$$

$$\boldsymbol{\sigma} = |\sigma_1^2, \sigma_2^2, \ldots, \sigma_0^2|$$

$$\boldsymbol{V}_1 = \begin{vmatrix} \boldsymbol{P}_1^{-1} & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \ldots & \ldots & \ldots & \ldots \\ \boldsymbol{0} & \boldsymbol{0} & \ldots & \boldsymbol{0} \end{vmatrix}, \boldsymbol{V}_2 = \begin{vmatrix} \boldsymbol{0} & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P}_2^{-1} & \ldots & \boldsymbol{0} \\ \ldots & \ldots & \ldots & \ldots \\ \boldsymbol{0} & \boldsymbol{0} & \ldots & \boldsymbol{0} \end{vmatrix}, \ldots$$

$$\boldsymbol{V}_0 = \begin{vmatrix} \boldsymbol{0} & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \ldots & \ldots & \ldots & \ldots \\ \boldsymbol{0} & \boldsymbol{0} & \ldots & \boldsymbol{P}_0^{-1} \end{vmatrix}$$

We call $\boldsymbol{X}$ the coefficient matrix on the left-hand side of (2.2) and obtain with (2.1), (2.3) and with

$$\boldsymbol{y} = |\boldsymbol{y}_1', \boldsymbol{y}_2', \ldots, \boldsymbol{y}_0'|', \quad \boldsymbol{e} = \boldsymbol{e}_1', \boldsymbol{e}_2', \ldots, \boldsymbol{e}_0'|'$$

instead of (2.2) the model in a abbreviated form

$$\boldsymbol{X\beta} = \boldsymbol{y} + \boldsymbol{e} \text{ with } D(\boldsymbol{y}|\boldsymbol{\sigma}) = \boldsymbol{P}^{-1} \tag{2.4}$$

The normal equations for the estimate $\hat{\boldsymbol{\beta}}$ of the unknown parameters $\boldsymbol{\beta}$ are given by

$$\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{X}'\boldsymbol{P}\boldsymbol{y} \tag{2.5}$$

We assume that rank $\boldsymbol{X} = q < u$ with u being the number of unknown parameters as defined above. Thus, $u - q$ is the rank deficiency due to the need for defining the datum. As the matrix $\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X}$ of normal equations in (2.5) is singular, we employ a generalized inverse for the solution. Let the rows of the $(u - q) \times u$ matrix $\boldsymbol{E}$

$$\boldsymbol{E} = |\boldsymbol{E}_1, \boldsymbol{E}_2, \ldots, \boldsymbol{E}_0, \boldsymbol{E}_f| \tag{2.6}$$

contain a basis for the null space of the coefficient matrix $\boldsymbol{X}$, that is

$$\boldsymbol{X}\boldsymbol{E}' = \boldsymbol{0} \tag{2.7}$$

In a deformation analysis the points with did not move establish the datum. This is accomplished by adding the constraints.

$$\boldsymbol{E}_f \boldsymbol{\beta}_f = \boldsymbol{0} \tag{2.8}$$

to the normal equations (2.5), which leads to a symmetrical reflexive generalized inverse $(\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X})_{rs}^{-}$ of the matrix of normal equations, see for instance Koch (1999, p. 186) or Koch (2000, p. 122). Using the fixed points only to define the datum avoids changes of the coordinate system from one time epoch to the next. However, in the observation equations (2.2) or (2.4) the fixed points have identical coordinates for all epochs which ensures a common coordinate system for all epochs. Instead of a symmetrical reflexive generalized inverse $(\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X})_{rs}^{-}$ we therefore may apply the pseudoinverse $(\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X})^{+}$ of the matrix of normal equations, which simplifies the analysis of the following chapters. The pseudoinverse is obtained by imposing the constraints

$$\boldsymbol{E\beta} = \boldsymbol{0} \tag{2.9}$$

Identical estimates are obtained with (2.8) and (2.9), if the coordinates estimated for $\boldsymbol{\beta}_i$ by (2.8) are introduced as approximate coordinates for the constraints (2.9), since the pseudoinverse minimizes the sum of squares of the differences between the adjusted and the approximate coordinates, see for instance Koch (1999, p. 191).

By adding the constraints (2.9) to the normal equations (2.5) we obtain in the expanded form

$$\begin{vmatrix} \frac{1}{\sigma_1^2}\boldsymbol{X}_1'\boldsymbol{P}_1\boldsymbol{X}_1 & \boldsymbol{0} & \ldots & \frac{1}{\sigma_1^2}\boldsymbol{X}_1'\boldsymbol{P}_1\boldsymbol{X}_{f1} & \boldsymbol{E}_1' \\ \boldsymbol{0} & \frac{1}{\sigma_2^2}\boldsymbol{X}_2'\boldsymbol{P}_2\boldsymbol{X}_2 & \ldots & \frac{1}{\sigma_2^2}\boldsymbol{X}_2'\boldsymbol{P}_2\boldsymbol{X}_{f2} & \boldsymbol{E}_2' \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \frac{1}{\sigma_1^2}\boldsymbol{X}_{f1}'\boldsymbol{P}_1\boldsymbol{X}_1 & \frac{1}{\sigma_2^2}\boldsymbol{X}_{f2}'\boldsymbol{P}_2\boldsymbol{X}_2 & \ldots & \boldsymbol{\Sigma}_{i=1}^{0}\,\frac{1}{\sigma_i^2}\boldsymbol{X}_{fi}'\boldsymbol{P}_i\boldsymbol{X}_{fi} & \boldsymbol{E}_f' \\ \boldsymbol{E}_1 & \boldsymbol{E}_2 & \ldots & \boldsymbol{E}_f & \boldsymbol{0} \end{vmatrix}$$

$$\begin{vmatrix} \hat{\boldsymbol{\beta}}_1 \\ \hat{\boldsymbol{\beta}}_2 \\ \dots \\ \hat{\boldsymbol{\beta}}_f \\ \mathbf{k} \end{vmatrix} = \begin{vmatrix} \frac{1}{\sigma_1^2}\boldsymbol{X}_1'\boldsymbol{P}_1\boldsymbol{y}_1 \\ \frac{1}{\sigma_2^2}\boldsymbol{X}_2'\boldsymbol{P}_2\boldsymbol{y}_2 \\ \dots \\ \boldsymbol{\Sigma}_{i=1}^{0}\frac{1}{\sigma_i^2}\boldsymbol{X}_{fi}'\boldsymbol{P}_i\boldsymbol{y}_i \\ \boldsymbol{0} \end{vmatrix} \tag{2.10}$$

The estimate $\hat{\boldsymbol{\beta}}$ of the unknown parameters $\boldsymbol{\beta}$ follows in the abbreviated form by

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X}) + \boldsymbol{X}'\boldsymbol{P}\boldsymbol{y} \tag{2.11}$$

with $(\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X})^+$ being the pseudoinverse of $\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X}$. The covariance matrix $D(\boldsymbol{\beta}|\boldsymbol{y}_1,\dots,\boldsymbol{y}_0) = \boldsymbol{\Sigma}$ of the unknown parameters $\boldsymbol{\beta}$ is given by, see for instance Koch (2000, p. 123),

$$D(\boldsymbol{\beta}|\boldsymbol{y}_1,\dots,\boldsymbol{y}_0) = (\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X})^+ = \boldsymbol{\Sigma} \tag{2.12}$$

In the following chapters we will take the pseudoinverse of this covariance matrix and obtain

$$((\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X})^+)^+ = \boldsymbol{\Sigma}^+ = \boldsymbol{X}'\boldsymbol{P}\boldsymbol{X} = \boldsymbol{N} \tag{2.13}$$

with $\boldsymbol{N}$ being the matrix of normal equations.
The unknown variance components σ_i^2 with $i \in \{1,\dots,o\}$ are estimated iteratively. One can either introduce approximate values so that the variance components are close to one and then iterate, or starting from approximate values for the variance components one iterates until the estimates converge. For the latter case the estimate $\hat{\sigma}_i^2$ of σ_i^2 is computed by, see for instance Koch (2000, p. 146).

$$\hat{\sigma}_i^2 = \hat{\boldsymbol{e}}_i'\boldsymbol{P}_i\hat{\boldsymbol{e}}_i/r_i \text{ for } i \in \{1,\dots,o\} \tag{2.14}$$

where $\hat{\boldsymbol{e}}_i$ denotes the vector of residuals obtained from (2.2)

$$\hat{\boldsymbol{e}}_i = \boldsymbol{X}_i\hat{\boldsymbol{\beta}}_i + \boldsymbol{X}_{fi}\hat{\boldsymbol{\beta}}_f - \boldsymbol{y}_i \tag{2.15}$$

and r_i the partial redundancy, that is the contribution of the observation vector $\boldsymbol{y}_i$ to the overall redundancy $n-(u-q)$ of the model (2.2) with $n = \boldsymbol{\Sigma}_{i=1}^{0} n_i$ being the number of all ovservations $\boldsymbol{y}_1,\dots,\boldsymbol{y}_0$ and u the number of unknown parameters $\boldsymbol{\beta}$. The partial redundancies r_i are computed by, see for instance Koch (2000, p. 146).

$$r_i = \operatorname{tr}(\boldsymbol{W}(\sigma_i^2\boldsymbol{V}_i) \text{ for } i \in \{1,\dots,0\} \tag{2.16}$$

with

$$\boldsymbol{W} = \boldsymbol{P} - \boldsymbol{P}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X})^+\boldsymbol{X}'\boldsymbol{P}$$

When combining different kinds of observations like $\boldsymbol{y}_i$ and $\boldsymbol{y}_j$ for $i \neq j$ their weight relation υ with

$$\upsilon = \sigma_i^2/\sigma_j^2 \text{ for } i \neq j \tag{2.17}$$

is of special interest. The posterior density function for υ is given by (Koch and Kusche 2002)

$$p(\upsilon|\boldsymbol{y}_i,\boldsymbol{y}_j) \propto (\frac{1}{\upsilon})^{\frac{r_i}{2}+1}(\frac{1}{2\upsilon}\hat{\boldsymbol{e}}_i'\boldsymbol{P}_i\hat{\boldsymbol{e}}_i + \frac{1}{2}\hat{\boldsymbol{e}}_j'\boldsymbol{P}_{\mathrm{j}}\hat{\boldsymbol{e}}_j)^{-\frac{r_i+r_j}{2}} \tag{2.18}$$

This density function leads to confidence intervals for the weight relation υ or to hypothesis tests for υ.

3 Regularization for the Second Step

In the second step of tha analysis the variations of the coordinates $\boldsymbol{\beta}_i$ of the moving points with respect to the fixed points $\boldsymbol{\beta}_f$ are expressed by f unknown random parameters $\boldsymbol{s}$ in the linear model

$$\boldsymbol{B}\boldsymbol{s} = \boldsymbol{\beta} + \boldsymbol{e}_{\boldsymbol{\beta}} \text{ with } D(\boldsymbol{\beta}) = \sigma_{\boldsymbol{\beta}}^2\boldsymbol{\Sigma} \tag{3.1}$$

where the $u \times f$ matrix $\boldsymbol{B}$ denotes the matrix of coefficients and $\sigma_{\boldsymbol{\beta}}^2$ the unknown variance factor of the covariance matrix $\boldsymbol{\Sigma}$ for $\boldsymbol{\beta}$ from (2.12). We assume rank $\boldsymbol{B} = f$.
An example for $\boldsymbol{B}\boldsymbol{s}$ is a transformation, for instance an affine transformation, with $\boldsymbol{s}$ containing the transformation parameters and $\boldsymbol{B}$ the coefficients of the transformation. An additional example for the unknown parameters $\boldsymbol{s}$ are the coordinates of the points of the network at the initial epoch and the veolocities of the coordinates. The velocities of the stable points, of course, will be small for such an example.
By introducing the estimates $\hat{\boldsymbol{\beta}}$ of the coordinates $\boldsymbol{\beta}$ from (2.11) as observations and by taking the pseudoinverse of the covariance matrix $\boldsymbol{\Sigma}$ according to (2.13) the estimates $\hat{\boldsymbol{s}}$ of the unknown parameters $\boldsymbol{s}$ follow from the normal equations

$$\boldsymbol{B}'\boldsymbol{N}\boldsymbol{B}\hat{\boldsymbol{s}} = \boldsymbol{B}'\boldsymbol{N}\hat{\boldsymbol{\beta}} \tag{3.2}$$

The matrix $\boldsymbol{N}$ is sinular, the matrix $\boldsymbol{B}'\boldsymbol{N}\boldsymbol{B}$ of normal equations is therefore also singular although the matrix $\boldsymbol{B}$ has full column rank. To remove the singularity prior information is assumed for the vector $\boldsymbol{s}$ of unknown parameters by

$$E(\boldsymbol{s}) = \boldsymbol{\mu} \text{ and } D(\boldsymbol{s}|\sigma_\mu^2) = \sigma_\mu^2\boldsymbol{P}_\mu^{-1} \tag{3.3}$$

In the following we will set $\boldsymbol{\mu} = \boldsymbol{0}$ since $\boldsymbol{s}$ contains the corrections of approximates values. The weight matrix $\boldsymbol{P}_\mu$ can be assumed to be diagonal and may be approximated by $\boldsymbol{P}_\mu = \boldsymbol{I}$. The variance factor σ_μ^2 however, is considered as unknown random parameter.
Prior information may be treated as an additional observation equation with error vector $\boldsymbol{e}_\mu$, see for instance Koch (2000, p. 119), so that we obtain with (3.3) instead of (3.1)

$$\begin{vmatrix} \boldsymbol{B} \\ \boldsymbol{I} \end{vmatrix}\boldsymbol{s} = \begin{vmatrix} \boldsymbol{\beta} + \boldsymbol{e}_{\boldsymbol{\beta}} \\ \boldsymbol{\mu} + \boldsymbol{e}_\mu \end{vmatrix} \tag{3.4}$$

with

$$D\left(\begin{vmatrix} \boldsymbol{\beta} \\ \boldsymbol{\mu} \end{vmatrix}|\sigma_{\boldsymbol{\beta}}^2,\sigma_\mu^2\right) = \sigma_{\boldsymbol{\beta}}^2\begin{vmatrix} \boldsymbol{\Sigma} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{vmatrix} + \sigma_\mu^2\begin{vmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P}_\mu^{-1} \end{vmatrix}$$

Since the variance factors $\sigma_{\boldsymbol{\beta}}^2$ and σ_μ^2 are considered as unknown random parameters, the comparison with (2.2) and (2.3) reveals that (3.4) is a linear model with the unknown variance components $\sigma_{\boldsymbol{\beta}}^2$ and σ_μ^2. As in (2.10) and (3.2) the estimate $\hat{\boldsymbol{s}}$ of the unknown parameters $\boldsymbol{s}$ follows from

$$(\frac{1}{\sigma_{\boldsymbol{\beta}}^2}\boldsymbol{B}'\boldsymbol{N}\boldsymbol{B} + \frac{1}{\sigma_\mu^2}\boldsymbol{P}_\mu)\hat{\boldsymbol{s}} = \frac{1}{\sigma_{\boldsymbol{\beta}}^2}\boldsymbol{B}'\boldsymbol{N}\hat{\boldsymbol{\beta}} + \frac{1}{\sigma_\mu^2}\boldsymbol{P}_\mu\boldsymbol{\mu} \tag{3.5}$$

By introducing the regularization parameter λ with

$$\lambda = \sigma_{\boldsymbol{\beta}}^2/\sigma_\mu^2 \tag{3.6}$$

and by setting

$$\boldsymbol{\mu} = \mathbf{0} \tag{3.7}$$

as mentioned above, we obtain what is known as Tikhonov regularization

$$(\boldsymbol{B}'\boldsymbol{N}\boldsymbol{B} + \lambda \boldsymbol{P}_\mu)\hat{\boldsymbol{s}} = \boldsymbol{B}'\boldsymbol{N}\hat{\boldsymbol{\beta}} \tag{3.8}$$

It can be further simplified if we define $\boldsymbol{P}_\mu = \boldsymbol{I}$ as already mentioned. The matrix of normal equations is now regular due to the regularization which is controlled by the regularization parameter λ. This parameter is estimated by the estimates of the variance components according to (2.14) with

$$\hat{\lambda} = \hat{\sigma}^2_{\beta}/\hat{\sigma}^2_{\mu} \tag{3.9}$$

The posterior distribution for λ is obtained by (2.18).
As already mentioned in the introduction, both steps of the analysis may be combined in a single step. To show this, we substiture (3.4) in (2.4) and find the model

$$\begin{vmatrix} \boldsymbol{X}\boldsymbol{B} \\ \boldsymbol{I} \end{vmatrix} \boldsymbol{s} = \begin{vmatrix} \boldsymbol{y} + \boldsymbol{e} \\ \boldsymbol{\mu} + \boldsymbol{e}_\mu \end{vmatrix} \text{ with } D\left(\begin{vmatrix} \boldsymbol{y} \\ \boldsymbol{\mu} \end{vmatrix} | \sigma \right) = \begin{vmatrix} \boldsymbol{P}^{-1} & \mathbf{0} \\ \mathbf{0} & \sigma^2_\mu \boldsymbol{P}^{-1}_\mu \end{vmatrix} \tag{3.10}$$

The normal equations for the estimate $\hat{\boldsymbol{s}}$ of the unknown parameters $\boldsymbol{s}$ follow with

$$(\boldsymbol{B}'\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X}\boldsymbol{B} + \frac{1}{\sigma^2_\mu}\boldsymbol{P}_\mu)\hat{\boldsymbol{s}} = \boldsymbol{B}'\boldsymbol{X}'\boldsymbol{P}\boldsymbol{y} + \frac{1}{\sigma^2_\mu}\boldsymbol{P}_\mu \boldsymbol{\mu} \tag{3.11}$$

Because of $\boldsymbol{N} = \boldsymbol{X}'\boldsymbol{P}\boldsymbol{X}$ from (2.13) the estimate $\hat{\boldsymbol{s}}$ from (3.5) agrees with $\hat{\boldsymbol{s}}$ from (3.11). However, the matrix $\boldsymbol{P}$ in (3.11) contains according to (2.3) the variance components $\sigma^2_1, \sigma^2_2, \ldots, \sigma^2_0$. Their common factor appears in (3.5) as $\sigma^2_{\boldsymbol{\beta}}$. The estimates of the variance components in (3.5) and (3.11) therefore have also to agree.

4 Model of the Extended Free Net Adjustment for the Second Step

We now apply the model of the extended free net adjustment (Papo 1985; Papo 1986). A $u \times 1$ vector $\boldsymbol{w}$ of unknown random parameters is therefore added to the linear model (3.1) to give

$$\boldsymbol{w} + \boldsymbol{B}\boldsymbol{s} = \boldsymbol{\beta} + \boldsymbol{e}_{\boldsymbol{\beta}} \text{ with } D(\boldsymbol{\beta}) = \sigma^2_{\boldsymbol{\beta}}\boldsymbol{\Sigma} \tag{4.1}$$

The dimensions and the meaning of the matrix $\boldsymbol{B}$ and the vector $\boldsymbol{s}$ are the same as in Chapter 3 and again rank $\boldsymbol{B} = f$. The unknown parameter vector $\boldsymbol{w}$ can therefore be interpreted such that it contains the coordinates which are obtained after applying the transformation $\boldsymbol{B}\boldsymbol{s}$ to the coordinate vector $\boldsymbol{\beta}$. By introducing the estimates $\hat{\boldsymbol{\beta}}$ and by taking the pseudoinverse of $\boldsymbol{\Sigma}$ as in Chapter 3 the estimates $\hat{\boldsymbol{w}}$ of $\boldsymbol{w}$ and $\hat{\boldsymbol{s}}$ of $\boldsymbol{s}$ follow from the normal equations

$$\begin{vmatrix} \boldsymbol{N} & \boldsymbol{N}\boldsymbol{B} \\ \boldsymbol{B}'\boldsymbol{N} & \boldsymbol{B}'\boldsymbol{N}\boldsymbol{B} \end{vmatrix} \begin{vmatrix} \hat{\boldsymbol{w}} \\ \hat{\boldsymbol{s}} \end{vmatrix} = \begin{vmatrix} \boldsymbol{N}\hat{\boldsymbol{\beta}} \\ \boldsymbol{B}'\boldsymbol{N}\hat{\boldsymbol{\beta}} \end{vmatrix} \tag{4.2}$$

Since $\boldsymbol{N}$ is singular, the matrix of normal equations in (4.2) is also singular. We will solve it by its pseudoinverse, as shown in the following.

For a one-step solution we substitute (4.1) in (2.4) and obtain

$$\boldsymbol{X}|\boldsymbol{I}, \boldsymbol{B}| \begin{vmatrix} \boldsymbol{w} \\ \boldsymbol{s} \end{vmatrix} = \boldsymbol{y} + \boldsymbol{e} \text{ with } D(\boldsymbol{y}|\boldsymbol{\sigma}) = \boldsymbol{P}^{-1} \tag{4.3}$$

Since $\boldsymbol{B}$ has full column rank, the matrices $\boldsymbol{X}$ and $\boldsymbol{X}|\boldsymbol{I}, \boldsymbol{B}|$ have the same rank deficiency. This is characteristic for the extended free net approach. As a consequence the shape of the adjusted free net does not change when adding the transformation parameters $\boldsymbol{s}$. This is the advantage of the extended free net adjustment (Koch and Papo 1985). The observation equations (4.3) represent a special model of the exstended net adjustment. They lead with (2.11), (2.13) and the identity in Koch (1999, p. 51) to the normal equations (4.2) because of

$$\begin{vmatrix} \boldsymbol{N} & \boldsymbol{N}\boldsymbol{B} \\ \boldsymbol{B}'\boldsymbol{N} & \boldsymbol{B}'\boldsymbol{N}\boldsymbol{B} \end{vmatrix} \begin{vmatrix} \hat{\boldsymbol{w}} \\ \hat{\boldsymbol{s}} \end{vmatrix} = \begin{vmatrix} \boldsymbol{X}'\boldsymbol{P}\boldsymbol{y} \\ \boldsymbol{B}'\boldsymbol{X}'\boldsymbol{P}\boldsymbol{y} \end{vmatrix} = \begin{vmatrix} \boldsymbol{N}\boldsymbol{N}^+\boldsymbol{X}'\boldsymbol{P}\boldsymbol{y} \\ \boldsymbol{B}'\boldsymbol{N}\boldsymbol{N}^+\boldsymbol{X}'\boldsymbol{P}\boldsymbol{y} \end{vmatrix}$$

$$= \begin{vmatrix} \boldsymbol{N}\hat{\boldsymbol{\beta}} \\ \boldsymbol{B}'\boldsymbol{N}\hat{\boldsymbol{\beta}} \end{vmatrix}$$

so that the estimates $\hat{\boldsymbol{s}}$ and $\hat{\boldsymbol{w}}$ of the first step and the second step agree.
A basis of the null space of the coefficient matrix in (4.3) is given with $\boldsymbol{E}$ from (2.6) by (Koch and Papo (1985)

$$\bar{\boldsymbol{E}}' = \begin{vmatrix} \boldsymbol{E}' & -\boldsymbol{B} \\ \mathbf{0} & \boldsymbol{I} \end{vmatrix} \tag{4.4}$$

The pseudoinverse of the matrix of normal equations in (4.2) obtained by (4.4) gives the estimates $\hat{\boldsymbol{s}}$ and $\hat{\boldsymbol{w}}$, see for instance Koch (1999, p. 186),

$$\begin{vmatrix} \hat{\boldsymbol{w}} \\ \hat{\boldsymbol{s}} \end{vmatrix} = \left(\begin{vmatrix} \boldsymbol{N} & \boldsymbol{N}\boldsymbol{B} \\ \boldsymbol{B}'\boldsymbol{N} & \boldsymbol{B}'\boldsymbol{N}\boldsymbol{B} \end{vmatrix} + \begin{vmatrix} \boldsymbol{E}'\boldsymbol{E} + \boldsymbol{B}\boldsymbol{B}' & -\boldsymbol{B} \\ -\boldsymbol{B}' & \boldsymbol{I} \end{vmatrix} \right)^{-1} \begin{vmatrix} \boldsymbol{N}\hat{\boldsymbol{\beta}} \\ \boldsymbol{B}'\boldsymbol{N}\hat{\boldsymbol{\beta}} \end{vmatrix} \tag{4.5}$$

By taking the inverse of the block matrix in (4.5), see for instance Koch (1999, p. 33), the explicit solutions for $\hat{\boldsymbol{s}}$ and $\hat{\boldsymbol{w}}$ are obtained. Symmetrical reflexive generalized inverses may as well be computed by means of (4.4)

5 Model of Prediction and Filtering for the Second Step

We add now system noise $\boldsymbol{w}$ to the model (3.1) to obtain

$$\boldsymbol{B}\boldsymbol{s} + \boldsymbol{w} = \boldsymbol{\beta} + \boldsymbol{e}_{\boldsymbol{\beta}} \text{ with } D(\boldsymbol{\beta}) = \sigma^2\boldsymbol{\Sigma} \tag{5.1}$$

and consider $\boldsymbol{w}$ as a $u \times 1$ random vector of unknown parameters. The unknown variance factor is called σ^2. The model (5.1) is interpreted as a special model of prediction and filtering (Koch 2000, p. 135). It means that the f unknown parameters $\boldsymbol{s}$ are the parameters of the trend. They have the same meaning as in Chapter 3. The unknown random noise $\boldsymbol{w}$ represents the signal. The $u \times f$ coefficient matrix $\boldsymbol{B}$ is again assumed to be of full column rank. No prior information is introduced for the parameters $\boldsymbol{s}$, however, prior information is assumed for $\boldsymbol{w}$ by

$$E(\boldsymbol{w}) = \mathbf{0} \text{ and } D(\boldsymbol{w}|\sigma^2) = \sigma^2\boldsymbol{\Sigma}_w \tag{5.2}$$

with $\boldsymbol{\Sigma}_w$ being a known positive definite matrix.

Under these assumptions the estimates $\hat{s}$ and $\hat{w}$ of s and w are given by (Koch 2000, p. 134–137).

$$\hat{s} = (B'(\Sigma_w + \Sigma)^{-1}B)^{-1}B'(\Sigma_w + \Sigma)^{-1}\hat{\beta} \quad (5.3)$$

$$\hat{w} = \Sigma_w(\Sigma_w + \Sigma)^{-1}(\hat{\beta} - B\hat{s}) \quad (5.4)$$

These estimates are identical with the estimates of the model of prediction and filtering of traditional statistics. If we set $\Sigma_w = 0$ and take with (2.13) the pseudoinverse of Σ, we obtain (3.2) instead of (5.3). Thus, introducing Σ_w causes a special kind of regularization.

The Bayesiann approach gives in addition the posterior density functions for s and w defined by the multivariate t-distributions

$$s|\hat{\beta} \sim t(\hat{s}, b_0(B'(\Sigma_w + \Sigma)^{-1}B)^{-1}/p_0, 2p_0) \quad (5.5)$$

$$w|\hat{\beta} \sim t(\hat{w},\ b_0\Sigma_{ww}/p_0,\ 2p_0) \quad (5.6)$$

with

$$\Sigma_{ww} = \Sigma_w - \Sigma_w(\Sigma_w + \Sigma)^{-1}\Sigma_w + (\Sigma_w - \Sigma_w(\Sigma_w + \Sigma)^{-1}\Sigma_w)$$
$$NB(B'NB - B'N(\Sigma_w - \Sigma_w(\Sigma_w + \Sigma)^{-1}\Sigma_w)NB)^{-}B'N$$
$$(\Sigma_w - \Sigma_w(\Sigma_w + \Sigma)^{-1}\Sigma_w) \quad (5.7)$$

These posterior density functions lead to confidence regions and hypothesis tests for s and w. The covariance matrices of s and w follow with

$$D(s|\hat{\beta}) = \hat{\sigma}^2(B'(\Sigma_w + \Sigma)^{-1}B)^{-1} \quad (5.8)$$

$$D(w|\hat{\beta}) = \hat{\sigma}^2\Sigma_{ww} \quad (5.9)$$

with Σ_{ww} from (5.7) and with $\hat{\sigma}^2$ being the Bayes estimate of the variance factor σ^2 given by

$$\hat{\sigma}^2 = b_0/(p_0 - 1) \quad (5.10)$$

The constants b_0 and p_0 are derived in (Koch 2000, p. 132, 136) based on the prior information $E(\sigma^2) = \sigma_p^2$ and $V(\sigma^2) = V_{\sigma^2}$ for the variance factor σ^2. However, no prior information for σ^2 may also be introduced by

$$\sigma_p^2 \to 0 \text{ and } 1/V_{\sigma^2} \to 0 \quad (5.11)$$

The constants b_0 and p_0 are then determined by

$$b_0 = \Omega/2 \quad (5.12)$$

$$p_0 = (u + 4)/2 \quad (5.13)$$

with u being the number of coordinates in β and Ω the weighted sum of squares of the residuals of model (5.1) together with (5.2)

$$\Omega = (\hat{\beta} - B\hat{s})'(\Sigma_w + \Sigma)^{-1}\Sigma_w(\Sigma_w + \Sigma)^{-1}(\hat{\beta} - B\hat{s}) +$$
$$(\hat{\beta} - B\hat{s} - \hat{w})'N(\hat{\beta} - B\hat{s} - \hat{w}) \quad (5.14)$$

We obtain the Bayes estimate $\hat{\sigma}^2$ from (5.10) with

$$\hat{\sigma}^2 = \Omega/(u + 2) \quad (5.15)$$

To derive the one-step solution for the model of prediction and filtering, (5.1) and (5.2) are substituted in (2.4) so that the model follows

$$XBs + Xw = y + e \text{ with } D(y|\sigma) = P^{-1} \text{ and } D(w) = \Sigma_w \quad (5.16)$$

Again, no prior information is introduced for s and prior information for w. The estimate $\hat{s}$ of s therefore follows with (Koch 2000, p. 134)

$$\hat{s} = (B'X'(X\Sigma_wX' + P^{-1})^{-1}XB)^{-1}B'X'(X\Sigma_wX' + P^{-1})^{-1}y \quad (5.17)$$

To show under which conditions (5.3) agrees with (5.17), we apply the matrix identity, see for instance Koch (1999, p. 34)

$$(X\Sigma_wX' + P^{-1})^{-1} = P - PX(\Sigma_w^{-1} + X'PX)^{-1}X'P \quad (5.18)$$

and obtain

$$X'(X\Sigma_wX' + P^{-1})^{-1}X = X'PX - X'PX(\Sigma_w^{-1} +$$
$$X'PX)^{-1}X'PX \quad (5.19)$$

At least for a regular matrix $X'PX$ of normal equations we obtain by applying (5.18) to the right-hand side of (5.19)

$$X'(X\Sigma_wX' + P^{-1})^{-1}X = (\Sigma_w + \Sigma)^{-1} \quad (5.20)$$

With the result we find instead of (5.17)

$$\hat{s} = (B'(\Sigma_w + \Sigma)^{-1}B)^{-1}B'X'(X\Sigma_wX' + P^{-1})^{-1}y \quad (5.21)$$

and with (2.4) instead of (5.3)

$$\hat{s} = B'(\Sigma_w + \Sigma)^{-1}B)^{-1}B'X'(X\Sigma_wX' + P^{-1})^{-1}(y + \hat{e}) \quad (5.22)$$

Thus, the estimates (5.3) and (5.17) agree, if $\hat{e} = 0$. However, this condition cannot be fulfilled in case of redundant observations.

6 Conclusions

Analysis of deformations in two consecutive steps is to be preferred over single-step procedures because it provides a clear-cut separation in terms of models and estimation. Determining the variance components of each individual batch of measurements is an essential part of the analysis at step one. This is particularly important in deformation analysis as measurements of different epochs may differ significantly in type and quality.

When applying the regularization approach at the second step of the analysis the estimates $\hat{s}$ are computed iteratively from (3.8) because the variance components have to be estimated iteratively from (3.9). If we set $P_\mu = I$ in (3.8), we do not have to specify anything in advance about the unknown parameters s. If we use the model of the extended free net adjustment, the shape of the net adjusted in the first step does not change when adding the additional unknown parameters s. In order to apply at the second step the approach of prediction and filtering, we have to specify the covariance matrix Σ_w if the system noise in (5.2) in order to estimate the unknown parameters s, but without iterations. The approach finally chosen for the second step should depend on the particular application and on the availability of prior information.

Literatur

Koch, K.R. (1985): Ein statistisches Auswerteverfahren für Deformationsmessungen. Allgemeine Vermessungs-Nachrichten, 92: 97–108

Koch, K.R. (1999): Parameter Estimation and Hypothesis Testing in Linear Models, 2nd Ed. Springer, Berlin

Koch, K.R. (2000): Einführung in die Bayes-Statistik. Springer, Berlin.

Koch, K.R.; Kusche, J. (2002): Regularization of geopotential determination from satellite data by variance components. J Geodesy, 76: 259–268

Koch, K.R.; Papo, H. (1985): Erweiterte freie Netzausgleichung. Z. Vermessungswesen, 110: 451–457

Papo, H.B. (1985): Extended free net adjustment constraints. Bulletin Geodesique, 59: 378–390

Papo, H.B. (1986): Extended Free Net Adjustment Constraints. NOAA Technical Report NOS 119 NGS 37, U.S. Department of Commerce, National Geodetic Survey, Rockville, Maryland

Papo, H.B.; Perelmuter, A. (1989): Datum of a dynamical network. Manuscripta Geodaetica, 14: 363–367

Papo, H.B.; Perelmuter, A. (1993): Two-step analysis of dynamical networks. Manuscripta Geodaetica, 18: 422–430

Anschrift des Verfassers:
Prof. Dr.-Ing. K.-R. Koch
Institut für Theoretische Geodäsie
Nussallee 17
53115 Bonn

Abstract

Deformations are detected in a two-step procedure using Bayesian statistics. At the first step the coordinates of points in a network are determined with respect to a chosen datum from measurements of different kinds at different epochs. The variance component of the measurements at each epoch is estimated together with the respective point coordinates. At the second step of the analysis the estimated variations of the point coordinates are modeled by a new set of unknown parameters. Three different approaches are proposed:

a) The regularization approach, where the singular matrix of normal equations is regularized. The posterior distribution of the regularization parameter is given.

b) The extented free net adjustment, where introducing the additional parameters does not change the shape of the network adjusted in the first step.

c) The prediction and filtering approach, where unknown system noise is introduced. Posterior distributions for the unknown parameters and for the system noise are given.

The choise of the approach will depend on the particular application. The above two steps can be combined into a single step so that the estimation of the parameters of the second step follows directly. To obtain identical results with the two-step solution certain conditions have to be observed.

Zusammenfassung

Mit der Bayes-Statistik werden Deformationen durch ein Verfahren bestimmt, das aus zwei Schritten besteht. Im ersten Schritt werden die Koordinaten von Punkten eines Netzes in bezug auf ein gewähltes Datum aus Messungen unterschiedlicher Natur und verschiedener Epochen geschätzt. In einem zweiten Schritt werden die Koordinatenänderungen durch weitere unbekannte Parameter ausgedrückt. Drei unterschiedliche Verfahren werden vorgeschlagen:

a) die Regularisierung, bei der die singuläre Normalgleichungsmatrix regularisiert wird. Die Posteriori-Verteilung für den Regulatisierungsparameter wird angegeben.

b) die erweiterte freie Netzausgleichung, bei der die Einführung der zusätzlichen unbekannten Parameter die Gestalt des im ersten Schritt ausgeglichenen Netzes nicht ändert.

c) die Prädiktion und Filterung, für die unbekanntes Systemrauschen eingeführt wird. Die Posteriori-Verteilungen für die unbekannten Parameter und das Systemrauschen werden angegeben.

Die Wahl des Verfahrens wird von der jeweiligen Anwendung abhängen. Die erwähnten beiden Schritte können zu einem Schritt zusammengefaßt werden, so daß die Parameter des zweiten Schritts direkt geschätzt werden. Um identische Ergebnisse mit der Lösung in zwei Schritten zu erhalten, müssen gewisse Bedingungen erfüllt sein.