Computational Methods in Ordinary Differential Equations

J. D. LAMBERT

Computational Methods in Ordinary Differential Equations

J. D. LAMBERT

Reader in Mathematics

University of Dundee, Scotland

JOHN WILEY & SONS LONDON NEW YORK SYDNEY TORONTO

Library of Congress Catalog Card No. 72–5718

ISBN 0 471 51194 3

Reprinted September 1976

Printed in Great Britain by J. W. Arrowsmith Ltd, Bristol

Introductory Mathematics for Scientists and Engineers

Foreword to the Series

The past few years have seen a steady increase in the courses of mathematics and computing provided for students undertaking higher education. This is partly due to high speed digital computation being more readily available and partly because so many disciplines now find mathematics an essential element of their curriculum. Many of the students, e.g. those of physics, chemistry, engineering, biology and economics, will be concerned with mathematics and computing mainly as tools, but tools with which they must acquire proficiency. On the other hand, courses for mathematicians must take cognizance of the existence of electronic computers. All these students may therefore study similar material though possibly at different stages of their careers—some, perhaps, will encounter it shortly after commencing their training while others may not come to grips with it until after graduating. This series is designed to cater for these differing requirements; some of the books are appropriate to the basic mathematical training of students in many disciplines while others, dealing with more specialized topics, are intended both for those for whom such topics are an essential ingredient of their course and for those who, although not specialists, find the need for a working knowledge of these areas. However, the presentation of all the books has been planned so as to demand the minimal mathematical equipment for the topics discussed. Instructors will therefore often be able to extract a shorter introductory course when a fuller treatment is not desired.

The authors have, in general, avoided the strict axiomatic approach which is favoured by some writers, but there has been no dilution of the standard of mathematical argument. Learning to follow and construct a logical sequence of ideas is one of the important attributes of courses in mathematics and computing.

While the authors' purpose has been to stress mathematical ideas which are central to applications and necessary for subsequent investigations, they have attempted, when appropriate, to convey some notion of the connection between a mathematical model and the real world. They have also taken account of the fact that most students now have access to electronic digital computers.

The careful explanation of difficult points and the provision of large numbers of worked examples and exercises should ensure the popularity of the books in this series with students and teachers alike.

D. S. JONES
Department of Mathematics
University of Dundee

Preface

Computational methods for ordinary differential equations, although constituting one of the older established areas of numerical analysis, have been the subject of a great deal of research in recent years. It is hoped that this book will, in addition to its primary purpose of serving as a text for advanced undergraduates, provide postgraduate students and general uses of numerical analysis with a readable account of these developments. The only prerequisites required of the reader are a sound course in calculus and some acquaintance with complex numbers, matrices, and vectors.

There is no general agreement on how the phrase 'numerical analysis' should be interpreted. Some see 'analysis' as the key word, and wish to embed the subject entirely in rigorous modern analysis. To others, 'numerical' is the vital word, and the algorithm the only respectable product. In this book I have tried to take a middle course between these two extremes. On the one hand, few theorems are stated (and even fewer proved), while, on the other, no programmes will be found. The approach is rather analogous to that of the classical applied mathematician, whose genuine interest in real problems does not detract from his delight in watching mathematics at work. Thus, most of the exercises and worked examples are intended to cast light (and, in some cases, doubt) on our interpretation of why numerical methods perform in the way that they do. It is hoped that the reader will supplement such exercises by programming and running the methods, discussed in the text, when applied to specific differential equations or systems of equations, preferably arising from real problems.

Much of the material of this book is based on lecture courses given to advanced undergraduate and early postgraduate students in the Universities of St. Andrews, Aberdeen, and Dundee. Chapters 2, 3, and 4 develop the study of linear multistep methods and Runge–Kutta methods in some detail, and culminate in some of the most efficient forms of these methods currently available. These two classes form a convenient basis for the development of concepts, such as weak stability, which are widely applicable. Chapter 5 is concerned with hybrid methods—a class whose computational potentialities have probably not yet been fully exploited.

Chapter 6 deals with the highly efficient class of extrapolation methods, while chapter 7 is concerned with the somewhat neglected area of special methods for problems which exhibit special features other than stiffness. Up to this point, in the interests of ease of exposition, only the single first-order differential equation is considered. The first half of chapter 8 gives an account of precisely how much of the previous work goes over unaltered to the problem of a system of first-order equations, how much needs modification, and how much is invalid; the second half is devoted to a full account of the problem of stiffness, and examines certain connections with the theory of stability of finite difference methods for partial differential equations. The last chapter is concerned with a special class of second-order differential equations. The emphasis throughout is on initial value problems, since direct techniques for boundary value problems lead to a large and separate area of numerical analysis. An exception is made in the case of the shooting method for two-point boundary value problems, which is described in the appendix.

Any book on this topic necessarily owes a large debt to the well known book by P. Henrici, *Discrete Variable Methods in Ordinary Differential Equations*. This book, together with the celebrated papers of G. Dahlquist, has played a unique rôle in the developments of the last decade. Henrici's approach is somewhat more rigorous than that of the present book, but I have purposely adopted a notation consistent with Henrici's, in order that the reader may the more easily make the transition to Henrici's book.

Many people have influenced the development of this book. Foremost among these is Professor A. R. Mitchell, who, over many years in the successive rôles of teacher, supervisor, colleague—and always as friend—has greatly influenced my attitude to numerical analysis. It was my good fortune that, during the preparation of this book, a year-long seminar on numerical analysis was held in the University of Dundee, with the generous support of the Science Research Council. This book was much influenced by useful discussions with many of the distinguished numerical analysts who took part in that seminar. These included Professors G. Dahlquist, C. W. Gear, W. B. Gragg, M. R. Osborne, H. J. Stetter, and, in particular, Professor T. E. Hull, who suggested several important improvements in the manuscript. My thanks are also due to Professors J. C. Butcher, J. D. Lawson, Drs. A. R. Gourlay, J. J. H. Miller, and Mr. S. T. Sigurdsson for useful discussions. I am also grateful to several past and present research students in the University of Dundee, and in particular to Mr. K. M. Lovell, for help with computational examples. Finally, I am indebted to Professor D. S. Jones for his useful comments on an early draft,

and to Mrs. M. L. Bennett for her accurate typing of part of the manuscript.

J. D. Lambert
Dundee, Scotland,
February, 1972.

Contents

1

Preliminaries

1.1 Notation

Throughtout this book we shall denote scalars by y, ϕ, etc., and (column) vectors by $\mathbf{y}$, $\boldsymbol{\phi}$, etc. The components of the m-dimensional vector $\mathbf{y}$ will be denoted by ${}^i y$, $i = 1, 2, \ldots, m$; that is, we can write $\mathbf{y} = [{}^1 y, {}^2 y, \ldots, {}^m y]^T$, where T denotes the transpose.

We shall repeatedly use the notation

$$f(x) = O(\phi(x)) \quad \text{as} \quad x \to x_0,$$

which means that there exists a positive constant K such that $|f(x)| \leqslant K|\phi(x)|$ for x sufficiently close to x_0. Our most frequent use of this notation will be in the context $f(h) = O(h^p)$ as $h \to 0$, where h is the steplength associated with some numerical method. So frequently shall we write this, that it will be impracticable always to include the phrase 'as $h \to 0$', which is consequently to be taken as read. However, it is important to guard against the temptation mentally to debase the notation $f(h) = O(h^p)$ to mean '$f(h)$ is roughly the same size as h^p, *whatever the size of* h'. Apparent discrepancies between theoretical estimates and numerical results frequently stem from a failure to realize that the notation $f(h) = O(h^p)$ carries with it the implication 'as $h \to 0$'.

The closed interval $a \leqslant x \leqslant b$ will be denoted by $[a, b]$, and the open interval $a < x < b$ by (a, b). We shall occasionally use the notation $y(x) \in C^m[a, b]$ to indicate that $y(x)$ possesses m continuous derivatives for $x \in [a, b]$.

1.2 Prerequisites

Remarkably little is required by way of prerequisites for the study of computational methods for ordinary differential equations. It is assumed that the reader is familiar with the following topics; suggested references are given.

(i) Introductory analysis, including the geometry of the complex plane. (Jones and Jordan.[82])

(ii) Elementary numerical analysis, including the finite difference operators Δ, ∇, and elementary interpolation and quadrature formulae. (Morris.[135])

(iii) For the material of chapter 8, elementary properties of vector and matrix norms. (Mitchell,[134] chapter 1.)

1.3 Initial value problems for first-order ordinary differential equations

A first-order differential equation $y' = f(x, y)$ may possess an infinite number of solutions. For example, the function $y(x) = Ce^{\lambda x}$ is, for any value of the constant C, a solution of the differential equation $y' = \lambda y$, where λ is a given constant. We can pick out any particular solution by prescribing an *initial condition*, $y(a) = \eta$. For the above example, the particular solution satisfying this initial condition is easily found to be $y(x) = \eta e^{\lambda(x-a)}$. We say that the differential equation together with an initial condition constitutes an *initial value problem*,

$$y' = f(x, y), \qquad y(a) = \eta. \tag{1}$$

The following theorem, whose proof may be found in Henrici,[67] states conditions on $f(x, y)$ which guarantee the existence of a unique solution of the initial value problem (1).

Theorem 1.1 Let $f(x, y)$ be defined and continuous for all points (x, y) in the region D defined by $a \leqslant x \leqslant b, -\infty < y < \infty$, a and b finite, and let there exist a constant L such that, for every x, y, y^* such that (x, y) and (x, y^*) are both in D,

$$|f(x, y) - f(x, y^*)| \leqslant L|y - y^*|. \tag{2}$$

Then, if η is any given number, there exists a unique solution $y(x)$ of the initial value problem (1), where $y(x)$ is continuous and differentiable for all (x, y) in D.

The requirement (2) is known as a *Lipschitz condition*, and the constant L as a *Lipschitz constant*. This condition may be thought of as being intermediate between differentiability and continuity, in the sense that

$f(x, y)$ continuously differentiable w.r.t. y for all (x, y) in D

$\Rightarrow f(x, y)$ satisfies a Lipschitz condition w.r.t. y for all (x, y) in D

$\Rightarrow f(x, y)$ continuous w.r.t. y for all (x, y) in D.

In particular, if $f(x, y)$ possesses a continuous derivative with respect to y for all (x, y) in D, then, by the mean value theorem,

$$f(x, y) - f(x, y^*) = \frac{\partial f(x, \bar{y})}{\partial y}(y - y^*),$$

where $\bar{y}$ is a point in the interior of the interval whose end-points are y and y^*, and (x, y) and (x, y^*) are both in D. Clearly, (2) is then satisfied if we choose

$$L = \sup_{(x,y)\in D} \left| \frac{\partial f(x, y)}{\partial y} \right|. \tag{3}$$

1.4 Initial value problems for systems of first-order ordinary differential equations

In many applications, we are faced, not with a single differential equation, but with a system of m simultaneous first-order equations in m dependent variables ${}^1y, {}^2y, \ldots, {}^my$. If each of these variables satisfies a given condition *at the same value a of x*, then we have an initial value problem for a first-order system, which we may write

$$\begin{aligned}
{}^1y' &= {}^1f(x, {}^1y, {}^2y, \ldots, {}^my), & {}^1y(a) &= {}^1\eta,\\
{}^2y' &= {}^2f(x, {}^1y, {}^2y, \ldots, {}^my), & {}^2y(a) &= {}^2\eta,\\
&\vdots & &\vdots\\
{}^my' &= {}^mf(x, {}^1y, {}^2y, \ldots, {}^my), & {}^my(a) &= {}^m\eta.
\end{aligned} \tag{4}$$

(If the ${}^iy, i = 1, 2, \ldots, m$, satisfy given conditions at *different* values $a, b, c, \ldots$ of x, then we have a *multipoint boundary value problem*; if there are just two different values a and b of x, then we have a *two-point boundary value problem*.) Introducing the vector notation

$$\mathbf{y} = [{}^1y, {}^2y, \ldots, {}^my]^T, \quad \mathbf{f} = [{}^1f, {}^2f, \ldots, {}^mf]^T = \mathbf{f}(x, \mathbf{y}),$$
$$\boldsymbol{\eta} = [{}^1\eta, {}^2\eta, \ldots, {}^m\eta]^T,$$

we may write the initial value problem (4) in the form

$$\mathbf{y}' = \mathbf{f}(x, \mathbf{y}), \quad \mathbf{y}(a) = \boldsymbol{\eta}. \tag{5}$$

Theorem 1.1 readily generalizes to give necessary conditions for the existence of a unique solution to (5); all that is required is that the region D now be defined by $a \leqslant x \leqslant b$, $-\infty < {}^iy < \infty$, $i = 1, 2, \ldots, m$, and (2) be replaced by the condition

$$\|\mathbf{f}(x, \mathbf{y}) - \mathbf{f}(x, \mathbf{y}^*)\| \leqslant L\|\mathbf{y} - \mathbf{y}^*\|, \tag{6}$$

where $(x, \mathbf{y})$ and $(x, \mathbf{y}^*)$ are in D, and $\|\,.\,\|$ denotes a vector norm (Mitchell,[134] chapter 1). In the case when each of the ${}^if(x, {}^1y, {}^2y, \ldots, {}^my)$, $i = 1, 2, \ldots, m$, possesses a continuous derivative with respect to each of the ${}^jy$, $j = 1, 2, \ldots, m$, then we may choose, analogously to (3),

$$L = \sup_{(x,\mathbf{y})\in D} \left\|\frac{\partial \mathbf{f}}{\partial \mathbf{y}}\right\|, \tag{7}$$

where $\partial\mathbf{f}/\partial\mathbf{y}$ is the *Jacobian* of $\mathbf{f}$ with respect to $\mathbf{y}$—that is, the $m \times m$ matrix whose i–jth element is $\partial\,{}^if(x, {}^1y, {}^2y, \ldots, {}^my)/\partial\,{}^jy$, and $\|\,.\,\|$ denotes a matrix norm subordinate to the vector norm employed in (6) (see Mitchell,[134] chapter 1).

1.5 Reduction of higher order differential equations to first-order systems

Let us consider an initial value problem involving an ordinary differential equation of order m, which can be written in the form†

$$y^{(m)} = f(x, y^{(0)}, y^{(1)}, \ldots, y^{(m-1)}), \qquad y^{(t)}(a) = \eta_t, \qquad t = 0, 1, \ldots, m-1. \tag{8}$$

We define new variables ${}^iy$, $i = 1, 2, \ldots, m$, as follows:

$$\begin{aligned}
{}^1y &\equiv \phantom{{}^1}y\ (\equiv y^{(0)}),\\
{}^2y &\equiv {}^1y'\ (\equiv y^{(1)}),\\
{}^3y &\equiv {}^2y'\ (\equiv y^{(2)}),\\
&\ \vdots \qquad \vdots\\
{}^my &\equiv {}^{m-1}y'\ (\equiv y^{(m-1)}).
\end{aligned}$$

Then, on writing ${}^i\eta$ for η_{i-1}, $i = 1, 2, \ldots, m$, the initial value problem (8) may be written as an initial value problem for a first-order system, namely

$$\begin{aligned}
{}^1y' &= {}^2y, & {}^1y(a) &= {}^1\eta,\\
{}^2y' &= {}^3y, & {}^2y(a) &= {}^2\eta,\\
&\ \vdots & &\ \vdots\\
{}^{m-1}y' &= {}^my, & {}^{m-1}y(a) &= {}^{m-1}\eta,\\
{}^my' &= f(x, {}^1y, {}^2y, \ldots, {}^my), & {}^my(a) &= {}^m\eta,
\end{aligned}$$

† Superscripts in round brackets indicate the order of higher derivatives; that is, $y^{(0)}(x) \equiv y(x)$, $y^{(1)}(x) \equiv y'(x)$, $y^{(2)}(x) \equiv y''(x)$, etc.

or,

$$\mathbf{y}' = \mathbf{f}(x, \mathbf{y}), \quad \mathbf{y}(a) = \boldsymbol{\eta},$$

where

$$\mathbf{f} = [{}^2y, {}^3y, \ldots, {}^my, f(x, {}^1y, {}^2y, \ldots, {}^my)]^T.$$

With certain exceptions to be discussed in chapter 9, our normal procedure for dealing with an initial value problem of the form (8) will be to reduce it to an initial value problem for an equivalent first-order system. Note that the differential equation appearing in (8) is not the most general differential equation of order m; *implicit differential equations* of the form

$$F(x, y^{(0)}, y^{(1)}, \ldots, y^{(m)}) = 0$$

are also possible. Using the technique described above, an initial value problem involving such an equation may be written in the form

$$\mathbf{F}(x, \mathbf{y}, \mathbf{y}') = \mathbf{0}, \quad \mathbf{y}(a) = \boldsymbol{\eta}.$$

There exist very few numerical methods which tackle this form of initial value problem directly, and we shall always assume that we are given an initial value problem in one of the forms (1), (5), or (8).

1.6 First-order linear systems with constant coefficients

The first-order system $\mathbf{y}' = \mathbf{f}(x, \mathbf{y})$, where $\mathbf{y}$ and $\mathbf{f}$ are m-dimensional vectors, is said to be *linear* if $\mathbf{f}(x, \mathbf{y}) = A(x)\mathbf{y} + \boldsymbol{\phi}(x)$, where $A(x)$ is an $m \times m$ matrix and $\boldsymbol{\phi}(x)$ an m-dimensional vector; if, in addition, $A(x) = A$, a constant matrix, the system is said to be *linear with constant coefficients.* In chapter 8, we shall require the general solution of such a system,

$$\mathbf{y}' = A\mathbf{y} + \boldsymbol{\phi}(x). \tag{9}$$

Let $\hat{\mathbf{y}}(x)$ be the general solution of the corresponding homogeneous system

$$\mathbf{y}' = A\mathbf{y}. \tag{10}$$

If $\boldsymbol{\psi}(x)$ is any particular solution of (9), then $\mathbf{y}(x) = \hat{\mathbf{y}}(x) + \boldsymbol{\psi}(x)$ is the general solution of (9). A set of solutions $\mathbf{y}_t(x)$, $t = 1, 2, \ldots, M$, of (10) is said to be *linearly independent* if $\sum_{t=1}^{M} a_t\mathbf{y}_t(x) \equiv 0$ implies $a_t = 0$, $t = 1, 2, \ldots, M$. A set of m linearly independent solutions $\hat{\mathbf{y}}_t(x)$, $t = 1, 2, \ldots, m$, of (10) is said to form a *fundamental system* of (10), and the most general solution of (10) may be written as a linear combination of the members of the fundamental system. It is easily seen that $\hat{\mathbf{y}}_t(x) = e^{\lambda_t x}\mathbf{c}_t$, where $\mathbf{c}_t$ is an m-dimensional vector, is a solution of (10) if $\lambda_t \mathbf{c}_t = A\mathbf{c}_t$, that is, if λ_t is an eigenvalue of A and $\mathbf{c}_t$ is the corresponding eigenvector. It will be sufficient for our purposes to consider only the case where A possesses *m distinct*, possibly complex, eigenvalues λ_t, $t = 1, 2, \ldots, m$. The corresponding eigenvectors $\mathbf{c}_t$, $t = 1, 2, \ldots, m$, are then linearly independent (Mitchell,[134] chapter 1), and it follows that the solutions $\hat{\mathbf{y}}_t(x) = e^{\lambda_t x}\mathbf{c}_t$, $t = 1, 2, \ldots, m$, form a fundamental system of (10), whose most general solution is thus of the form $\sum_{t=1}^{m} k_t e^{\lambda_t x}\mathbf{c}_t$, where the k_t, $t = 1, 2, \ldots, m$ are arbitrary constants. The most general solution of (9) is then

$$\mathbf{y}(x) = \sum_{t=1}^{m} k_t e^{\lambda_t x}\mathbf{c}_t + \boldsymbol{\psi}(x). \tag{11}$$

We can now find the solution of the initial value problem

$$\mathbf{y}' = A\mathbf{y} + \boldsymbol{\phi}(x), \qquad \mathbf{y}(a) = \boldsymbol{\eta} \tag{12}$$

under the assumptions that A has m distinct eigenvalues, and that we know a particular solution $\boldsymbol{\psi}(x)$ of (9). By (11), the general solution of (9) satisfies the initial condition given in (12) if

$$\boldsymbol{\eta} - \boldsymbol{\psi}(a) = \sum_{t=1}^{m} k_t e^{\lambda_t a}\mathbf{c}_t. \tag{13}$$

Since the vectors $\mathbf{c}_t$, $t = 1, 2, \ldots, m$, form a basis of the m-dimensional vector space (Mitchell,[134] chapter 1), we may express $\boldsymbol{\eta} - \boldsymbol{\psi}(a)$ uniquely in the form

$$\boldsymbol{\eta} - \boldsymbol{\psi}(a) = \sum_{t=1}^{m} \kappa_t \mathbf{c}_t. \tag{14}$$

On comparing (13) with (14), we see that (11) is a solution of (12) if we choose $k_t = \kappa_t e^{-\lambda_t a}$. The solution of (12) is thus

$$\mathbf{y}(x) = \sum_{t=1}^{m} \kappa_t e^{\lambda_t (x-a)}\mathbf{c}_t + \boldsymbol{\psi}(x).$$

For a fuller treatment of first-order linear systems with constant coefficients, the reader is referred to Hurewicz.[78]

Example 1 Solve the initial value problem $\mathbf{y}' = A\mathbf{y}$, $\mathbf{y}(0) = [1, 0, -1]^T$, *where*

$$A = \begin{bmatrix} -21 & 19 & -20 \\ 19 & -21 & 20 \\ 40 & -40 & -40 \end{bmatrix}.$$