Programs and Algorithms of Numerical Matematics 17
J. Chleboun, P. Přikryl, K. Segeth, J. Šístek, T. Vejchodský (Eds.)
Institute of Mathematics AS CR, Prague 2015

IDENTIFICATION OF PARAMETERS IN INITIAL VALUE PROBLEMS FOR ORDINARY DIFFERENTIAL EQUATIONS

Jan Chleboun, Karel Mikeš

Faculty of Civil Engineering, Czech Technical University in Prague
Thákurova 7, 166 29 Prague 6, Czech Republic
chleboun@mat.fsv.cvut.cz, karel.mikes.1@fsv.cvut.cz

Abstract

Scalar parameter values as well as initial condition values are to be identified in initial value problems for ordinary differential equations (ODE). To achieve this goal, computer algebra tools are combined with numerical tools in the MATLAB® environment. The best fit is obtained through the minimization of the summed squares of the difference between measured data and ODE solution. The minimization is based on a gradient algorithm where the gradient of the summed squares is calculated either numerically or via auxiliary initial value problems. In the latter case, the MATLAB® Symbolic Math Toolbox™ is used to derive the expressions that define the auxiliary problems and to transform them into MATLAB® routines.

1. Introduction

This work was initiated by [3], where parameter identification is performed by an artificial neural network algorithm. A question arose, whether a more traditional method could be effective in solving the identification problem. By a more traditional method, we mean the minimization of a relevant cost function by a gradient-based minimization algorithm.

Parameter identification is a common task in chemistry, biology, and engineering. If the underlying problem is not ill-posed, parameters can be identified by a straightforward method, see, for instance, [5], a short report providing the reader with an easy introduction to the subject, or a more advanced applications [1, 6]. Let us emphasize that we do not consider data polluted by noise, though it is a common difficulty in practice, see [4].

2. Identification problems

Cement hydration. The cement hydration process is modeled by the following initial value problem (IVP) presented in [3]

$$\frac{\mathrm{d}\alpha}{\mathrm{d}t}(t) = B_1 \left(\frac{B_2}{\alpha_\infty} + \alpha(t) \right) (\alpha_\infty - \alpha(t)) \exp\left(\frac{\eta}{\alpha_\infty} \alpha(t) \right) C, \tag{1}$$

$$\alpha(0) = 0, \tag{2}$$

where α is the time dependent degree of hydration and α_∞ stands for its limit value, B_1 and B_2 are coefficients dependent on the cement chemical composition, η represents the microdiffusion of free water, and $C \approx 2 \times 10^{-7}$ is a known constant, see [3].

It is assumed that

$$(\alpha_\infty, B_1, B_2, \eta) \in I_\alpha = [0.7, 1.0] \times [10^6, 10^7] \times [10^{-6}, 10^{-3}] \times [-12, -2]. \tag{3}$$

Generalized Van der Pol oscillator. Let us consider the following nonlinear IVP

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} = (c_1 - c_2 y^2) \frac{\mathrm{d}y}{\mathrm{d}t} - c_3 y, \tag{4}$$

$$y(0) = c_4, \quad \frac{\mathrm{d}y}{\mathrm{d}t}(0) = c_5, \tag{5}$$

where c_1, c_2, and c_3 are positive parameters, and c_4, c_5 are real parameters. If $c_1 = c_2 = c_3 = 1$, then we get the Van der Pol oscillator. It is assumed

$$(c_1, c_2, c_3, c_4, c_5) \in I_C = [0.5, 3]^3 \times [1, 3] \times [-1, 1]. \tag{6}$$

In both IVPs, the values of parameters are to be identified through m_i, that is, the measurements of either the hydration at time points $t_i \in [0, T_\alpha]$, $i = 1, 2, \ldots, n_\alpha$, or the measurements of the position $y(t_i)$ at $t_i \in [0, T_C]$, $i = 1, 2, \ldots, n_C$.

The identification problem: Find $\widehat{p} \in I$ such that

$$\widehat{p} = \arg\min_{p \in I} \Psi(p), \tag{7}$$

where

$$\Psi(p) = \sum_{i=1}^{n} w_i (m_i - u(t_i))^2 \tag{8}$$

and either $I \equiv I_\alpha$, $n \equiv n_\alpha$, and $u \equiv \alpha$ solves (1)–(2), or $I \equiv I_C$, $n \equiv n_C$, and $u \equiv y$ solves (4)–(5). The positive weighting factors w_i are also problem dependent and enable to increase or decrease the importance of some measurements.

3. Sensitivity analysis

To employ a gradient method for the minimization of Ψ, the gradient of Ψ with respect to the components of $p \in I$ is necessary. The partial derivatives of Ψ can be approximated by the numerical differentiation of Ψ, or by solving auxiliary problems. In the latter case, since

$$\frac{\partial \Psi}{\partial p_j} = 2 \sum_{i=1}^{n} w_i \left(m_i - u(t_i)\right) u'_{p_j}(t_i), \tag{9}$$

where p_j is a component of p and u'_{p_j} stands for the derivative of the state solution with respect to p_j, we have to find functions u'_{p_j} as the solutions of auxiliary IVPs.

According to [2], the derivatives of the state solution α with respect to α_∞, B_1, B_2, and η exist and they are solutions of

$$\frac{\mathrm{d}v}{\mathrm{d}t}(t) = g(t)v(t) + q(t), \tag{10}$$

$$v(0) = 0, \tag{11}$$

where the function g originates from the right-hand side of the state equation (1) differentiated w.r.t. the symbol α, whereas the derivative w.r.t. a parameter from the set $\{\alpha_\infty, B_1, B_2, \eta\}$ results in the function q.

To arrive at an IVP analogous to (10)–(11), we rewrite (4)–(5) into a system of first order equations

$$\frac{\mathrm{d}y_1}{\mathrm{d}t} = y_2, \tag{12}$$

$$\frac{\mathrm{d}y_2}{\mathrm{d}t} = (c_1 - c_2 y_1^2) y_2 - c_3 y_1, \tag{13}$$

$$y_1(0) = c_4, \quad y_2(0) = c_5. \tag{14}$$

By differentiating (12)–(14) w.r.t. y_1, y_2 and the parameters, we obtain the following parallel to (10)–(11)

$$\begin{pmatrix} \mathrm{d}v_1/\mathrm{d}t \\ \mathrm{d}v_2/\mathrm{d}t \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -2c_2 y_1 y_2 - c_3 & c_1 - c_2 y_1^2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} 0 \\ \omega \end{pmatrix}, \tag{15}$$

$$v_1(0) = \theta_1, \quad v_2(0) = \theta_2, \tag{16}$$

where $\theta_1 = 0 = \theta_2$ and $\omega = y_2$ if the derivative of the state solution $y \equiv y_1$ w.r.t. c_1 is to be calculated, $\omega = -y_1^2 y_2$ and $\omega = -y_1$ if we differentiate w.r.t. c_2 and c_3, respectively. If the derivative of y with respect to the initial conditions is sought, then $\omega = 0$ in (15) and $\theta_1 = 1$, $\theta_2 = 0$ in (16) if we differentiate w.r.t. c_4, or $\theta_1 = 0$, $\theta_2 = 1$ if we are interested in the sensitivity to c_5. Details in [2, Chapter 13 and 14].

To summarize, let us recall that for each parameter $\alpha_\infty, B_1, B_2, \eta$ (or $c_1, \ldots, c_5$), we infer and solve (10)–(11) (or (15)–(16)). After substituting v (or v_1) for u'_{p_j} in (9), we obtain one component of the gradient of Ψ.

The derivation of the expressions appearing on the right-hand side of (10) is easy for the generalized Van der Pol equation, see (15), but more laborious for the hydration problem. Nevertheless, it is effortlessly performed by the MATLAB® Symbolic Math Toolbox™ (we used its R2012b version), namely by its functions `diff` and `matlabFunction`. The latter converts symbolic expressions to MATLAB® functions.

4. Minimization

To minimize (8), the MATLAB® R2012b Optimization Toolbox™ `fmincon` function was used. It is designed for constrained minimization, see (3) and (6). The cost function gradient can be calculated by a black-box numerical differentiation, or by a code delivered by the user. We tried both approaches, and applied the sensitivity analysis approach explained in Section 3 in the latter.

Since parameter identification is a global minimization problem and `fmincon` is a tool for local minimization, optimization runs starting from different initial points belonging to I_α or I_c, see (3) and (6), were necessary to increase the chance of finding a global minimum.

5. Results, observations, and conclusions

In both problems, the weights w_i were chosen as equal.

Cement hydration. Figure 1 shows the graphs of the derivatives of the state solution α determined by $(0.7, 5 \times 10^6, 5 \times 10^{-4}, -2.5) \in I_\alpha$ with respect to the parameters. We observe a high sensitivity to B_2 and a low sensitivity to B_1. Moreover, the peak sensitivity occurs in a neighborhood of $t = 25$ and is, except for the case of α_∞, strongly localized. We deduce that the most important measurements are those made in between, say, $t = 5$ and $t = 50$ or $t = 100$. The state solution rapidly increases in $[0, 50]$, see Figure 2 (left), where the best fit to a set of 23719 real-world measurements of the cement hydration process (1)–(2) is depicted (time, t, in hours).

Generalized Van der Pol oscillator. Examples of the derivatives of $y \equiv y_1$, see (4) and (12), at $c_1 = c_2 = c_3 = 1$, $c_4 = 2$, and $c_5 = 0$ are depicted in Figure 3. The state solution y as well as some of its derivatives are periodic for a range of parameters, but the amplitude of the other derivatives is increasing, which might decrease the accuracy of the approximate expansion of y (w.r.t. the parameters) at times far from the initial time. Figure 2 (right) shows the initial solution y for the above values $c_1, \dots, c_5$, also seven points obtained via "measurements" derived from the state solution determined by parameters that are to be re-identified, and the state solution determined by the identified parameters.

Let us present a few observations and conclusions. The coupling of symbolic and numerical computation substantially reduces the amount of problem-dependent user-written code. Although the derivatives of the state solution w.r.t. the parameters reveal the sensitivity of the state solution to the perturbation of the parameters and are beneficial in the evaluation of (9) and, if possible, in the placement of the times of

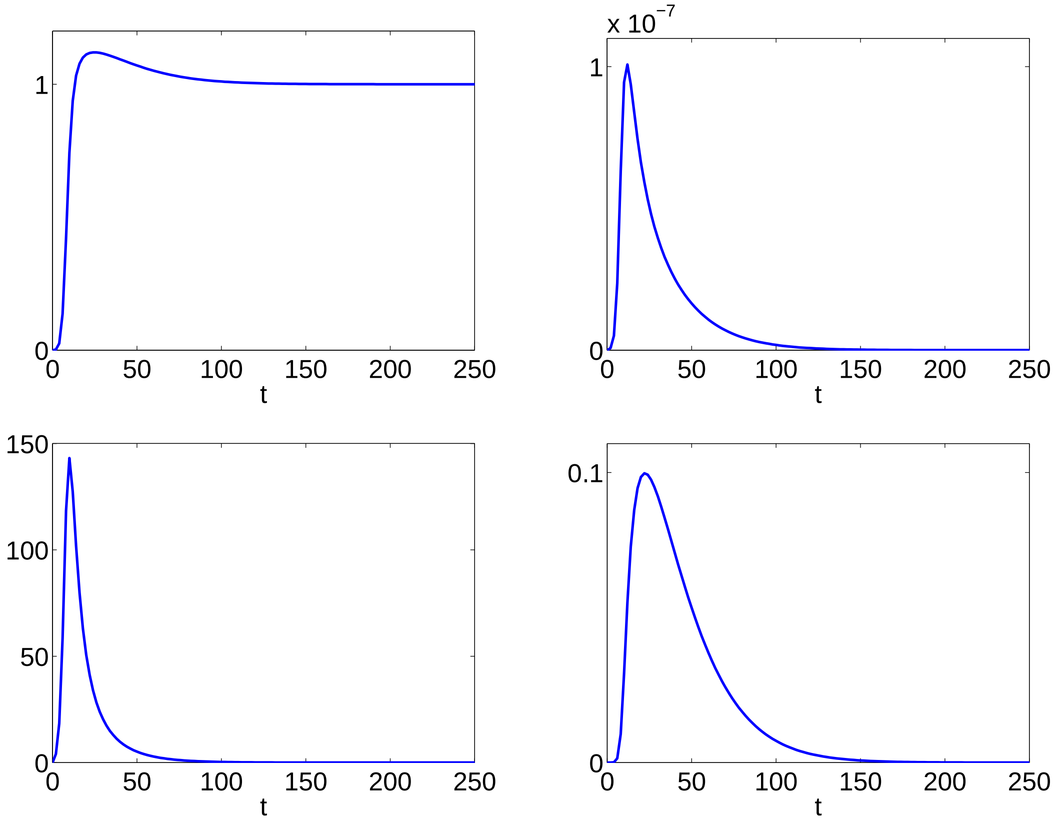

Figure 1: The derivative of α w.r.t. α_∞ (top left), B_1 (top right), B_2 (bottom left), η (bottom right).

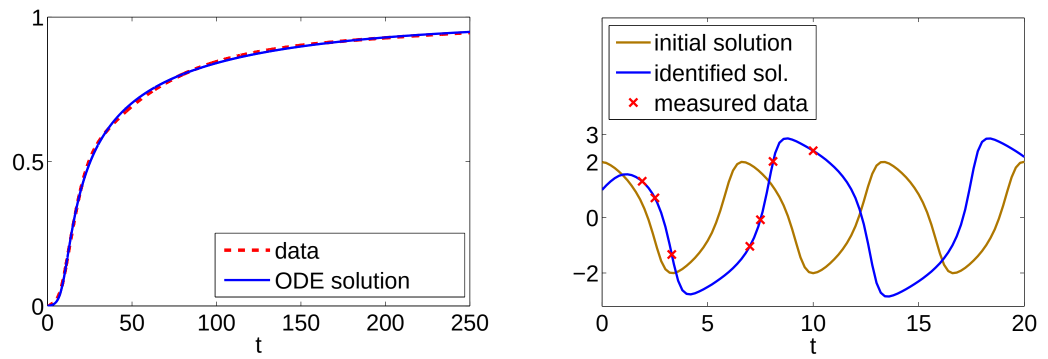

Figure 2: Identified solution to (1)–(2) (left), and to (4)–(5) (right).

measurements, their calculation slows the minimization process. Indeed, numerical differentiation turned out to be quite fast and accurate and might be considered the method of first choice in `fmincon` if the parameter identification is the only goal of calculation. In any case, however, the adjoint equation technique is worth considering to speed up the minimization process.

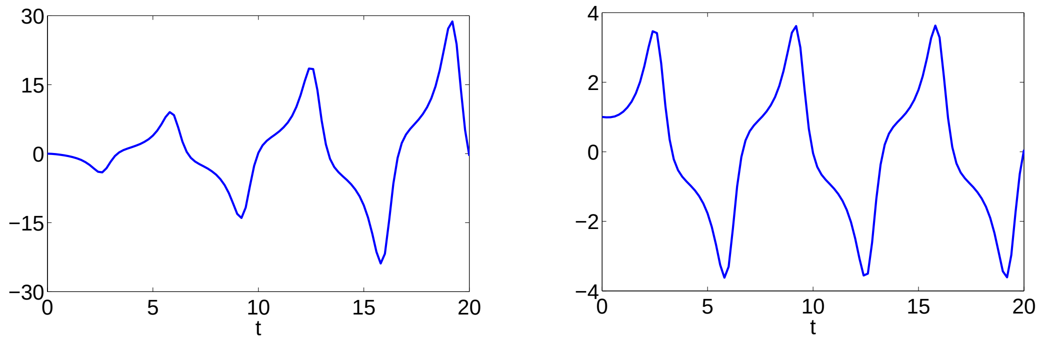

Figure 3: The derivative of y w.r.t. c_3 (left) and c_4 (right).

Acknowledgements

This work was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS14/003/OHK1/1T/11. The authors also wish to thank Karel Hájek and Ondřej Petlík for their assistance in the project and an anonymous referee for valuable comments.

References

[1] Babadzanjanz, L. K., Boyle, J. A., Sarkissian, D. R., Zhu, J.: Parameter identification for oscillating chemical reactions modelled by systems of ordinary differential equations. J. Comp. Methods Sci. Eng. **3** (2003), 223–232.

[2] Kurzweil, J.: *Ordinary differential equations: introduction to the theory of ordinary differential equations in the real domain.* Studies in Applied Mechanics, vol. 13, Elsevier Science Ltd, Amsterdam, 1986.

[3] Mareš, T.: Artificial neural networks in calibration of nonlinear models. Report, Dept. of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, 2012.

[4] Müller, T. G., Noykova, N., Gyllenberg, M., Timmer, J.: Parameter identification in dynamical models of anaerobic waste water treatment. Math. Biosci. **177/178** (2002), 147–160.

[5] Munster, D.: Parameter identification: a comparison of methods. Report, Dept. of Mathematics, College of Science, Virginia Tech, 2009.

[6] Wang, J., Ye, J., Yin, H., Feng, E., Wang, L.: Sensitivity analysis and identification of kinetic parameters in batch fermentation of glycerol. J. Comput. Appl. Math. **236** (2012), 2268–2276.