European Symposium on Computer Aided Process Engineering – 15
L. Puigjaner and A. Espuña (Editors)

Global Optimization of Ordinary Differential Equations Models

Angelo Lucia*, Meghan L. Bellows and Leah M. Octavio
Department of Chemical Engineering, University of Rhode Island, Kingston, RI 02881

Abstract

This paper contains a study of the global optimization of mathematical models described by ordinary differential equations. It is shown that recently developed terrain methods can be used with direct numerical integration to find minima, saddle points and other important problem solving information for ordinary differential equations models. Necessary partial derivative information can be computed as a by-product of the integration. A parameter estimation example for a power law rate model for hydrogen production in fuel cell applications is used to illustrate the optimization methodology.

Keywords: Global optimization, integral path methods, sensitivity analysis

1. Introduction

The global optimization of models described by ordinary differential equations is an important and challenging task. Typical examples in which optimization of differential models are important include process control and parameter estimation problems (Beigler, 1990; Vassiliadis & Floudas, 1997). In most applications these models are generally comprised of nonlinear differential equations (or differential/algebraic systems) and are sufficiently complicated to make the direct use of Newton-like methods for optimization challenging. That is, first and second partial derivatives of the objective function with respect to the unknown variables are often embedded within the results of the evolution of the model dynamics and, as a result, are difficult to compute. Moreover these derivatives can be subject to a number of inaccuracies including those from integration error, rounding, and truncation.

2. Problem Statement

The problem we are interested in is finding the

$$\text{global min } \phi[z(p), p] \text{ subject to } z' = f[z, p, t] \text{ on } D \text{ for } 0 \leq t \leq T \qquad (1)$$

where the objective function, ϕ, and the right hand side of the differential equations, f, are three times continuously differentiable with respect to the unknown variables, p, and the dependent variables in the differential equations, z. The independent variable in the differential equations, t, can be time or distance and is normally subject to lower and upper limits of integration. We assume that the number of differential equations is m,

* Author to whom correspondence should be addressed: lucia@egr.uri.edu

that the number of unknown variables is n, and that the domain, D, is defined by bounds on the z and/or p variables. It is important to understand that our computational goals go beyond just finding the global minimum. We are actually interested in finding all physically relevant minima and saddle points of Eq. 1. One way to find minima, saddles and singular points is to use the recently developed terrain (or integral path) methods of Lucia and co-workers (see, Lucia and Yang, 2003 or Lucia, DiMaggio and Depa, 2004). Terrain methods are a class of methods that move from one stationary point to another by following valleys and/or ridges and to do this they require first and second partial derivative information. However, when the constraints are given by differential equations, partial derivative information is difficult to obtain analytically because it is coupled to the integration process. Thus the accuracy of partial derivatives can be strongly affected by the integration process (i.e., integration step size, rounding error and truncation).

3. Partial Derivative Information

The partial derivative information that we require for global optimization is $g_p\phi$ and $g_{pp}{}^2\phi$, which are given by

$$g_p\phi = J_z{}^T g_z\phi + g_p\phi \tag{2}$$

$$g_{pp}{}^2\phi = J_z J_z{}^T + A + g_{pp}{}^2\phi \tag{3}$$

where J_z is the (n x m) Jacobian matrix of the dependent variables in the differential equations with respect to the unknown optimization variables, and where $J_z J_z{}^T$, $g_{pp}{}^2\phi$ and A are (n x n) matrices that represent appropriate parts of the Hessian matrix. The notation used in Eqs. 2 and 3 is intended to reflect the recursive nature of derivatives in dynamical situations and not meant to be confusing. In many practical applications, ϕ is not an explicit function of the unknown variables and therefore the last term in Eqs. 2 and 3 are often zero. Note that the Jacobian matrix, J_z, depends on the right hand side of the differential equations in Eq. 1. However, in practice, J_z will also depend on the algorithm selected for integration (i.e., forward Euler, Runge-Kutta, Gear's method, etc.) and the integration step size. To see this, note that

$$z(t) = \int_0^T f[z, p, t]\, dt \tag{4}$$

but that the trajectory z(t) is not known explicitly. Therefore, numerical integration plays an important role in defining z(t) and, as a result, impacts J_z. To illustrate the point, consider simple forward Euler integration. Application of the implicit function theorem gives the recursion relationship

$$J_z = [I + hG_z]J_z + hG_p \tag{5}$$

where h is the integration step size and G_z and G_p are the (m x m) and (m x n) Jacobian matrices of f[z(p), p] with respect to z and p respectively. Equation 5 clearly shows that the Jacobian matrix, J_z, at any value of the independent variable, t, is a function of the trajectory, z(t), the integration algorithm, and the step size. Thus if there is integration error, the accuracy of J_z can deteriorate over the integration process. Moreover, the term $G_z J_z$ can also introduce error if the variables in the differential equations are poorly scaled, even if the integration is accurate. On balance, however, Eq. 5 does show that given an integration algorithm and good step size control, accurate derivative information for use in optimization can be computed as a by-product of integration.

4. Numerical Example

Consider a parameter estimation problem for determining the reaction rate constants for a fuel cell application. In particular, let the reaction system be the dry reforming and reverse water-gas shift reactions for the production of hydrogen from methane given by

$$CH_4 + CO_2 = 2CO + 2H_2 \quad (6)$$
$$CO_2 + H_2 = CO + H_2O \quad (7)$$

The objectives of this study were to

1) Determine optimal parameters for a proposed model under isothermal conditions by minimizing the chi-squared error between the model-predicted and experimental conversions for CH_4 and CO_2,
2) Repeat the optimization for several temperatures to obtain Arrhenius plots, and
3) Validate the model from the Arrhenius plots.

4.1 Rate Model

The rate model in this example was defined by power law expressions in terms of the partial pressures of the species involved in the reactions and given by

$$r_{RF} = k_{RF}P^{a}_{CH4}\,P^{b}_{CO2}[1 - (P^{2}_{CO}P^{2}_{H2}/\,K_{EQRF}P_{CH4}P_{CO2})] \quad (8)$$
$$r_{RWGS} = k_{RWGS}P^{c}_{H2}P^{d}_{CO2}[1 - (P_{CO}P_{H2O}/K_{EQRWGS}P_{H2}P_{CO2})] \quad (9)$$

where P_i denotes the partial pressure of the ith chemical species. K_{EQRF} and K_{EQRWGS} are the equilibrium constants for reforming and reverse water-gas shift and given by

$$K_{EQRF} = \exp[-28657.2/T + 6.091\ln(T) - 0.008168T/2 + 2.164\times10^{-6}\,T^2/6 + 126100/(2T^2) - 6.88612] \quad (10)$$

$$K_{EQRWGS} = \exp[-5872.37/T - 1.86\ln(T) + 0.00054T/2 + 116400/(2T^2) + 18.0133] \quad (11)$$

4.2 Differential Equations Model

The set of ordinary differential equations that describe the change in the component flow rates as a function of plug flow reactor length are given by

$$dF_{CH4}/dx = -(L_{cat}/r_L)r_{RF} \quad (12)$$
$$dF_{CO2}/dx = -(L_{cat}/r_L)[r_{RF} + r_{RWGS}] \quad (13)$$
$$dF_{CO}/dx = (L_{cat}/r_L)[2r_{RF} + r_{RWGS}] \quad (14)$$
$$dF_{H2}/dx = (L_{cat}/r_L)[2r_{RF} - r_{RWGS}] \quad (15)$$
$$dF_{H2O}/dx = (L_{cat}/r_L)r_{RWGS} \quad (16)$$

where L_{cat} is the catalyst loading, r_L is the reactor length, and x is axial distance along the reactor. Partial pressures are related to species flow rates by $P_i = [F_i/\Sigma F_j + F_{Ar}]P$, where F_{Ar} is argon flow and P is the system pressure. Here, $L_{cat} = 0.030$ or 0.015 g of Rh/Al_2O_3 /cm of catalyst bed, P varied from 1 to 20 atm, the reactor length was $r_L = 0.01$ cm, and initial conditions (or inlet flows) ranged from

$$1.02\times10^{-5}\ \text{mol s}^{-1} \le F_{Ar}(0) \le 2.05\times10^{-5} \le F_{CH4}(0),\ F_{CO2}(0) \le 2.73\times10^{-5}\ \text{mol s}^{-1} \quad (17)$$

$$3.41\times10^{-6}\ \text{mol s}^{-1} \le F_{CO}(0),\ F_{H2}(0) \le 6.82\times10^{-6}\ \text{mol s}^{-1} \quad (18)$$

Forward Euler integration with a fixed step size of $h = 10^{-5}$ cm. was used because it gave accurate results and required far less work than other integrators.

4.3 Bounds on Variables

The rate constants k_{RF} and k_{RWGS} and the exponents a, b, c, and d in the power law expressions are the optimization variables and subject to the bounds

$0.000001 \text{ mol s}^{-1} \text{ g}^{-1}_{cat} \leq k_{RF}, k_{RWGS} \leq 0.1 \text{ mol s}^{-1} \text{ g}^{-1}_{cat}$ and $-2 \leq a, b, c, d \leq 2$ (19)

Bounds on the dependent variables were

$0 < F_{CH4} \leq F_{CH4}(0); 0 < F_{CO2} \leq F_{CO2}(0); 0 < F_{CO} \leq F_{CO}(0) + 2F_{CH4}(0) + F_{CO2}(0)$ (20)

$0 < F_{H2} \leq 2F_{CH4}(0); 0 < F_{H2O} \leq F_{CO2}(0)$ (21)

4.4 Objective Function

The objective function used in the optimization studies was the chi-squared function

$$\chi^2 = \Sigma [X_{i,k}^{exp}(T) - X_{i,k}^{calc}(T)]^2 / X_{i,k}^{calc}(T) \quad \text{for } i = CH_4, CO_2 \qquad (22)$$

where X is conversion, k is an index for the experimental data points and i is either methane or carbon dioxide. The experimental data can be found in Lee et al. (2004).

4.5 Partial Derivatives

The derivatives $(\partial\chi^2/\partial p_i)$ and $(\partial^2\chi^2/\partial p_i \partial p_j)$ were computed from Eqs. 2, 3, and 5 and the partial derivatives of the right hand sides of Eqs. 12 through 16 with respect to the F_i's and with respect to k_{RF}, $k_{RWGS,}$ a, b, c, and d.

4.6 Numerical Results

Here we present results for the optimal parameter estimation of the reactor model using the terrain method of Lucia and co-workers (2003, 2004).

4.6.1 Two-Parameter Optimizations

Table 1 gives a summary of the numerical results for some initial two-parameter optimization studies performed with respect to k_{RF} and k_{RWGS}. In this case, the exponents in the rate expressions were fixed at a = 1, b = 0, c = 1, and d = 1.

Table 1. Two-parameter optimization of a plug flow reactor model

Parameter		Temperature			
		600 C	650 C	700 C	750 C
Global Min					
	k_{RF}	7.54468E-04	1.20973E-03	2.25248E-03	4.22166E-03
	k_{RWGS}	8.27328E-05	3.05630E-03	4.05462E-03	3.31441E-02
	χ^2	1.15142E+00	8.10080E-01	7.85881E-01	4.53196E-01
Saddle Point					
	k_{RF}	7.11976E-04	1.44883E-03	2.69970E-03	
	k_{RWGS}	4.14636E-04	4.02310E-04	4.20091E-04	
	χ^2	1.31652E+00	9.42163E-01	9.16545E-01	
Local Min					
	k_{RF}	5.78257E-04	1.50823E-03	2.79656E-03	
	k_{RWGS}	2.55940E-03	8.39438E-05	9.03480E-05	
	χ^2	1.18697E+00	8.41705E-01	8.36243E-01	

Note that there are multiple stationary points to three of the two-parameter optimizations. Figure 1 shows the geometry of the surface as a function of the two rate constants for 650 C. Note that the χ^2 surface has a reasonably narrow valley very close to the boundary of the feasible region. This is typical for these problems.

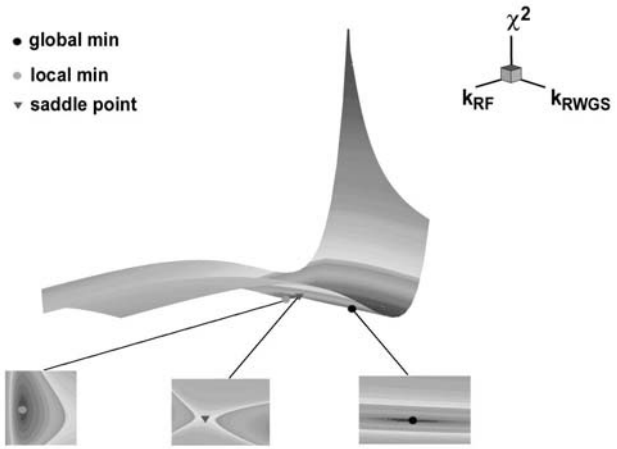

Figure 1. Two-parameter χ^2 surface for a power law kinetics model at 650 C

4.6.2 Six-Parameter Optimizations

We also performed optimization calculations in which all six parameters were permitted to vary. These numerical results are shown in Table 2.

Table 2. Six-parameter optimization of a plug flow reactor model

Parameter	Temperature			
	600 C	650 C	700 C	750 C
k_{RF}	1.10680E-04	2.36939E-04	3.70103E-04	5.76126E-04
k_{RWGS}	7.20348E-05	1.08613E-04	1.19161E-04	2.05944E-04
a	-0.350805	-0.265953	-0.205533	-0.107881
b	0.0372041	0.112096	0.0943085	0.0975606
c	-0.777327	-0.728368	-1.18516	-0.713833
d	1.23106	1.05197	1.66133	1.18313
χ^2	0.166254	0.156352	0.185106	0.194572

The six-parameter optimizations resulted in significantly better χ^2 fits of the experimental data than the two-parameter optimizations did. Also the multiplicity in the stationary points that appeared in the two-parameter optimizations completely disappeared in the six-parameter optimizations. Figure 2 shows a comparison of experimental versus calculated conversions for 700 C for the six-parameter case. Note

that the fit of the data is quite good with a correlation coefficient of 0.97. Similar good agreement was obtained for all temperatures.

4.7 Model Validation

It is also important to validate the model using Arrhenius plots in order to verify the temperature functionality of the reaction parameters. In this case, we were interested in the temperature behavior of the rate constants, k_{RF} and k_{RWGS}. Figure 3 shows Arrhenius plots for the calculated values of k_{RF} and k_{RWGS} in Table 2. Note that k_{RF} and k_{RWGS} increase with increasing temperature and that the calculated results fall on or very close to the lines in Fig. 3, suggesting that the optimization results correctly capture the temperature behavior of the rate constants.

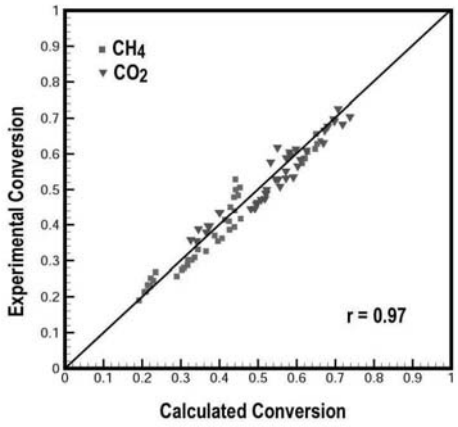

Figure 2. Optimal fit of conversion data at 700 C

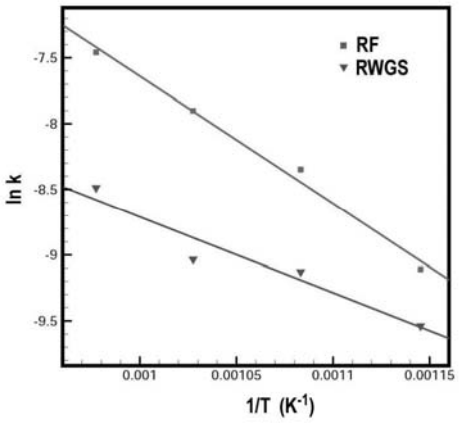

Figure 3. Arrhenius plots of rate constants

5. Conclusions

Direct numerical integration coupled with the terrain method for the global optimization was applied to a reactor model described by ordinary differential equations. Numerical results clearly show that the terrain method provides a reliable means of global optimization of differential equation models.

References

Biegler, L.T., 1990, Strategies for simultaneous solution and optimization of differential-algebraic systems. In Foundations of Computer Aided Process Design, J.J. Siirola, I.E. Grossmann and G. Stephanopoulos, Eds. Elsevier, New York.

Lee, D., S.T. Oyama and A. Lucia, 2004, Kinetics of the dry reforming of methane on a Rh/Al2O3 catalyst, manuscript in preparation.

Lucia, A. and F. Yang, 2003, Multivariable terrain methods, AIChE J.,49, 2553.

Lucia, A., P.A. DiMaggio and P. Depa, 2004, A geometric terrain methodology for global optimization, J. Global Optim., 29, 297.

Vassiliadis, V.S. and C.A. Floudas, 1997, The modified barrier function approach for large-scale optimization, Comput. Chem. Engng., 8, 855.

Acknowledgements

The authors would like to thank the National Science Foundation for financial support of this work under Grant No. CTS-0113091.