Mathematical Biosciences 233 (2011) 19–31

Contents lists available at ScienceDirect

Mathematical Biosciences

journal homepage: www.elsevier.com/locate/mbs

Finding identifiable parameter combinations in nonlinear ODE models and the rational reparameterization of their input–output equations

Nicolette Meshkat [a,*], Chris Anderson [a], Joseph J. DiStefano III [b]

[a] UCLA, Department of Mathematics, United States
[b] UCLA, Department of Computer Science, United States

ARTICLE INFO

Article history:
Received 19 November 2010
Received in revised form 30 May 2011
Accepted 2 June 2011
Available online 18 July 2011

Keywords:
Identifiability
Differential algebra
Gröbner Basis
Reparameterization

ABSTRACT

When examining the structural identifiability properties of dynamic system models, some parameters can take on an infinite number of values and yet yield identical input–output data. These parameters and the model are then said to be unidentifiable. Finding identifiable combinations of parameters with which to reparameterize the model provides a means for quantitatively analyzing the model and computing solutions in terms of the combinations. In this paper, we revisit and explore the properties of an algorithm for finding identifiable parameter combinations using Gröbner Bases and prove useful theoretical properties of these parameter combinations. We prove a set of M algebraically independent identifiable parameter combinations can be found using this algorithm and that there exists a unique rational reparameterization of the input–output equations over these parameter combinations. We also demonstrate application of the procedure to a nonlinear biomodel.

1. Introduction

Parameter identifiability analysis for dynamic system ODE models addresses the question of which unknown parameters can be quantified from given input–output data. *Unidentifiable* parameters can take on an uncountably infinite number of values and yet result in identical input–output data. In such cases, the model and its parameter vector $\boldsymbol{p}$ are underdetermined with respect to the input–output data. This indeterminacy can be removed by finding the 'simplest' combinations of parameters that take on a unique or finite number of values, which are then used as candidates to reparameterize the model, rendering it *identifiable*. Thus the question becomes, how can identifiable parameter combinations be found?

This question has been partially answered for several model classes, under limited conditions. Evans and Chappell [1] and Gunn et al. [2] adapt the Taylor series approach of Pohjanpalo [3] to find locally identifiable combinations. Chappell and Gunn [4] use the similarity transformation approach to generate locally identifiable reparameterizations. Thus, with these methods identifiability can only be guaranteed (at least) locally. The problem of finding identifiable parameter combinations has also been addressed using differential algebra methods, as Denis-Vidal et al. [5] and Boulier [6] find globally identifiable combinations of parameters using an "inspection" method as discussed later in this paper. However, as shown by Meshkat et al. [7], this method is difficult to implement as a fully automated computational procedure.

In [7], an algorithm was outlined for finding the 'simplest' set of globally identifiable parameter combinations for a practical class of nonlinear ODE models. This algorithm extended the method of Saccomani et al. [8] using a variation of the Gröbner Basis approach. In this paper, we address several questions that arose in [7] regarding properties of the identifiable parameter combinations found, including algebraic independence and the existence of a rational reparameterization of the input–output equations derived from the original nonlinear model. Although a rational reparameterization of the original nonlinear model cannot always be done (as shown in [1]), we prove here that a unique rational reparameterization of the input–output equations can always be found over algebraically independent parameter combinations. In addition to being useful in quantifying the model and exercising its solutions, we will show that the ability to rationally reparameterize the input–output equations leads to a rigorous proof of identifiability.

2. Nonlinear ODE model

The general form of the models under consideration is:

$$\begin{aligned}\dot{\boldsymbol{x}}(t,\boldsymbol{p}) &= \boldsymbol{f}(\boldsymbol{x}(t,\boldsymbol{p}),\boldsymbol{u}(t),t;\boldsymbol{p}), \quad t\in[t_0,T]\\ \boldsymbol{y}(t,\boldsymbol{p}) &= \boldsymbol{g}(\boldsymbol{x}(t,\boldsymbol{p});\boldsymbol{p})\end{aligned} \tag{2.1}$$

Here $\boldsymbol{x}$ is a n-dimensional state variable, $\boldsymbol{p}$ is a P-dimensional parameter vector, $\boldsymbol{u}$ is the r-dimensional input vector, and $\boldsymbol{y}$ is the m-dimensional output vector. We assume $\boldsymbol{f}$ and $\boldsymbol{g}$ are rational

* Corresponding author.
E-mail address: nmeshkat@math.ucla.edu (N. Meshkat).

0025-5564/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.mbs.2011.06.001

polynomial functions of their arguments. Also, constraints reflecting known relationships among parameters, states, and/or inputs are assumed to be already included in (2.1), because they generally affect identifiability properties [9]. For example, $\boldsymbol{p} \geqslant \boldsymbol{0}$ is common.

3. Identifiability and the differential algebra approach

The question of *a priori structural identifiability* concerns finding one or more sets of solutions for the unknown parameters of a model from noise-free experimental data. Structural identifiability is a necessary condition for finding parameter values in the real "noisy" data problem, often called the *numerical identifiability* problem.

Structural identifiability can be expressed as an injectivity condition, as in [8]. Let $\boldsymbol{y} = \boldsymbol{\Phi}(\boldsymbol{p},\boldsymbol{u})$ be the input–output map determined from (2.1) by eliminating the state variable $\boldsymbol{x}$. Consider the equation $\boldsymbol{\Phi}(\boldsymbol{p},\boldsymbol{u}) = \boldsymbol{\Phi}(\boldsymbol{p}^*,\boldsymbol{u})$, where $\boldsymbol{p}^*$ is an arbitrary point in parameter space and $\boldsymbol{u}$ is the input function. If there exists only one solution $\boldsymbol{p} = \boldsymbol{p}^*$, then this corresponds to global identifiability. If there exists finitely many distinct solutions for $\boldsymbol{p}$, then this corresponds to local identifiability. Infinitely many solutions for $\boldsymbol{p}$ corresponds to unidentifiability.

The *a priori structural identifiability* problem can be solved using the differential algebra approach of Saccomani et al. [8], which follows methods developed by Ljung and Glad [10] and Ollivier [11,12]. Their program, DAISY, can be used to automatically check global identifiability of nonlinear dynamic models [13]. We note that DAISY has been applied to mostly low-dimensional systems, where a "pseudo-randomly" generated numerical value for $\boldsymbol{p}^*$ is used to speed up the computation process [13]. We summarize their approach below. A detailed description can be found in [7,13].

Using Ritt's pseudodivision algorithm, an input–output map can be determined in implicit form. The result of the pseudodivision algorithm is called the *characteristic set* [11]. Since the ideal generated by (2.1) is a prime ideal [14], the characteristic set is a "minimal" set of differential polynomials which generate the same differential ideal as the ideal generated by (2.1) [13]. The first m equations of the characteristic set are those independent of the state variables, and form the *input–output relations* [13]:

$$\boldsymbol{\Psi}(\boldsymbol{y},\boldsymbol{u},\boldsymbol{p}) = \boldsymbol{0}. \tag{3.1}$$

The characteristic set is in general non-unique, but the coefficients of the input–output equations can be fixed uniquely by normalizing the equations to make them monic [13].

The m equations of the input–output relations $\boldsymbol{\Psi}(\boldsymbol{y},\boldsymbol{u},\boldsymbol{p}) = \boldsymbol{0}$ are polynomial equations in the variables $\boldsymbol{u},\dot{\boldsymbol{u}},\ddot{\boldsymbol{u}},\ldots,\boldsymbol{y},\dot{\boldsymbol{y}},\ddot{\boldsymbol{y}},\ldots$ with rational coefficients in the elements of the parameter vector $\boldsymbol{p}$. Specifically, these equations involve polynomials from the differential ring $\mathbb{R}(\boldsymbol{p})[\boldsymbol{u},\boldsymbol{y}]$, where $\mathbb{R}(\boldsymbol{p})$ is the field of rational functions over the real numbers in the parameter vector $\boldsymbol{p}$. For each equation, we can write $\Psi_j(\boldsymbol{y},\boldsymbol{u},\boldsymbol{p}) = \sum_k c_{jk}(\boldsymbol{p})\psi_{jk}(\boldsymbol{u},\boldsymbol{y})$, where $c_{jk}(\boldsymbol{p})$ is a rational function in the parameter vector $\boldsymbol{p}$ and $\psi_{jk}(\boldsymbol{u},\boldsymbol{y})$ is a monomial function in the variables $\boldsymbol{u},\dot{\boldsymbol{u}},\ddot{\boldsymbol{u}},\ldots,\boldsymbol{y},\dot{\boldsymbol{y}},\ddot{\boldsymbol{y}},\ldots$, etc. We call $c_{jk}(\boldsymbol{p})$ the coefficients of the input–output equations, and for convenience we reindex the coefficients as $c_i(\boldsymbol{p})$, where $1 \leqslant i \leqslant l$, l is the total number of coefficients, and $l \geqslant M$, defined below.

To form an injectivity condition, we set $\boldsymbol{\Psi}(\boldsymbol{y},\boldsymbol{u},\boldsymbol{p}) = \boldsymbol{\Psi}(\boldsymbol{y},\boldsymbol{u},\boldsymbol{p}^*)$. Then global identifiability becomes injectivity of the map $\boldsymbol{c}(\boldsymbol{p})$ [13]. That is, identifiability is determined by the equations

$$\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*) \tag{3.2}$$

for arbitrary $\boldsymbol{p}^*$ [13]. Thus, the model (2.1) is *a priori* globally identifiable if and only if $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ implies $\boldsymbol{p} = \boldsymbol{p}^*$ for arbitrary $\boldsymbol{p}^*$ [13]. The equations $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ are called the *exhaustive summary* [11].

If there are finitely many distinct solutions for $\boldsymbol{p}$, then the model (2.1) is locally identifiable. The model (2.1) is unidentifiable if there are infinitely many solutions for $\boldsymbol{p}$, that is, the solution for $\boldsymbol{p}$ is expressed in terms of one or more free variables. Thus, determining structural identifiability is reduced to the nature of the solutions to $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$, which is typically solved by finding a Gröbner Basis and using elimination [13].

4. Some methods for finding identifiable parameter combinations

We focus on the case when (3.2) has infinitely many solutions (unidentifiability) in this paper. Unidentifiable models cannot be quantified from input–output data. A useful alternative is to find identifiable parameter combinations which can always be determined from input–output data, and attempt to reparameterize our model (2.1) in terms of these new parameters. Before we revisit our method for finding identifiable parameter combinations [7], we briefly present two other methods for finding identifiable parameter combinations using the differential algebra approach. Both procedures rely on using the exhaustive summary $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ to find parameter combinations that are either uniquely or finitely determined by $\boldsymbol{p}^*$.

Definition. Let s be the number of free parameters, defined as the total number of parameters P minus the number of components M in any branch of the solution vector to $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$. This number is also defined as the dimension of a variety, and will be discussed in Section 9.1.

That is, there are s free parameters and M "non-free" parameters, where $P = M + s$. Sometimes identifiable combinations can easily be found directly from the solutions to the equations $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$, by algebraically manipulating their solutions to form $M = P - s$ parameter combinations in terms of $\boldsymbol{p}^*$ only. In other words, find solutions of the form $g(\boldsymbol{p}) = g(\boldsymbol{p}^*)$. For example, in the Nonlinear 2-Compartment Model in [7], the solution to $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ is of the following form, where $\boldsymbol{p} = \{k_{21},k_{12},V_M,K_M,k_{02},c_1,b_1\}$ and $\boldsymbol{p}^* = \{\alpha,\beta,\gamma,\delta,\epsilon,\zeta,\eta\}$:

$$V_M = \frac{\gamma\zeta}{c_1}$$
$$k_{21} = \alpha$$
$$k_{12} = \beta$$
$$b_1 = \frac{\zeta\eta}{c_1}$$
$$k_{02} = \epsilon$$
$$K_M = \frac{\delta\zeta}{c_1}$$

Then clearly $\{c_1V_M,k_{21},k_{12},b_1c_1,k_{02},c_1K_M\}$ are uniquely determined by $\boldsymbol{p}^*$ because we can move the parameter vector $\boldsymbol{p}$ all to one side of the equation. To verify global identifiability, one would then reparameterize $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{c}(\boldsymbol{p}^*)$ over these parameter combinations $\{c_1V_M,k_{21},k_{12},b_1c_1,k_{02},c_1K_M\}$ and check the injectivity condition.

However, this ability to "move all parameters to one side of the equation" and thus "decouple" our parameter solution cannot always easily be done, as demonstrated in the Linear 2-Compartment Model below [7], where $\boldsymbol{p} = \{k_{01},k_{02},k_{12},k_{21},v\}$ and $\boldsymbol{p}^* = \{\alpha,\beta,\gamma,\delta,\epsilon\}$:

$$k_{21} = \alpha + \delta - \frac{\alpha k_{02} - \alpha\beta + \delta k_{02} - \beta\delta - \alpha\gamma}{k_{02} - \beta - \gamma}$$
$$k_{01} = \frac{\alpha k_{02} - \alpha\beta + \delta k_{02} - \beta\delta - \alpha\gamma}{k_{02} - \beta - \gamma}$$
$$k_{12} = -k_{02} + \beta + \gamma$$
$$v = \epsilon$$

Here we see that it takes more effort to find the uniquely determined parameter combinations $\{v,k_{12}k_{21},k_{02} + k_{12},k_{01} + k_{21}\}$.