Reliable uncertainty evaluation for ODE parameter estimation – a comparison

S. Eichstädt and C. Elster
Physikalisch-Technische Bundesanstalt, Abbestr. 2-12, 10587 Berlin, Germany

Sascha.Eichstaedt@ptb.de

Abstract. The traceability of measurements to SI units requires a traceable calibration of the measurement devices employed. In the calibration for time-dependent measurements the mathematical model typically consists of a system of ODEs with constant parameters. The calibration then requires the estimation of these parameters from measurements of corresponding trajectories, and the assignment of reliable uncertainties to the obtained parameter estimates. Many approaches to parameter estimation in ODEs are available. However, the evaluation of a reliable uncertainty associated with the corresponding parameter estimates is challenging. The reasons are, for instance, the existence of many local minima, numerical instabilities or the non-identifiability of parameters from the available measurement data. Here we discuss some general approaches to ODE parameter estimation and demonstrate practical issues for the example of calibrating force sensors from shock force measurements.

1. Introduction

Metrology, the science of measurement, is concerned with the establishment of measurement units, the realization of measurement standards and the transfer of these standards to industry by means of calibrations. This requires the harmonized and reliable treatment of measurement uncertainties. To this end, the Guide to the Expression of Uncertainty in Measurement (GUM) [1] and its supplementary documents, established by a number of international standard bodies, provides a framework for the evaluation and interpretation of uncertainties. It expresses uncertain knowledge about the value of a quantity in terms of a probability density function (PDF). The estimate of the value of a quantity and its associated uncertainty are then defined as the expectation and the standard deviation of the associated PDF, respectively. The propagation of uncertainties is carried out on the basis of a change-of-variables transformation of probability measures [2]. Provided that all sources of uncertainty are accounted for, and provided that the model that relates all relevant quantities to the measurand (the quantity of interest) is adequate, the result is a reliable characterization of the state of knowledge about the measurand.

The treatment of time-dependent measurements is a topic of growing importance in metrology. The calibration of sensors employed in time-dependent measurements typically requires the estimation of the parameters of an ordinary differential equation (ODE) from observed trajectories. Many approaches for the estimation of ODE parameters can be found in the literature, see, e.g., [3-8]. However, a reliable evaluation of the uncertainties associated with the obtained result is still challenging. In this paper we review some of the methods applied in ODE parameter estimation and discuss practical issues in the context of the calibration of a force transducer.
2. Problem specification

Let us assume that the time evolution of the state of the measurement system is modelled by

\[ x(t) = h(x(t), t, \theta) \quad x(t_0) = x_0, \]  

(1)

with \( x: \mathbb{R} \rightarrow \mathbb{R}^m \) and the ODE parameter vector \( \theta \in \mathbb{R}^k \), and let us consider the discrete-time observations

\[ y[n] = g(x(t), \theta)[n] + \eta[n], \]  

(2)

at time instants \( t_n \) corrupted by the measurement noise process \( \eta \). The parameter vector \( \theta \) may be separated into sub-vectors \( \theta = (\theta^{(1)}, \theta^{(2)})^T \), where \( \theta^{(1)} \) contains the parameters of interest, i.e., the measurand, whereas \( \theta^{(2)} \) is a vector of additional parameters. It is assumed that knowledge about the parameter vector \( \theta^{(2)} \) is available in terms of an associated PDF \( p_{\theta^{(2)}} \). Dynamic calibration in accordance with GUM requires the determination of a mapping \( f \) so that

\[ \theta^{(1)} = f(y, x_0, \theta^{(2)}). \]  

(3)

The goal is then to determine the joint PDF for \( y, x_0 \) and \( \theta^{(2)} \) and – in using (3) – to calculate the PDF \( p_{\theta^{(1)}} \) associated with the measurand \( \theta^{(1)} \) that models the state of knowledge about the measurand.

3. Estimation methods

The estimation of the parameter vector \( \theta^{(1)} \) from equations (1) and (2) in a least-squares sense is carried out by minimizing

\[ J(\theta^{(1)}) = \| y - G(x_0, \theta^{(1)}, \theta^{(2)}) \|^2_W, \]  

(4)

with the weighting matrix \( W \), observations \( y = (y[0], ..., y[N])^T \) and \( G \) a function or procedure that calculates the estimated observations for given values of \( \theta \) and the initial conditions \( x_0 \). The function \( G \) typically relies on a numerical ODE solver or some (smoothed) interpolation of \( x \) and its derivative(s) to solve equation (1) for given \( x_0 \) and \( \theta \). It is well known that the solution of the initial value problem (1) can be very unstable for unknown or only partly known \( x(t) \) and \( \theta \). To this end, for instance, the use of multiple shooting methods is advocated in the literature [2].

The maximum-likelihood estimate of the value of the measurand \( \theta^{(1)} \) is calculated as

\[ \hat{\theta}^{(1)} = \text{arg max}_{\theta^{(1)}} \{ l(\theta^{(1)}; y) \}, \]  

(5)

where \( l(\cdot) \) denotes the likelihood function associated with (2) for fixed \( \theta^{(2)} \) and \( x_0 \). Note that in order to calculate the likelihood, a distribution of the measurement errors in (2) needs to be specified.

As a measure for the reliability of the parameter estimates obtained from (4) or (5), typically a certain matrix \( C_{\theta^{(1)}} \) which is related to the inverse Fisher information matrix is employed. Under certain regularity conditions and assumptions about the measurement noise process \( \eta \), the matrix \( C_{\theta^{(1)}} \) can be employed to calculate confidence intervals for the individual parameters \( \theta_k^{(1)} \). Prior knowledge about the parameters \( \theta^{(2)} \) and \( x_0 \) can be taken into account in (5) by carrying out constraint optimization. This is related to assigning corresponding uniform distributions to these parameters to model the state of knowledge about their value.
In a sequential estimation approach the $y[n]$ are considered as univariate quantities and the estimation of $\theta$ is carried out sequentially in time, assuming that the system model satisfies the Markov property. The classical example of a sequential parameter estimation method is the Kalman filter augmented by the artificial state equations $\frac{d}{dt}q^{(1)}_k = 0$. In the Kalman filter approach also an error-covariance matrix is estimated as a means of an uncertainty associated with the estimation result. However, the incorporation of uncertain prior knowledge about the parameter vector $\theta^{(2)}$ is challenging. Moreover, the derivation of higher moments and thereby the sought characterization of the PDF $p_{\theta^{(1)}}$ from the Kalman filter result requires normality of the noise processes and certain regularity assumptions for the model functions which may limit the applicability of this approach for metrological applications.

A Bayesian inference for ODE parameter estimation allows taking into account uncertain knowledge about the influencing parameters $\theta^{(2)}$ and initial conditions $x_0$ as well as prior knowledge about the measurand. The characterization of uncertain knowledge by means of a state-of-knowledge PDF in metrology is in accordance with a Bayesian point of view. Hence, the application of Bayesian inference to obtain a characterization of the measurand from the observations is a natural choice. However, the equivalence of the propagation of uncertainty in metrology to the result of a Bayesian inference is not necessarily satisfied. Moreover, the posterior PDF is typically obtained by carrying out a Markov chain Monte Carlo (MCMC) method which can be very challenging. A requirement in a Bayesian inference is the definition of prior distributions. In the case of no real prior knowledge about the parameters, either so-called non-informative priors have to be derived analytically or a careful analysis of the impact of the prior choice has to be carried out [9].

4. Practical considerations

As an example application we consider a dynamic force calibration using impact force measurements [10]. In the calibration experiment a shock force is generated by a collinear collision of two mass blocks. The force transducer under test is mounted on a mass body which is at rest initially. The second mass block is accelerated and collides with the transducer and its attached mass. The movement of the two mass blocks is observed by means of laser-Doppler interferometers. The impact of the colliding masses is modelled as a Hertzian contact $h(\cdot, k_H)$ with an unknown damping factor $k_H$. The adaptation of the sensor to the second mass $m_2$ is modelled as a damper-spring system with unknown damping $d_m$ and stiffness $k_m$. The time evolution of the system state is modelled by

$$
\begin{align*}
  m_1 \ddot{x}_1 &= -h(x_1 - x_h, k_H) \\
  m_h \ddot{x}_h &= h(x_1 - x_h, k_H) - k_s \cdot (x_h - x_h) - d_s \cdot (\dot{x}_h - \dot{x}_h) \\
  m_b \ddot{x}_b &= k_s \cdot (x_h - x_h) + d_s \cdot (\dot{x}_h - \dot{x}_h) - k_m \cdot (x_b - x_2) - d_m \cdot (\dot{x}_b - \dot{x}_b) \\
  m_2 \ddot{x}_2 &= k_m \cdot (x_b - x_2) + d_m \cdot (\dot{x}_b - \dot{x}_b)
\end{align*}
$$

(7)

where $x_1, x_2$ denote the position of the two mass blocks and $x_h, x_b$ the position of the sensor's head mass and base mass, respectively. The observations are

$$
y = \begin{pmatrix} x_1 \\ x_2 \\ g(\ddot{x}_b, \ddot{x}_h, k_s, d_s, \rho) \end{pmatrix} + \eta,
$$

(8)

where $g(\cdot)$ denotes the output signal of the sensor under test. The measurand is the vector $\theta^{(1)} = (k_s, d_s, \rho)$ of parameters of the dynamic model of the sensor. The system model (7) requires knowledge of the parameters $\theta^{(2)} = (m_1, m_2, m_h, m_b, k_H, k_m, d_m)$, which model stiffness and damping of the adaption of the sensor to the mass at rest and the damping of the Hertzian contact. These parameters cannot be inferred from other experiments, but have a significant influence on the set of admissible parameter estimates $\theta^{(1)}$. As a consequence, parameter estimates may be ambiguous which poses a significant challenge for the subsequent identification. In principle, any acceleration of
the mass blocks is translated to a force sensed by the transducer under test. This would allow a transformation of model equations (7) and (8) to a linear second-order ODE. However, in practice the mass blocks show significant modal oscillations induced by their geometry which for the impact mass are not translated to a force after the impact.

The above-mentioned difficulties cannot be solved by experimental investigations alone. Likewise, these challenges cannot be overcome by improving on the identification methods only. Future research will thus need to combine experiments with well-defined changes in the measurement system and adapted mathematical and statistical identification methods. That means, the estimation will be based on a whole set of measurements instead of using just a single measurement, utilizing the fact that the sensor parameters should remain invariant.

5. Summary
The estimation of ODE parameters is a usual yet challenging task in many applications, and many mathematical and statistical methods can be found in the literature. In metrology, the goal is not primarily an estimation method with high computational efficiency, but rather one with high reliability. This requires the reliable evaluation of uncertainties associated with the estimation result. The evaluation of uncertainties in accordance with the guidelines in metrology requires taking into account uncertainties associated with all influencing parameters. We discussed some general estimation approaches and their ability regarding the evaluation of uncertainties. In addition we presented challenges for the example of a force sensor calibration. We conclude that, for this example, a reliable estimation of the sensor model parameters requires an identification approach that combines calibration experiments with varying experimental conditions and a corresponding mathematical and statistical approach to identify the sensor parameters from such sets of calibration data.

References
[1] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP and OIML "Guide to the Expression of Uncertainty in Measurement" International Organization for Standardization, Geneva, Switzerland 1995
[2] S. Eichstädt, "Analysis of Dynamic Measurements – Evaluation of dynamic measurement uncertainty" PhD Thesis, PTB-Report IT-16 2012
[3] M. Pfeifer and J. Timmer "Parameter estimation in ordinary differential equations for biochemical processes using the method of multiple shooting", IET Syst. Biol., 2007, Nr. 1, 78 – 88
[4] A. Lj. Juloski, W. Heemels and G. Ferrari-Trecate, "Data-based Hybrid Modelling of the Component Placement Process in the Pick-and-Place Machines", Control Engin. Practice, 2004, Nr. 12, 1241 – 1252
[5] J. Williams and Z. Kalogiratou, "Least squares and Chebyshev fitting for parameter estimation in ODEs", Adv. in Comp. Mathematics, 1993, Nr. 1, 357 – 366
[6] A. Poyton, M. Varziri, K. McAuley, P. McLellan and J. Ramsay, "Parameter estimation in continuous-time dynamic models using principal differential analysis", Computers & Chemic. Engin. 2006, Nr. 30, 698 – 708
[7] J. Ramsay, G. Hooker, D. Campbell and J. Cao, "Parameter estimation for differential equations: A generalized smoothing approach", Journal of the Royal Statistical Soc., 2007, Nr. 69, 741 – 796
[8] O. Ghamesi, M. Lindsey, T. Yang, N. Nguyen, Y. Huang and Y.-F. Jin, "Bayesian parameter estimation for nonlinear modelling of biological pathways", BIOCOMP 2010, Las Vegas, NV, USA
[9] R. E. Kass and L. Wasserman "The selection of prior distributions by formal rules" J. Am. Stat. Assoc. 1996, Nr. 91 1343-70
[10] M. Kobusch, T. Bruns, "The new impact force machine at PTB", XVII IMEKO World Congress 2003, Dubrovnik, Croata