Journal of the Taiwan Institute of Chemical Engineers 57 (2015) 26–35

Contents lists available at ScienceDirect

Journal of the Taiwan Institute of Chemical Engineers

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

Two-stage parameter estimation applied to ordinary differential equation models

Jyh-Shyong Chang*, Chia-Chi Li, Wei-Ling Liu, Jin-Han Deng

Department of Chemical Engineering, Tatung University, 40 Chungshan North Road, 3rd Sec., Taipei, Taiwan

ARTICLE INFO

Article history:
Received 20 December 2014
Revised 20 April 2015
Accepted 3 May 2015
Available online 30 May 2015

Keywords:
Ordinary differential equation models
Parameter estimation
Principal differential analysis

ABSTRACT

A two-stage (TS) parameter estimation method for identifying ordinary differential equation (ODE) models, which minimizes the squared residuals between the derivatives of the fitted values of available measurements and the rate changes of the states, is presented. The model is an alternative to the traditional method, in which the measured squared residuals and measurements of state are used to obtain near-optimal estimated parameters for ODEs in the first stage of the TS method. Therefore, the problem of ill-conditioning, encountered in integrating the system equations, can be ignored. The second stage of the TS method minimizes the squared residuals in the model's traditional integrated form by using the estimated parameter values from the first stage as initial guesses. Three examples were examined, and the results confirmed the utility of the developed TS parameter estimation method for easy programming and low computation time. The first example is a semi-batch reactor model of toluene hydrogenation, while a continuous stirred-tank reactor model is used for the second example. Finally, a batch-process model focused on the synthesis of glycerol ether synthesis from glycerol and tert-butyl alcohol is examined.

1. Introduction

A fundamental task in engineering and scientific research is the extraction of information from raw data. Accurate parameter estimation is crucially important for setting boundary conditions and estimating the constants that are crucial to in the mathematical models that we use for process simulation [1]. Obtaining good parameter values requires reliable parameter estimation techniques. The models that describe underlying process behavior, within defined initial and boundary conditions may be presented as algebraic, differential, or integral equations. Among the models, ordinary differential equations (ODEs) or differential-algebraic equations (DAEs) are widely used to model dynamic processes in natural science disciplines spanning biology to engineering [2]. This article focuses on developing an efficient and easy-to-use parameter estimation algorithm using continuous-time ODEs or DAEs. The scripts lsqnonlin and scripts ode45 that are used to solve nonlinear least-squares (*i.e.* nonlinear data-fitting) problems and initial value problems for ordinary differential equations respectively, constitute the central parts of the developed parameter estimation algorithm.

The most often used statistical techniques used to obtain the parameters of an input—output model adopt least-squares (linear or non-linear) regression, involving minimization of the objective function with respect to known parameters, which is usually the sum of the squared differences, *i.e.* between measured values and the model's predictions. Conventional optimization methods are commonly placed in one of two categories, *i.e.* direct search or gradient based methods. The methods are contrasted in that gradient based search methods require derivatives of the objective function, while direct search methods are derivative free. The Gauss-Newton method, a gradient based method, is an iterative nonlinear least-squares (NLS) procedure applicable to parameter estimation in ODE/DAE models with no analytical solutions. ODEs, or DAEs, when used to simulate dynamic experiments use guesses for the initial parameter values [3]. Linga et al. [4], compared the direct search optimization method of Luus and Jaakola (LJ) [5] and the Gauss-Newton (GN) method [6], which can be used to solve four parameter optimization problems, commonly encountered in the sciences, in *e.g.* chemical and biochemical processes, which are currently described by ordinary differential equation based models. It was found that in contrast to the initial guesses having only a minimal influence on the resulting optimized parameters, the Gauss-Newton algorithm was sensitive to the values chosen.

A new optimization algorithm has been created to compute a set of parameter estimates for the model that should give better predictions—based on a comparison of model predictions and measured responses. Sensitivity equations, solved along with the model ODEs, give the Jacobian of the response variables with respect to the

* Corresponding author Tel.: +886 2 1822928-6266; fax: +886 2 5861939.
E-mail address: jschang@ttu.edu.tw (J.-S. Chang).

http://dx.doi.org/10.1016/j.jtice.2015.05.004
1876-1070/

parameters [7], so that the 'optimizer' can determine the correct search direction to refine the parameters' values [4]. Alternatives, using either a numerical Jacobian, or a direct-search optimizer, require additional dynamic simulations, together with perturbations for each parameter. The ODEs (or DAEs) are solved numerically, subsequent to the determination of new parameter values being determined by the optimizer, using the updated parameters. The derived information is then compared with observed data—iteration involving parameter updating and the computation of numerical solutions continues until the parameters' convergence criteria are satisfied; or, until there is no additional improvement in the objective function. In terms of computation, such iterative methods can be time consuming, with much of computational effort being devoted to the repeated solution of the differential equations. With respect to the unknown values of the initial states, one needs to estimate them along with the model parameters. Furthermore, numerical integration of the differential equations can get stuck at a parameter domain, resulting in the cessation of the parameter estimation process. Similar problems can occur in the application of the direct method.

Varah [8] developed an alternative parameter estimation technique, not requiring repeated solution of the ODEs, which is based on the works of Swartz and Bremermann [9] and Benson [10]. In this approach, together with a related technique, known as 'principal differential analysis' (PDA) [11], discrete measurements of the output variables, y, are fitted empirically using splines that can be differentiated with respect to time to give estimated time-derivative curves. This time-derivative information is substituted into the ODEs, thereby converting the estimation of parameters, from a relatively difficult dynamic optimization problem, into a comparatively simple algebraic optimization operation that can be solved using either linear, or nonlinear, least-squares methods depending on the linearity of the ODEs in the parameters. PDA techniques [8,11] differ from the often used nonlinear least-squares methods for dynamic models, and from an early spline-based method [12] and its iterative extension as given by Madar et al. [13]. In PDA, the parameter values are selected to minimize squared residuals in the differential form of the model, rather than in the traditional integrated form. Ramsay and Silverman [14] and Poyton et al. [3] extended Varah's method by iterating the two steps, and replacing the previous iteration's 'roughness penalty' with a penalty on the model's differential form using the last minimizing value of the estimated parameters $\boldsymbol{k}$. They found that the PDA approach, converged quickly to give estimates of both $\boldsymbol{x}$ (states or outputs of ODEs) and $\boldsymbol{k}$ (unknown structural parameters) giving substantially improved bias and precision. Ramsay et al. [15] described a method that used 'noisy measurements' on a subset of variables in order to estimate parameters that define a system of non-linear differential equations. Their approach used modified data smoothing methods together with a generalized profiled estimation. A developed multi-criterion optimization problem based on the inner, middle and outer criteria for obtaining estimates and confidence intervals of the parameters showed them to have low bias and good coverage properties. Ramsay et al. [15] adopted collocation methods that express the approximation $\tilde{x}_i$ of x_i in terms of the basis function $(\varphi_i(t))$ expansion $\tilde{x}_i(t) = \boldsymbol{c'}_i\varphi_i(t)$. However, the number of the parameters $\boldsymbol{c}$ is potentially orders of magnitude larger than the number of structural parameters $\boldsymbol{k}$; thus adding to the complexity of the parameter estimation algorithm.

The advantage in applying the PDA method to obtain near optimal estimated parameters for ODEs is that the parameter values are selected to minimize squared residuals in the differential form of the model and not its traditional integrated form. Therefore, the problem of ill-conditioning encountered in integrating the system equations can be eliminated. Hence a two-stage (TS) parameter estimation for differential equations is proposed, which includes a first stage to obtain near optimal estimated parameters by the PDA method as the initial guesses for the second stage. In the second stage, parameter estimation for differential equations by the well-established non-linear regression framework [16] requiring integration of the system equations is then performed. Using simulation experiments, the two-stage parameter estimation for differential equations, developed in this work, can provide the user with an efficient and easy-to-use parameter estimation algorithm. The idea of employing the estimated parameters, obtained by minimizing the squared residuals in the differential form of the underlying model, as an initial guess for the conventional approach has already been introduced in the literatures [17–19]. Further, many two-stage parameter estimation (or identification) methods have been proposed in the system identification field [20–26].

2. Two-stage parameter estimation method

2.1. Initial parameter estimation for ODEs using the PDA method (stage I)

Considering processes described by ODE models of the following form.

$$\frac{d\boldsymbol{x}(t)}{dt} = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}, \boldsymbol{k}) \tag{1}$$

$$\boldsymbol{y}(t) = \boldsymbol{C}\boldsymbol{x}(t) + \boldsymbol{\varepsilon} \tag{2}$$

Where: $\boldsymbol{k} = [k_1, k_2, \ldots, k_p]^{\mathrm{T}}$ is a p-dimensional vector of parameters with unknown numerical values; $\boldsymbol{x}(t) = [x_1, x_2, \ldots, x_n]^{\mathrm{T}}$ is an n-dimensional vector of state variables; $\boldsymbol{x}_0 = [x_{10}, x_{20}, \ldots, x_{n0}]^{\mathrm{T}}$ is an n-dimensional vector representing the initial conditions for the state variables which are assumed to be known precisely; $\boldsymbol{u} = [u_1, u_2, \ldots, u_r]^{\mathrm{T}}$ is an r-dimensional vector of manipulated variables, either estimated, or definite if the numerical values are known *e.g.* from measurements; $\boldsymbol{f} = [f_1, f_2, \ldots, f_n]^{\mathrm{T}}$ is an n-dimensional vector function of known form (the differential equations); while $\boldsymbol{y} = [y_1, y_2, \ldots, y_m]^{\mathrm{T}}$ is the m-dimensional output vector, *i.e.*, the set of variables that is measured experimentally; $\boldsymbol{C}$ is the $m \times n$ observation matrix, $\boldsymbol{y} = \boldsymbol{C}\boldsymbol{x} + \boldsymbol{\varepsilon}$; and $\boldsymbol{\varepsilon}$ is a random vector.

Another important consideration is how to select the appropriate objective function. In general, the unknown parameter vector $\boldsymbol{k}$ is found by minimizing a scalar function, often termed as the objective function, which is a measure of the overall difference between the model-calculated-values and the measured values.

A suitable objective function to be to minimized is given by the equation,

$$S(\boldsymbol{k}) = \left[\sum_{i=1}^{N} [\hat{\boldsymbol{y}}_i - \boldsymbol{y}(t_i, \boldsymbol{k})]^{\mathrm{T}} \boldsymbol{Q}_i [\hat{\boldsymbol{y}}_i - \boldsymbol{y}(t_i, \boldsymbol{k})]\right] \tag{3}$$

Where: $S(\boldsymbol{k})$ represents the objective function, $\boldsymbol{k} = [k_1, k_2, \ldots, k_p]^{\mathrm{T}}$ is the p-dimensional vector representing parameters with unknown values, $\hat{\boldsymbol{y}}_i$ are the experimentally measured values, $\boldsymbol{y}(t_i, \boldsymbol{k})$ are the model calculated output values, Q_i is an $m \times m$ user specified weighting matrix, and N is the number of experimentally measured values for each run. The weighting matrix Q_i should be chosen, such that the parameter estimates have appropriate or desirable statistical properties. If the error terms $(\boldsymbol{\varepsilon}_i)$ are normally distributed with zero as the mean and with a known covariance matrix $\sum_i$, then Q_i is the inverse of this covariance matrix [6].

Instead of adopting Eq. 3 as the objective function, the PDA approach [3] suggested an alternative objective function to be minimized:

$$S(\boldsymbol{k}) = \boldsymbol{\lambda}^{\mathrm{T}} \int \| (\boldsymbol{C}\dot{\tilde{\boldsymbol{x}}} - \boldsymbol{C} f(\tilde{\boldsymbol{x}}(t), \boldsymbol{u}, \boldsymbol{k})) \|^2 dt \tag{4}$$

where: λ is the weighting vector of the ODEs. The discrete measurements $\hat{y}_i$ can be fitted into a smooth function of time $\tilde{y}_i(t)$ by