Online identification of nonlinearities in time dependent differential equations

P Kuegler

1 RICAM, Austrian Academy of Sciences, Altenbergerstrasse 69, 4040 Linz, Austria
E-mail: philipp.kuegler@jku.at

Abstract. Parameter identification in time dependent differential equations from time course observations related to the physical state can be understood as a nonlinear inverse and ill-posed problem and appears in a variety of applications in science and engineering. Online identification means to identify model parameters at the same time as the data are collected. Such algorithms become necessary whenever the model is needed in order to support decisions that have to be taken during the operation of the real system. In this paper we focus on the online identification of nonlinearities, i.e., functions that depend on (components of) the state variable described by the underlying differential equation. We apply a recently suggested method which is based on a nonlinear implicitly defined operator, works both for ODEs and time-dependent PDEs and renders typical restrictive assumptions such as full state observability and data differentiation unnecessary. Numerical illustrations are given.

1. Introduction

Abstract dynamical systems

\[ \begin{align*}
  u_t(t) &= f(u, q, t), \\
  u(0) &= u_0,
\end{align*} \]  

(1)

such as systems of ordinary differential equations or time-dependent partial differential equations play an important role in the modeling of instationary processes in science and engineering. Often, these models contain parameters \( q \) that are functions of the state variable \( u \) and/or its gradient. For this scenario, modelling examples from heat conduction, porous media flow and continuum mechanics can be found in [14]. For the purpose of illustrations we will consider the nonlinear PDE

\[ \begin{align*}
  u_t(x, t) - \nabla(q(|\nabla u|) \nabla u(x, t)) &= r(x, t), \quad x \in \Omega, \quad t > 0, \\
  u(x, t) &= 0, \quad x \in \partial\Omega, \quad t > 0, \\
  u(x, 0) &= u_0(x), \quad x \in \Omega
\end{align*} \]  

(2)

as a reference example for (1), which is nonlinear due to a model parameter \( q \) depending on \( |\nabla u| \).

In the context of (1), parameter identification is the task of indirectly determining model parameters \( q \) from observations in the frequency or time-domain related to the dynamical system. In
case of time-domain data the problem of parameter identification can be treated either off-line or online - depending on the application one has in mind. In off-line identification one first observes the dynamical system over a period of time and only then uses the data collected for the determination of the parameters. In online identification the parameters have to be identified simultaneously to the evolution of the real system and the data collection process. Parameter guesses at initial time have to be continuously improved since accurate parameter values are needed for making decisions while the system is in operation, e.g., as an input to control algorithms.

In the purely finite-dimensional case, i.e., (1) representing an ODE system with constant parameter vector \( q \), the literature on online parameter identification algorithms, see [1], [9], [13], [11] focuses on linear time invariant (and special nonlinear) dynamical systems that can be parameterized in a linear way based on data filtering. To the best of our knowledge, the online identification of nonlinearities in ODEs so far has not been considered. In case of (1) representing time-dependent PDEs, only a few papers on online identification of infinite dimensional parameters, i.e., functions of space or state variables, can be found in the literature, see [2], [3], [5], [12]. However, these techniques require full state observations of \( u \) and in many cases - including the identification of conductivity parameters - even observations of spatial derivatives. In [4], the authors consider "the elimination of these restrictions" as a "formidable challenge".

In [10] we suggested and analyzed an online identification method for (1) in finite and infinite dimensions that works without the typical assumptions such as full state observability, data filtering or differentiability or on structural properties of (1). This is achieved by the use of a nonlinear parameter-to-output operator that is only implicitly defined via the underlying dynamical system. The central idea is to drive the update of the current parameter guess by reducing the mismatch between the current data and the simulated output while only incorporating derivative information of the operator with respect to the current value of the parameter guess (instead of also using perturbations with respect to past values).

The purpose of this paper is to demonstrate the applicability of our method to the online identification of nonlinearities both in ODEs and PDEs. The paper is organized as follows. The next section sets the stage for our further discussion while the derivation of our method is shortly reviewed in Section 3. Section 4 contains numerical illustrations by means of a finite dimensional toy problem as well as of the nonlinear PDE (2), introduced in [3] as a reference problem. Special challenges in the context of identification of nonlinearities will be highlighted.

2. Preliminaries

We introduce the following notation and setup. Let the physical state \( u(t) \) and the model parameter \( q \) of (1) belong to Hilbert spaces \( U \) and \( Q \). We assume that a system output

\[
y(t) = h(u(t), q, t)
\]

with \( y(t) \) belonging to another Hilbert space \( Y \) can be observed over time and the corresponding data, possibly affected by measurement errors, are denoted by \( z(t) \in Y \). Furthermore, we suppose that the integration of (1) and evaluation of (3) is well-posed in the sense that for all parameters \( q \) belonging to an admissible set \( Q_{\text{ad}} \subseteq Q \) a unique solution \( u^q \) of (1) with corresponding output \( y^q \) exists such that \( u^q(t) \in U \) and \( y^q(t) \in Y \) holds for all \( t \). Furthermore, we assume that (1), (3) is differentiable with respect to the parameter \( q \in Q_{\text{ad}} \), that is for any...
perturbation \( p \in Q \) the linearized problem
\[
\begin{align*}
v_i(t) &= f_u(u^q(t), q, t)v(t) + f_q(u^q(t), q, t)p \\
v(0) &= 0
\end{align*}
\]
admits a unique solution \( v^{q,p} \) with \( v^{q,p}(t) \in U \) for all \( t \) and the corresponding output
\[
\begin{align*}
w^{q,p}(t) &= h_u(u^q(t), q, t)v^{q,p}(t) + h_q(u^q(t), q, t)p \\
\end{align*}
\]
satisfies \( w^{q,p}(t) \in Y \). Therein, \( v^{q,p} \) describes the sensitivity of \( u^q \) with respect to a perturbation of \( q \) in direction \( p \), see, e.g., [8] for details. Though these assumptions of course impose certain restrictions on the right-hand side \( f \) in (1), e.g., differentiability with respect to \( u \) and \( q \), see [6] or [16], we do not require \( f \) to have any special structural properties such as linearity or bilinearity with respect to its arguments.

Addressing the online parameter identification problem, we recall that neither existence, uniqueness nor stability of the solution to an inverse problem is guaranteed, see [7]. In this paper, we simply assume - for the case of exact data \( z(t) \in Y \) - the existence of a solution \( q_* \in Q_{ad} \) with
\[
z(t) = y^{q_*}(t)
\]
and shall discuss the issue of identifiability of nonlinearities and stability with respect to data noise.

With online identification meant to be performed simultaneously to the evolution of the dynamical system modelled by (1), (3), the goal of any online technique (in case of exact data) is to establish convergence of the current parameter guess \( \hat{q}(t) \in Q \) at time \( t \) towards \( q_* \) as time evolves, i.e.,
\[
\hat{q}(t) \to q_* \text{ as } t \to \infty.
\]
To this end we emphasize a central concept in the design and analysis of online identification methods, common to the methods discussed in the citations given and also followed in this paper. Though online parameter identification algorithms in practice are often also used to track time-varying parameters - all available convergence results are based on the assumption that the parameter \( q_* \in Q \) does not explicitly depend on the time variable. The reason is that the time derivative \( e_t(t) \) of the parameter estimation error
\[
e(t) := \hat{q}(t) - q_*
\]
then equals the time-derivative \( \hat{q}_t(t) \) of the parameter estimate, which tremendously simplifies theoretical matters, see for instance [15], [3]. The use of parameters constant with respect to time in the analysis is often also motivated by considering the plant dynamics to be much faster than those of the parameter. This does not exclude the case of parameter functions \( q_* \in Q \) that depend on (parts of) the time-varying state variable, e.g., \( q_* = q_*([\nabla u(x, t)]) \) as in (2). However, with \( \hat{q}(t) \in Q \) denoting the estimate of \( q_* \) at time \( t \), it is obvious that the parameter estimate \( \hat{q} \) depends on the time variable as a mapping from the time domain into the parameter space \( Q \).
Of course, for any fixed \( t \) the estimate \( \hat{q}(t) \in Q \) also shows the functional dependencies dictated by the definition of the parameter space \( Q \). For instance, consider (2) with a true parameter function
\[
q_*(\tau) = 0.9(1 - 0.5e^{-0.5\tau^2}), \ \tau \in I
\]
that satisfies \( q_* \in Q = H^1(I) \), where \( I \) denotes some positive real interval. Then, \( \hat{q}(t) \in Q \) corresponds to an \( H^1(I) \)-function for any \( t \) and convergence in (5) has to be understood with respect to the \( H^1(I) \)-norm.
3. Online Estimation Method based on a Nonlinear Operator

As in any online method the basic idea of our approach is to compare the data $z(t)$ at time $t$ with a predicted output $\hat{y}(t)$ and to use the deviation between them in order to improve the current parameter guess $\hat{q}(t)$. Our method involves a nonlinear parameter-to-output operator

$$F(\cdot, s) : Q_\rho \to Y, \hat{q} \to \hat{y}(s),$$

which is implicitly defined via integration of

$$\begin{align*}
\hat{u}_t(t) &= f(\hat{u}(t), \hat{q}(t), t), \quad \hat{u}(0) = u_0, \\
\hat{y}(t) &= h(\hat{u}(t), \hat{q}(t), t),
\end{align*}$$

up to time $s$. As opposed to (1), the system (9) now is considered for time varying parameters $\hat{q}$ belonging to the domain

$$Q_\rho := \{\text{time-varying } \hat{q} \mid \hat{q}(t) \in Q \land \|\hat{q}(t) - q_0\| < \rho, \ t \geq 0\}$$

of $F(\cdot, s)$ with some appropriate $\rho > 0$. Note that future values $\hat{q}(s)$ with $s > t$ have no impact on $\hat{y}(s)$ and a restriction of the domain of $F(\cdot, s)$ to the open ball $B_\rho(q_0)$ around $q_0 \in Q$ simply means to integrate (9) with a time-constant parameter, i.e.,

$$F(q, t) \equiv F(\hat{q}, t), \quad q \in B_\rho(q_0), \quad \hat{q} \in Q_\rho \text{ with } \hat{q}(s) \equiv q, \ s \leq t.$$  

As a key feature of our method, updates of current parameter guesses will only involve derivatives of $F$ with respect to the current guess $\hat{q}(t) \in Q$ (instead of derivatives with respect to $\hat{q} \in Q_\rho$). To this end, we introduce (with slight notational misuse of the superscript $'$ since usually reserved for the Fréchet derivative) for $\hat{q} \in Q_\rho$ the linear operator

$$F'(\hat{q}, s) : Q \to Y, p \to \hat{w}(s),$$

evaluated via integration of

$$\begin{align*}
\hat{v}_t(t) &= f_u(\hat{u}(t), \hat{q}(t), t)\hat{v}(t) + f_q(\hat{u}(t), \hat{q}(t), t)p, \quad \hat{v}(0) = 0, \\
\hat{w}(t) &= h_u(\hat{u}(t), \hat{q}(t), t)\hat{v}(t) + h_q(\hat{u}(t), \hat{q}(t), t)p.
\end{align*}$$

Note that its domain is only given by $Q$ instead of $Q_\rho$, i.e., only time-constant perturbations $p \in Q$ in (13) are considered. Our derivative approach is motivated by the fact that in online identification one is only interested in correction and update of the current parameter guess $\hat{q}(t) \in Q$, but not in improving past parameter guesses $\hat{q}(s)$ with $s < t$ (for which a time varying perturbation would be needed).

The concept chosen for the comparison between the data and the output actually is the concept of the total prediction error with exponential data forgetting. More precisely, we consider the functional

$$J(\hat{q}(t), t) = \int_0^t e^{-\beta(t-s)}\|z(s) - F(\hat{q}(t), s)\|_Y^2 \, ds + e^{-\beta t}(\hat{q}(t) - q_0)^T G_0^{-1}(\hat{q}(t) - q_0).$$

The first term measures the mismatch between all data up to current time $t$ and all output predictions obtained using the current parameter guess $\hat{q}(t)$ (here, $F(\hat{q}(t), s)$ has to be understood in the sense of (11) with a time-constant parameter $\hat{q}(t) \in B_\rho(q_0)$). For strictly positive $\beta$ the
exponential term in the integral acts as a forgetting factor, attaching less importance to the data the further they lie in the past. The purpose of the second term is to avoid that \( \hat{q} \) drifts too far away from the initial parameter guess \( q_0 \) during the initial phase of the identification process where the amount of available data still is limited. Therein, \( G_0 \) denotes a linear, selfadjoint positive definite operator on \( Q \).

Looking at the first order necessary condition for a minimizer of (15) and choosing an Euler method for its solution, then the online parameter identification method

\[
\begin{align*}
\dot{\hat{q}}(t) &= G(t)F'(\hat{q}, t) (z(t) - F(\hat{q}, t)), \quad \hat{q}(0) = q_0, \\
G_t(t) &= \beta G(t) - G(t) \left[ F'(\hat{q}, t)^* F'(\hat{q}, t) + \beta \bar{G}^{-1} \right] G(t), \quad G(0) = G_0.
\end{align*}
\]

(16) can be motivated. The update law (16) is defined by means of the adjoint of (12) and the gain operator \( G(t) : Q \rightarrow Q \). In the update law (17) for the latter, the linear, selfadjoint positive definite operator \( \bar{G} \) acts as an upper bound on the gain.

Convergence (5) of the method (16), (17) is shown for exact data in [10] under the following main assumption. It is required that the system (1) is permanently excited and that the output (3) gives rise to a data set rich enough. The abstract formulation of this persistance of excitation condition is

\[
\exists \gamma, \bar{t} > 0, \forall t \in R^+ \int_t^{t+\bar{t}} F'(\hat{q}, s)^* F'(\hat{q}, s) ds \geq \gamma I, \quad \hat{q} \in Q, \rho,
\]

which can be understood as an extension of the ideas used in, e.g., [15], [9].

4. Numerical Experiments

In this section, we apply the method (16), (17) to the online identification of nonlinearities. For a basic illustration we choose a simple ordinary differential equation, whereas the second example deals with the partial differential equation (2).

4.1. Toy Example

We consider the nonlinear ordinary differential equation

\[
u_t = 4 - c(t) \sin(u) - q_\star(u) + \cos\left(\frac{\pi}{2} - u\right), \quad u(0) = 0
\]

(18) and the online identification of the true parameter

\[
q_\star(\tau) = 2 + 3\tau + \tau^2
\]

(19) from observations of the system output \( y(t) = u(t) \). The excitation of the system is handled by the function \( c(t) \) for which we will choose between the two alternatives

\[
c(t) = 3 \quad \text{and} \quad c(t) = 3 \sin(0.05t).
\]

(20) Obviously, the numerical realization of the identification of infinite dimensional parameters requires their discretization. In the context of nonlinearities, i.e., state dependent parameter functions, one faces the problem that the range taken by their arguments may not be known a-priori. Since online identification is performed simultaneously to the data collection process, relevant
parameter domains for the discretization have to be narrowed down by anticipations about the time evolution of the state variable and possible adapted during the computations. However, for the sake of straightforwardness of our example, we fix the parameter space \( Q = H^1([0, 4.5]) \) with the interval \([0, 4.5]\) then sufficiently large for covering the ranges of the solution \( u_q(t) \) of (18) and of state estimates \( \hat{u} \) of (9) in both cases of (20) for all \( t > 0 \). For its discretization we choose standard linear ansatz functions defined on a uniform mesh for \([0, 4.5]\) consisting of \( m \) nodes.

Figures 1 and 2 show the results obtained by application of the method (16), (17) with the setting \( \beta = 0.1, \, \bar{G} = 10 \cdot I, \, G_0 = I, \, m = 20 \) and initial guess \( \hat{q}(\tau, 0) = q_0(\tau) \equiv 0 \) (similar results are obtained for similar setting). Figure 1 reflects the situation for the choice \( c(t) = 3 \) in (20). Then, the system (18) is not persistently excited and the true state \( u^{q_\ast}(t) \) quickly assumes the constant value \( u^{q_\ast} \equiv 0.375 \). Hence, the true parameter (19) can only be reconstructed for \( \tau = 0.375 \), which is achieved by the estimator \( \hat{q}(\tau, t) \) as \( t \) increases. For other values of \( \tau \) the data contain no information and the shape of the estimates is determined by the initial guess \( q_0 \) and smoothing properties of the space \( Q \).

For the choice \( c(t) = 3 \sin(0.05t) \) the state \( u^{q_\ast}(t) \) actually is driven into an oscillatory behaviour with an amplitude between 0.3755 and 1.3567. Now, the data contain enough information to recover the parameter (19) on the subinterval \( I_\ast = [0.3755, 1.3567] \) and convergence of the online estimator \( \hat{q}(t) \) towards \( q_\ast \) is given on \( I_\ast \) as \( t \to \infty \), see Figure 1 for the estimates \( \hat{q}(\tau, t_i) \) at times \( t_i = 0, 10 \) and 600.

### 4.2. Nonlinear PDE Example

In our second numerical illustration we consider the case where both the parameter space \( Q \) and the state space \( U \) are infinite dimensional. Motivated by an example presented in [3] we focus on the online identification of the parameter function (7) in the nonlinear PDE (2) for \( \Omega = [0, 1] \) with initial state \( u_0(x) = 0 \) and a source term

\[
r(x, t) = (\sin(4\pi t) + 0.001t^2) \cdot \chi_{[0.0, 0.5]}.
\]

(21)

Here, \( \chi_{[0.0, 0.5]} \) denotes the characteristic function of the interval \([0, 0.5]\). As opposed to the example of Section 4.1, the parameter \( q_\ast \) to be identified now does not depend on the state variable \( u^{q_\ast} \).
itself but on the absolute value of its gradient, i.e.,

\[ q_* = q_* (|\nabla u^{q_*}|). \]

The exact data in our example either result from full state observations, i.e.,

\[ z(x, t) = u^{q_0}(x, t), \quad x \in [0, 1], \quad t > 0, \tag{22} \]

or from partial state observations on the right half of the domain, i.e.,

\[ z(x, t) = u^{q_*}(x, t), \quad x \in [0.5, 1], \quad t > 0, \tag{23} \]

where \( u^{q_0} \) denotes the unique solution of (2) corresponding to the true parameter (7). The initial parameter guess is chosen as the constant function \( \hat{q}(\tau, 0) = 1 \). As before, the relevant parameter domain for the discretization of \( Q \) is not known a-priori, again we for simplicity rely on choosing an interval sufficiently large, here \( I = [0, 4] \). Both \( Q = H^1(I) \) and \( U = H^1([0, 1]) \) are discretized by means of linear ansatz functions, where in the latter case the homogeneous Dirichlet boundary conditions are taken into account. Those functions are also used for the discretization of the output space \( Z = L^2([0, 1]) \) or \( Z = L^2([0.5, 1]) \), respectively.

Figures 3 and 4 show the results obtained by the online method (16), (17) with

\[ \beta = 1, \quad G_0 = I, \quad \hat{G} = 100I, \quad n = 16, \quad m = 16, \tag{24} \]

where \( m \) and \( n \) denote the number of gridpoints used in the discretization of the spaces \( Q \) and \( U \). The case of full state observations is illustrated in Figure 3, which shows that the parameter estimate \( \hat{q} \) has converged towards \( q_* \) at time \( t = 100 \). This indicates that the system (2) is sufficiently excited by the source term (21). If exact data are available, one could also apply the method discussed in [3] (the quality of the results than are similar as above). However, the technique in [3] relies on the evaluation of the estimate \( \hat{q} \) by means of the data, i.e.

\[ \hat{q}(|\nabla z(x, t)|), \]

also showing why full state and exact data are necessary. As opposed to, our method (16), (17) builds on the evaluation of the estimate \( \hat{q} \) by means of the predicted state variable, i.e.,

\[ \hat{q}(|\nabla u^{\hat{q}}(x, t)|), \]
then also allowing for partial state observations and for noisy data.

Figure 4 shows the situation for the partial state observations (23). While the convergence property of the estimates is no longer given on all of the parameter domain $I$ due to the reduced amount of data, still a reliable online approximation of $q_\ast$ is obtained. In fact, the parameter estimates not only track the data on $[0, 0.5]$ but also follow the full state on $[0, 1]$ (not shown).

5. Conclusions and Outlook

In this paper we focused on the online identification of nonlinearities in finite and infinite dimensional time-dependent differential equations. To this end we utilized a method based on a nonlinear implicitly defined parameter-to-output operator and demonstrated its applicability by means of two different numerical examples. The advantage of the method is that it renders standard requirements such as full state observability, data differentiability or on structural properties of the underlying dynamical system unnecessary.

Several issues call for further research. It is well-known in the context of parameter identification that the case of noisy data $z^\delta$ requires special care due to the ill-posed nature of the problem. In online identification, the most popular technique for avoiding parameter divergence due to data noise is the leakage approach, see, e.g., [9] or [5], which means to extend (16) according to

$$\dot{\bar{q}}(t) = G(t) F'(\bar{q}, t)^\dagger (z^\delta - F(\bar{q}, t)) - \sigma(t) G(t) (\bar{q}(t) - q_0),$$

with some positive function $\sigma(t)$. While first numerical tests with spatially varying parameters give confidence in the use of (25), it is an open question if (25) renders a regularization method in the strict sense of the definition and how the regularization function $\sigma$ should be chosen.

From the numerical point of view, the actual integration of (16) (or (25)), (17) poses future research challenges. So far, we only used the MATLAB built in ODE solver package, however, a tremendous reduction of the computational cost is expected from especially tailored integration codes and would be certainly needed for online real world applications.

References

[1] K.J. Aström, Theory and applications of adaptive control, Automatica 75 (1983), 471-486.
[2] H.T. Banks, K. Kunisch, Estimation Techniques for Distributed Parameter Systems, Birkhäuser, 1989.
[3] J. Baumeister, W. Scondo, M.A. Demetriou, I.G. Rosen, Online parameter estimation for infinite-dimensional dynamical systems, SIAM J. Control Optim. 35, (1997), 678-713.
[4] M. Böhm, M.A. Demetriou, S. Reich, I.G. Rosen, Model reference adaptive control of distributed parameter systems, SIAM J. Control Optim. 36, (1998), 33-81.
[5] M.A. Demetriou, I.G. Rosen, On-line robust parameter identification for parabolic systems, Int. J. Adapt. Control and Signal Processing, 15, (2001), 615-631.
[6] P. Deuflhard, F. Bornemann, Scientific Computing with Ordinary Differential Equations, Springer, 2002.
[7] H.W. Engl, M. Hanke, A. Neubauer, Regularization of Inverse Problems, Kluwer Academic Publishers, 1996.
[8] A.C. Hindmarsh et al., SUNDIALS: Suite of nonlinear and differential/algebraic equation solvers, ACM Transactions on Mathematical Software (TOMS) archive 31, (2005), 363 - 396.
[9] P. Ioannou, J. Sun, Robust Adaptive Control, Prentice Hall, 1996, electronic copy at http://www-rcf.usc.edu/~ioannou/RobustAdaptiveControl.htm.
[10] P. Kuegler, Online parameter identification in time dependent differential equations as a nonlinear inverse problem, submitted to EJAM, 2007.
[11] L. Ljung, S. Gunnarsson, Adaptation and tracking in system identification - a survey, Automatica 26, (1990), 7-21.
[12] Y. Orlov, J. Bentsman, Adaptive distributed parameter systems identification with enforceable identifiability conditions and reduced-order spatial differentiation, IEEE Transactions on Automatic Control 45 (2000), 203 – 216.

[13] S. Sastry, M. Bodson, Adaptive Control - Stability, Convergence and Robustness, Prentice Hall, 1989, electronic copy available at http://www.ece.utah.edu/~bodson/acscr/.

[14] R.E. Showalter, Monotone Operators in Banach Space and Nonlinear Partial Differential Equations, American Mathematical Society, 1996.

[15] J.E. Slotine, W. Li, Applied Nonlinear Control, Prentice Hall, New Jersey, 1991.

[16] R. Temam, Infinite-Dimensional Dynamical Systems in Mechanics and Physics, Applied Mathematical Sciences 68, Springer, 1988.