Journal of Geodesy (1998) 72: 436–441

Journal of Geodesy

Robust Kalman filter for rank deficient observation models

K. R. Koch, Y. Yang

Institute for Theoretical Geodesy of the University of Bonn, Nussallee 17, D-53115 Bonn, Germany
Tel.: + 49 228 73 2626; fax: + 49 228 73 3029; e-mail: koch@theor.geod.uni-bonn.de

Received: 8 November 1996 / Accepted: 11 February 1998

Abstract. A robust Kalman filter is derived for rank deficient observation models. The datum for the Kalman filter is introduced at the zero epoch by the choice of a generalized inverse. The robust filter is obtained by Bayesian statistics and by applying a robust M-estimate. Outliers are not only looked for in the observations but also in the updated parameters. The ability of the robust Kalman filter to detect outliers is demonstrated by an example.

Key words. Kalman filter · Robust estimation · Rank deficiency

1 Introduction

Parameters in a dynamic system are generally estimated by the Kalman filter, see for instance Arent et al. (1992), Schwarz (1983). Like all parameter estimations which minimize the sum of squares of the residuals the Kalman filter is sensitive to outliers. Thus, a robust Kalman filter should be applied if data contaminated by outliers are to be processed. This problem was solved by Masreliez and Martin (1977), who applied heavy-tailed Gaussian and non-Gaussian distributions to account for outliers. A more efficient robust Kalman filter based on polynomial interpolation was developed by Tsai and Kurz (1983). Wang and Kubik (1993) derived a robust Kalman filter by the variance-inflation model. To the normal distribution containing the variances for the observations a second normal distribution is added which has larger variances than the first, thus accommodating outliers. A different approach of rendering the Kalman filter robust was presented by Teunissen (1990), who uses a recursive testing procedure to eliminate outliers. Finally, Schaffrin (1995) proposed a look-ahead filter which takes some future observations for the update and therefore has smaller mean square error than the common Kalman filter.

In this paper we will render robust the Kalman filter by following Huber (1964) and by applying his robust M-estimate (Huber 1981, p. 43) which is now considered a breakthrough in statistics (Hampel 1992). The big advantage of Huber's M-estimate is its simple derivation and easy implementation by an iterative downweighting of the outlying observations within the method of least squares. In the robust Kalman filter derived here outliers are not only downweighted in the observations but also in the updated parameters, thus checking the dynamic system. This procedure has similarities to the wave algorithm of Salychev and Schaffrin (1992), which controls the linear dynamic system by a non-random impulse vector. In addition, by modifying the M-estimate, methods for detecting outliers in leverage points are discussed by applying the method of severe downweighting given by Koch (1996).

We use Bayesian statistics for the derivation of the robust Kalman filter similar to the approach of Koch (1990, p. 92) and Yang (1991). The resulting posterior density may then be used to compute confidence regions for the unknown parameters and test hypotheses, as was shown for the robust estimation in linear models by Koch and Yang (1998).

If the linear model for the observations connected with the Kalman filter has a rank deficiency, the datum for the Kalman filter needs to be established. Stelzer and Papo (1994) therefore minimized the norm of the unknown parameters in addition to the norm of the residuals and introduced linear constraints to maintain a consistent datum through all epochs of observations. They arrived at filter equations which are similar to the common Kalman filter. In the following the datum is introduced at the zero epoch by the choice of a generalized inverse and then propagated by the dynamic system through all epochs of observations thus obtaining a datum common to all epochs.

Correspondence to: K. R. Koch

A free planar network established by distance measurements for the detection of movements is used to test the robust Kalman filter.

2 Kalman filter with arbitrary datum definitions

Let the linear dynamic system be given by

$$\beta_{k+1} = \mathbf{\Phi}(k+1,k)\beta_k + \boldsymbol{w}_k \tag{2.1}$$

with

$$E(\boldsymbol{w}_k) = \mathbf{0}, \quad D(\boldsymbol{w}_k) = \boldsymbol{Q}_k, \quad k \in \{1,\ldots,N-1\}$$

where β_k denotes the $u \times 1$ vector of unknown random parameters, the so-called state vector, at the epoch $k, \mathbf{\Phi}(k+1,k)$ the $u \times u$ transition matrix with $\mathrm{rk}\mathbf{\Phi}(k+1,k) = u$, $\boldsymbol{w}_k$ the $u \times 1$ vector of random disturbances with the expected value $E(\boldsymbol{w}_k) = \mathbf{0}$ and the positive definite covariance matrix $D(\boldsymbol{w}_k) = \boldsymbol{Q}_k$. Let the $n_k \times 1$ vector $\boldsymbol{y}_k$ of observations at the epoch k establish the linear model not of full rank

$$E(\boldsymbol{y}_k|\beta_k) = \boldsymbol{X}_k\beta_k \tag{2.2}$$

with

$$\mathrm{rk}(\boldsymbol{X}_k) = q < u, \boldsymbol{D}(\boldsymbol{y}_k) = \sigma^2\boldsymbol{P}_k^{-1}$$

where $\boldsymbol{X}_k$ denotes the $n_k \times u$ matrix of coefficients, $\boldsymbol{P}_k$ the $n_k \times n_k$ positive definite weight matrix of $\boldsymbol{y}_k$ and σ^2 the variance factor which is assumed to be known.

Let the random disturbances $\boldsymbol{w}_k$ and $\boldsymbol{w}_{k+1}$, the observations $\boldsymbol{y}_k$ and $\boldsymbol{y}_{k+1}$ for $k \neq k+1$ and $\boldsymbol{w}_k$ and $\boldsymbol{y}_k$ be independent and normally distributed

$$\boldsymbol{w}_k \sim \boldsymbol{N}(\mathbf{0}, \boldsymbol{Q}_k) \tag{2.3}$$

$$\boldsymbol{y}_k|\beta_k \sim \boldsymbol{N}(\boldsymbol{X}_k\beta_k, \sigma^2\boldsymbol{P}_k^{-1}) \tag{2.4}$$

The Kalman filter estimates the unknown parameters recursively and is therefore well suited to a derivation by Bayesian statistics, which will be applied in the following. The Kalman filter starts with the estimate $\hat{\beta}_{1,0}$ of β_1 and its covariance matrix $\mathbf{\Sigma}_{1,0}$, which are given by prior information. As will be shown, the matrix $\mathbf{\Sigma}_{1,0}$ is positive definite. The first index in $\hat{\beta}_{1,0}$ and $\mathbf{\Sigma}_{1,0}$ refers to the epoch of β_1 and the second index to the epoch of the observations $\boldsymbol{y}_0$, by which the prior information is introduced.

The prior distribution of β_k given the observation vectors $\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1}$ is derived by recursively applying Bayes' theorem (Koch 1990, p. 94)

$$\beta_k|\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{k-1} \sim \boldsymbol{N}(\hat{\beta}_{k,k-1}, \mathbf{\Sigma}_{k,k-1}) \tag{2.5}$$

with

$$\mathbf{\Sigma}_{k,k-1} = \sigma^2\boldsymbol{N}_{k,k-1}^{-1} \tag{2.6}$$

where $\hat{\beta}_{k,k-1}$ denotes the estimate of β_k using the data up to epoch $k-1$, $\mathbf{\Sigma}_{k,k-1}$ its positive definite covariance matrix and $\boldsymbol{N}_{k,k-1}$ its positive definite matrix of normal equations.

The posterior density of β_k follows with Bayes' theorem and Eq. (2.4) by

$$\begin{aligned} p(\beta_k|\boldsymbol{y}_1,\ldots,\boldsymbol{y}_k) & \\ \propto \exp\Big\{-\frac{1}{2\sigma^2}[(\beta_k - \hat{\beta}_{k,k-1})'\boldsymbol{N}_{k,k-1}(\beta_k - \hat{\beta}_{k,k-1}) & \\ + (\boldsymbol{y}_k - \boldsymbol{X}_k\beta_k)'\boldsymbol{P}_k(\boldsymbol{y}_k - \boldsymbol{X}_k\beta_k)]\Big\} & \end{aligned} \tag{2.7}$$

By differentiating the exponent with respect to β_k and by setting the derivatives equal to zero we obtain the MAP (maximum a posteriori) estimate $\hat{\beta}_{k,k}$ of β_k which is identical with the Bayes estimate

$$\hat{\beta}_{k,k} = (\boldsymbol{X}_k'\boldsymbol{P}_k\boldsymbol{X}_k + \boldsymbol{N}_{k,k-1})^{-1}(\boldsymbol{X}_k'\boldsymbol{P}_k\boldsymbol{y}_k + \boldsymbol{N}_{k,k-1}\hat{\beta}_{k,k-1}) \tag{2.8}$$

and its covariance matrix $\mathbf{\Sigma}_{k,k}$

$$\mathbf{\Sigma}_{k,k} = \sigma^2(\boldsymbol{X}_k'\boldsymbol{P}_k\boldsymbol{X}_k + \boldsymbol{N}_{k,k-1})^{-1} \tag{2.9}$$

Although $\boldsymbol{X}_k'\boldsymbol{P}_k\boldsymbol{X}_k$ has $\mathrm{rk}(\boldsymbol{X}_k'\boldsymbol{P}_k\boldsymbol{X}_k) = q$ because of Eq. (2.2), $\mathbf{\Sigma}_{k,k}$ is positive definite since $\boldsymbol{N}_{k,k-1}$ is positive definite.

The estimate $\hat{\beta}_{k,k-1}$ and its covariance matrix $\mathbf{\Sigma}_{k,k-1}$ are obtained by updating $\hat{\beta}_{k-1,k-1}$ and $\mathbf{\Sigma}_{k-1,k-1}$ by means of the dynamic system given by Eq. (2.1):

$$\hat{\beta}_{k,k-1} = \mathbf{\Phi}(k,k-1)\hat{\beta}_{k-1,k-1} \tag{2.10}$$

$$\mathbf{\Sigma}_{k,k-1} = \mathbf{\Phi}(k,k-1)\mathbf{\Sigma}_{k-1,k-1}\mathbf{\Phi}'(k,k-1) + \boldsymbol{Q}_{k-1} \tag{2.11}$$

These relations follow by the theorem for the linear transformation of normally distributed random variables.

Equations (2.8)–(2.11) establish the Kalman filter in a form which is computationally inefficient. We therefore apply two matrix identities to Eqs. (2.8) and (2.9), see for instance Koch (1988, p. 39), and obtain with Eq. (2.6)

$$\hat{\beta}_{k,k} = \hat{\beta}_{k,k-1} + \boldsymbol{F}_k(\boldsymbol{y}_k - \boldsymbol{X}_k\hat{\beta}_{k,k-1}) \tag{2.12}$$

$$\boldsymbol{F}_k = \mathbf{\Sigma}_{k,k-1}\boldsymbol{X}_k{}'(\boldsymbol{X}_k\mathbf{\Sigma}_{k,k-1}\boldsymbol{X}_k{}' + \sigma^2\boldsymbol{P}_k^{-1})^{-1} \tag{2.13}$$

$$\mathbf{\Sigma}_{k,k} = (\boldsymbol{I} - \boldsymbol{F}_k\boldsymbol{X}_k)\mathbf{\Sigma}_{k,k-1} \tag{2.14}$$

Equations (2.12)–(2.14) together with Eqs. (2.10) and (2.11) establish the well-known Kalman filter.

Since the observation model of Eq. (2.2) is not of full rank, a datum needs to be established. It will be defined at the zero epoch from which the prior information results. Let us assume that Eqs. (2.1) and (2.2) are also valid for the zero epoch. Unbiasedly estimable parameters are found by a projection and the estimates $\hat{\beta}_{0,0}$ of the parameters β_0 of the zero epoch are given by Koch (1990, p. 79)

$$\hat{\beta}_{0,0} = (\boldsymbol{X}_0'\boldsymbol{P}_0\boldsymbol{X}_0)^-_{rs}\boldsymbol{X}_0'\boldsymbol{P}_0\boldsymbol{y}_0 \tag{2.15}$$

with their covariance matrix

$$\mathbf{\Sigma}_{0,0} = \sigma^2(\boldsymbol{X}_0'\boldsymbol{P}_0\boldsymbol{X}_0)^-_{rs} \tag{2.16}$$

which is positive semidefinite. The choice of the symmetrical reflexive generalized inverse $(\boldsymbol{X}_0'\boldsymbol{P}_0\boldsymbol{X}_0)^-_{rs}$ of the matrix $\boldsymbol{X}_0'\boldsymbol{P}_0\boldsymbol{X}_0$ of normal equations establishes the datum. By applying Eqs. (2.10) and (2.11) the prior information $\hat{\beta}_{1,0}$ and $\mathbf{\Sigma}_{1,0}$ is obtained by which the Kalman filter starts as already mentioned. Since $\boldsymbol{Q}_0$ in Eq. (2.11) is positive definite, $\mathbf{\Sigma}_{1,0}$ is also positive definite, although $\mathbf{\Sigma}_{0,0}$ is positive semidefinite. Thus, the datum is propagated through all epochs of observations by Eqs. (2.10) and (2.11) after it has been defined at the zero epoch.

3 Robust Kalman filter

Robust parameter estimation is well founded only for independent observations. We therefore assume for $\boldsymbol{P}_k$ in Eq. (2.2)

$$\boldsymbol{P}_k = \operatorname{diag}(p_{1k}, \ldots, p_{n_k k}) \tag{3.1}$$

We will not only look for outliers in the observations $\boldsymbol{y}_k$ but also in the parameters $\hat{\beta}_{k,k-1}$, thus checking the dynamic system. We therefore apply the Cholesky factorization to the positive definite matrix $\boldsymbol{N}_{k,k-1}$

$$\boldsymbol{N}_{k,k-1} = \boldsymbol{G}_k\boldsymbol{G}_k' \tag{3.2}$$

where $\boldsymbol{G}_k$ denotes a $u \times u$ regular lower triangular matrix. Thus, the posterior density of β_k follows from Eq. (2.7) by

$$\begin{aligned} &p(\beta_k|\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k) \\ &\quad\propto \exp\{-\frac{1}{2\sigma^2}[(\boldsymbol{G}_k'\beta_k - \boldsymbol{G}_k'\hat{\beta}_{k,k-1})'(\boldsymbol{G}_k'\beta_k - \boldsymbol{G}_k'\hat{\beta}_{k,k-1}) \\ &\quad + (\boldsymbol{y}_k - \boldsymbol{X}_k\beta_k)'\boldsymbol{P}_k(\boldsymbol{y}_k - \boldsymbol{X}_k\beta_k)]\} \end{aligned}$$

or

$$\begin{aligned} &p(\beta_k|\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k) \\ &\propto \exp\left\{-\frac{1}{2\sigma^2}\left[\left(\begin{vmatrix} \boldsymbol{y}_k \\ \boldsymbol{G}_k'\hat{\beta}_{k,k-1} \end{vmatrix} - \begin{vmatrix} \boldsymbol{X}_k \\ \boldsymbol{G}_k' \end{vmatrix}\beta_k\right)' \right.\right. \\ &\left.\left. \begin{vmatrix} \boldsymbol{P}_k & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I} \end{vmatrix}\left(\begin{vmatrix} \boldsymbol{y}_k \\ \boldsymbol{G}_k'\hat{\beta}_{k,k-1} \end{vmatrix} - \begin{vmatrix} \boldsymbol{X}_k \\ \boldsymbol{G}_k' \end{vmatrix}\beta_k\right)\right]\right\} \end{aligned}$$

and finally

$$\begin{aligned} &p(\beta_k|\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k) \\ &\quad\propto \exp\{-\frac{1}{2\sigma^2}[(\tilde{\boldsymbol{y}} - \tilde{\boldsymbol{X}}\beta_k)'\tilde{\boldsymbol{P}}(\tilde{\boldsymbol{y}} - \tilde{\boldsymbol{X}}\beta_k)]\} \end{aligned} \tag{3.3}$$

Instead of the normal distribution a distribution contaminated by outliers is assumed. Thus, the M-estimate instead of the MAP-estimate is applied which means not

$$\sum_{i=1}^{m} \tilde{p}_i\tilde{e}_i^2 \to \text{minimum}$$

with $m = n_k + u$, $\tilde{\boldsymbol{e}} = \tilde{\boldsymbol{X}}\beta_k - \tilde{\boldsymbol{y}} = (\tilde{e}_i)$ and $\tilde{\boldsymbol{P}} = \operatorname{diag}(\tilde{p}_i)$, but

$$\sum_{i=1}^{m} \rho\left(\frac{\sqrt{\tilde{p}_i}\tilde{e}_i}{\sigma}\right) \to \text{minimum}$$

where $\rho(\ldots)$ denotes the score function. With $\beta_k = (\beta_l)$, $\tilde{\boldsymbol{X}} = (x_{il})$, $\psi(\ldots)$ the derivative of $\rho(\ldots)$ and

$$\frac{\partial}{\partial\beta_l}\rho\left(\frac{\sqrt{\tilde{p}_i}\tilde{e}_i}{\sigma}\right) = \psi\left(\frac{\sqrt{\tilde{p}_i}\tilde{e}_i}{\sigma}\right)\frac{\sqrt{\tilde{p}_i}x_{il}}{\sigma}, \; l \in \{1, \ldots, u\}$$

we obtain the M-estimate

$$\frac{1}{\sigma}\sum_{i=1}^{m}\sqrt{\tilde{p}_i}\psi\left(\frac{\sqrt{\tilde{p}_i}\hat{\tilde{e}}_i}{\sigma}\right)x_{il} = 0, \; l \in \{1, \ldots, u\}$$

with $\hat{\tilde{\boldsymbol{e}}} = \tilde{\boldsymbol{X}}\hat{\beta}_{k,k} - \tilde{\boldsymbol{y}} = (\hat{\tilde{e}}_i)$.
By introducing the equivalent weights

$$w_i = \tilde{p}_i\psi\left(\frac{\sqrt{\tilde{p}_i}\hat{\tilde{e}}_i}{\sigma}\right)\bigg/\left(\frac{\sqrt{\tilde{p}_i}\hat{\tilde{e}}_i}{\sigma}\right) \tag{3.4}$$

we find

$$\frac{1}{\sigma^2}\sum_{i=1}^{m} w_i\hat{\tilde{e}}_i x_{il} = 0$$

or with $\boldsymbol{W} = \operatorname{diag}(w_1, \ldots, w_m)$

$$\tilde{\boldsymbol{X}}'\boldsymbol{W}\hat{\tilde{\boldsymbol{e}}} = \mathbf{0}$$

and finally

$$\tilde{\boldsymbol{X}}'\boldsymbol{W}\tilde{\boldsymbol{X}}\hat{\beta}_{k,k} = \tilde{\boldsymbol{X}}'\boldsymbol{W}\tilde{\boldsymbol{y}} \tag{3.5}$$

We apply the "least informative" density given by Huber (1981, p. 71) and obtain

$$\begin{aligned} \psi\left(\frac{\sqrt{\tilde{p}_i}\hat{\tilde{e}}_i}{\sigma}\right) &= \frac{\sqrt{\tilde{p}_i}\hat{\tilde{e}}_i}{\sigma} \quad &\text{for} \quad \frac{\sqrt{\tilde{p}_i}}{\sigma}|\hat{\tilde{e}}_i| \le c \\ \psi\left(\frac{\sqrt{\tilde{p}_i}\hat{\tilde{e}}_i}{\sigma}\right) &= c\frac{\hat{\tilde{e}}_i}{|\hat{\tilde{e}}_i|} \quad &\text{for} \quad \frac{\sqrt{\tilde{p}_i}}{\sigma}|\hat{\tilde{e}}_i| > c \end{aligned}$$

where generally $c = 1.5$. The equivalent weights follow from Eq. (3.4) by

$$\begin{aligned} w_i &= \tilde{p}_i \quad &\text{for} \quad \sqrt{\tilde{p}_i}|\hat{\tilde{e}}_i| \le \sigma c \\ w_i &= \frac{c\sigma\sqrt{\tilde{p}_i}}{|\hat{\tilde{e}}_i|} \quad &\text{for} \quad \sqrt{\tilde{p}_i}|\hat{\tilde{e}}_i| > \sigma c \end{aligned} \tag{3.6}$$

With

$$\hat{\tilde{\boldsymbol{e}}} = \begin{vmatrix} \hat{\boldsymbol{e}}_{ky} \\ \hat{\boldsymbol{e}}_{k\beta} \end{vmatrix} = \begin{vmatrix} \boldsymbol{X}_k \\ \boldsymbol{G}_k' \end{vmatrix}\hat{\beta}_{k,k} - \begin{vmatrix} \boldsymbol{y}_k \\ \boldsymbol{G}_k'\hat{\beta}_{k,k-1} \end{vmatrix}$$

and

$$\boldsymbol{W} = \begin{vmatrix} \boldsymbol{W}_{ky} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{W}_{k\beta} \end{vmatrix}$$

we obtain from Eq. (3.5) the robust estimate $\hat{\beta}_{k,k}$

$$\hat{\beta}_{k,k} = (\boldsymbol{X}_k' \boldsymbol{W}_{ky} \boldsymbol{X}_k + \boldsymbol{G}_k \boldsymbol{W}_{k\beta} \boldsymbol{G}_k')^{-1} (\boldsymbol{X}_k' \boldsymbol{W}_{ky} \boldsymbol{y}_k + \boldsymbol{G}_k \boldsymbol{W}_{k\beta} \boldsymbol{G}_k' \hat{\beta}_{k,k-1}) \tag{3.7}$$

The equivalent weights for the observations $\boldsymbol{y}_k$ in Eq. (3.7) result with

$$\boldsymbol{W}_{ky} = \mathrm{diag}(w_{iy}), \quad \hat{\boldsymbol{e}}_{ik} = (\hat{e}_{ik})$$

from Eqs. (3.1) and (3.6) by

$$\begin{array}{lllll} w_{iy} & = & p_{ik} & \text{for} & \sqrt{p}_{ik}|\hat{e}_{ik}| \le \sigma c \\ w_{iy} & = & \frac{c\sigma\sqrt{p}_{ik}}{|\hat{e}_{ik}|} & \text{for} & \sqrt{p}_{ik}|\hat{e}_{ik}| > \sigma c \end{array} \tag{3.8}$$

The equivalent weights for the updated and transformed parameters $\boldsymbol{G}_k' \hat{\beta}_{k,k-1}$ are obtained with

$$\boldsymbol{W}_{k\beta} = \mathrm{diag}(w_{i\beta}), \quad \hat{\boldsymbol{e}}_{k\beta} = (\hat{e}_{i\beta})$$

from Eq. (3.6) by

$$\begin{array}{lllll} w_{i\beta} & = & 1 & \text{for} & |\hat{e}_{i\beta}| \le \sigma c \\ w_{i\beta} & = & \frac{c\sigma}{|\hat{e}_{i\beta}|} & \text{for} & |\hat{e}_{i\beta}| > \sigma c \end{array} \tag{3.9}$$

The two matrix identities already mentioned are applied to Eq. (3.7) and lead together with Eqs. (2.6), (2.10) and (2.11) to

$$\hat{\beta}_{k,k} = \hat{\beta}_{k,k-1} + \boldsymbol{F}_k(\boldsymbol{y}_k - \boldsymbol{X}_k \hat{\beta}_{k,k-1}) \tag{3.10}$$

$$\boldsymbol{F}_k = \tilde{\boldsymbol{\Sigma}}_{k,k-1} \boldsymbol{X}_k' (\boldsymbol{X}_k \tilde{\boldsymbol{\Sigma}}_{k,k-1} \boldsymbol{X}_k' + \sigma^2 \boldsymbol{W}_{ky}^{-1})^{-1} \tag{3.11}$$

$$\boldsymbol{\Sigma}_{k,k} = (\boldsymbol{I} - \boldsymbol{F}_k \boldsymbol{X}_k) \tilde{\boldsymbol{\Sigma}}_{k,k-1} \tag{3.12}$$

$$\tilde{\boldsymbol{\Sigma}}_{k,k-1} = \sigma^2 (\boldsymbol{G}_k \boldsymbol{W}_{k\beta} \boldsymbol{G}_k')^{-1} \tag{3.13}$$

$$\hat{\beta}_{k,k-1} = \boldsymbol{\Phi}(k, k-1) \hat{\beta}_{k-1,k-1} \tag{3.14}$$

$$\boldsymbol{\Sigma}_{k,k-1} = \boldsymbol{\Phi}(k, k-1) \boldsymbol{\Sigma}_{k-1,k-1} \boldsymbol{\Phi}'(k, k-1) + \boldsymbol{Q}_{k-1} \tag{3.15}$$

$$\boldsymbol{G}_k \boldsymbol{G}_k' = \boldsymbol{N}_{k,k-1} = \sigma^2 \boldsymbol{\Sigma}_{k,k-1}^{-1} \tag{3.16}$$

This is the robust Kalman filter. After updating the estimate $\hat{\beta}_{k-1,k-1}$ and its covariance matrix $\boldsymbol{\Sigma}_{k-1,k-1}$ by Eqs. (3.14) and (3.15), $\boldsymbol{N}_{k,k-1}$ is computed and decomposed by Eq. (3.16). The estimate $\hat{\beta}_{k,k}$ is then iteratively computed by Eqs. (3.10)–(3.13) such that outliers in the observations $\boldsymbol{y}_k$ and in the updated and transformed parameters $\boldsymbol{G}_k' \hat{\beta}_{k,k-1}$ are downweighted according to Eqs. (3.8) and (3.9). If $\hat{\beta}_{k,k}^{(m+1)}$ denotes the estimate of iteration $m+1$, we obtain from Eqs. (3.10)–(3.13)

$$\hat{\beta}_{k,k}^{(m+1)} = \hat{\beta}_{k,k-1} + \boldsymbol{F}_k^{(m)} (\boldsymbol{y}_k - \boldsymbol{X}_k \hat{\beta}_{k,k-1})$$

$$\boldsymbol{F}_k^{(m)} = \tilde{\boldsymbol{\Sigma}}_{k,k-1}^{(m)} \boldsymbol{X}_k' \left(\boldsymbol{X}_k \tilde{\boldsymbol{\Sigma}}_{k,k-1}^{(m)} \boldsymbol{X}_k' + \sigma^2 \left(\boldsymbol{W}_{ky}^{(m)} \right)^{-1} \right)^{-1}$$

$$\boldsymbol{\Sigma}_{k,k}^{(m+1)} = \left(\boldsymbol{I} - \boldsymbol{F}_k^{(m)} \boldsymbol{X}_k \right) \tilde{\boldsymbol{\Sigma}}_{k,k-1}^{(m)}$$

$$\tilde{\boldsymbol{\Sigma}}_{k,k-1}^{(m)} = \sigma^2 \left(\boldsymbol{G}_k' \boldsymbol{W}_{k\beta}^{(m)} \boldsymbol{G}_k \right)^{-1} \tag{3.17}$$

If outliers are looked for only in the observations $\boldsymbol{y}_k$, then $\boldsymbol{W}_{k\beta} = \boldsymbol{I}$ and the iterations for the robust Kalman filter follow with

$$\hat{\beta}_{k,k}^{(m+1)} = \hat{\beta}_{k,k-1} + \boldsymbol{F}_k^{(m)} \left(\boldsymbol{y}_k - \boldsymbol{X}_k \hat{\beta}_{k,k-1} \right)$$

$$\boldsymbol{F}_k^{(m)} = \boldsymbol{\Sigma}_{k,k-1} \boldsymbol{X}_k' \left(\boldsymbol{X}_k \boldsymbol{\Sigma}_{k,k-1} \boldsymbol{X}_k' + \sigma^2 \left(\boldsymbol{W}_{ky}^{(m)} \right)^{-1} \right)^{-1}$$

$$\boldsymbol{\Sigma}_{k,k}^{(m+1)} = \left(\boldsymbol{I} - \boldsymbol{F}_k^{(m)} \boldsymbol{X}_k \right) \boldsymbol{\Sigma}_{k,k-1} \tag{3.18}$$

together with the updates in Eqs. (3.14) and (3.15). The equivalent weights $\boldsymbol{W}_{ky}$ are determined by Eq. (3.8).

If leverage points exist in the data, the downweighting is not sufficient to detect outliers in these points. A more drastic downweighting is attained by the equivalent weights (Koch 1996)

$$\begin{array}{lllll} w_{iy} & = & p_{ik} r_i^{p/2} / r_D & \text{for} & \sqrt{p}_{ik}|\hat{e}_{ik}| \le c\sigma r_i^{p/2} / r_D \\ w_{iy} & = & \frac{c\sigma\sqrt{p}_{ik} r_i^p}{r_D^2 |\hat{e}_{ik}|} & \text{for} & \sqrt{p}_{ik}|\hat{e}_{ik}| > c\sigma r_i^{p/2} / r_D \end{array} \tag{3.19}$$

with r_i being the redundancy number

$$r_i = \left(\boldsymbol{I} - \boldsymbol{X}_k \left(\boldsymbol{X}_k' \boldsymbol{P}_k \boldsymbol{X}_k \right)^{-1} \boldsymbol{X}_k' \boldsymbol{P}_k \right)_{ii}$$

and

$$r_D = \frac{1}{n_k} \sum_{i=1}^{n_k} r_i^{p/2} \tag{3.20}$$

A good choice for $p/2$ is $p/2 = 8$. But a smaller value may be used if some redundancy numbers turn out to be small. An outlier search by the weights in Eq. (3.19) has to be preceded by a search with Eq. (3.8), since the weights in Eq. (3.19) prevent outliers to be detected which are found by Eq. (3.8) (Koch 1996).

4 Test computations

A planar free network established by distance measurements for the detection of movements of ten points is used to test the robust Kalman filter. The rank deficiency $u - q$ of the model given by Eq. (2.2) is $u - q = 3$. Four epochs of observations are given each consisting of 46 measurements of about 3 to 8 km in length. The standard deviations of the independent observations vary between 4 and 9 mm as functions of the length of the distances. The weights in Eq. (2.2) are chosen such that $\sigma^2 = 1$. The observational setup is not changed between the epochs. A zero epoch of distance measurements is also available. It is used to compute the prior information $\hat{\beta}_{0,0}$ on the unknown coordinates and its covariance matrix $\mathbf{\Sigma}_{0,0}$ according to Eqs. (2.15) and (2.16). The datum for the planar network is established by choosing the pseudo-inverse as symmetrical reflexive generalized inverse in Eqs. (2.15) and (2.16).

To make a sensible choice for the covariance matrix $\boldsymbol{Q}_k$ of the random disturbances $\boldsymbol{w}_k$ in Eq. (2.1), one has to keep in mind that in this example the robust Kalman filter should detect deformations. Setting $\boldsymbol{Q}_k = \mathbf{0}$ would correspond to a recursive estimation. With each epoch the covariance matrix $\mathbf{\Sigma}_{k,k-1}$ would yield smaller variances and movements could not be detected anymore. $\mathbf{\Sigma}_{k,k-1}$ should therefore reflect not only the variances and covariances of the coordinates of epoch $k-1$, but also of the previous epoch $k-2$, hence

$$\mathbf{\Sigma}_{k,k-1} = \mathbf{\Phi}(k,k-1)(\mathbf{\Sigma}_{k-1,k-1} + \operatorname{diag}\mathbf{\Sigma}_{k-1,k-2})\mathbf{\Phi}'(k,k-1) \tag{4.1}$$

where diag $\mathbf{\Sigma}_{k-1,k-2}$ means the diagonal matrix with its diagonal elements from $\mathbf{\Sigma}_{k-1,k-2}$, and therefore in comparison with Eqs. (2.11) or (3.15)

$$\boldsymbol{Q}_{k-1} = \mathbf{\Phi}(k,k-1)\operatorname{diag}\mathbf{\Sigma}_{k-1,k-2}\mathbf{\Phi}'(k,k-1) \tag{4.2}$$

Distortions or movements of the whole network are very unlikely, hence we set

$$\mathbf{\Phi}(k+1,k) = \boldsymbol{I} \tag{4.3}$$

in Eqs. (2.10) and (2.11) and in Eqs. (3.14) and (3.15). However, movements of single points are expected. We have $k-1=0$ at the zero epoch so that the matrix $\mathbf{\Sigma}_{0,-1}$ is not available in Eqs. (4.1) and (4.2). We therefore set

$$\mathbf{\Sigma}_{0,-1} = \mathbf{\Sigma}_{0,0} \tag{4.4}$$

where $\mathbf{\Sigma}_{0,0}$, as already mentioned, is the pseudo-inverse which establishes the datum, and obtain by Eq. (4.1) the positive definite covariance matrix $\mathbf{\Sigma}_{1,0}$.

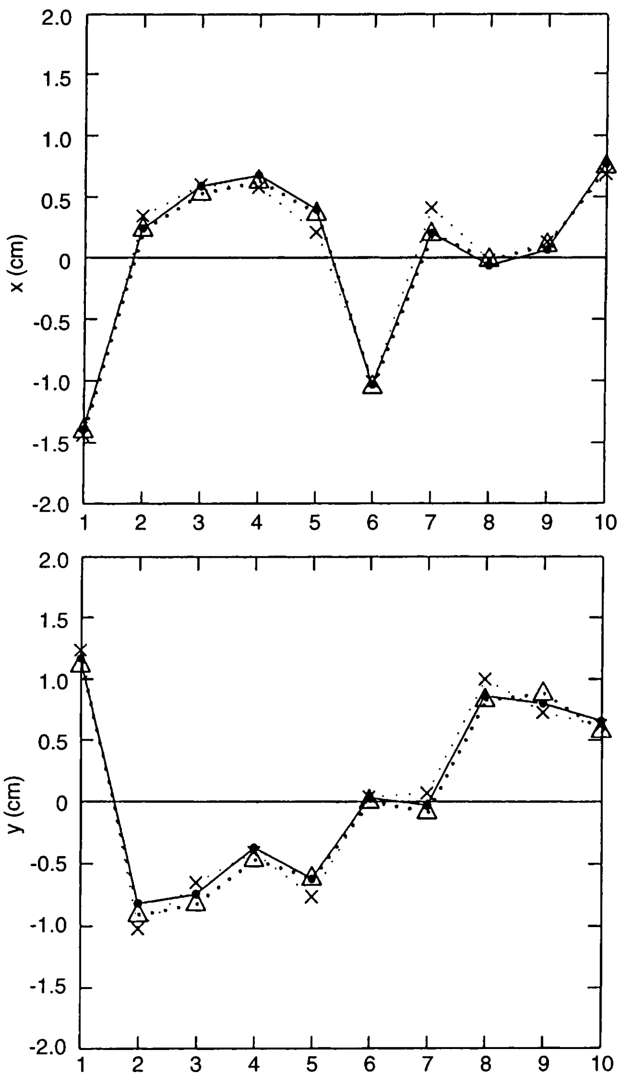

Fig. 1. Sum of movements in x (top) and y (bottom) over four epochs at ten points, ×: Kalman filter, ●: robust Kalman filter for outliers in the observations, △: robust Kalman filter for outliers in the observations and in the updated and transformed parameters

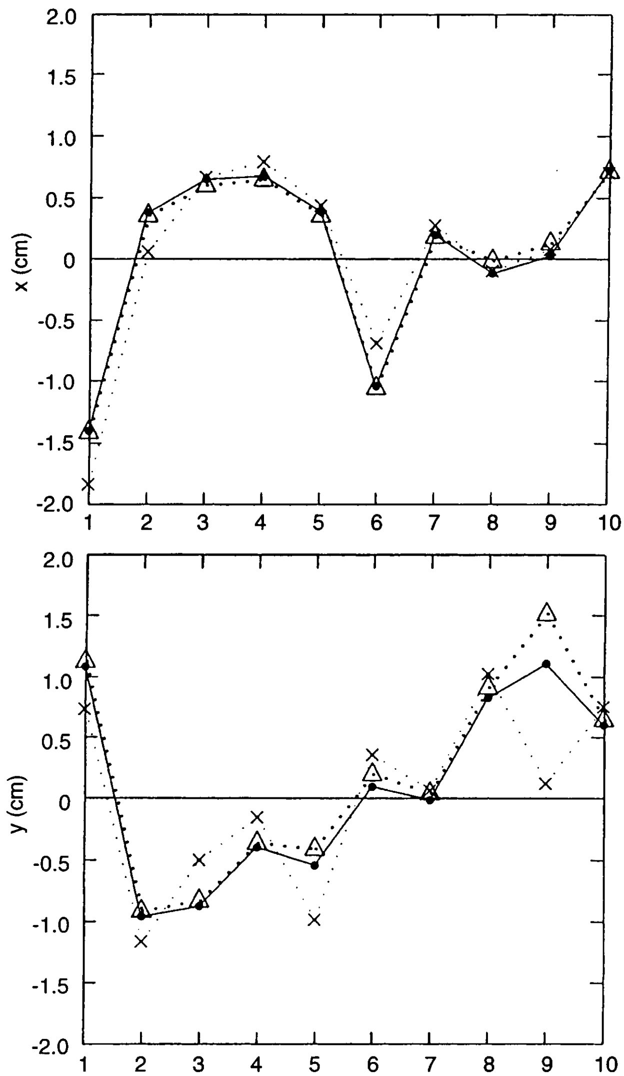

Fig. 2. Sum of movements in x (top) and y (bottom) over four epochs at ten points after introducing 10 additional outliers, ×: Kalman filter, ●: robust Kalman filter for outliers in the observations, △: robust Kalman filter for outliers in the observations and in the updated and transformed parameters

The four epochs of distance measurements have been analyzed by the Kalman filter of Eqs. (2.10)–(2.14), the robust Kalman filter of Eq. (3.18) for outliers in the observations and the robust Kalman filter of Eqs. (3.10)–(3.16) for outliers in the observations and the updated and transformed parameters. The equivalent weights in Eqs. (3.8) and (3.9) were used by putting $c = 1.5$. The movements of the ten points over the four epochs are of interest. The sum of those movements expressed in *x*- and *y*-coordinates is therefore shown in Fig. 1. The sum is obtained by forming the coordinate differences between the last epoch and the zero epoch, i.e., $\hat{\beta}_{4,4} - \hat{\beta}_{0,0}$.

The results of the Kalman filter and the two robust Kalman filters agree very well, the results of the two robust Kalman filters are nearly identical. No serious outliers are therefore present in the data. Nevertheless, a total of 34 observations were downweighted by the robust Kalman filter for outliers in the observations and 36 observations and 6 updated coordinates in the robust Kalman filter for outliers in the observations and in the updated coordinates. Ten outliers of the magnitude ± 3.5 cm were then added to the observations of epoch 3. The results are again expressed by the sum of movements of the ten points over the four epochs and given in Fig. 2.

The results of the Kalman filter now differ from the results of the two robust Kalman filters. Both robust filters identify the ten outliers in epoch 3. The results of these two filters agree very well, although the outliers in the observations affect the coordinates so that outliers are also found in the coordinates updated for epoch 4. The robust Kalman filters eliminate the outliers very well, since the sum of movements given by these filters and shown in Fig. 2 are very similar to the movements shown in Fig. 1.

Acknowledgements. We are indebted to Prof. B. Schaffrin for his valuable comments.

References

Arent N, Hückelheim G, Koch KR (1992) Method for obtaining geoid undulations from satellite altimetry data by a quasi-geostrophic model of the sea surface topography. Manuscr Geod 17: 174–185

Hampel FR (1992) Introduction to Huber (1964) robust estimation of a location parameter. In: Kotz S, Johnson NL (eds) Breakthroughs in statistics, vol II. Springer, Berlin Heidelberg New York, pp 479–491

Huber PJ (1964) Robust estimation of a location parameter. Ann Math Stat 35: 73–101

Huber PJ (1981) Robust statistics. Wiley, New York

Koch KR (1988) Parameter estimation and hypothesis testing in linear models. Springer, Berlin Heidelberg New York

Koch KR (1990) Bayesian inference with geodetic applications. Springer, Berlin Heidelberg New York

Koch KR (1996) Robuste Parameterschätzung. Allg Vermessungs-Nachr 103: 1–18

Koch KR, Yang Y (1998) Konfidenzbereiche und Hypothesentests für robuste Parameterschätzungen. Z Vermessungswesen 123: 20–26

Masreliez CJ, Martin RD (1977) Robust Bayesian estimation for the linear model and robustifying the Kalman filter. IEEE Trans Autom Control 22: 361–371

Salychev O, Schaffrin B (1992) New filter approaches for GPS/INS integration. In: Kumar M, Fell PJ (eds) Proc Sixth Int Geod Symp Satellite Positioning, vol II. Ohio State University, Columbus, pp 670–680

Schaffrin B (1995) On some alternatives to Kalman filtering. In: Sanso F (ed) Geodetic theory today. Springer, Berlin Heidelberg New York, pp 235–245

Schwarz KP (1983) Inertial surveying and geodesy. Rev Geophys Space Phys 21: 878–890

Stelzer D, Papo HB (1994) Kalman filter, smoothing and datum of dynamic reference systems. Manuscr Geod 19: 172–179

Teunissen PJG (1990) Quality control in integrated navigation systems. IEEE Trans Aerosp Electr Syst 5: 35–41

Tsai C, Kurz L (1983) An adaptive robustizing approach to Kalman filtering. Automatica 19: 279–288

Wang YJ, Kubik KK (1993) Robust Kalman filter and its geodetic applications. Manuscr Geod 18: 349–354

Yang Y (1991) Robust Bayesian estimation. Bull Geod 65: 145–150