Recursive Least-Squares Estimation in Case of Interval Observation Data

H. Kutterer[1)], and I. Neumann[2)]
[1)] *Geodetic Institute, Leibniz University Hannover,*
D-30167 Hannover, Germany, kutterer@gih.uni-hannover.de
[2)] *Institute of Geodesy - Geodetic Laboratory, University FAF Munich,*
D-85579 Neubiberg, Germany, ingo.neumann@unibw.de

Abstract: In the engineering sciences, observation uncertainty often consists of two main types: random variability due to uncontrollable external effects, and imprecision due to remaining systematic errors in the data. Interval mathematics is well-suited to treat this second type of uncertainty in, e. g., interval-mathematical extensions of the least-squares estimation procedure if the set-theoretical overestimation is avoided (Schön and Kutterer, 2005). Overestimation means that the true range of parameter values representing both a mean value and imprecision is only quantified by rough, meaningless upper bounds. If recursively formulated estimation algorithms are used for better efficiency, overestimation becomes a key problem. This is the case in state-space estimation which is relevant in real-time applications and which is essentially based on recursions. Hence, overestimation has to be analyzed thoroughly to minimize its impact on the range of the estimated parameters. This paper is based on previous work (Kutterer and Neumann, 2009) which is extended regarding the particular modeling of the interval uncertainty of the observations. Besides a naïve approach, observation imprecision models using physically meaningful influence parameters are considered; see, e. g., Schön and Kutterer (2006). The impact of possible overestimation due to the respective models is rigorously avoided. In addition, the recursion algorithm is reformulated yielding an increased efficiency. In order to illustrate and discuss the theoretical results a damped harmonic oscillation is presented as a typical recursive estimation example in Geodesy.

Keywords: Interval mathematics, imprecision, recursive parameter estimation, overestimation, least-squares, damped harmonic oscillation.

1. Introduction

State-space estimation is an important task in many engineering disciplines. It is typically based on a compact recursive reformulation of the classical least-squares estimation of the parameters which describe the system state. This reformulation reflects the optimal combination of the most recent parameter estimate and of newly available observation data; it is equivalent to a least-squares parameter estimation which uses all available data. However, through the recursive formulation it allows a more efficient update of the estimated values which makes it well-suited for real-time applications. Conventionally, the real-time capability of a process or algorithm, respectively, means that the results are available without any delay when they are required within the process.

In a system-theoretical framework also physical knowledge about the dynamic system state can be available in terms of a system of differential equations. In this case a state-space filter such as the well-

4th International Workshop on Reliable Engineering Computing (REC 2010)
Edited by Michael Beer, Rafi L. Muhanna and Robert L. Mullen
Copyright © 2010 Professional Activities Centre, National University of Singapore.
ISBN: 978-981-08-5118-7. *Published by* Research Publishing Services.
doi:10.3850/978-981-08-5118-7_012

known Kalman filter is used which extends the concept of state-space estimation as it combines predicted system information from the solution of the set of differential equations and additional, newly available observation data (Gelb, 1974). As a special case of state-space filtering, state-space estimation considers the same parameter vector through all recursion steps; nevertheless the estimated values will vary. Moreover, time is not the relevant quantity but the observation index. This allows some convenient features such as the efficient elimination of observation data which are considered as outliers. In any case, the state-space can comprise parameters which are system-immanent and not directly observable.

It is common practice to assess the uncertainty of the observation data in a stochastic framework, only. This means that the observation errors are modeled as random variables and vectors, respectively. This type of uncertainty is called random variability. Classical models in parameter estimation refer to expectation vectors and variance-covariance matrices as first and second moments of the random distribution of the observation. Other approaches based on the Maximum-Likelihood estimation take the complete random distribution into account. In case of non-normal distribution numerical approximation techniques such as Monte-Carlo sampling procedures are applied for the derivation of the densities of the estimated parameters as well as of derived quantities and measures (Koch, 2007).

However, there are more sources of uncertainty in the data than just random errors. Actually, depending on the particular application unknown deterministic effects can introduce a significant level of uncertainty. Such effects are also known as systematic errors which are typically reduced or even eliminated by a mixture of different techniques if an adequate observation configuration was implemented: (i) modification of the observation values using physical or geometrical correction models, (ii) linear combinations of the original observations such as observation differences which can reduce synchronization errors or atmospherically induced run-time differences in distance observations, (iii) dedicated parameterization of the effect in the observation equations. Since none of these techniques is rigorously capable to eliminate an unknown deterministic effect completely or to determine its value, this effect has to be modeled accordingly. Here, interval mathematics is used as theoretical background introducing intervals and interval vectors as additional uncertain quantities. This second type of uncertainty is called imprecision.

The joint assessment of random variability and imprecision of observation data in least-squares estimation has been treated in a number of publications. However, the consideration of recursive state-space estimation has to treat the overestimation problem of interval-mathematical evaluations in a more elaborated way than in classical estimation. Overestimation is caused by, e. g., (hidden) dependencies between interval quantities and it is visible in interval-mathematical properties like, e. g., sub-distributivity. A further problem is caused by the interval inclusion of the range of values of a linear mapping of a vector consisting of interval data which usually generates additional values; see, e. g., Schön and Kutterer (2005) for a discussion of the two- and three-dimensional case. Since recursive formulations particularly exploit such dependencies for the sake of a compact and efficient notation a significant overestimation is expected.

This study is based on previous work on the interval and fuzzy extension of the Kalman filter (Kutterer and Neumann, 2009). Here, two main differences have to be mentioned. First, the approach is simplified as the system-state parameters are considered as static quantities which do not change with time (or forces). Second, the efficiency of the derivation of the measures of the imprecision of the estimated parameters is increased due to a new formulation. The uncertainty of the observation data is formulated in a comprehensive way referring to physically meaningful deterministic influence parameters.

The paper is organized as follows. In Section 2 least-squares parameter estimation is reviewed whereas in Section 3 the recursive formulation is introduced and discussed. In Section 4 the applied model of imprecision is motivated and described. Section 5 provides the interval formulation of the interval-mathematical extension of recursive least-squares state-space estimation. In Section 6 the recursive estimation of state-space parameters based on the observation of a damped harmonic oscillation is discussed as an illustrative example. Section 7 concludes the paper.

2. Least-Squares Parameter Estimation in Linear Models

Recursive least-squares state-space estimation is based on the reformulation of the least-squares estimation using all available observation data; see, e. g., Koch (1999). The model with observation equations is considered in the following. It is a typical linear model which is also known as Gauss-Markov model. It consists of a functional part

$$E\left(\mathbf{l}\right) = \mathbf{A}\mathbf{x} \tag{1}$$

which relates the expectation vector $E(\mathbf{l})$ of the $n \times 1$ -dimensional vector $\mathbf{l}$ of the observations with a linear combination of the unknown $u \times 1$ -dimensional vector of the parameters $\mathbf{x}$ with $n \geq u$. The $n \times u$ -dimensional matrix $\mathbf{A}$ is called configuration matrix or design matrix, respectively. Note that the matrix $\mathbf{A}$ can be either column-regular or column-singular. The difference $r = n - u$ (or $r = n - u + d$ in case of column-singular models with d the rank deficiency) is called redundancy; it quantifies the degree of over-determination of the linear estimation model.

In case of an originally non-linear model a linearization based on a multidimensional Taylor series expansion of the $n \times 1$ vector-valued function $\mathbf{f}$ is derived as

$$E\left(\mathbf{l}\right) = \mathbf{f}\left(\mathbf{x}\right) \approx \mathbf{f}\left(\mathbf{x}_0\right) + \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}_0}\left(\mathbf{x} - \mathbf{x}_0\right)$$

$$\Leftrightarrow \quad E\left(\mathbf{l} - \mathbf{f}\left(\mathbf{x}_0\right)\right) \approx \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}_0}\left(\mathbf{x} - \mathbf{x}_0\right)$$

which yields a fully analogous representation to Eq. (1) if the "$\approx$" sign is neglected:

$$E\left(\Delta\mathbf{l}\right) = \mathbf{A}\Delta\mathbf{x}, \quad \text{with } \Delta\mathbf{l} := \mathbf{l} - \mathbf{f}\left(\mathbf{x}_0\right), \Delta\mathbf{x} := \mathbf{x} - \mathbf{x}_0, \mathbf{A} := \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}_0}. \tag{2}$$

For the sake of a simpler representation only the linear case according to Eq. (1) is discussed in the following. Typically, the functional model part is given through the residual equations

$$\mathbf{v} = \mathbf{A}\mathbf{x} - \mathbf{l} \quad \text{with } \mathbf{v} = E\left(\mathbf{l}\right) - \mathbf{l}. \tag{3}$$

The Gauss-Markov model also comprises a second model part which refers to uncertainty in terms of the regular variance-covariance matrix (vcm) of the observations $\mathbf{\Sigma}_{ll}$ and residuals $\mathbf{\Sigma}_{vv}$, respectively, as

$$V\left(\mathbf{l}\right) = \mathbf{\Sigma}_{ll} = \mathbf{\Sigma}_{vv} = \sigma_0^2\mathbf{Q}_{ll} = \sigma_0^2\mathbf{P}^{-1} \tag{4}$$

with the (theoretical) variance of the unit weight σ_0^2, the cofactor matrix of the observations $\mathbf{Q}_{ll}$ and the weight matrix of the observations $\mathbf{P} = \mathbf{Q}_{ll}^{-1}$.

The unknown vector of parameters is estimated based on the principle of weighted least-squares via the normal equations systems

$$\mathbf{A}^T\mathbf{P}\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^T\mathbf{P}\mathbf{l} \tag{5}$$

as

$$\hat{\mathbf{x}} = \left(\mathbf{A}^T\mathbf{P}\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{P}\mathbf{l} \tag{6}$$

for a column-regular design matrix **A**. In case of a column-singular design matrix a generalized matrix inverse is used leading to

$$\hat{\mathbf{x}} = \left(\mathbf{A}^T\mathbf{P}\mathbf{A}\right)^{-}\mathbf{A}^T\mathbf{P}\mathbf{l}. \tag{7}$$

The cofactor matrix and the vcm of the estimated parameters are derived by the law of variance propagation as

$$\mathbf{Q}_{\hat{x}\hat{x}} = \left(\mathbf{A}^T\mathbf{P}\mathbf{A}\right)^{-1} \text{ and } \Sigma_{\hat{x}\hat{x}} = \sigma_0^2\mathbf{Q}_{\hat{x}\hat{x}}, \tag{8}$$

and

$$\mathbf{Q}_{\hat{x}\hat{x}} = \left(\mathbf{A}^T\mathbf{P}\mathbf{A}\right)^{-} \text{ and } \Sigma_{\hat{x}\hat{x}} = \sigma_0^2\mathbf{Q}_{\hat{x}\hat{x}}, \tag{9}$$

respectively. Note that there are several other quantities of interest such as the estimated vectors of observations $\hat{\mathbf{l}}$ and residuals $\hat{\mathbf{v}}$, the corresponding cofactor matrices and vcms, and the estimated value of the variance of the unit weight

$$\hat{\sigma}_0^2 = \frac{\hat{\mathbf{v}}^T\mathbf{P}\hat{\mathbf{v}}}{r}. \tag{10}$$

Due to the restricted space these quantities are not treated in this paper. The discussion is limited to the recursive estimation of the parameter vector and on the determination of its vcm.

3. Recursive Parameter Estimation in Linear Models

The idea behind recursive parameter estimation is the optimal combination of the most recent estimated parameter vector and of observation data which were not included in the previous estimation due to, e. g., their later availability. This is a typical situation in continuously operating monitoring systems where the state of the considered object is observed repeatedly in defined intervals. The set of parameter vector components can be understood as state-space representation. With each newly incoming set of observations the estimated state of the object is updated as a basis for further analysis and possibly required decisions such as, e. g., in alarm systems. Note that the algorithms presented here just rely on the indices of the observation data which are not necessarily related with time. Hence, by introducing negative weights it is also possible to eliminate observation data from the estimation which is required in case of erroneous data.

This combination is considered as optimal in the meaning of the least-squares principle. Thus, the required equations are derived from the equations given in Section 2. The observation vector is separated into two parts, the first one containing the set of all old observations $\mathbf{l}^{i-1}$ and the second one containing the new observations $\mathbf{l}^{i}$. The residual vector **v**, the design matrix **A** and the weight matrix **P** are divided into corresponding parts according to

$$\begin{bmatrix} \mathbf{v}^{i-1} \\ \mathbf{v}^{i} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^{i-1} \\ \mathbf{A}^{i} \end{bmatrix}\mathbf{x} - \begin{bmatrix} \mathbf{l}^{i-1} \\ \mathbf{l}^{i} \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} \mathbf{P}^{i-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}^{i} \end{bmatrix} \tag{11}$$