Stage-wise Parameter Estimation for Stiff Differential Equations

The use of singular perturbation approximation for estimating parameters in stiff ordinary differential equations is illustrated. A sequential parameter estimation approach, in which the difficult-to-estimate parameters are introduced in higher-order terms of the perturbation solution, is developed and illustrated. This approach is shown to provide excellent initial estimates, to be computationally efficient, and to be insensitive to initial guesses for the parameters. Aspects of data requirements for this problem estimation are discussed.

V. SRINIVASAN and R. C. AIKEN

Department of Chemical Engineering
University of Utah
Salt Lake City, UT 84112

SCOPE

Ordinary differential equation (ODE) models are used for numerical simulation in many scientific and engineering applications. Successful simulation of a process must be preceded by the solution of the so-called inverse problem: given measured data and an associated theoretical model, determine parameters in the model that give the "best fit." This is a complex problem: when the model equations are stiff, the computational and experimental complexity increases manyfold. Typically, a parameter estimation procedure would require numerical integration of the stiff system many times during an estimation attempt. Reliable estimation of parameters requires careful design of the experiment to acquire data in the required regions.

The present study offers a parameter estimation technique for stiff ODE's based on the singular perturbation approximation to the stiff system. There is an equivalence between stiff and singularly perturbed equations (Aiken and Lapidus, 1974) which is exploited in developing this technique. It is numerically very efficient, stable, and accurate. Details of the algorithm are presented here. The algorithm is applied to a problem of estimating rate constants in an enzyme-catalyzed reaction.

CONCLUSIONS AND SIGNIFICANCE

Recognition of the equivalence of stiff and singularly perturbed equations makes available new techniques for parameter estimation in stiff ODE's. In this study, use is made of singular perturbation approximation to develop a sequential estimation procedure for parameter estimation in stiff ODE's. This technique is numerically very efficient, stable, and accurate. A unique feature is the ability to link model detail to experimental detail and the decoupling of the estimation process into stiff and nonstiff parts. This approach is applicable to numerous problems in chemical engineering and other disciplines.

PARAMETER ESTIMATION IN ODE's

Consider the exact mathematical model:

$$\frac{dy}{d\tau} = h(y, \tau, k),\ y\ (\tau_0). \tag{1}$$

The experiments constitute a set of consecutive $\hat{y}_\mu$ observations as times passes $(\tau_\mu / \mu = 1,2, \ldots\ldots\ldots, n)$. The estimation problem consists of finding k, defined as k^*, that when used in Eq. 1 gives agreement with the noisy experimental data, the best in some sense.

For a known k, we can compute the residual

$$e_\mu = \hat{y}_\mu - y. \tag{2}$$

We choose the k^* that minimizes some function of the residual, the objective function. The most widely used objective function is the weighted least squares:

$$\Phi(\mathbf{e}_\mu) = \Phi(k) = \sum_{a=1}^{m}\sum_{\mu=1}^{n} \omega_{a\mu}\{e_{\mu a}(k)\}^2. \tag{3}$$

The minimization algorithms for analytical objective functions (Bard, 1974) still apply with increasing computations at each stage. Each evaluation of the objective function requires the numerical integration of Eq. 1 for the prescribed k.

Minimization methods requiring the gradient of the objective function $g = \partial\Phi/\partial k$, are generally more reliable (Bard, 1970). These methods require the evaluation of g at each iteration. This is done either by computing the finite difference approximation to g at each step or by setting up and numerically solving the sensi-

tivity equations for the system along with the model equations to yield the gradient. The computational effort involved in the two methods is roughly the same. Greater accuracy is obtained by using the sensitivity equations, but much more effort is required to prepare a problem.

If the original model equations are stiff, then the sensitivity equations are also stiff because

$$\frac{\partial\left[\frac{d}{dt}\left(\frac{\partial y}{\partial k}\right)\right]}{\partial\left(\frac{\partial y}{\partial k}\right)} = \frac{\partial h}{\partial y},$$

i.e., the model equations and the sensitivity equations have the same Jacobian (Gear, 1971), so they share the same eigenvalues.

ENZYME KINETIC MODEL

The stiff ODE system used to develop and illustrate the sequential parameter estimation procedure is the simplest enzymatic reaction:

$$s^* + e^* \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} c^* \xrightarrow{k_2} p^* + e^*.$$

The experimental conditions at the start of the reaction ($t^* = 0$) are $s^* = s^*(o)$, $e^* = e^*(o)$, and $c^* = p^* = o$. Two independent equations are required to model the system (Heineken et al., 1967):

$$\frac{ds^*}{dt^*} = -k_1 e^*(o) s^* + (k_1 s^* + k_{-1})\, c^* \tag{4a}$$

$$\frac{dc^*}{dt^*} = k_1 e^*(o) s^* - (k_1 s^* + k_{-1} + k_2)\, c^*. \tag{4b}$$

Two commonly occurring, independent conditions make Eq. 4 stiff. In almost all enzyme kinetic experiments the substrate is, at least initially, much more plentiful than the enzyme, i.e., $e^*(o) \ll s^*(o)$, and the product formation reaction is relatively slower than the other reactions. The former condition is predominant in most practical systems. Both these conditions lead to the Michaelis-Menten (MM) approximation (Srinivasan, 1983). Other conditions on the kinetic rate constants lead to simpler approximations, very different from the MM approximation. The use of the MM approximation in such situations leads to misleading parameter estimates (Srinivasan and Aiken, 1984b).

SINGULAR PERTURBATION APPROXIMATION TO THE ENZYME KINETIC MODEL

Following the approach of Lin and Segel (1974), which is based on the treatment of Heineken et al. (1967), Eq. 4 can be written in dimensionless form as:

$$\frac{ds}{dt} = -s + (s + \varkappa - \lambda)\, c, \tag{5a}$$

$$\epsilon \frac{dc}{dt} = s - (s + \varkappa)\, c, \tag{5b}$$

$$s(o) = 1,\ c(o) = o, \tag{5c}$$

where

$$s = s^*/s^*(o),\ c = c^*/e^*(o),\ \text{and}\ t = k_1 e^*(o) t^*.$$

The zero-order uniform approximation to Eq. 5 is:

$$\frac{ds_0}{dt} = \frac{-\lambda s_0}{\varkappa + s_0}, \tag{6a}$$

$$c_0 = \frac{s_0}{s_0 + \varkappa} - \frac{\exp\{-(\varkappa + 1)t/\epsilon\}}{(\varkappa + 1)}. \tag{6b}$$

The uniform first-order approximation for s is:

$$s_1 = s_0 + s' + \frac{(1 + \varkappa - \lambda)\exp\{-(\varkappa + 1)t/\epsilon\}}{(\varkappa + 1)}, \tag{6c}$$

where

$$s' = \frac{s_0}{s_0 + \varkappa}\left[\frac{\varkappa - \lambda}{\varkappa} \ln\left\{\frac{\varkappa + s_0}{(1 + \varkappa)s_0}\right\} - \frac{\varkappa - \lambda + s_0}{\varkappa + s_0}\right].$$

The above expression for s (and c) is an asymptotic approximation to the exact solution $s^{(\epsilon)}$ (and $c^{(\epsilon)}$) in the sense that $|s^{(\epsilon)}(t,\epsilon) - s(t,\epsilon)| < a\epsilon^2$ for $o < t < T$, for sufficiently small ϵ. The quantity a is independent of t and ϵ, but it may depend on T.

Equation 6a is the dimensionless form of the Michaelis-Menten approximation:

$$\frac{ds^*}{dt^*} = \frac{-k_2 e^*(o) s^*}{K_M + s^*}. \tag{7}$$

STIFF ESTIMATION FEATURES

Stiff estimation is defined informally as estimation of parameters in stiff ODE's. The variables in a stiff system that display sharp transients are called stiff variables. The parameters that are directly involved in producing sharp transients or large absolute-value eigenvalues are defined as stiff parameters. In most practical systems this definition is obvious. In the enzyme kinetic system that we consider here, c^* is the stiff variable, k_1 and k_{-1} are the stiff parameters.

The singular perturbation approximation simplifies the model from a stiff ODE set to one nonstiff differential equation (6a) and two algebraic equations (6b and 6c). In an iterative parameter estimation procedure this leads to significant savings on computation time. Equation 6 also sheds some light on data requirements for reliable parameter estimation. The exponential terms in Eqs. 6b and 6c decay rapidly to zero after a brief initial period. Data in this transient region seems essential, as otherwise the estimation procedure will not be sensitive to the exponential terms.

Equations 6a and 6c indicate that all the parameters in the system can be estimated from substrate concentration measurements only; but this is infeasible because of stringent data accuracy requirements (Srinivasan and Aiken, 1983a). Also the singular perturbation approximation of a given order is much more accurate for the stiff variable than it is for the nonstiff variable. So it is possible to use a lower-order approximation for the stiff variable than for the nonstiff variable, thus saving additional computation. In this example, the zero-order approximation has been used for predicting c^* and the first-order approximation for predicting s^*. The complex concentration residual was scaled in the estimation procedure to bring it to the same order of magnitude as the substrate concentration residual.

DATA SIMULATION AND ESTIMATION SOFTWARE

The concentration vs. time profile for s^* and c^* was simulated by solving Eq. 4 with LSODE (Hindmarsh, 1980). Table 1 lists the exact rate constant values used for data simulation. Table 2 in-

TABLE 1. EXACT RATE CONSTANT VALUES

Parameter	Exact Value
k_1	1.0E8
k_{-1}	1.0E3
k_2	1.0E1

TABLE 2. DATA SETS SIMULATED

Index	$\epsilon = e^*(o)/s^*(o)$	Noise
A	1.0E−3	none
B	1.0E−3	up to 5%
C	1.0E−3	up to 10%

dexes the data sets simulated. Reference to data sets in subsequent sections is consistent with this indexing scheme.

A nonlinear weighted least-squares regression program from the BMDP statistical software package (Ralston, 1981) was used for parameter estimation. This program was linked to the very reliable stiff ODE solver LSODE by a set of subroutines (Srinivasan, 1983). All computations were performed in double precision on the University of Utah UNIVAC 1100/60 system.

RESULTS AND DISCUSSION

Use of Zero-order Singular Perturbation Approximation

All three parameters appear in Eqs. 6a and 6b, which are the zero-order singular perturbation approximation. It is possible to estimate all three rate constants using this approximation. The quality of the estimates will depend on how well the approximation models Eq. 4 and on the extent of experimental error in the data.

Table 3 shows that for accurate data and a stiff system the zero-order approximation yields excellent estimates. However, the quality of estimates for the stiff variables deteriorates rapidly as the data become noisy; this also is evident from Table 3. The estimates can be improved by using a higher-order approximation. The zero-order approximation for the stiff variable is quite accurate, as noted earlier. We have therefore used the first-order approximation (Eq. 6c) only for predicting the substrate concentration, which is the nonstiff variable.

Singular Perturbation Approximation vs. Stiff Exact Model

For the same data set C, with identical initial guesses, Table 4 compares the two approaches. The estimates, residuals, and standard deviations are almost identical. However, the computation time for the singular perturbation approach is about 33% of that required if the stiff dynamic model is used. For larger systems, the saving in computation time is likely to be more significant. Similar results were obtained with the other data sets.

TABLE 3. RESULTS FOR ZERO-ORDER APPROXIMATION

Initial Guess	Exact	Estimate	σ
	Exact Data Set A		
1.0E7	1.0E8	9.9265E7	2.1886E5
1.0E2	1.0E3	9.9289E2	2.2923
1.0	1.0E1	1.0015E1	5.1250E−3
	Noisy Data Set C		
1.0E7	1.0E8	2.979E7	1.298E7
1.0E2	1.0E3	2.852E2	1.374E2
1.0	1.0E1	1.014E1	8.375E−1

TABLE 4. COMPARISON OF THE TWO APPROACHES USING DATA SET C

		Eq. 4	Eq. 6
Residuals at convergence		5.578E−19	5.579E−19
CPU time used, s		13.378	4.196
Estimates	k_1	1.026E8	1.026E8
	k_{-1}	1.025E3	1.026E3
	k_2	1.017E1	1.017E1
σ	k_1	3.758E7	3.767E7
	k_{-1}	3.932E2	3.941E2
	k_2	8.298E−1	8.300E−1

Transient Region Information

Irrespective of whether the dynamic model or the singular perturbation approximation is used, whether the data is exact or noisy, points in the transient region are necessary for reliable stiff parameter estimates. An attempt to estimate all three rate constants using Eq. 4 and post-transient-region data from set A yielded results that are summarized in Table 5. The initial guesses are very close to the exact value, yet the estimates for the stiff parameters are poor. More important, the estimates are misleading because the statistics seem to indicate a good fit. Table 6 shows an attempt to estimate all three rate constants using Eq. 6 and post-transient-region data from set A. The initial guesses for all three parameters are now lower than the exact values by a factor of 10. While the estimates for the stiff parameters change enormously, the residual sum of squares changes very marginally.

A stiff system is sensitive to the stiff parameters only in the transient region. In practical terms this means that outside this region one may be able to change the stiff parameters enormously without affecting the estimation procedure to any significant extent.

Initial Guesses for Stiff Systems

All the parameter estimation runs were initiated with guesses that were lower than the exact value by a factor of 10 (unless otherwise stated). Similar results were reproduced with lower starting guesses at the expense of more computing time. Any attempt to initiate the iteration with guesses for the stiff parameters slightly higher than their exact values converged to nonsensical estimates similar to the situation in Table 6. The initial guess for the nonstiff rate constant k_2 created no such problem.

A higher initial guess for the stiff parameter starts off the itera-

TABLE 5. EFFECT OF USING POST-TRANSIENT REGION DATA

	Initial	Estimate	σ
k_1	9.0E7	5.31E7	2.98E5
k_{-1}	9.0E2	5.26E2	3.01
k_2	9.0	1.00E1	1.17E−5

TABLE 6. EFFECT OF USING POST-TRANSIENT REGION DATA

Iteration No.	RSS*	k_1	k_{-1}	k_2
0	3.34266E−11	1.0000E+7	1.0000E+2	1.0000E+0
10	4.65494E−20	5.6017E+7	5.5604E+2	1.0003E+1
15	2.98500E−20	4.5075E+9	4.5538E+4	1.0003E+1
20	2.98037E−20	5.3357E+10	5.3915E+5	1.0003E+1
25	2.98027E−20	1.2741E+13	1.2874E+8	1.0003E+1
30	2.98007E−20	2.3632E+17	2.3880E+12	1.0003E+1
40	2.97487E−20	1.3864E+29	1.4010E+14	1.0003E+1

* Residual Sum of Squares.

Table 7. Nonstiff Parameter Estimates

Parameter	Run	Initial	Exact	Estimate	σ
k_2	a	1.0+00	1.00E+1	1.081E+1	3.240+00
	b	1.0+02	1.00E+1	1.081E+1	3.242+00
K_M	a	1.0−02	1.01E−5	1.135E−5	5.292−06
	b	1.0+00	1.01E−5	1.135E−5	5.294−06

tion with a stiffer system; the transient region for such a system is narrower. So the initial guesses and the data could correspond to a stiffer system with post-transient-region data.

Sequential Estimation Procedure

The singular perturbation approach offers a sound theoretical basis for sequentially estimating parameters in a stiff system. The estimation of stiff parameters is decoupled from the nonstiff parameter estimation part. Such a two-step search, initially estimating the nonstiff parameters and subsequently the stiff parameters, does not suffer from the problems associated with initial guesses for the stiff parameters. This procedure is illustrated in Figure 1.

We now apply this sequential estimation procedure to our system using data set *B*. The first stage is the estimation of the MM parameters in Eq. 7 using data on substrate concentration alone. This part gives us an estimate for the nonstiff parameter k_2. Table 7 shows that this stage proceeds uneventfully even with arbitrary initial guesses. The estimate obtained for k_2 is 10.8. In the next stage k_2 is kept fixed at this value and k_1 and k_{-1} are estimated using Eq. 6 and data on substrate and complex concentrations. This stage converges to the same estimates for k_1 and k_{-1} whether the initial guess is higher or lower than the exact value. This is shown in Table 8. The next stage is simply a refinement of the existing estimates, which supply the initial guesses for this stage. The exact model can be used for this stage and the procedure will converge in very few iterations. Table 9 summarizes the results for the refinement stage.

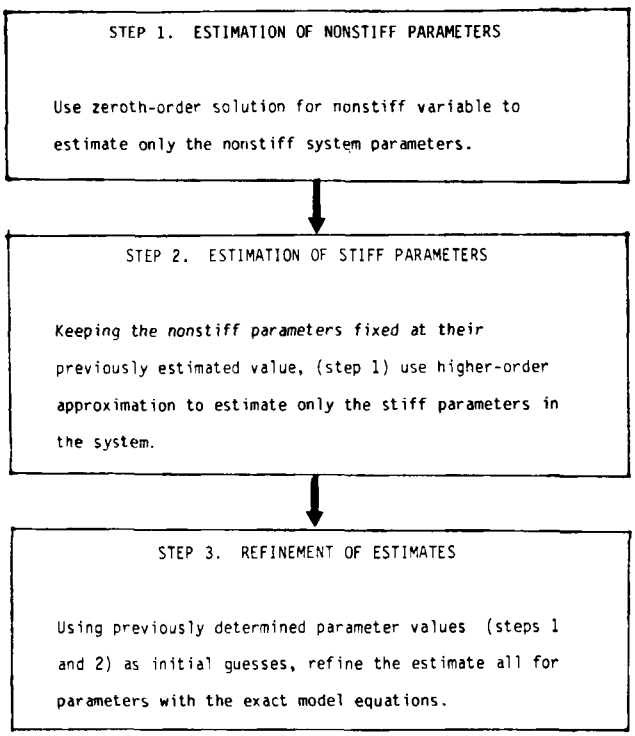

Figure 1. Sequential estimation procedure.

Table 8. Stiff Parameter Estimates

Parameter	Run	Initial	Exact	Estimate	σ
k_1	a	1.0+07	1.0+08	1.0088+08	1.787+07
	b	1.0+09			
k_{-1}	a	1.0+02	1.0+03	1.0091+03	1.870+02
	b	1.0+04			

Table 9. Refined Estimates

Parameter	k_1	k_{-1}	k_2
Initial	1.0090+08	1.0090+03	1.087+01
Estimate	1.0087+08	1.0088+03	1.007+01
σ	1.8224+07	1.9076+02	4.1427−01

The sequential estimation procedure increases model detail progressively. By decoupling the estimation procedure into two parts, instability with respect to initial guesses for stiff parameters is avoided. Substituting the singular perturbation approximation for the exact model in the early stages saves considerable computation time.

NOTATION

c = complex concentration, dimensionless
e = enzyme concentration, dimensionless
$\boldsymbol{e}$ = residual vector of dimension m
$\boldsymbol{g}$ = gradient of the objective function, dimension l
$\boldsymbol{h}$ = a general vector function of dimension m
$\boldsymbol{k}$ = vector of model parameters of dimension l
K_M = $(k_{-1} + k_2)/k_1$, the Michaelis constant
k_1, k_{-1}, k_2 = kinetic rate constants
l = number of parameters in the model
m = number of equations in the model
n = number of experimental observations
p = reaction product concentration, dimensionless
s = substrate concentration, dimensionless
t = time, dimensionless
$\boldsymbol{y}$ = vector of dependent variables of dimensions m

Greek Letters

Φ = objective function
ϵ = $e^*(o)/s^*(o)$, a small parameter
$\varkappa$ = $(k_{-1}+k_2/k_1 s^*(o)$
λ = $k_2/k_1 s^*(o)$
τ = independent variable
ω = weighting matrix for the residual
σ = standard deviation

Subscripts

0 = zero-order expansion in ϵ
1 = first-order expansion in ϵ
μ = counter for number of experimental observations

Superscripts

$*$ = dimensional variable
$\hat{}$ = observed value of the variable
(e) = exact solution

LITERATURE CITED

Aiken, R. C., and L. Lapidus, "An Effective Numerical Integration Method for Typical Stiff Systems,"*AIChE J.*, **20**, 368 (1974).

Bard, Y., "Comparison of Gradient Methods for the Solution of Nonlinear Parameter Estimation Problems," *SIAM J. Numer. Anal.*, **7**, 157 (1970).

——— , *Nonlinear Parameter Estimation*, Academic Press, New York, 83–100 (1974).

Gear, C. W., *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall, Englewood Cliffs, NJ, 228 (1971).

Heineken, F. G., H. M. Tsuchiya, and R. Aris, "On the Mathematical Status of the Pseudo-Steady State Hypothesis of Biochemical Kinetics," *Math. Biosc.*, **1**, 95 (1967).

Hindmarsh, A. C., "LSODE and LSODI, Two New Initial-Value Ordinary Differential Equation Solvers," *ACM-SIGNUM Newsletter*, **15**, 10 (1980).

Lin, C. C., and L. A. Segel, *Mathematics Applied to Deterministic Problems in the Natural Sciences*, Macmillan, New York, 308–311 (1974).

Ralston, M. L., "Derivative Free Nonlinear Regression," *BMDP Statistical Software User Manual*, W. Dixon, Ed., Dept. of Biomathematics, Univ. of California, Los Angeles (1981).

Srinivasan, V., "Analysis of Detailed Dynamic Systems," M. S. Thesis, Univ. of Utah (1983).

Srinivasan, V., and R. C. Aiken, "Singular Perturbation Approach to Stiff Estimation," *Mathematical Modelling in Science and Technology*, J. R. Avola et. al., Eds., Pergamon Press, 154–159 (1984a).

Srinivasan, V., and R. C. Aiken, "Parameter Estimation in Approximate Enzyme Kinetic Model," *Math. Biosci.*, **3**, (May, 1984b).

Manuscript received May 17, 1984, and revision received Feb. 21, 1985.