INT. J. SYSTEMS SCI., 1979, VOL. 10, NO. 4, 395–407

On a new approach to parameter estimation by the method of sensitivity functions

V. VEMURI† and A. RAEFSKY†

This paper investigates the advantages of combining the sensitivity analysis method of parameter estimation with a new computational method for the solution of systems of ordinary differential equations. It is shown that the new method allows one to take advantage of the fact that the sensitivity equations have the same structure as the model equations.

1. Introduction

The general identification problem, in its wider sense, includes all techniques devised to determine and characterize a model from measurements performed on signals entering and leaving a system and from possible additional knowledge of the structure and behaviour of the system. In a restricted sense, identification is the process of finding the unknown characteristics, say the parameters, of a system from measured values of input–output data. In this restricted case the problem is a parameter estimation problem.

The theory of sensitivity functions and its applicability to the parameter identification or estimation problem has been established and used with success for over a decade (Tomovic and Vukobratovic 1972). The purpose of this paper is to investigate the advantages of combining the sensitivity analysis method with a new computational method for the solution of systems of ordinary differential equations (Raefsky and Vemuri 1978). The resulting new approach appears to have some advantages in solving the parameter identification problem.

2. Approximation of ordinary differential equations

A first step in developing a numerical approximation to the solution of systems of ordinary differential equations is to note that any differential equation is an algebraic combination of differentiation operators and functions. Therefore if one could find a numerical representation of these operators and functions, one would then be able to proceed in the development of an approximation of the differential equation. The approximation which we shall develop yields a rectangular matrix equation, which can then be solved with standard routines to find its general solution (Stewart 1973). The general solution of the matrix equation will correspond to the general solution of the given differential equation (Eisemann 1973).

The operator d/dx is the most basic operator in an ordinary differential equation. To develop a rectangular matrix analogue to this operator, we

---

Received 20 October 1977 ; revision received 28 March 1978.
† Department of Computer Science, State University of New York at Binghamton, Binghamton, New York 13901, U.S.A.

follow a procedure that was first suggested by Csendes *et al.* (1973). Toward this end, let

$$y(x) = \sum_{i=0}^{n} y_i^{(n)} l_i^{(n)}(x) \tag{1}$$

where the $l_i^{(n)}$ are the nth order Lagrange interpolation polynomials defined on the interval

$$I = [x_0^{(n)}, x_1^{(n)}, x_2^{(n)}, \ldots, x_n^{(n)}] \triangleq \{x^{(n)}\} \quad \text{and} \quad h = x_n^{(n)} - x_0^{(n)}$$

and y_i is the value of $y(x)$ at $x = x_i$.

Now, using eqn. (1), we see

$$\frac{d}{dx}(y(x)) = \frac{d}{dx} \sum_{k=0}^{n} y_k^{(n)} l_k^{(n)}(x) \tag{2}$$

Alternatively, one can regard dy/dx as another function $g(x)$ and approximate it as

$$\frac{dy}{dx} = g(x) = \sum_{j=0}^{n-1} g_j^{(n-1)} l_j^{(n-1)}(x) = \sum_{j=0}^{n-1} (y_j^{(n-1)})' l_j^{(n-1)}(x) \tag{3}$$

where $l_j^{(n-1)}(x)$ are the $(n-1)$th order Lagrange interpolation polynomials, and

$$(y_j^{(n-1)})' = \frac{d}{dx}(yx_j^{(n-1)})$$

Notice the set of points $\{x^{(n)}\}$ need not coincide with the set $\{x^{(n-1)}\}$.

Equations (2) and (3) imply that

$$\sum_{k=0}^{n} y_k^{(n)} \frac{dl_k^{(n)}(x)}{dx} = \sum_{i=0}^{n-1} g_i^{(n-1)} l_i^{(n-1)}(x) \tag{4}$$

Since the Lagrangian polynomials have the property, that (Ralston, 1965)

$$l_i^{(n)}(x_j^{(n)}) = \delta_{ij} \tag{5}$$

where $\{x_j^{(n)}\}$, $j = 0, 1, 2, \ldots, n$ are the set of $(n+1)$ points the nth order Lagrange polynomials interpolate on. Therefore, eqns. (4) and (5) give

$$g_i^{(n-1)} = (y_i^{(n-1)})' = \sum_{k=0}^{n} y_k^{(n)} \frac{d}{dx} l_k^{(n)} x_i^{(n-1)}, \quad i = 0, 1, \ldots, n-1 \tag{6 a}$$

or in vector-matrix notation

$$\mathbf{g} = \mathbf{y}' = \mathbf{D}^{(n)}\mathbf{y} \tag{6 b}$$

where

$$\mathbf{y}' = (y'_0, y'_1, \ldots, y'_{n-1})^{\mathrm{T}}$$
$$\mathbf{y} = (y_0, y_1, \ldots, y_n)^{\mathrm{T}}$$

and $\mathbf{D}^{(n)}$ is a n by $n+1$ matrix with elements

$$d^{(n)}_{i+1,j+1} = \frac{d}{dx} l_j^{(n)}(x_i^{(n-1)}), \quad i = 0, 1, \ldots, n-1, \quad j = 0, 1, \ldots, n \tag{7}$$

The matrices $\mathbf{D}^{(n)}$ are called *differentiation matrices* of order n. Thus, the derivative of $y(x)$ at the discrete set of points $x_j^{(n-1)}$, $j=0, 1, \ldots, n-1$, is obtained by using the $\mathbf{D}^{(n)}$ as operators on the set of discrete values, $y_0, y_1, \ldots, y_n$. That is, to differentiate the vector $\mathbf{y}$ we just perform the matrix multiplication $\mathbf{D}^{(n)}\mathbf{y}$. The numerical values of the differentiation matrices have been evaluated for polynomials ranging from first to fourth order, assuming the points $\{x_i^{(n)}\}$ and $\{x_i^{(n-1)}\}$ are taken to be equi-spaced, and are presented in Raefsky and Vemuri (1978).

It will be noticed that the rank of each differentiation matrix is one less than the dimension of its domain and the row sums of all matrices are zero. This is a necessary condition if $y(x)$ is a constant, so that the result of differentiation of a constant is zero. Thus the nullspace of a $\mathbf{D}^{(n)}$ matrix corresponds to the nullspace of d/dx.

In a similar manner, if an operator contains a function of x, the function may be discretized by multiplying the vector of coefficients of $y(x)$ by a diagonal matrix containing the values of the function at the interpolation points.

An additional type of matrix is needed to maintain algebraic consistency in the approximation of ordinary differential equation. The need arises whenever two expressions of differing order are to be added. Consequently, a projection matrix is defined to map higher order polynomials into lower order ones so that the least squares norm is a minimum. These matrices shall be called the *projection matrices*, $\mathbf{P}^{(n)}$ (Raefsky and Vemuri 1978).

For example the differential equation

$$\frac{dy}{dx}+xy=x, \quad 0 \leqslant x \leqslant 1 \tag{8}$$

can be discretized by letting $y(x)$ be approximated by, say a quadratic polynomial. Then eqn. (8) reduces to

$$\mathbf{D}^{(2)}\mathbf{y}+\mathbf{P}^{(2)}\mathbf{x}\mathbf{y}=\mathbf{x} \tag{9}$$

or

$$\mathbf{A}\mathbf{y}=\mathbf{x} \tag{10}$$

Equation (10) is a rectangular matrix equation, which can be solved by generalized matrix inversion. The general solution of eqn. (10) will correspond to the general solution of eqn. (8). The method of using this approach in conjunction with sensitivity analysis will be demonstrated in § 4.

3. Sensitivity analysis

Consider a dynamic system represented by a differential equation

$$\dot{\mathbf{x}}=\mathbf{f}(\mathbf{x}, t, \lambda), \quad \mathbf{x}(0)=\mathbf{x}_0 \tag{11}$$

where $\mathbf{f}=(f_1, f_2, \ldots, f_n)^{\mathrm{T}}$ may characterize an nth order system of ordinary differential equations, $x=(x_1, x_2, \ldots, x_n)^{\mathrm{T}}$, the state vector of order n and $\lambda=(\lambda_0, \lambda_1, \ldots, \lambda_n)^{\mathrm{T}}$, is a parameter vector.

To facilitate the definition of a sensitivity coefficient (Graupe 1972), let a simplified version of eqn. (11) be

$$\dot{x}=f(x, t, \lambda)\ ; \quad x(0)=x_0 \tag{12}$$

We are interested in determining the sensitivity of x, the solution of (12), to a perturbation in the value of the parameter λ. Differentiating (12) with respect to λ, we find

$$\frac{\partial}{\partial \lambda}\left(\frac{\partial \dot{x}}{\partial \lambda}\right)=\frac{\partial f}{\partial x}\frac{\partial x}{\partial \lambda}+\frac{\partial f}{\partial \lambda} \tag{13}$$

Equation (13) is valid only if λ is independent of t. If x is continuous and differentiable in both t and λ, the order of differentiation in (13) can be interchanged to yield

$$\frac{\partial}{\partial t}\left(\frac{\partial x}{\partial \lambda}\right)=\frac{\partial f}{\partial x}\frac{\partial x}{\partial \lambda}+\frac{\partial f}{\partial \lambda} \tag{14}$$

Defining a sensitivity function $u(x, t, \lambda)$ as

$$u=u(x, t, \lambda)=\frac{\partial x(t, \lambda)}{\partial \lambda} \tag{15}$$

eqn. (14) becomes

$$\dot{u}-\left(\frac{\partial f}{\partial x}\right)u=\frac{\partial f}{\partial \lambda}\ ; \quad u(0)=u_0=0 \tag{16}$$

Equation (16) is called the sensitivity equation. In general, if there are m parameters, as in (11), there will be m sensitivity equations whose solutions correspond to the m sensitivity functions. The sensitivity equation plays an important role in the study of dynamic systems, and it is useful to list some of its important properties (Tomovic and Vukobratovic 1972) :

(1) all the sensitivity equations are of the same order as the original equation ;

(2) the sensitivity equations are always linear, regardless of the linearity or non-linearity of the original system ;

(3) if the original equation is linear, then the structure of the sensitivity equation is identical to the original system ;

(4) if the original equation is non-linear, then the structure of the sensitivity equation is identical to those equations generated by the quasi-linearization process (Bellman and Kalaba 1965).

4. The sensitivity equation in parameter identification

To develop the parameter identification algorithm, which is the main concern of this paper, we define the dynamic system of n first order differential equations, with m parameters $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_m$ to be

$$\dot{\mathbf{x}}=\mathbf{f}(\mathbf{x}, t, \lambda) \tag{17}$$

with boundary conditions $\mathbf{x}(0)=\mathbf{x}_0$, $\mathbf{x}(1)=\mathbf{x}_e$.

We shall assume the form of the vector functions $\mathbf{f}$, is known. The idea is to construct a model, characterized by

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}, t, \mathbf{p}) \tag{18}$$

such that the model response $\mathbf{y}$ and the dynamic response $\mathbf{x}$ of the actual system are close to each other in some acceptable sense. In eqn. (18) the vector

$$\mathbf{p} = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_m)^{\mathrm{T}} \tag{19}$$

represents the model parameters, as opposed to $\boldsymbol{\lambda}$, which represents the unknown system parameters.

If we define the error vector $\mathbf{e}(t, \mathbf{p})$ as

$$\mathbf{e}(t, p) = \mathbf{y}(t, \mathbf{p}) - \mathbf{x}(t) \tag{20}$$

then a convenient measure of the closeness of $\mathbf{y}(t, \mathbf{p})$ to $\mathbf{x}(t)$ is to minimize

$$J(\mathbf{p}) = \int_{t_0}^{t_e} \mathbf{e}^{\mathrm{T}}\mathbf{e}\, dt \tag{21}$$

In eqn. (21), the argument under the integral sign is a scalar-valued function of time, and J is a scalar function. The value of J is a function of the parameters $\mathbf{p}$. Therefore, it can be minimized by adjusting the parameter $\mathbf{p}$. If we can determine how J varies in the vicinity of $\mathbf{p}$, then it can become possible to find a direction in which to change $\mathbf{p}$ so as to reduce J in an optimum fashion.

Computationally, this can be accomplished by the steepest descent method which attempts to modify the parameter vector by iteratively updating the vector used in the preceding step of the calculation. In other words, the value of $\mathbf{p}$ at the $(i+1)$th iteration is given in terms of its value at the ith iteration as

$$\mathbf{p}^{i+1} = \mathbf{p}^i + \Delta\mathbf{p} \tag{22}$$

where the superscript i indicates the iteration, and $\Delta\mathbf{p}$ is given by

$$\Delta\mathbf{p} = -k \operatorname{grad} J \tag{23}$$

In this method, computation of grad J is perhaps the biggest computational task. In some cases a good deal of ingenuity is required to arrive at the best method of evaluating the gradient. With this in mind, let us compute the ith component of eqn. (22).

$$\Delta p_i = -k \sum_j \frac{\partial}{\partial p_i} \int_{t_0}^{t_e} e_j^2\, dt, \quad i = 1, 2, \ldots, m \tag{24}$$

$$= -2k \sum_j \int_{t_0}^{t_e} e_j \frac{\partial e_j}{\partial p_i}\, dt \tag{25}$$

$$= -2k \sum_j \int_{t_0}^{t_e} e_j \frac{\partial y_j}{\partial p_i}\, dt \tag{26}$$

Equation (26) follows from eqn. (25) by virtue of the relation in eqn. (20).

Now, using the definition of the sensitivity function,

$$u_i = \frac{\partial y_j}{\partial p_i} \tag{27}$$

eqn. (26) can be rewritten as

$$\Delta p_i = -2k \sum_j \int_{t_0}^{t_e} e_j u_i \, dt \tag{28}$$

$i = 1, 2, \ldots, m$, the sum j is over the number of points at which the error function is evaluated. Therefore, if the sensitivity functions u_i are known, the steepest descent method described above can be implemented by computing Δp_i, using eqn. (28). The sensitivity functions u_i are evaluated by solving the sensitivity equations.

The modified sensitivity algorithm can now be expressed as follows :

(1) Assign a nominal value $\mathbf{p}^0$ to the parameter vector $\mathbf{p}$. This choice may be arbitrary and may be governed by physical considerations.

(2) With the current value of $\mathbf{p}^i$, solve the model equations, with the new method described in § 2. If the equation is linear, the process of solving the sensitivity equations by the new method is relatively easy ; the reason being that the set of sensitivity equations can be solved, with the original linear equation, by augmenting the right-hand sides of each sensitivity equation to the augmented matrix used to solve the model equation (Raefsky and Vemuri 1978). This can be done because the model equation and the sensitivity equations have the same forms and only differ in the right-hand sides.

If the equation is non-linear, the set of sensitivity equations can be solved, with the linear equation obtained through the quasi-linearization process, by augmenting the right-hand sides of each sensitivity equation to the augmented matrix used to solve the linearized model equations. This can be done because the linearized model equation and the sensitivity equations have the same forms and only differ in the right-hand side.

(3) Using a numerical integration technique, compute the gradient of the criterion function J, with respect to the parameters, i.e. compute eqn. (28).

(4) Update the parameter vector according to eqn. (22), until

$$|\mathbf{p}^{i+1} - \mathbf{p}^i| < \boldsymbol{\epsilon}$$

where $\boldsymbol{\epsilon}$ is a tolerance defined by the user.

(5) If the test in step (4) is not satisfied, go to step (2). If the above test is satisfied, take the values $\mathbf{p}^{i+1}$ as the estimate of the unknown vector.

5. Examples

Example 1. *Linear equation*

To illustrate the procedure for linear equations, consider the following differential equation

$$\left.\begin{aligned} \dot{x} &= k_{11}x + k_{12}y \\ \dot{y} &= k_{21}x + k_{22}y \end{aligned}\right\} \tag{29}$$

with the measured outputs on x and y as shown in Table 1.

The problem is to identify the parameters $k_{11}, k_{12}, k_{21}, k_{22}$ from the data in Table 1. Problems of this kind occur in studies of pharmacokinetics using compartmental models (Jacquez 1972). First, we assume the initial values of the parameters $k_{11}, k_{12}, k_{21}, k_{22}$ as being equal to zero. The time domain [0, 1] is divided into 2 elements of length, $h = 0{\cdot}5$. Equation (29) is approximated in each element by assuming

$$\left.\begin{aligned} x^{(4)}(t) &= \sum_{i=0}^{4} x_i^{(4)} l_i^{(4)}(t) \\ y^{(4)}(t) &= \sum_{i=0}^{4} y_i^{(4)} l_i^{(4)}(t) \end{aligned}\right\} \quad (30)$$

Therefore, our approximation to eqn. (29) is

$$\left.\begin{aligned} \mathbf{D}^{(4)}\mathbf{x}^{(4)} - h\mathbf{P}^{(4)}\mathbf{F}_{11}{}^{(4)}\mathbf{x}^{(4)} - h\mathbf{P}^{(4)}\mathbf{F}_{12}{}^{(4)}\mathbf{y}^{(4)} = \mathbf{0} \\ \mathbf{D}^{(4)}\mathbf{y}^{(4)} - h\mathbf{P}^{(4)}\mathbf{F}_{21}{}^{(4)}\mathbf{x}^{(4)} - h\mathbf{P}^{(4)}\mathbf{F}_{22}{}^{(4)}\mathbf{y}^{(4)} = \mathbf{0} \end{aligned}\right\} \quad (31)$$

where the matrices $\mathbf{D}^{(4)}$, $\mathbf{P}^{(4)}$ are defined in Raefsky and Vemuri (1978). The matrices $\mathbf{F}_{11}{}^{(4)}$, $\mathbf{F}_{12}{}^{(4)}$, $\mathbf{F}_{21}{}^{(4)}$, $\mathbf{F}_{22}{}^{(4)}$ are 5×5 diagonal matrices, with constant entries $k_{11}, k_{12}, k_{21}, k_{22}$ respectively on each diagonal.

Time	x	y
0·00000000D 00	0·00000000D 00	0·10000000D 01
0·12500000D 00	−0·14200846D 00	0·11420085D 01
0·25000000D 00	−0·32434973D 00	0·13243497D 01
0·37500000D 00	−0·55847901D 00	0·15584790D 01
0·50000000D 00	−0·85910497D 00	0·18591050D 01
0·62500000D 00	−0·12451138D 01	0·22451138D 01
0·75000000D 00	−0·17407556D 01	0·27407556D 01
0·87500000D 00	−0·23771682D 01	0·33771682D 01
0·10000000D 01	−0·31943326D 01	0·41943326D 01

Table 1. Measured data.

Equation (31) is rewritten in matrix form as

$$\mathbf{A}\mathbf{q}_1 = \mathbf{b}_1$$

where $\mathbf{A}$ is the matrix

$$\mathbf{A} = \left[\begin{array}{c:c} \mathbf{D}^{(4)} - h\mathbf{P}^{(4)}\mathbf{F}_{11}{}^{(4)} & -h\mathbf{P}^{(4)}\mathbf{F}_{12}{}^{(4)} \\ \hdashline -h\mathbf{P}^{(4)}\mathbf{F}_{21}{}^{(4)} & \mathbf{D}^{(4)} - h\mathbf{P}^{(4)}\mathbf{F}_{22}{}^{(4)} \end{array}\right] \quad (32)$$

$\mathbf{q}_1$ and $\mathbf{b}_1$ are the vectors $(\mathbf{x}^{(4)} : \mathbf{y}^{(4)})^{\mathrm{T}}$, $(\mathbf{0} : \mathbf{0})^{\mathrm{T}}$ respectively. Now the sensitivity equations for the parameters : $k_{11}, k_{12}, k_{21}, k_{22}$ are found to be

$$\left.\begin{aligned} \dot{u}_{11} &= k_{11}u_{11} + k_{12}u_{21} + x \\ \dot{u}_{21} &= k_{21}u_{11} + k_{22}u_{21} \\ \text{where} \quad \frac{\partial x}{\partial k_{11}} &= u_{11}, \quad \frac{\partial y}{\partial k_{11}} = u_{21} \end{aligned}\right\} \tag{33 a}$$

$$\left.\begin{aligned} \dot{u}_{12} &= k_{11}u_{12} + k_{12}u_{22} + y \\ \dot{u}_{22} &= k_{21}u_{12} + k_{22}u_{22} \\ \text{where} \quad \frac{\partial x}{\partial k_{12}} &= u_{12}, \quad \frac{\partial y}{\partial k_{12}} = u_{22} \end{aligned}\right\} \tag{33 b}$$

$$\left.\begin{aligned} \dot{u}_{13} &= k_{11}u_{13} + k_{12}u_{23} \\ \dot{u}_{23} &= k_{21}u_{13} + k_{22}u_{23} + x \\ \text{where} \quad \frac{\partial y}{\partial k_{21}} &= u_{23}, \quad \frac{\partial x}{\partial k_{21}} = u_{13} \end{aligned}\right\} \tag{33 c}$$

$$\left.\begin{aligned} \dot{u}_{14} &= k_{11}u_{14} + k_{12}u_{24} \\ \dot{u}_{24} &= k_{21}u_{14} + k_{22}u_{24} + y \\ \text{where} \quad \frac{\partial x}{\partial k_{22}} &= u_{14}, \quad \frac{\partial y}{\partial k_{22}} = u_{24} \end{aligned}\right\} \tag{33 d}$$

As can be seen from the above equations, the sensitivity equations and the original eqn. (29) have the same form except for the right-hand sides.

Therefore, each of the eqns. (33) can be approximated by the same discrete analogue as eqn. (29), i.e. the above set of equations become

$$\mathbf{A}\mathbf{q}_2 = \mathbf{b}_2 \tag{34 a}$$

where

$$\mathbf{q}_2 = (\mathbf{u}_{11}{}^{(4)} : \mathbf{u}_{21}{}^{(4)})^{\mathrm{T}}, \quad \mathbf{b}_2 = (h\mathbf{P}^{(4)}\mathbf{x}^{(4)} : \mathbf{0})^{\mathrm{T}}$$

and

$$\mathbf{A}\mathbf{q}_3 = \mathbf{b}_3 \tag{34 b}$$

where

$$\mathbf{q}_3 = (\mathbf{u}_{12}{}^{(4)} : \mathbf{u}_{22}{}^{(4)})^{\mathrm{T}}, \quad \mathbf{b}_3 = (h\mathbf{P}^{(4)}\mathbf{y}^{(4)} : \mathbf{0})^{\mathrm{T}}$$

and

$$\mathbf{A}\mathbf{q}_4 = \mathbf{b}_4 \tag{34 c}$$

where

$$\mathbf{q}_4 = (\mathbf{u}_{13}{}^{(4)} : \mathbf{u}_{23}{}^{(4)})^{\mathrm{T}}, \quad \mathbf{b}_4 = (\mathbf{0} : h\mathbf{P}^{(4)}\mathbf{x}^{(4)})^{\mathrm{T}}$$

and

$$\mathbf{A}\mathbf{q}_5 = \mathbf{b}_5 \tag{34 d}$$

where

$$\mathbf{q}_5 = (\mathbf{u}_{14}{}^{(4)} : \mathbf{u}_{24}{}^{(4)})^{\mathrm{T}}, \quad \mathbf{b}_5 = (\mathbf{0} : h\mathbf{P}^{(4)}\mathbf{y}^{(4)})^{\mathrm{T}}$$

The matrix $\mathbf{A}$ in the above equations is defined in eqn. (32).

The solutions to eqns. (31) and (34) can be found simultaneously from the augmented matrix

$$\{\mathbf{A} : \mathbf{b}_1 : \mathbf{b}_2 : \mathbf{b}_3 : \mathbf{b}_4 : \mathbf{b}_5\} \tag{35}$$

i.e. the matrix formed by augmenting the matrix $\mathbf{A}$ with each of the right-hand sides of eqns. (31) and (34).

Therefore, in the ith element the solutions to eqns. (31) and (34) are

$$\mathbf{q}_1 = \mathbf{A}^+\mathbf{b}_1 + \mathbf{N}\mathbf{z}_1 \tag{36 a}$$

$$\mathbf{q}_2 = \mathbf{A}^+\mathbf{b}_2 + \mathbf{N}\mathbf{z}_2 \tag{36 b}$$

$$\mathbf{q}_3 = \mathbf{A}^+\mathbf{b}_3 + \mathbf{N}\mathbf{z}_3 \tag{36 c}$$

$$\mathbf{q}_4 = \mathbf{A}^+\mathbf{b}_4 + \mathbf{N}\mathbf{z}_4 \tag{36 d}$$

$$\mathbf{q}_5 = \mathbf{A}^+\mathbf{b}_5 + \mathbf{N}\mathbf{z}_5 \tag{36 e}$$

where the vectors $\mathbf{q}_j$ and $\mathbf{b}_j$ ($j = 1, 2, \ldots, 5$) are the same as above. The vectors $\mathbf{z}_j$ ($j = 1, \ldots, 5$) are the vectors of arbitrary constants in the general solutions.

Now to eliminate the arbitrary constants for each of the eqns. (36) in each of the i elements, we must apply the interelement continuity conditions and the five sets of boundary conditions

$$\left.\begin{aligned} x(0) &= 0 \\ x(1) &= -0{\cdot}319\,433\,6 \end{aligned}\right\} \tag{37 a}$$

$$\left.\begin{aligned} u_{11}(0) &= 0 \\ u_{11}(1) &= 0 \end{aligned}\right\} \tag{37 b}$$

$$\left.\begin{aligned} u_{12}(0) &= 0 \\ u_{12}(1) &= 0 \end{aligned}\right\} \tag{37 c}$$

$$\left.\begin{aligned} u_{13}(0) &= 0 \\ u_{13}(1) &= 0 \end{aligned}\right\} \tag{37 d}$$

$$\left.\begin{aligned} u_{14}(0) &= 0 \\ u_{14}(1) &= 0 \end{aligned}\right\} \tag{37 e}$$

where the values of $x(0)$ and $x(1)$ in eqn. (37 a) are read directly from Table 1.

For eqn. (31), the application of interelement continuity and boundary conditions (37 a) results in the matrix equation

$$\mathbf{B}\mathbf{a}_1 = \mathbf{g}_1 \tag{38 a}$$

The matrix $\mathbf{B}$ and the vector $\mathbf{g}_1$ contain the $2n$ (where n is the number of elements in the domain) interelement continuity and boundary conditions. The vector $\mathbf{a}_1$ will be the vector of $2n$ arbitrary constants in the general solution to eqn. (29).

In a similar manner the matrix equations

$$\mathbf{B}\mathbf{a}_2 = \mathbf{g}_2 \tag{38 b}$$

$$\mathbf{B}\mathbf{a}_3 = \mathbf{g}_3 \tag{38 c}$$

$$\mathbf{B}\mathbf{a}_4 = \mathbf{g}_4 \tag{38 d}$$

$$\mathbf{B}\mathbf{a}_5 = \mathbf{g}_5 \tag{38 e}$$

can be defined for each of the sensitivity eqns. (33).

Again the important point to note is that each of the eqns. (38) involve the same matrix $\mathbf{B}$. The vectors $\mathbf{a}_j$ $(j = 2, 3, 4, 5)$ are the vectors of arbitrary constants in the general solutions to the sensitivity eqns. (33) respectively. The values of the constants can be found simultaneously from the augmented matrix

$$\{\mathbf{B} : \mathbf{g}_1 : \mathbf{g}_2 : \mathbf{g}_3 : \mathbf{g}_4 : \mathbf{g}_5\} \tag{39}$$

Once the values of the constants are found, they are eliminated from the general solutions.

Now that we have the solutions of the model equation and the sensitivity equation, we can calculate the value of the error vector (eqn. (20)) and carry out step (3) in § 4. We then update the parameter vector and test for convergence, i.e. step (4). We repeat the above procedure until, we achieve convergence for the parameter p^i, $i = 1, 2, \ldots$.

Iteration number n	Value of k_{11}	Value of k_{12}	Value of k_{21}	Value of k_{22}
0	0	0	0	0
3	0·264792	−0·164845	−0·368245	0·195728
6	0·683642	−0·792365	−0·783028	0·596736
9	0·837053	−0·916395	−0·927375	0·834672
12	0·998571	−0·999129	−1·00469	1·002716

Table 2. The values of the parameters for eqn. (29).

The values of the parameters for eqn. (29) with iterations are shown in Table 2. The exact values are

$$\left.\begin{aligned} k_{11} &= 1{\cdot}0 \\ k_{12} &= -1{\cdot}0 \\ k_{21} &= -1{\cdot}0 \\ k_{22} &= 1{\cdot}0 \end{aligned}\right\} \tag{40}$$

The solution to eqn. (29) after the 12th iteration is shown in Table 3. The C.P.U. time under an Extended H Fortran compiler on I.B.M. 370/158 system was 1·8 seconds.

Time	x	y
0·000	−0·00000000	0·99993891
0·125	−0·14201727	1·14195618
0·250	−0·32434332	1·32428223
0·375	−0·55845037	1·55838928
0·500	−0·85908838	1·55838928
0·625	−1·24513131	2·24507023
0·750	−1·74074490	2·74068381
0·875	−2·37711380	3·37705271
1·000	−3·19433260	4·19427151

Table 3. Solution of eqn. (29) after the 12th iteration.

Example 2. *A non-linear equation*

To illustrate the procedure for non-linear equations, consider Van der Pol's equation

$$\ddot{x}+\lambda(x^2-1)\dot{x}+x=0 \tag{41}$$

with measured data on x as shown on Table 4.

Time	x
0	1·0
0·5	0·8604301689
1·0	−1·448279578
1·2	−2·004112696
1·5	−1·983369609
2·0	−1·948005012
3·0	−1·878810212
4·0	−1·802578908
5·0	−1·719109739
6·0	−1·62273326
7·0	−1·521198856
8·0	−1·34134059
9·0	−1·220283685
10·0	−0·7772658874

Table 4. Observed data on x for the Van der Pol's equation.

Since, eqn. (41) is non-linear, we must linearize the equation to get the model equation, i.e. the linearized version of eqn. (41) is

$$\left.\begin{aligned} \dot{u}^{k+1}&=v^{k+1}\\ \dot{v}^{k+1}&=-(2\lambda^k v^k u^k+1)u^{k+1}-\lambda^k((u^k)^2-1)v^{k+1}+2\lambda^k v^k(u^k)^2,\\ &\qquad k=0,1,2,\ldots \end{aligned}\right\} \tag{42}$$

where the value of λ^k is obtained from eqn. (22). The sensitivity equation for λ is

$$\dot{w}_1^{k+1} = w_2^{k+1}$$

$$\dot{w}_2^{k+1} = -(2\lambda^k u^k v^k + 1) w_1^{k+1} - \lambda^k((u^k)^2 - 1) w_2^{k+1} + ((u^k)^2 - 1) v^k \tag{43}$$

where

$$w_1 = \frac{\partial u}{\partial \lambda} \quad \text{and} \quad w_2 = \frac{\partial v}{\partial \lambda}$$

Notice once again that both eqns. (42) and (43) are of the same form.

We can find the value of λ by taking eqn. (42) as our model equation for the kth iteration and eqn. (43) as our sensitivity equation for the kth iteration. We then proceed in exactly the same manner as we did for the linear problem.

The value of the parameter λ for eqn. (41) was found to be 10·00034 after seven iterations. The results with iterations are shown in Table 5. The C.P.U. time under an Extended H Fortran compiler on I.B.M. 370/158 system was 2·95 seconds.

Iterations	λ
1	0·5
2	3·64318
3	5·71842
4	8·93415
5	9·13416
6	9·99863
7	10·00034

Table 5. Convergence of λ with iterations.

6. Discussion

This paper demonstrates the feasibility of using a new computational method for systems of ordinary differential equations in combination with sensitivity analysis to estimate the parameters in systems of ordinary differential equations. The method allows one to take advantage of the fact that the sensitivity equations have the same structure as the dynamic equations. Therefore, the model equations and the sensitivity equations can be solved simultaneously, in the same solution procedure.

In general to estimate m parameters, an augmented matrix is formed from the rectangular matrix approximation to the homogeneous equation, and the $m+1$ right-hand sides of the model and sensitivity equations.

References

Bellman, R. E., and Kalaba, R. E., 1965, *Quasilinearization and Nonlinear Boundary Value Problems* (New York : American Elsevier).

Csendes, Z., Gopinath, A., and Silvester, P., 1973, *The Mathematics of Finite Elements and Applications*, edited by J. R. Whiteman (London: Academic Press), pp. 189–99.

Eisemann, K., 1973, *I.E.E.E. Trans. circuit Theory*, **20**, 481.

Graupe, D., 1972, *Identification of Systems* (New York: Van Nostrand Reinhold).

Jacquez, J. A., 1972, *Compartmental Analysis in Biology and Medicine* (New York: American Elsevier).

Raefsky, A., and Vemuri, V., 1978, *Int. J. Comput. Elect. Engng* (in the press).

Ralston, A., 1965, *A First Course in Numerical Analysis* (New York: McGraw-Hill).

Stewart, G. W., 1973, *Introduction to Matrix Computations* (New York: Academic Press).

Tomovic, R., and Vukobratovic, M., 1972, *General Sensitivity Theory* (New York: American Elsevier).