Title
APPLICATION OF LINEAR PROGRAMING TO THE NUMERICAL SOLUTION OF LINEAR DIFFERENTIAL EQUATIONS

Author
Young, Jonathan D.

Publication Date
1962-03-01

UCRL 10110

University of California

Ernest O. Lawrence Radiation Laboratory

APPLICATION OF LINEAR PROGRAMMING TO THE NUMERICAL SOLUTION OF LINEAR DIFFERENTIAL EQUATIONS

Berkeley, California

DISCLAIMER

Research and Development

UCRL-10110
UC-32 Mathematics and Computers
TID-4500 (17th Ed.)

UNIVERSITY OF CALIFORNIA

Lawrence Radiation Laboratory
Berkeley, California

Contract No. W-7405-eng-48

APPLICATION OF LINEAR PROGRAMMING TO THE NUMERICAL SOLUTION OF LINEAR DIFFERENTIAL EQUATIONS

Jonathan D. Young

(Thesis)

March 1, 1962

Printed in USA. Price $2.00. Available from the
Office of Technical Services
U. S. Department of Commerce
Washington 25, D.C.

APPLICATION OF LINEAR PROGRAMMING TO THE NUMERICAL SOLUTION OF LINEAR DIFFERENTIAL EQUATIONS

Contents

APPLICATION OF LINEAR PROGRAMMING TO THE NUMERICAL SOLUTION OF LINEAR DIFFERENTIAL EQUATIONS

Jonathan D. Young

Lawrence Radiation Laboratory
University of California
Berkeley, California

March 1, 1962

ABSTRACT

The linear programming procedure is applied to obtain an optimal solution to an overdetermined system of linear inequalities that are derived from the localization of the differential equation to some set of discrete points, from the prescribed conditions, and from the application of approximation formulas.

The method is applied to linear ordinary differential equations and systems and to first-and second-order partial differential equations and systems in two independent variables. Extensions are suggested. Many examples illustrating the method are provided.

I. INTRODUCTION

A linear program is a representation of a decision situation as a mathematical model in which the objective is to optimize a linear function of nonnegative variable subject to linear restraints. Although the most extensive development in procedure and application has taken place in the last twenty years (Ref. 1, p. 3), the theoretical origins may be traced well into the nineteenth century. The minimax theory, which is essential to linear programming methods, arises from the concept of using inequalities to determine relative maxima (Fourier) and from the geometric treatment of inequalities as properties of convex sets (Minkowski).

The mathematical approach to game theory was initiated by von Neumann in 1928.[2] Weyl's exposition on the properties of convex polyhedra followed in 1935.[3] The economic application of minimax theory to games of strategy appeared in 1944 in the well-known treatise, "Theory of Games and Economic Behaviour," by von Neumann and Morgenstern.[4] The theory was applied with varying success to more specific problems; foremost among these were the transportation problem (1941) of Hitchcock[5] and the diet problem (1945) of Stigler.[6]

In 1947 a group of mathematicians investigated the possibility of devising mathematical approaches to the complex planning and programming problems of the U. S. Air Force. The linear analogue was proposed by George B. Dantzig. Project SCOOP was organized to exploit this idea.[7] A major result of the work on this project was the development by Dantzig of a systematic procedure for the linear programming computation. The algorithm that he constructed is called the Simplex method.[8] The rapid progress in the design and use of electronic high-speed computers made it possible to apply the algorithm to larger and larger models.

In June, 1949, several papers were presented at the Conference on Linear Programming held at the University of Chicago. These were edited by T. C. Koopmans and his associates and published by the Cowles Commission.[9] Symposia on linear programming were held in 1951 and

1955. The published proceedings of these meetings show the rapid progress which took place in that period.[10, 11] In 1953 the first introductory text on linear programming appeared.[12]

The Rand Corporation, in a joint effort with various industrial interests, has been most acitve in the improvement of computational procedures and the preparation of computer codes for linear programming. The work of Orchard-Hays and his associates has been extremely important in this endeavour.[13, 14, 15]

The comprehensive bibliography by Riley and Gass (1958) presents at least a thousand abstracts of articles and books on linear programming and related subjects.[1] Some of these are of a theoretical nature, others deal with computational procedures; but the greatest number make more or less specific application of the method to such diverse fields as industry, military planning, and agriculture.

Little has been done toward applying linear programming to a class of problems phrased in mathematical language per se. We offer here a novel effort in this direction which we hope will be interesting and useful. Our purpose is to develop a numerical method for solving linear differential equations by means of linear programming and to investigate the validity and practicality of such a method. The exposition stresses procedure and illustration rather than abstract rigor and generality.

Our subject brings together two mathematical areas, linear programming and differential equations, which have had very little common background. It seems that some readers may be familiar with one of these areas but somewhat unfamiliar with the other. For this reason we discuss each of them independently, linear programming in Chapter II, and differential equations in Chapter III. The material offered in each case is limited to that which is useful to our purpose. For more complete information on either topic, the reader may resort to the many references.

The two topics are brought together in Chapter IV, in which we construct linear programs for solving linear differential equations.

Our conclusions on the merits of the proposed method are given in Chapter V. Many examples are given in the Appendix.

The author wishes to particularly acknowledge the assistance and encouragement of D. H. Lehmer and the constructive suggestions of P. L. Chambré in connection with the preparation of this thesis.

II. FUNDAMENTALS OF LINEAR PROGRAMMING

Preliminary Discussion

Decision situations present the possibility of engaging in various activities at varying levels in order to best attain some specified objective. They are exemplified in the theory of games and in operations analysis, but in fact they are found in almost every area of human action. Some of these situations have been found to be readily amenable to mathematical representation or approximation. The resulting model is characterized by many possible solutions, among which an optimal selection must be made. Often this choice is not unique. Programming means the construction and solution of the mathematical system that adequately represents the given situation.

The linear program has been found to be a most useful and convenient form of mathematical representation. In the linear program, the variables are restricted to nonnegative values and are subject to other linear restraints. The objective is to optimize (either maximize or minimize) a linear function of the variables. There will, of course, be no solution if the restraints are contradictory. Linear programming will usually involve

a. the given description of the decision situation,
b. an adequate linear representation,
c. the linear program,
d. the computation leading to
    (1). a demonstration that no solution exists, or
    (2). a determination of an optimal solution,
e. the interpretation of the result.

Simple examples illustrating the above procedure are given in the Appendix, and many others may be found in the references (e.g., Refs. 8, 12, 16, 17).

General Linear Program

The most general form for a linear program is:

Optimize the objective function,

$$z = \Sigma\, c_j x_j, \qquad \text{(A1)}$$

subject to the restraints,

$$\Sigma\, a_{ij} x_j = b_i; \quad i = 1(1)\, m_1\text{-}1, \qquad \text{(B1)}$$

$$\Sigma\, a_{ij} x_j \leqslant b_i; \quad i = m_1(1)\, m_2\text{-}1, \qquad \text{(B2)}$$

$$\Sigma\, a_{ij} x_j \geqslant b_i; \quad i = m_2(1)\, m, \qquad \text{(B3)}$$

$$x_j \geqslant 0; \quad j = 1(1)\, n_1. \qquad \text{(C1)}$$

If the required optimization is to minimize z, we can choose instead to maximize - z. The order of the inequalities (B3) can be reversed to that of (B2) by changing the signs of the terms. With these formal operations and the slight obvious redefinition, the program can be written in the form:

Maximize the function,

$$z = \Sigma\, c_j x_j\,, \qquad \text{(A1')}$$

subject to

$$\Sigma\, a_{ij} x_j = b_i; \quad i = 1\ (1)\ m_1\text{-}1, \qquad \text{(B1)}$$

$$\Sigma\, a_{ij} x_j \leqslant b_i; \quad i = m_1(1)\, m, \qquad \text{(B2')}$$

$$x_j \geqslant 0; \quad j = 1\ (1)\ n_1. \qquad \text{(C1)}$$

Into each of the inequalities (B2') we can introduce a nonnegative slack variable with unit coefficient, and replace the inequality by an equality. All the b_i can be made nonnegative by changes of signs of the terms in those equations where necessary. With these formal changes, we can write our program in matrix-vector notation as: a standard linear program.

$$\text{Maximize} \quad z = CX, \tag{A}$$

$$\text{subject to} \quad AX = B; \quad B \geqslant 0, \tag{B}$$

$$\text{and} \quad X \geqslant 0, \tag{C}$$

where X (which now includes the slack variables) and C (many of whose components may be zero) now have n ($n \geqslant n_1$) components; the matrix A is $m \times n$, and B has m components.

In the linear system (B) let r be the rank of A, and s be the rank of the augmented matrix A, B. The following cases may occur:

$$r = s = m \leqslant n, \tag{a}$$

$$r = s < m, \tag{b}$$

$$r < s. \tag{c}$$

For case (b) the system (B) has redundant equations; for (c) the system is inconsistent. The effect of these conditions will appear in later developments. We confine our attention to case (a) in which the system (B) has at least one solution.

For case (a) the matrix A has at least one set of m linearly independent column vectors. Any such set of vectors is called a basis. Suppose, for convenience, that the first m columns of A constitute a basis, then a solution to (B) may be readily obtained by setting

$$x_j = 0, \qquad j = m + 1(1)\, n,$$

and solving the (possibly) abbreviated system for x_j; $j = 1(1)\, m$. Such a solution corresponding to any choice of basis is called a basic solution. The number of basic solutions is finite; there is at least one, and there are no more than $\binom{n}{m}$, the number of combinations of n things taken m at a time.

Any solution to (B) which also satisfies (C) is said to be feasible. If there is no solution to the system (B) that satisfies restraints (C), we say our program is infeasible. [This statement is applicable to case (c) above.]

If the linear function x defined in (A) can be made arbitrarily large for a feasible solution, we day our objective is unbounded. Otherwise there is at least one feasible solution for which z has a maximum value. Any such solution is called an optimal solution.

The following may be readily proved:

Lemma 1. The set of feasible solutions is a convex set. (Ref. 16)

Theorem 1. If there is a feasible solution, there is a basic feasible solution. (Ref. 18)

Lemma 2. The set of optimal solutions is convex. (Ref. 18)

Theorem 2. If there is an optimal solution, there is a basic optimal solution. (Ref. 18)

The above theorems enable us to look for an optimal solution in the finite set of basic solutions. The following theorem, for which we later give a constructive proof, makes it possible for us to search for an optimal solution along some direction of greatest ascent.

Theorem 3. With a known nonnegative basic solution, it is possible to select successive basic solutions in such a way that these solutions are nonnegative, and the objective function does not decrease. The order of selection may not be unique.

Duality of Linear Models

Analogous to the duality of points and planes in projective geometry, there are linear models that involve a duality of variables and linear restraints. Thus the dual of our standard linear program,

$$\begin{aligned} &\text{maximize} && z &&= CX, \\ &\text{subject to} && AX &&= B, \\ &\text{and} && X &&\geqslant 0, \end{aligned}$$

is an unrestricted (as to sign of the variables) linear model,

$$\begin{aligned} &\text{minimize} && w &&= BU, \\ &\text{subject to} && A'U &&\geqslant C, \end{aligned}$$

with no restriction on the sign of the components of U.

We have the following theorem on duality (for proof, see Ref. 19.

Theorem 4. If either of the dual models has an optimal solution, then so does the other; further, the optimal objective values are equal, that is, max z = min w.

The power of this theorem is immediately apparent. If our decision situation can be represented as the unrestricted linear model above, then its dual is a standard linear program whose computation gives coincidentally the unrestricted solution we are seeking. This procedure is illustrated in the Appendix.

The Simplex Algorithm

The simplex algorithm is a method of steepest ascent for solving the standard linear program,

maximize	$z = CX,$	(A)
subject to	$AX = B;\quad B \geqslant 0,$	(B)
with	$X \geqslant 0,$	(C)

where X has n components, and B has m.

For an initial feasible basis, we should like to have foremost an $m \times m$ identity matrix; then the corresponding basic solution would be X = B, 0. It is unlikely that our model will have the above highly desirable form, but we can modify it in a rather formal manner by introducing m artificial variables, assigning them indices j, j = n + 1(1) n + m in the now expanded vector X (with $n^* = m + n$ components), and including them one at a time in Eq. (B), so that the now expanded coefficient matrix of (B) has the form I_m, A. Because any eventual optimal solution in which the artificial variables appeared with nonzero values would be meaningless, we also must include them in the objective function with predominantly negative, but unspecified, coefficients that will insure (if possible) that they are forced to "nonbasic" status or to zero values. The introduction of these artificial variables and the elimination of their corresponding column vectors

from the basis is called Phase I of the Simplex process. Aside from the convenience, the introduction of these artificial variables has presented us with an initial system (B') replacing (B), which cannot be redundant or inconsistent and for which there is an obvious nonnegative basic solution.

We are now in position to assume the hypothesis of Theorem 3. Let X' be any current nonnegative basic solution, and suppose for convenience $X' = X^*, 0$ where $X^* = x'_1, \cdots x'_m$. The coefficient matrix has the form Q, Q_0, where Q is our current basis. Abbreviating our model by discarding zero terms, we have

$$z^* = C^*X^* \qquad (C^* = c_1 \cdots, c_m),$$

and

$$QX^* = B.$$

We assume the complete coefficient matrix to consist of column vectors P_j, $j = 1(1)\, n^* m$, and compute

$$z_j = CC^*Q^{-1} P_j, \qquad j = 1(1)\, n + m,$$

and let

$$w_j = z_j - c_j, \qquad j = 1(1)\, n + m.$$

If all these $w_j \geq 0$, z has attained a maximum value. If all the artificial variables have zero values, we already have an optimal solution. If some nonzero artificial variables still appear at this optimum, the program is infeasible.

If some w_j are negative, then for some k, $w_k = \min w_j$. For the corresponding column vector P_k of the matrix A, we compute

$$X_k = Q^{-1}P_k, \qquad X_k: x_{ik};\ i = 1(1)\, m,$$

and

$$R_i = x'_i / x_{ik}, \text{ for } x_{ik} \neq 0 \text{ and } i = 1(1)m.$$

If all the $R_i \leq 0$, our objective is unbounded, otherwise some of the R_i are nonnegative and, for some ℓ,

$$R_\ell = \min R_i \geqslant 0,$$

and the vector P_ℓ replaces P_k in the basis. The solution X'' corresponding to this new basis will be

$$\begin{aligned} x''_j &= x'_j - R_\ell x_{jk}, \qquad j = 1(1)\ m, \text{ except for } \ell, \\ x''_\ell &= 0, \\ x''_k &= R_\ell, \\ x''_j &= 0, \qquad j = m+1(1)\ n+m, \text{ except for } k, \end{aligned}$$

and is obviously nonnegative. The new value for the linear function z is

$$z'' = z^* - R_\ell w_k \geqslant z^*,$$

and we have proved Theorem 3 of page 7. We now note that neither the choice of w_k or R_ℓ need have been unique. The sequence of basic solutions and the ultimate numerical values of the variables in the optimal solution may depend on the arbitrary choices made when "ties" occurred. In common practice a "tie" is resolved by selecting the lowest qualifying index.

Computer Codes

The computation involved in the Simplex algorithm is intricate but well defined and highly repetitive. Because of the latter properties it may be readily codified for high-speed computers.

Such codes require only that the linear program be input in some specified form, and direct the computer through the necessary process to an optimal solution or to a demonstration of infeasibility or unboundedness.

The code used by the computation of the linear programs which appear in the Appendix was the IBM-704 code SCROL (Ref. 30) which is available from the "SHARE" organization of IBM machine users. Recently, (Nov. 1961) a linear program code, LP-90, has been released by the same organization for use on the IBM-7090. For other computers, there are doubtless similar codes available or soon to be released.

III. LINEAR DIFFERENTIAL EQUATIONS, ORDINARY AND PARTIAL

Preliminary Discussion

The mathematical model chosen to represent a given situation need not be limited to algebraic processes. When the known relations involve rates of change of one or more quantities with respect to one or more others, the process of differentiation (or integration) is implicit to the representation, and the model may consist of or contain differential equations.

A differential equation is an equation connecting certain independent variables $s, t, \cdots$, certain (unknown) functions $u, v, \cdots$ of these variables, and certain derivatives of these functions with respect to those variables (Ref. 20, p. 2). The order of a differential equation is that of its highest-ordered derivative. A differential system consists of as many simultaneous differential equations as there are unknown functions.

A linear differential equation is linear in the unknown functions and their derivatives; that is, these occur to the first degree only and not as higher powers or products (Ref. 21, p. 3). In a linear differential system all the equations are linear. As for those known functions of the independent variables that appear as coefficients or "constant terms", we assume that they are numerically defined, single-valued and continuous, and that those appearing as coefficients of the highest-order derivative terms do not vanish.

An ordinary differential equation relates one independent variable s, one dependent variable $u(s)$ and the (ordinary) derivatives of u with respect to s. An ordinary differential system relates one independent variable s, two or more dependent variables $u(s), v(s), \cdots$ and their ordinary derivatives with respect to s.

A partial differential equation relates two or more independent variables $s, t, \cdots$, one dependent variable $u(s, t, \cdots)$ and the partial

derivatives of u with respect to s, t, ··· . A partial differential system relates two or more independent variables, s, t, ··· two or more dependent variables u(s, t, ···), v(s, t, ···), ··· and the partial derivatives of u, v, ··· with respect to s, t, ··· .

A solution to a differential equation is an extraction (free of derivatives) of the intrinsic relation between the unknown dependent and independent variables. The existence of a continuous solution cannot always be predicted. A functional solution to a differential equation is an equation that does not involve derivatives, that relates the unknown function(s) and independent variable(s), and that can be subjected to differentiation and algebraic processes to yield the differential equation. As differentiation with respect to a variable results in the elimination of constants or functions not involving that variable, a functional solution may contain arbitrary constants and/or arbitrary functions. Such a solution is called a general solution.

A particular solution is a functional solution that does not involve arbitrary constants or functions. If sufficient conditions are prescribed, it may be possible to specify the arbitrary constants or functions in a general solution and obtain thereby a particular solution that satisfies both the prescribed conditions and the differential equation.

The Differential Problem

A differential problem consists of a differential equation or system, together with any additional conditions that may be prescribed. The solution to the differential problem must satisfy the differential equation or system and the prescribed conditions. A differential problem has a unique solution for some domain if, for each point of this domain, there is one and only one solution. Even a particular solution need not be unique since it need not be single-valued. In a linear differential problem both conditions and equations are linear in unknown functions and derivatives.

A well-posed differential problem admits of a unique continuous solution on the domain of interest. The question of the existence of a unique solution to a differential problem has been investigated extensively. A comprehensive treatment of this subject is beyond the scope of this work. We shall limit ourselves to the statement of two important theorems. For further information, the reader is advised to consult Refs. 22 and 23.

Theorem 5. If in the linear ordinary differential equation

$$a_0(s)\, u + a_1(s)\, u' + \cdots + a_J(s)\, u^{(J)} = c(s)$$

the functions a_0, a_1, $\cdots$, a_J and c are continuous for the interval domain, $\underline{s} \leq s \leq \overline{s}$, and the function a_J does not vanish therein, then there exists a unique solution u that, together with its first $J-1$ derivatives, is continuous on $(\underline{s}, \overline{s})$ and that satisfies the prescribed conditions

$$u(s_0) = u_0,$$
$$u'(s_0) = u'_0,$$

$$u^{(J)}(s_0) = u_0,$$

where s_0 is a point of $(\underline{s}, \overline{s})$. (Ref. 21, p. 73.)

Theorem 6. (Kowalewski) If $g(t)$ and all its derivatives are continuous for $|t-t_0| < \delta$, if s_0 is a given number and $u_0 = g(t_0)$ and $u_{s0} = g'(t_0)$, and if $f(s, t, u, u_t)$ and all its partial derivatives are continuous in a region defined by

$$|s - s_0| < \delta, \quad |t - t_0| < \delta, \quad |u_t - u_{t0}| < \delta,$$

then there exists a unique function $\phi(s, t)$ so that

(a) $\phi(s, t)$ and all its partial derivatives are continuous in a region defined by

$$|s - s_0| < \delta_1, \qquad |t - t_0| < \delta_2;$$

(b) For all (s, t) in this region, $u = \phi(s, t)$ is a solution of the equation, $u_s = f(s, t, u, u_t)$;

(c) For all values of t in the interval $|t-t_0| < \delta_2$,

$$\phi(s_0, t) = g(t);$$

(ref. 24, p. 49; ref. 23, p. 32 through 36).

Limited Solution of a Differential Problem

In most practical work with differential problems, only a limited or tabulated solution is required. Such a solution consists of a table of values of the unknown function(s) for a relatively small number of specified values of the independent variable(s).

Sometimes a general solution or solutions can be found for the differential equation or system. If the problem is well-posed, it may be possible then to specify the arbitrary constants or functions so that the prescribed conditions are satisfied. The resulting particular solution is perforce unique. Then, by substitution of the specified values of the independent variable(s), we can obtain the required limited solution. Even here the work of finding the numerical values may be arduous, involving the use of tables, interpolation, and perhaps integration. Usually the numerical values cannot be expressed precisely in terminating decimals; of necessity then the results are approximations to the number of decimal places employed (Ref. 25, p. 13, 19). When a limited solution is attempted by the procedure outlined in this paragraph, we say we have used an analytic approach.

The numerical approach may be used to obtain a limited solution to a differential problem without recourse to any functional solution. In this approach we attempt to solve the differential problem in toto by numerical methods; that is, the prescribed conditions are involved at the very start (Ref. 26, p. 3). If the problem is not well-posed we cannot hope for any result from this approach other than perhaps to learn this sad fact. The numerical procedures, if successful, will give us a satisfactory approximation of the limited solution desired.

Even for a limited solution, the analytic approach is usually attempted first. Its advantages are that the amount of computation required may be slight and that limited solutions for other prescribed conditions may be obtained by respecifying the arbitrary constants or functions. When the analytic approach fails because there is no known analytic method applicable, or when it leads to tedious computation, we resort to numerical methods that have more widespread application and that encompass all the necessary computation.

In the numerical approach, a mesh is constructed over the domain. The set of mesh points must include all the points for which the limited solution was required and some or all (if the number is finite) of the points at which conditions are prescribed. Approximation formulas are used that relate the values of the unknown function(s) and derivatives at adjacent points of the mesh. For example, if the value of an unknown function and some of its derivatives are prescribed, or can be computed at one mesh point, it may be possible to approximate the value of the unknown function at an adjacent point by means of a Taylor expansion truncated at the highest-known derivative. The truncation error may be reduced by iterative techniques or by use of formulas of greater sophistication.

Generally, these formulas have the property that they approximate the solution by some polynomial, and the degree of the polynomial for which the method is exact is a fairly good measure of the validity of the method. Often (but not always) the approximation may be improved by refinement of the mesh. Descriptions of the many numerical methods is beyond our scope, and the reader is advised to

consult Ref. 25, 27, 28.

Linear Model Approximation of a Linear Differential Problem

A linear differential equation applied at a specific mesh point becomes a linear algebraic equation in which the unknowns are solution and derivative values at that point. If we consider only prescribed conditions which are linear in solution and derivative values (this is almost invariably the case), then a prescribed condition at a mesh point becomes a linear algebraic equation in solution and derivative values at this point. Most approximation formulas have the property that they are linear in solution and derivative values for two or more mesh points, hence they are linear algebraic equations in these values when applied to specific points. Using the linear algebraic equations available from the three sources, it is possible to construct an algebraic system in which the unknowns are solution and derivative values for the mesh points. Fox gives a complete discussion of this technique for the boundary-condition problem for an ordinary differential equation.[26]

In Chapter IV, we construct such an algebraic system, which can be made overdeterminate. We allow for error in the approximation formulas by introducing the maximum absolute value of the error as a new variable, which we wish to minimize. The result is an unrestricted (as to sign of variables) linear model whose dual is a standard linear model (See pp. 7, 8).

IV. LINEAR PROGRAMS FOR WELL-POSED LINEAR DIFFERENTIAL PROBLEMS, ORDINARY AND PARTIAL

Preliminary Discussion

The linear program method, which we describe, applies very generally to well-posed linear differential problems for which a limited solution is required. It is a numerical method and concerns itself with solution and derivative values for a finite set of values (mesh points) of the independent variable(s). The set of mesh points must include those points for which a limited solution is required and a sufficient number of the points at which conditions are prescribed. The method represents the discretized differential problem by a linear model consisting of a system of linear inequalities whose variables denote the maximum error of approximation and the values of the solution and derivatives at the mesh points. The maximum absolute error is to be minimized. The foregoing model is the dual of a linear program whose computation provides optimal values for its dual variables which indicate the error of approximation and which approximate the solution and its derivatives at the mesh points. The required limited solution appears among those values.

Despite the generality of the method, we shall follow the traditional pattern in the succeeding sections by constructing programs for linear problems involving, first, an ordinary differential equation, than an ordinary system, then a partial differential equation and, finally, a partial differential system. The number of variables in the linear model is a rapidly increasing function of the number of mesh points, the order of the equations involved, the number of unknown functions, and the number of independent variables; hence there is a practical limit on the complexity of the problem which can be treated. For this reason, and because of the dearth of information as to what constitute well-posed problems, in the case of partial differential problems, we shall limit ourselves to equations of second order in two unknown functions of two independent variables.

All superscripts and subscripts appearing in the equations of this section are indices except except for the superscripts associated with the mesh-step quantities, h or h, these superscripts are exponents.

Linear Ordinary Differential Equation Problem

The linear program method applies rather universally to well-posed problems for a linear ordinary differential equation of reasonable order, provided the number of mesh points is not excessively large. No special treatment is required for subclasses as to order, homogeneity or type of prescribed conditions, except that the latter must be numerical conditions on linear combinations of the unknown function and derivatives at specified points.

We have a differential equation that holds on some closed interval S of an independent variable s and that has the form

$$a_0(s)u + a_1(s)u' + \cdots + a_J(s)u^{(J)} = c(s), \tag{1}$$

where

u is an unknown function of s, the superscripts indicate derivatives, J is the order of the equation, the a_j and c are numerically defined and continuous on S, and a_j does not vanish in S.

For a well-posed problem (Theorem 5, Chapter III), we must have J independent conditions prescribed at a point or points s^ℓ of S. The conditions must have the form

$$a_0^\ell u_\ell + a_1^\ell u'_\ell + \cdots + a_{J-1}^\ell u_\ell^{(J-1)} = c^\ell; \ell = 1(1)\, J, \tag{2}$$

where

the a_j^ℓ and c^ℓ are real numbers, the superscript ℓ indices the conditions, and the subscript ℓ indicates value at s^ℓ.

A limited solution is required consisting of a table of solution values at a finite number of points s^* of S. We consider a set of K mesh points s_ℓ in S. This set includes, but is not limited to, the points s^ℓ and s^*. For a nontrivial problem, there must be at least two mesh points. The linear-program method does <u>not</u> require a uniform step-length in the mesh.

The differential problem discretized to the points s_k involves $KJ + K$ unknown quantities, namely, the numerical values of the solution and its first J derivatives at the K mesh points. In Eq. (2), we have J linear relations for some of these unknowns. An additional K linear equations may be obtained by applying differential Eq. (1) at the mesh points to obtain

$$a_0^k u_k + a_1^k u_k' + \cdots + a_J^k u_k^{(J)} = c^k; \quad k = 1(1)K, \tag{3}$$

where

the a_j^k and c^k are the real numerical values at s_k, and the subscript k indicates value at s_k.

As yet we have not represented the intrinsic relation between a function and its derivatives. If the solution u is an analytic function of s, a Taylor expansion provides a linear relation for solution and derivative values at adjacent points of the mesh. However, we have not assumed that u is analytic, nor could we deal numerically with the possible infinitude of terms if it were. If the solution u is a polynomial, the Taylor expansion terminates with the highest-ordered nonzero derivative, but here we have gained finiteness of terms at the expense of even greater restriction on the nature of the solution.

We assume only that the solution function is continuous and has continuous derivatives up to the order $J-1$. (See Theorem 5, p. 13). In general then, by numerical methods, we can only hope to approximate the solution by some polynomial. It is well known that a continuous function such as u can be approximated by a polynomial over a closed interval such as S. In fact we are limited to polynomials of degree J, since we have no knowledge of higher-ordered derivatives. For a sufficiently small closed interval, such as any of the $[s_k, s_{k+1}]$, it is possible that the solution function u can be approximated by a polynomial of degree J or less. We shall make $K-1$ such approximations and use over each mesh-step a J^{th} degree polynomial approximation to the solution, u.

For the set of consecutive mesh points s_k, s_{k+1}, $k = 1(1)\,K-1$, we write the forward (from s_k to s_{k+1}) Taylor expansions:

$$\begin{aligned}
&u_k + h_k u'_k + \cdots + h_k^J u_k^{(J)}/J! - u_{k+1} = E_{1k}, \\
&h_k u'_k + h_k^2 u''_k + \cdots + h_k^J u k^{(J)}/(J-1)! - h_k u'_{k+1} = E_{2k}, \\
&\qquad\vdots \qquad\qquad\qquad\qquad\qquad\qquad k = 1(1)K-1, \\
&\qquad\vdots \qquad\qquad\qquad\qquad\qquad\qquad (4) \\
&h_k^{J-1} u_k^{(J-1)} + h_k^J u_k^{(J)} - h_k^{J-1} u_{k+1}^{(J-1)} = E_{Jk},
\end{aligned}$$

where

$h_k = s_{k+1} - s_k$, and the E_{jk} represent remainder (or error) terms. We have inserted the common factor h_k^j in the expansions for each derivative (j) so that all the remainder terms are of the order h_k^{J+1}, if u is, in fact, an analytic function.

We can also write the backward (from s_{k+1} to s_k) expansions:

$$\begin{aligned}
&u_{k+1} + d_k u'_{k+1} + \cdots + d_k^J u_{k+1}^{(J)}/J! - u_k = F_{1k}, \\
&d_k u'_{k+1} + d_k^2 u''_{k+1} + \cdots + d_k^J u_{k+1}^{(J)}/(J-1)! - d_k u'_k = F_{2k}, \\
&\qquad\vdots \qquad\qquad\qquad\qquad\qquad\qquad k - 1(1)K-1, \\
&\qquad\vdots \qquad\qquad\qquad\qquad\qquad\qquad (5) \\
&d_k^{J-1} u_{k+1}^{(J-1)} + d_k^J u_{k+1}^{(J)} - d_k^{J-1} u_k^{(J-1)} = F_{Jk},
\end{aligned}$$

where

$d_k = s_k - s_{k+1} = -h_k$, and the F_{jk} represent error terms of the order h_k^{J+1}.

We now think of Eq. (4) and (5) as being homogenous (zero right-hand sides) but involving possible errors. Then the linear system consisting of the "homogenized" equations (4) and (5) and Eq. (2) and (3) has

$J(K-1) + J(K-1) + K + J$ or $2KJ + K - J$ equations,

$KJ + K$ unknowns.

For $K \geq 2$, there are obviously more equations than unknowns. In order to admit the approximation errors, we consider the system to be overdetermined. Such an overdetermined system can be solved optimally by linear programming (see Example 3 in the Appendix).

We define the Chebyshev norm of the errors,

$$u_0 \equiv \max\,(|E_{jk}|,\ |F_{jk}|), \tag{6}$$

and seek to minimize u_0. (7)

For reasons that will be apparent, we formally replace Eq. (2) and (3) by inequalities,

$$\begin{aligned} a_0^{\ell}u_{\ell} + a_1^{\ell}u'_{\ell} + \cdots + a_{J-1}^{\ell}u_{\ell}^{(J-1)} &\geq c^{\ell} \\ -a_0^{\ell}u_{\ell} - a_1^{\ell}u'_{\ell} - \cdots - a_{J-1}^{\ell}u_{\ell}^{(J-1)} &\geq -c^{\ell} \end{aligned} \qquad \ell = 1(1)\,J, \tag{2'}$$

$$\begin{aligned} a_0^{k}u_{k} + a_1^{k}u'_{k} + \cdots + a_J^{k}u_{k}^{(J)} &\geq c^{k} \\ -a_0^{k}u_{k} - a_1^{k}u'_{k} - \cdots - a_J^{k}u_{k}^{(J)} &\geq -c^{k} \end{aligned} \qquad k = 1(1)\,K. \tag{3'}$$

We can deduce the following inequalities from our definition of u_0 in Eq. (6):

$$u_0 \pm E_{jk} \geq 0, \tag{6'}$$

$$u_0 \pm F_{jk} \geq 0, \tag{6''}$$

and then replace the E_{jk} and F_{jk} by their equivalents from Eqs. (4) and (5) to obtain

$$\begin{aligned} u_0 \pm (u_k + h_k u'_k + \cdots + h_k^J u_k^{(J)}/\,J! - u_{k+1}) &\geq 0, \\ &\vdots \\ u_0 \pm (h_k^{J-1}u_k^{(J-1)} + h_k^J u_k^{(J)} - h^{J-1}u_{k+1}^{(J-1)}) &\geq 0; \end{aligned} \tag{4'}$$

$$
\begin{aligned}
u_0 \pm (u_{k+1} + d_k u'_{k+1} + \cdots + d_k u^{(J)}_{k+1}/J! - u_k) &\geqslant 0, \\
&\vdots \\
u_0 \pm (d_k^{J-1} u^{(J-1)}_{k+1} + d_k u^{(J)}_{k+1} - d_k^{J-1} u_k^{(J-1)}) &\geqslant 0.
\end{aligned}
\tag{5'}
$$

If we now use (7) as our objective and the inequalities (2'), (3'), (4'), and (5') as restraints, we have a linear model in matrix-vector notation of the form,

$$
\begin{aligned}
&\text{minimize} \quad w = BU, \\
&\text{subject to} \quad AU \geqslant C, \\
&\text{with no restriction on the sign of} \quad U,
\end{aligned}
$$

where
the vector U consists of u_0 and solution and derivative values, the vector $B = (1, 0, \cdots, 0)$, the matrix A is the coefficient matrix for the inequalities, and the vector C is the constant terms from the inequalities.

The above linear model is the dual of a standard linear program (see p. 7) of the form,

$$
\begin{aligned}
&\text{maximize} \quad z = CX, \\
&\text{subject to} \quad A'X = B \quad \text{and} \quad X \geqslant 0,
\end{aligned}
$$

where
the matrix A' is the transpose of A, the variables X are for our purposes purely formal.

The computation of the linear program gives by duality the optimal values for the components of U; i.e., for the error and for the solution and derivatives up to the order J at each of the K mesh points.

If the solution u is in fact a polynomial of degree J or less, then there is a certainty that the value zero will be attained by u_0 in the linear program. Therefore, the linear program method gives us an exact result for such a polynomial, and we have completely established the following:

Theorem 7. If a well-posed linear problem for a linear ordinary differential equation of order J on a closed interval S is discretized for a mesh of two or more points in S, these mesh points including all points at which a limited solution is required and all points at which the linear conditions are prescribed, then there is a linear program whose computation minimizes approximation errors, gives the maximum absolute value of such errors, and provides approximate values for the solution and its first J derivatives at each of the mesh points. The approximation includes the required limited solution and is exact for a polynomial solution of degree J or less.

Except for a mesh of only two points, the last statement of the theorem does not fully reflect the accuracy of the approximation. For $K \geqslant 3$, we are actually approximating the solution by polynomial arcs rather than by a single polynomial.

As is the case with many numerical methods, improvements of the proposed procedure may suggest themselves. Other than mentioning the more obvious ones of mesh-refinement, of other approximation formulas and, where possible, of differentiation of the original equation, we shall leave variations of the procedure to the ingenuity and imagination of the reader.

Example 4 in the Appendix illustrates the linear program method as applied to a linear ordinary partial differential equation.

Linear Ordinary Differential System Problem

The procedure outlined in the previous section is readily extended to a well-posed linear problem for an ordinary differential system. There are, of course, practical limits on the number of equations, their orders, and the number of mesh points. All of the equations are assumed to hold on some common closed interval S of the independent variable s.

For the sake of notational simplicity, we shall limit our exposition to a linear system of two differential equations of second order. The extension to more elaborate systems is obvious.

We consider the linear ordinary differential system,

$$a_{10}(s)u + a_{11}(s)u' + a_{13}(s)v + a_{14}(s)v' + a_{15}(s)v'' = c_1(s); \quad (8)$$

$$a_{20}(s)u + a_{21}(s)u' + a_{22}(s)u'' + a_{23}(s)v + a_{24}(s)v' = c_2(s), \quad (9)$$

where
u and v are unknown functions of s, u', v', u'' and v'' are ordinary derivatives with respect to s, the a_{ij} and c_i are numerically defined and continuous on S, and a_{15} and a_{22} do not vanish on S.

For a well-posed linear problem, we must have four independent local conditions prescribed at points s^ℓ of S:

$$a^\ell_{10}u_\ell + a^\ell_{11}u'_\ell + a^\ell_{13}v_\ell + a_{14}v'_\ell = c^\ell; \quad \ell = 1(1)4, \quad (10)$$

where
the a^ℓ_{ij} and c^ℓ are real numbers, the superscript ℓ indices the conditions, and the subscript ℓ denotes the function value at s^ℓ.

We are required to find a table of values for u and v at a finite number of points s^* of S. We consider a set of at least two mesh points s_k, $k = 1(1)K$, in S which includes the points s^ℓ and s^*. The problem as discretized to the mesh points involves $6K$ unknown quantities, namely, the values $u_k, v_k, u'_k, v'_k, u''_k$, and v''_k at the K points s_k. When the differential equations (8) and (9) are applied at the mesh points, we obtain:

$$a^k_{10}u_k + a^k_{11}u'_k + a^k_{13}v_k + a^k_{14}v'_k + a^k_{15}v''_k = c^k_1; k=1(1)K; \quad (11)$$

$$a^k_{20}u_k + a^k_{21}u'_k + a^k_{22}u''_k + a^k_{23}v_k + a^k_{24}v'_k = c^k_2; k=1(1)K, \quad (12)$$

where
the a^k_{ij} and c^k_i are numerical values at s_k.

For the consecutive points s_k, s_{k+1}, $k = 1)1)K-1$, we can write the Taylor expansions from s_k to s_{k+1}:

$$\begin{aligned} u_k + h_k u'_k + h_k^2 u''_k/2 - u_{k+1} &= E_{1k}; \; k = 1(1)K-1, \\ h_k u'_k + h_k^2 u''_k - h_k u'_{k+1} &= E_{2k}; \; k = 1(1)K-1; \end{aligned} \quad (13)$$

$$v_k + h_k v'_k + h_k^2 v''_k/2 - v_{k+1} = E_{3k};\ k = 1(1)K-1,$$
$$h_k v' + h_k^2 v'' - h_k v'_{k-1} = E_{4k};\ k = 1(1)K-1, \tag{14}$$

and the expansions from s_{k+1} to s_k:

$$u_k - u_{k+1} + h_k u'_{k+1} - h_k^2 u''_{k+1}/2 = F_{1k};\ k = 1(1)K-1,$$
$$h_k u'_k - h_k u'_{k+1} + h_k^2 u''_{k+1} = F_{2k};\ k = 1(1)K-1; \tag{15}$$

$$v_k - v_{k+1} + h_k v'_{k+1} - h_k^2 v''_{k+1}/2 = F_{3k};\ k = 1(1)K-1,$$
$$v_k - v_{k+1} + h_k v'_{k+1} - h_k^2 v''_{k+1}/2 = F_{3k};\ k = 1(1)k-1,$$
$$h_k v'_k = h_k v'_{k+1} + h_k^2 v''_{k+1} = F_{4k};\ k = 1(1)K-1, \tag{16}$$

where

$h_k = s_{k+1} - s_k$, and the E_{jk} and F_{jk} represent error terms.

As in the previous section, we define

$$u_0 = \max\ (|E_{jk}|, |F_{jk}|), \tag{17}$$

and seek to minimize u_0. (18)

We formally replace Eq. (10), (11), and (12) by the inequalities,

$$a^{\ell}_{10} u_{\ell} + a^{\ell}_{11} u'_{\ell} + a^{\ell}_{13} v_{\ell} + a^{\ell}_{14} v'_{\ell} \geq c^{\ell};\ \ell = 1(1)\,4,$$
$$-a^{\ell}_{10} u_{\ell} - a^{\ell}_{11} u'_{\ell} - a^{\ell}_{13} v_{\ell} - a^{\ell}_{14} v'_{\ell} \geq -c^{\ell};\ \ell = 1(1)\,4; \tag{10'}$$

$$a^{k}_{10} u_k + a^{k}_{11} u'_k + a^{k}_{13} v_k + a^{k}_{14} v'_k + a^{k}_{15} v''_k \geq c^{k}_1;\ k = 1(1)\,K,$$
$$-a^{k}_{10} u_k - a^{k}_{11} u'_k - a^{k}_{13} v_k - a^{k}_{14} v'_k - a^{k}_{15} v''_k \geq -c^{k}_1;\ k = 1(1)\,K; \tag{11'}$$

$$a^{k}_{20} u_k + a^{k}_{21} u'_k + a^{k}_{22} u''_k + a^{k}_{23} v_k + a^{k}_{24} v'_k \geq c^{k}_2,\ k = 1(1)\,K,$$
$$-a^{k}_{20} u_k - a^{k}_{21} u'_k - a^{k}_{22} u''_k - a^{k}_{24} v_k - a^{k}_{24} v'_k - -c^{k}_2;\ k = 1(1)\,K. \tag{12'}$$

From our definition (17), we have

$$u_0 \pm E_{jk} \geq 0,$$
$$u_0 \pm F_{jk} \geq 0,$$

which we can use with Eq. (13), (14), (15), and (16) to obtain

$$u_0 \pm (u_k + h_k u'_k + h_k^2 u''_k/2 - u_{k+1}) \geq 0;\ k = 1(1)K-1,$$
$$u_0 \pm (h_k u'_k + h_k^2 u''_k - h_k u'_{k+1}) \geq 0;\ k = 1(1)K-1; \tag{13'}$$

$$u_0 \pm (v_k + h_k v'_k + h_k^2 v''_k/2 - v_{k+1}) \geq 0;\ k = 1(1)K-1,$$
$$u_0 \pm (h_k v'_k + h_k^2 v''_k - h_k v'_{k+1}) \geq 0;\ k = 1(1)K-1; \tag{14'}$$

$$u_0 \pm (u_k - u_{k+1} + h_k u'_{k+1} - h_k^2 u''_{k+1}/2) \geq 0;\ k = 1(1)K-1,$$
$$u_0 \pm (h_k u'_k - h\, u'_{k+1} + h_k^2 u''_{k+1}) \geq 0;\ k = 1(1)K-1; \tag{15'}$$

$$u_0 \pm (v_k - v_{k+1} - h_k v'_{k+1} + h_k^2 v''_{k+1}/2) \geq 0;\ k = 1(1)K-1,$$
$$u_0 \pm (h_k v'_k - h_k v'_{k-1} + h_k^2 v''_{k-1}) \geq 0;\ k = 1(1)K-1, \tag{16'}$$

The inequalities (10'), (11'), (12'), (13'), (14'), (15'), and (16'), with the objective (18), constitute a linear model which is the dual of a standard linear program. By virture of duality, the computation of the linear program gives us coincidentally the optimal values for the maximum absolute error and for the solutions and their derivatives at the mesh points.

If the solution functions u and v are in fact parabolas the value of zero is attained by u_0. In this case, the linear program method gives an exact result for the solutions. For the problem described, we have established:

Theorem 8. If a well-posed linear problem for a linear ordinary differential system on a closed interval S is discretized for a mesh of two or more points in S, these mesh points including all points at which linear conditions are prescribed, then there is a linear program which gives the maximum absolute error of approximation and the

approximate values of the unknown solution functions and some of their derivatives at each of the mesh points. For a solution function whose Jth derivative appears in the system, the approximation is exact if this function is in fact a polynomial of degree J.

Example 5 in the Appendix illustrates the application of the linear program to a system of linear ordinary differential equations.

Linear Partial Differential Equation Problem--Two Independent Variables, First-Order Linear Partial Differential Equation

We consider a partial differential equation which holds on a closed rectangular domain R and has the form,

$$a^o(s,t)u + a^s(s,t)u_s + a^t(s,t)u_t = c(s,t), \tag{19}$$

where

u is an unknown function of (s, t), u_s and u_t are first partial derivatives, the a^a and c are numerically defined and continuous on R, and a^s and a^t do not vanish on R.

Because we have two unknown derivative functions in Eq. (19), we say this is a "two-condition" problem. The conditions may be prescribed on various point sets. We shall cover the possibilities rather generally for lattice points (s_k, t_k), $k = 1(1)\mathcal{K}$, and $k = 1(1)K$, by

(1) On any line $t = t_k$ considered, we prescribe two independent local conditions at a point or points (s^l, t_k) of R. These conditions have the form,

$$a^o_{lk}u^{lk} + a^t_{lk}u^{lk}_t = c_{lk};\ l = 1, 2,\ k = 1(1)K \geq \mathcal{K}, \tag{20}$$

where

the a^a_{lk} and c_{lk} are real numbers.

(2) On any line $s = s_k$ considered, we prescribe two independent local conditions at a point or points (s_k, t^l) of R. These conditions have the form:

$$a^o_{kl}u^{kl} + a^s_{kl}u^{kl}_s = c_{kl};\ l = 1, 2,\ k = 1(1)\mathcal{K} \geq K, \tag{21}$$

where
the $a^a_{\ell l}$ and $c_{\ell l}$ are real numbers.

(3) On any line $t - t_k$ considered, we prescribe condition (20) for $\ell = 1$, and on any line $s = s_\ell$ considered, we prescribe condition (21) for $\ell = 2$.

We are required to find a limited solution consisting of numerical values of u for some finite set of points (s^*, t^*) in R. We consider a set of $\mathcal{K}$ K lattice points, (s_ℓ, t_k), $\ell = 1(1)\mathcal{K}$, $k = 1(1)K$, $\mathcal{K} \geq 2$, $K \geq 2$. This set must include all the points (s^*, t^*), and the lines $s = s^*$ and/or $t = t^*$ must be considered under the pertinent prescribed-condition case. Further, the closed rectangular domain $R^*: s_1 \leq s \leq s_{\mathcal{K}}, t_1 \leq t \leq t_K$ must lie entirely in R, and the set of lattice points must include pertinent points (s^l, t_k) and/or (s_ℓ, t^l), for which conditions are prescribed.

The differential equation (19) applied at each of the lattice points yields $\mathcal{K}$ K algebraic equations:

$$a^o_{\ell k} u^{\ell k} + a^s_{\ell k} u^{\ell k}_s + a^t_{\ell k} u^{\ell k}_t = c_{\ell k}; \quad \ell = 1(1)\mathcal{K}, \; k = 1(1)K, \tag{22}$$

where
the $a^a_{\ell k}$ and $c_{\ell k}$ are real numerical values at (s_ℓ, t_k), and the superscripts ℓk indicate function values at (s_ℓ, t_k).

On every line $t = t_k$, $k = 1(1)K$, for consecutive values s_ℓ, $s_{\ell+1}$; for $\ell = 1(1)\mathcal{K}-1$, construct the Taylor expansions:

$$u^{\ell k} + h_\ell u^{\ell k}_s - u^{\ell+1, k} = \mathcal{E}_{\ell k}; \quad \ell = 1(1)\mathcal{K}-1, \; k = 1(1)K; \tag{23}$$

$$u^{\ell k} - u^{\ell+1, k} + h_\ell u^{\ell+1, k}_s = \mathcal{F}_{\ell k}; \quad \ell = 1(1)\mathcal{K}-1, \; k = 1(1)K, \tag{24}$$

where
$h_\ell = s_{\ell+1} - s_\ell$, and the $\mathcal{E}_{\ell k}$ and $\mathcal{F}_{\ell k}$ are error terms.

Similarly, for every line $s = s_\ell$, $\ell = 1(1)\mathcal{K}$ and for t_k, $k = 1(1)K-1$, we obtain Taylor expansions,

$$u^{\ell k} + h_k u_t^{\ell k} - u^{\ell, k-1} = E_{\ell k}; \ell = 1(1)\mathcal{L}, \; k = 1(1)K-1; \quad (25)$$

$$u^{\ell k} - u^{\ell, k+1} + h_k u_t^{\ell, k+1} = F_{\ell k}; \ell = 1(1)\mathcal{L}, \; k = 1(1)K-1, \quad (26)$$

where

$h_k = t_{k+1} - t_k$, and the $E_{\ell k}$ and $F_{\ell k}$ are error terms.

We now proceed as in the previous sections. We regard Eq. (23), (24), (25), and (26) as being homogeneous but subject to some errors which we wish to minimize. When we use these equations along with (20) and/or (21) as pertinent and (22), we have an overdetermined system whose unknowns are solution and first partial derivative values at the lattice points.

We define

$$u_0 = \max\left(|\mathcal{E}_{\ell k}|, |\mathcal{F}_{\ell k}|, |E_{\ell k}|, |F_{\ell k}|\right),$$

and convert the equations of the previous paragraph to linear inequalities ($\geq$) involving u_0 and the solution and derivative values. Our objective is to minimize u_0.

Computation of the linear program, for which the linear model of the previous paragraph is the dual, gives, by virtue of duality, an optimal value for u_0 and approximations for u, u_s and u_t at each of the lattice-points.

For the cases considered, we have established the following:

Theorem 9. If a well-posed linear problem for a linear first-order partial differential equation with two independent variables on a closed rectangle is discretized for a nontrivial rectilinear lattice, then there is a linear program whose computation minimizes the error of approximation and provides approximate values for the solution function and its first partial derivatives at the lattice points.

The procedure described above is illustrated in Example 6 in the Appendix. Possible improvements on the approximation and modifications for other domains and conditions we leave to the ingenuity of the reader.

Second-Order Linear Partial Differential Equation

For second-order (or higher) linear partial differential equations in even two independent variables, there is a lack of comprehensive theory on the well-posed problem. The equations are traditionally subclassified, and even within the subclassification it is common to consider very special problems. From the physical point of view, this specialization is a natural one arising from the situation for which the differential problem is a mathematical model. For the mathematician, the specialization is useful in that the special properties of the problem may make it possible to decide on a domain of solution, to determine the nature of the prescribed conditions, to make assumptions concerning the solution, and even, in some cases, to devise an analytic approach. We shall return to special problems later.

We now consider a general linear second-order partial differential equation in the two independent variables s and t that holds on a closed rectangular domain R: $S \le s \le T$ and has the form,

$$a^{0}u + a^{s}u_{s} + a^{t}u_{t} + a^{ss}u_{ss} + a^{st}u_{st} + a^{tt}u_{tt} = c, \qquad (27)$$

where

u is an unknown function of s and t, u_s, u_t, u_{ss}, u_{st} and u_{ss} are its partial derivatives, a^{o}, a^{s}, a^{t}, a^{ss}, a^{tt} and c are numerically defined, continuous functions of s, t on R, and a^{ss}, a^{st} and a^{tt} do not vanish on R.

Because we have five unknown derivative functions in Eq. (27), we say this is a "five-condition" problem. Conditions may be prescribed at various point sets. For a discretization of the problem for lattice points (s_ℓ, t_k), $\ell = 1(1)\mathcal{L}$; $k = 1(1)K$, in R, we cover the possibilities rather generally by the following cases:

1. On the lines $t = t_k$ considered, we prescribe five independent local(ized) conditions at points $(s_{\ell k}, t_k)$ of the form,

$$a^{o}_{\ell 1}u^{\ell k} + a^{s}_{\ell k}u^{\ell k}_{s} + a^{t}_{\ell k}u^{\ell k}_{t} + a^{st}_{\ell k}u^{\ell k}_{st} + a^{tt}_{\ell k}u^{\ell k}_{tt} = c_{\ell k};$$

$$\ell k = 1(1)5, \quad k = 1(1)K \ge \mathcal{K}. \qquad (28)$$

2. On the lines $s = s_k$ considered, we prescribe five independent local(ized) conditions at points (s_k, t_{kl}) of the form,

$$a^{o}_{kl}u^{kl} + a^{s}_{kl}u^{kl}_{s} + a^{t}_{kl}u^{kl}_{t} + a^{ss}_{kl}u^{kl}_{ss} + a^{st}_{kl}u^{kl}_{st} = c_{kl};$$

$$k = 1(1)\mathcal{K},\ kl = 1(1)5, \qquad (29)$$

$$\mathcal{K} \geq K.$$

3. We prescribe five conditions from (28) and (29).

The set of lattice-points (s_k, t_k) considered must include all the points (s^*, t^*) at which a limited solution is required and must include pertinent prescription points (s_{lk}, t_k) and/or (s_k, t_{kl}).

The differential equation (27) applied at the lattice points yields

$$a^{o}_{kk}u^{kk} + a^{s}_{kk}u^{kk}_{s} + a^{t}_{kk}u^{kk}_{t} + a^{ss}_{kk}u^{kk}_{ss} + a^{st}_{kk}u^{kk}_{st} + a^{tt}_{kk}u^{kk}_{tt} = c_{kk}. \qquad (30)$$

Along each line $t = t_k$, $k = 1(1)K$, for consecutive values s_k, s_{k+1}, we have the following Taylor expansions:

$$u^{kk} + h_k u^{kk}_{s} + h^2_k u^{kk}_{ss}/2 - u^{k+1,k} = \mathcal{E}^{o}_{k};\ k = 1(1)\mathcal{K}-1, \qquad (31)$$

$$h_k u^{kk}_{s} + h^2_k u^{kk}_{ss} - h_k u^{k+1,k}_{s} = \mathcal{E}^{s}_{kk};\ k = 1(1)\mathcal{K}-1; \qquad (32)$$

$$h_k u^{kk}_{t} + h^2_k u^{kk}_{st} - h_k u^{k+1,k}_{t} = \mathcal{E}^{t}_{kk};\ k = 1(1)\mathcal{K}-1; \qquad (33)$$

$$u^{kk} - u^{k+1,k} + h_k u^{k+1,k}_{s} - h^2_k u^{k+1,k}_{ss}/2 = \mathcal{F}^{o}_{kk};\ k = 1(1)\mathcal{K}-1; \qquad (34)$$

$$h_k u^{k}_{s} - h_k u^{-1,k}_{s} + h^2_k u^{-1,k}_{ss} = \mathcal{F}^{s}_{kk};\ k = 1(1)\mathcal{K}-1; \qquad (35)$$

$$h_k u^{k}_{t} - h_k u^{-1,k}_{t} + h^2_k u^{-1,k}_{st} = \mathcal{F}^{t}_{kk};\ k = 1(1)\mathcal{K}-1, \qquad (36)$$

where

$$h_k = s_{k+1} - s_k.$$

Similarly, along the lines $s = s_\ell$, $\ell = 1(1)\mathcal{L}$, we have:

$$u^{\ell k} + h_k u_t^{\ell k} + h_k^2 u_{tt}^{\ell k}/2 - u^{\ell, k+1} = E^o_{\ell k};\ k = 1(1)K-1; \tag{37}$$

$$h_k u_s^{\ell k} + h_k^2 u_{st}^{\ell k} - h_k u_s^{\ell, k+1} = E^s_{\ell k};\ k = 1(1)K-1; \tag{38}$$

$$h_k u_t^{\ell k} + h_k^2 u_{tt}^{\ell k} - h_k u_t^{\ell, k+1} = E^t_{\ell k};\ k = 1(1)K-1; \tag{39}$$

$$u^{\ell k} - u^{\ell, k+1} + h_k u_t^{\ell, k+1} - h_k^2 u_{tt}^{\ell, k+1}/2 = F^o_{\ell k};\ k = 1(1)K-1; \tag{39}$$

$$u^{\ell k} - u^{\ell, k+1} + h_k u_t^{\ell, k+1} - h_k^2 u_{tt}^{\ell, k+1}/2 = F^o_{\ell k};\ k = 1(1)K-1; \tag{40}$$

$$h_k u_s^{\ell k} - h_k u_s^{\ell, k+1} + h_k^2 u_{st}^{\ell, k+1} = F^s_{\ell k};\ k = 1(1)K-1; \tag{41}$$

$$h_k u_t^{\ell k} - h_k u_t^{\ell, k+1} + h_k^2 u_{tt}^{\ell, k-1} = F^t_{\ell k};\ k = 1(1)K-1, \tag{42}$$

where

$h_k = t_{k+1} - t_k$.

As before, we regard Eq. (31) to (42) as being homogeneous but subject to errors, which we wish to minimize. These equations along with Eq. (28) and/or (29) as pertinent, together with Eq. (30), constitute an overdetermined linear algebraic system whose unknowns are the values of u, u_s, u_t, u_{ss}, u_{st} and u_{tt} at the lattice points.

We define

$$u_0 = \max\left(|\mathcal{E}_k|, |\mathcal{F}_{\ell k}|, |E_{\ell k}|, |F_{\ell k}|\right),$$

and convert to a system of inequalities ($\geq$) involving u_0 and the solution and derivative values. Our objective is to minimize u_0.

Computation of the linear program for which the foregoing linear model is the dual gives, by the duality principle, the optimal value for u_0 and approximate values for u, u_s, u_t, u_{ss}, u_{st} and u_{tt} at the (s_ℓ, t_k).

For the cases considered, we have established:

Theorem 10. If a well-posed linear problem for a linear second-order partial differential equation in two independent variables on a closed rectangle is discretized for a nontrivial rectilinear lattice, then there is a linear program whose computation minimizes the error of approximation and provides approximate values for the solution function and its first and second partial derivatives at the lattice points.

We leave refinements and improvements on the foregoing to the reader and turn now to subclassification of the second-order equation.

A linear second-order differential equation in two independent variables, whose discriminant H defined by

$$H = (a^{st})^2 - 4a^{ss}a^{tt},$$

does not change sign on the two-dimensional domain being considered, may be classified as parabolic, hyperbolic, or elliptic.

Parabolic Equations. If H is everywhere zero on the domain, the equation is said to be parabolic. An equation of sufficient generality for this class is the following (Ref. 27, p. 107, Ref. 29, p. 73):

$$a^{o}u + a^{s}u_{s} - u_{t} + u_{ss} = c. \tag{43}$$

Problems involving this equation may be "three-condition" on a rectangle whose sides are parallel to the coordinate axes. Care must be exercised in the discretization. A safe rule to follow is to require that every point (s_{ℓ}, t_{k}), the grid step in t in either direction, be less than or equal to twice the square of the grid step in s in either direction (Ref. 27, p. 92ff.). This precaution will insure that errors do not compound to render the numerical solution unstable. Enumeration of all the possible conditioning which would provide a well-posed problem will not be attempted here. Three possibilities may occur: (1) the problem may be underconditioned; (2) the problem may be properly conditioned; and (3) the problem may be overconditioned. In (1) the linear program

method will find a solution which satisfies all the prescribed conditions. Such a solution is not likely to be unique, nor is it likely to be much of an approximation to the true solution sought. In (2), the conditioning may be barely sufficient, so that approximation formulas must be treated as exact equations, and there is no opportunity for minimization of approximation error. In this instance, the solution may be found by solving a linear system of algebraic equations for its unique solution. Otherwise the linear program method may be used to minimize the approximation errors. In (3) it is unlikely that the conditioning will be precisely consistent because of errors in measurement or in rounding of numerical values for functions. The linear program method then can be used either by allowing tolerances on the numerical values, or by allowing these errors to be minimized by the program. Without one or the other of the above, the linear program will have an unbounded solution because of the inconsistency.

In many problems, particularly where uniform grid steps are used, it is more accurate and efficient to employ approximation formulas other than truncated Taylor series. Central finite-difference approximations, for example, may be used.

The numerical solution of several parabolic problems is illustrated in Example 7 of the Appendix. As with most numerical methods, the user must exercise some ingenuity in applying this method to a specific problem.

Hyperbolic Equations. If H is everywhere strictly positive on the domain being considered, the equation is said to be hyperbolic. An equation of sufficient generality for this class is the following (Ref. 29, p. 73):

$$a^{o}u + a^{s}u_{s} + a^{t}u_{t} + u_{ss} + u_{tt} = c. \tag{44}$$

Problems on this equation may be "three-condition," or even "two-condition," on domains bounded by appropriate characteristic curves of the equation. Conditioning should not be along any line or curve whose

direction at any point of the domain is the same as that of a characteristic curve through the same point.

The only safe rule for discretization for any conditioning is the use of uniform equal grid steps in s and t. Whether it is safe to use a larger grid step for one variable than the other depends on the conditioning (Ref. 27, l. 18).

The remarks on conditioning given in connection with parabolic equations apply here. For illustrations on the solution of hyperbolic problems see Example 8 in the Appendix.

If the discriminant H is strictly negative on the two-dimensional domain considered, then the linear second-order partial diffential equation in two independent variables is said to be elliptic. For such equations, we can, without loss of generality, restrict ourselves to (see Ref. 29, p. 72):

$$a^o u + a^s u_s + a^t u_t + u_{ss} + u_{tt} = c. \tag{45}$$

Since the partial derivative u_{st} does not appear, a problem involving Eq. (45) may be a "four-condition" problem, even without limiting the considered domain to the rectangle R: SXT.

We consider a closed two-dimensional domain D whose boundary consists of the closed continuous curve(s) Γ. We denote by D^o the open region D less Γ. From the historical and practical standpoint, the problems involving the elliptic differential equation are those in which conditions are prescribed along Γ, while the equation is assumed to hold only on D^o. Traditionally, the problems treated have been highly special as to the equation, domain, and conditions. We shall attempt a rather general treatment.

By means of lines $s = s_k$ and $t = t_k$, we construct a rectilinear grid over the domain D. We speak of portions of these lines which lie in D as being included segments, and insist that each of these have two points in common with Γ and that each be intersected by at least one orthogonal included segment at a point in D^o. We discretize the problem for the set of points consisting of (a) the intersections of

included segments with Γ (called boundary points), and (b) the intersections of included segments with each other (called lattice points). We seek a limited solution consisting of values for u, u_s, u_t, u_{ss} and u_{tt} at the lattice-points subject to linear local(ized) conditions at the boundary points. These conditions may have the form,

$$a_\ell^o u^\ell + a_\ell^s u_s^\ell + a_\ell^t u_t^\ell = c_\ell; \quad \ell = 1(1)2L, \tag{46}$$

where

a_ℓ^o, a_ℓ^s, a_ℓ^t and c_ℓ are real numbers, nonzero, u^ℓ, u_s^ℓ and u_t^ℓ represent values at boundary points, ℓ indices the local(ized) conditions, and L is the number of included segments.

Each included segment has at least one lattice point, and about each such point we can write Taylor expansions along the segment for u and a first partial derivative (either u_s or u_t, according to the direction of the segment) at the two adjacent discrete points (either lattice points or boundary points). Each included segments has two boundary points and from each of these we can write a (first-order) Taylor expansion for u at the adjacent lattice point. These expansions, considered as homogeneous equations but subject to approximation error, together with Eq. (46) and Eq. (45) applied at the lattice points, provide us with a linear system whose unknowns are the values of u, u_s and u_t at the boundary points and of u, u_s, u_t, u_{ss} and u_{tt} at the lattice points. With the objective of minimizing the maximum-error magnitude, we have a linear model which is the dual of a linear program. Computation of this linear program will minimize the error and give approximate values for aforementioned unknowns.

Many problems involving elliptic equations do not have boundary conditions prescribed that meet the nonzero "coefficient" requirement imposed on Eq. (46). These cases may require special treatment in setting up the Taylor expansions. For example, if the value of u is prescribed on the boundary, we have localized conditions of the form

$$u_\ell = c_\ell; \quad \ell = 1(1)2L. \tag{47}$$

For this problem, the Taylor expansions from the boundary points should be omitted, the numerical value c_ℓ can be substituted for u_ℓ wherever it appears in the other expansions, and the result will be an overdetermined linear system obtained from the expansions and from Eq. (45) applied at the lattice points whose unknowns are the values of u, u_s, u_t, u_{ss} and u_{tt} at the lattice points only. Similar modification can be made in other cases. The solution of problems involving elliptic equations is illustrated in Example 9 of the Appendix.

Problems for Linear Partial Differential Systems

The procedure of the previous section can be extended to some problems involving a system of linear partial differential equations. There is, of course, a practical limit on the number and orders of the differential equations. In the extension, the constructed linear system involves, as unknowns, the values of all the solution functions and some of their partial derivatives at some set of discrete points. Obviously a solution can be attempted only on a domain having common propriety for all the equations. Any general discussion of necessary conditioning for a well-posed problem is beyond our scope. Even without this question, a detailed description of the linear program method for a problem of any generality would be lengthy, notationally difficult, and, in a great measure, repetitive. For these reasons, we content ourselves with a simple illustration in Example 10 of the Appendix. Following this example as a guide, the ready may apply the method to other, not-too-elaborate problems involving linear partial differential systems.

Use of Computers

The formulations in the preceding sections, frightening in their literal expression, may be readily codified for high-speed computers. Rather general but simple codes can be written which will convert a minimum amount of problem information into suitable input for an existing linear programming code. Some such codes have been written by the author and used in connection with the examples given in the Appendix. Preparatory codes together with available linear program codes such as SCROL (Ref. 30) reduce the human participation in the computation to the preparation of a few data and control cards.

V. CONCLUSIONS

In this chapter, we attempt to draw some conclusions with respect to the merits of the linear program method for the numerical solution of linear differential problems. The following properties are considered: stability, accuracy, convenience, and versatility. Special features of the method, which are not possessed by other methods, are discussed.

Stability is a property of the approximation formulas used, of the problem itself, and of the discretization. A problem is said to be stable for a particular approximation if the small errors, which are almost certain to appear in rounding and truncation, have relatively small effect on the final result. For a thorough discussion of stability (in the case of difference equations), the reader is referred to Forsythe and Wasow (Ref. 27, pp. 29 to 35). Because the linear program method may employ several approximation formulas on a particular problem (see Examples 7, 8, 9, 10 in the Appendix), it inherits their stability attributes. In this method, it is possible to choose the approximation formulas least likely to be unstable, or to use one approximation formula to offset the instability of another. Unfortunately, and contrary to a too popular belief, finer discretization or the use of higher-order approximation formulas do not always improve stability. Unlike most numerical methods that are satisfied with a determined numerical approximation, the linear program method operates on an overdetermined situation. In the set of approximations employed, it need not force compliance with any approximation formula, but rather can minimize departures from several. This advantage of the linear program method is not present in other non-iterative methods for the numerical solution of differential problems.

Unlike stability, accuracy is concerned with the inception of errors rather than with their growth. It is generally assessed for a small domain containing a few points of discretization. Sometimes accuracy is expressed by a statement that the numerical approximation is exact for a polynomial of some specified degree. In other cases an error term is formulated. The linear program method derives its accuracy from the approximation formulas it employs. In the previous

discussion nonuniform mesh or grid steps have been stressed. Flexibility in discretization is desirable in many problems. However, if the mesh or grid steps are uniform, the linear program method is usually able to effect a better approximation. For example, for an ordinary differential equation of order J, by use of only backward and forward Taylor Expansions with $\pm u_0$ replacing all terms above the containing the Jth derivative, the linear program method gives an approximation which is exact if the solution is a polynomial of degree J+1. This is because the error term u_0 is forced to take the absolute value of $h^{J+1}u^{(J+1)}/(J+1)!$, which is a constant for such a polynomial. Again the use of diverse approximation formulas makes possible the reduction of the error that would be produced by any one of them.

The linear program method is not very convenient to use. The construction of an overdeterminate system is certain to require more formulation than would be the case for a determinate system. The fact that the linear model evolved, which is the dual of a standard linear program, must consist of linear inequalities ($\geq$) in effect doubles the amount of formulation. Thus the linear model that represents a fairly simple problem may appear too expansive to be practical (see Example 4, Appendix). Certainly the linear program method should not be applied where much simpler but entirely satisfactory methods are available. Computer time is another factor which should be considered under convenience. The computation of a linear program of any size is likely to be more time-consuming than a simple step-by-step noniterative method. For ordinary differential problems with initial conditions, the latter should be used unless, of course, they lack sufficient accuracy or are unstable. For boundary condition problems on ordinary differential equations, the linear program method is probably no more awkward than other existing methods (e.g., Ref. 26, pp. 67 through 79). The result that the linear program method supplies approximations both for derivatives and for solutions may be sufficiently useful to offset some or all of the inconvenience. Except for the very rare problems for which an analytic

approach is possible, all methods for solving partial differential equations are tedious. Usually the solution of a system of linear algebraic equations is necessary. Where this is the case, the linear program method is not much more elaborate or expensive in computer time.

The linear program method is restricted to linear differential problems. Aside from this admittedly serious limitation, the linear program method can be applied very generally to well-posed problems on linear ordinary, or partial differential equations or systems. It can be applied to any linear differential problem for which there is an applicable existing method. It can, in fact, incorporate two or more such methods, and apply them, not independently to small sets of mesh or lattice points but simultaneously to all such sets over the entire domain.

The linear program method is unique in its ability to use discrete measured data for prescribed conditions. No other method which relies on a determined linear model can allow for errors in measurement or satisfy an overconditioned problem with minimum discrepancy. Examples 7 and 10 on partial differential equations illustrate this power.

Another advantage of the linear program method not common in other methods is its ability to utilize a nonuniform mesh or grid. This property can be important where measurements to provide conditioning data can not be made uniformly, as on the boundaries of the domain for the elliptic partial differential equation in Example 10.

In the last analysis the merits of any numerical method must be determined in extensive application to a great diversity of problems. This work must be suggestive rather than conclusive. The linear program method presents great opportunity for the exercise of ingenuity. Final judgment on its value is relegated to those who may enjoy experimentation in a fresh approach to the very important and often perplexing task of approximating the solution of a well-posed differential problem.

APPENDIX

Example 1. Illustration of Simplex Computation

A company has three warehouses. Because of remodeling, it is necessary to transfer a certain quantity of material from Warehouse No. 3 to the other warehouses. The movement of material is to be made at least cost.

Formulation

Activities:

(1) Movement of material from Warehouse No. 3 to Warehouse No. 1,

(2) Movement of material from Warehouse No. 3 to Warehouse No. 2.

Variables:

x_1 Quantity moved from No. 3 to No. 1,

x_2 Quantity moved from No. 3 to No. 2,

y Total cost to be minimized.

Constants:

b_1 Available storage at No. 1,

b_2 Available storage at No. 2,

b_3 Quantity to be moved from No. 3,

c_1 Cost to move unit quantity from No. 3 to No. 1,

c_2 Cost to move unit quantity from No. 3 to No. 2.

Relations:

$y = c_1x_1 + c_2x_2$ to be minimized,

$x_1 \leqslant b_1$,

$x_2 \leqslant b_2$,

$x_1 + x_2 \geqslant b_3$,

$x_1,\ x_2 \geqslant 0$.

We introduce nonnegative slack variables x_3, x_4, and x_6 to obtain equality, and an artificial variable x_5 to construct an initial feasible basis. The latter must also be included in the cost function with a predominantly large cost coefficient M. With a new objective function z (z = - y) we have the following:

Standard Model:

$$\begin{aligned} &\text{maximize} && z = -c_1x_1 - c_2x_2 - Mx_5, \\ &\text{subject to} && x_3 + x_1 = b_1, \\ & && x_4 + x_2 = b_2, \\ & && x_5 - x_6 + x_1 + x_2 = b_3, \\ & && x_i \geq 0; \quad i = 1, \cdots, 6. \end{aligned}$$

Implementation

Suppose $c_1 = \$5$ and $c_2 = \$10$, then we have for our objective

$$\text{maximize} \qquad z = -5x_1 - 10x_2 - Mx_5.$$

We consider the following cases:

(a) $b_1 = 100$; $b_2 = 200$; $b_3 = 400$, that is,

$$\begin{aligned} x_3 + x_1 &= 100, \\ x_4 + x_2 &= 200, \\ x_5 - x_6 + x_1 + x_2 &= 400, \\ x_i \geq 0; \quad i &= 1, \ldots, 6; \end{aligned}$$

(b) $b_1 = 200$; $b_2 = 300$; $b_3 = 400$, that is,

$$\begin{aligned} x_3 + x_1 &= 200, \\ x_4 + x_2 &= 300, \\ x_5 - x_6 + x_1 + x_2 &= 400, \\ x_i \geq 0; \quad i &= 1, \ldots, 6. \end{aligned}$$

Computation

The computation involved in the Simplex process is more easily followed if the data is arranged in tableau form.

(a) Computation for case (a) above:

Tableau I

	c_j	0	0	-M	0	-5	-10	
c_i	Basis	P_3	P_4	P_5	P_6	P_1	P_2	B
0	P_3	1	0	0	0	1*	0	100
0	P_4	0	0	0	0	0	1	200
-M	P_5	0	1	-1	-1	1	1	400
	z_j	0	0	-M	M	-M	-M	
	$w_j=z_j-c_j$	0	0	M	M	5-M	10-M	

Basic Solution

$x_3 = 100$

$x_4 = 200$

$x_5 = 400$

$x_1, x_2, x_6 = 0$

Since -M is predominantly negative, we have

$w_1 = \min w_j < 0$; P_1 is to enter the basis, and

$R_3 = \min R_i \geq 0$; P_3 is to be replaced. The new basis is P_1, P_4, P_5. The pivot element is indicated by an asterisk.

Tableau II is constructed as follows: The first row is designated as P_1 and computed by dividing the corresponding elements of the first row (P_3) of Tableau I by the pivot element; other elements, a'_{ij} (row P_i, column P_j) are computed from a_{ij} (row P_i, column P_j of I), a_{i1} (row P_i, column P_1 of I), and a'_{1j} (row P_1, column P_j) by

$$a'_{ij} = a_{ij} - a_{i1} a'_{1j} .$$

This computation, which seems quite tedious even for the simple problem under discussion, is usally done on high-speed computers.

We now have:

Tableau II

	c_j	0	0	-M	0	-5	-10	
c_i	Basis	P_3	P_4	P_5	P_6	P_1	P_2	B
-5	P_1	1	0	0	0	1	0	100
0	P_4	0	1	0	0	0	1	200
-M	P_5	-1	0	1	-1	0	1	300
	z_j	M -5	0	-M	M	-5	-M	
	w_j	M -5	0	0	M	0	10 -M	

Basic Solution

$x_1 = 100$

$x_4 = 200$

$x_5 = 300$

$x_2, x_3, x_6 = 0$

$w_2 = \min w_j < 0$ and $R_4 = \min R_i$; P_2 replaces P_4 in the basis.

Tableau III is derived from Tableau II by the process previously described.

Tableau III

	c_j	0	0	-M	0	-5	-10	
c_i	Basis	P_3	P_4	P_5	P_6	P_1	P_2	B
-5	P_1	1	0	0	0	1	0	100
-10	P_2	0	1	0	0	0	1	200
-M	P_5	-1	-1	1	-1	0	0	100
	z_j	M -5	M -10	-M	M	-5	-10	
	w_j	M -5	M -10		M	0	0	

Basic Solution

$x_1 = 100$

$x_2 = 200$

$x_5 = 100$

$x_3, x_4, x_6 = 0$

Now all $w_j \geq 0$, hence no new vector can enter the basis; however, the vector P_5, which corresponds to the artificial variable x_5, is still in the basis. This model is infeasible.

(b) Computation for case (b):

Tableau I

	c_j	0	0	-M	0	-5	-10	
c_i	Basis	P_3	P_4	P_5	P_6	P_1	P_2	B
0	P_3	1	0	0	0	1	0	200
0	P_4	0	1	0	0	0	1	300
-M	P_5	0	0	1	-1	1	1	400
	z_j	0	0	-M	M	-M	-M	
	w_j	0	0	0	M	5 -M	10 -M	

Basic Solution

$x_3 = 200$

$x_4 = 300$

$x_5 = 400$

$x_1, x_2, x_6 = 0$

$w_1 = \min\ w_j < 0$; P_1 to enter basis,

$R_3 = \min\ R_i \geqslant 0$; P_3 to be replaced.

New basis is P_1, P_4, P_5.

Tableau II

	c	0	0	-M	0	-5	-10	
c	Basis	P_3	P_4	P_5	P_6	P_1	P_2	B
-5	P_1	1	0	0	0	1	0	200
-10	P_4	0	1	0	0	0	1	300
-M	P_5	-1	0	1	-1	0	1	200
	z_j	M -5	0	-M	M	-5	-M	
	w_j	M -5	0	0	M	0	10 -M	

Basic Solution

$x_1 = 200$

$x_4 = 300$

$x_5 = 200$

$x_2, x_3, x_6 = 0$

$w_2 = \min\ w_j < 0$; $R_5 = \min\ R_i > 0$; P_2 replaces P_5 in the basis.

New basis is P_1, P_4, P_2.

	c_j	0	0	-M	0	-5	-10	
c_i	Basis	P_3	P_4	P_5	P_6	P_1	P_2	B
-5	P_1	1	0	0	0	1	0	200
0	P_4	0	1	-1	1	0	0	100
-10	P_2	-1	0	1	-1	0	1	200
	z_j	10	0	-10	10	0	0	
	w_j	10	0	M -10	10	5	10	

Basic Solution

$x_1 = 200$

$x_4 = 100$

$x_2 = 200$

$x_3, x_5, x_6 = 0$

All $w_j \geq 0$ and P_5 is not in the basis. This is an optimal solution: $z = -5(200) - 10(200) = -3000$.

Interpretation

(a) The space available at Warehouse No. 1 and No. 2 is not sufficient to store the material that must be moved from Warehouse No. 3.

(b) Of the 400 units of material which must be moved from Warehouse No. 3, 200 should be moved to No. 1, and 200 should be moved to No. 2. The total cost will be $3000.

Example 2. Illustration of a Use of Duality

We are given three points (a_1, b_1), (a_2, b_2), and (a_3, b_3), and are required to fit a straight line to them. We assume the points are not collinear, and that $a_1 < a_2 < a_3$.

Formulation

The desired straight line will be determined by its ordinate u_1 at abscissa a_1 and by its slope s from the equation,

$$u = u_1 + s(v - a_1),$$

where v is the variable abscissa

The error of the fit can be measured by

$$e = \max \left| u_i - b_i \right|, \quad i = 1, 2, 3.$$

Variables:

u_1 ordinate on the line at a_1,
s slope of the line,
e maximum absolute error.

Constants:

b_1 given ordinate at a_1,
b_2 given ordinate at a_2,
b_3 given ordinate at a_3,
h_2 $a_2 - a_1$,
h_3 $a_3 - a_1$.

Relations:

$w = e$ to be minimized, subject to

$u_1 - e \leqslant b_1$,
$u_1 + e \geqslant b_1$,
$u_1 + h_2 s - e \leqslant b_2$,
$u_1 + h_2 s + e \geqslant b_2$,
$u_1 + h_3 s - e \leqslant b_3$,
$u_1 + h_3 s + e \geqslant b_3$.

No restriction on the sign of u_1 or s.

By appropriate changes of sign we reverse the order of some of the inequalities and obtain

$$\begin{aligned} &\text{minimize} && w = e, \\ &\text{subject to} && -u_1 + e \geqslant -b_1, \\ & && u_1 + e \geqslant b_1, \\ & && -u_1 - h_2 s + e \geqslant -b_2, \\ & && u_1 + h_2 s + e \geqslant b_2, \\ & && -u_1 - h_3 s + e \geqslant -b_3, \\ & && u_1 + h_3 s + e \geqslant b_3, \end{aligned}$$

which is in the form of Model II of our discussion of duality. The dual model for the above is

$$
\begin{aligned}
\text{maximize} \quad & z = -b_1x_1 + b_1x_2 - b_2x_3 + b_2x_4 - b_3x_5 + b_3x_6, \\
\text{subject to} \quad & -x_1 + x_2 - x_3 + x_4 - x_5 + x_6 = 0, \\
& -h_2x_3 + h_2x_4 - h_3x_5 + h_3x_6 = 0, \\
& x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 1, \\
& x_i \geq 0; \quad i = 1, \cdots, 6.
\end{aligned}
$$

We introduce the artificial variables, x_7, x_8 and x_9 to obtain an initial feasible basis. Then our model becomes standard:

$$
\begin{aligned}
\text{maximize} \quad & z = -Mx_7 - Mx_8 - b_1x_1 + b_1x_2 - b_2x_3 + b_2x_4 \\
& \qquad\qquad - b_3x_5 + b_3x_6, \\
\text{subject to} \quad & x_7 - x_1 + x_2 - x_3 + x_4 - x_5 + x_6 = 0, \\
& x_8 - h_2x_3, \; h_2x_4 - h_3x_5 + h_3x_6 = 0, \\
& x_9 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 1, \\
& x_i \geq 0; \quad i = 1, \cdots 9,
\end{aligned}
$$

with M a predominantly positive number.

Implementation

Suppose the three given points are (1, 2), (2, 4) and (3, 5), then we have: $b_1 = 2$, $b_2 = 4$, $b_3 = 5$, $h_2 = 1$, $h_3 = 2$. We now have

$$
\begin{aligned}
\text{maximize} \quad & z = -Mx_7 - Mx_8 - Mx_9 - 2x_1 + 2x_2 - 4x_3 + 4x_4 - 5x_5 \\
& \qquad\qquad + 5x_6, \\
\text{subject to} \quad & x_7 - x_1 + x_2 - x_3 + x_4 - x_5 + x_6 = 0, \\
& x_8 - 2x_3 + 2x_4 - 3x_5 + 3x_6 = 0, \\
& x_9 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 1, \\
& x_i \geq 0; \quad i = 1, \cdots 9.
\end{aligned}
$$

Computation

The Simplex algorithm applied to the above gives the following:

Final Tableau

	c	-M	-M	-M	-2	2	-4	4	-5	5	
c_i	Basis	P_7	P_8	P_9	P_1	P_2	P_3	P_4	P_5	P_6	B
		0.5	0	0.5	0	1	0	1	0	1	0.5
		0.75	0.5	0.25	1	-.5	0.5	0	0	0.5	0.25
		0.25	-.5	0.25	0	0.5	0.5	0	1	-.5	0.25
	z_j	2.25	1.5	0.25	-2	2.5	-3.5	4	-5	5.5	
	w_j	2.26 +M	1.5 +M	0.25 +M	0	.5	.5	0	0		

Basic Solution

$x_4 = 0.5$

$x_1 = 0.25$

$x_5 = 0.25$

All w_j are nonnegative, artificial variables eliminated, this is an optimal solution. To obtain the values of u_1, s and e, we take the w_1, w_2 and w_3 without the M, which has served its purpose. We then have

$$u_1 = 2.25; \quad s = 1.5; \text{ and } \quad e = 0.25,$$

as the optimal solution for the line fit.

Interpretation

The Chebyshev approximation for a line fit is the line through the point (1, 2.5) with the slope 1.5. It should be noted that a different fit, namely, the least-square fit, is usually preferred.

Exampe 3. An Overdetermined Linear System

We are given three linear equations in two unknowns, which are inconsistent due to errors in measurements by which the right-hand sides were obtained. We wish to find values of the unknowns which will fit all three equations within an error of minimum magnitude.

Formulation

Let r_i represent the error on b_i, $i = 1, 2, 3$, and let

$$u_3 = \max |r_i|.$$

Variables:

u_1 the first unknown,
u_2 the second unknown,
u_3 maximum magnitude of error.

Constants:

a_{ij} coefficients in given system, $i = 1, 2, 3$; $j = 1, 2$,
b_i right-hand sides, $i = 1, 2, 3$.

Relations:

$$\text{minimize} \quad w = u_3,$$
$$\text{subject to} \quad a_{i1} u_1 + a_{i2} u_2 + u_3 \geq b_i, \quad i = 1, 2, 3,$$
$$-a_{i1} u_1 - a_{i2} u_2 + u_3 \geq -b_i, \quad i = 1, 2, 3.$$

The standard model, which is the dual of above, is

$$\text{maximize} \quad z = -Mx_7 - Mx_8 - Mx_9 + b_1 x_1 + b_2 x_2 + b_3 x_3 - b_1 x_4 - b_2 x_5 - b_3 x_6,$$
$$\text{subject to} \quad x_7 + a_{11} x_1 + a_{21} x_1 + a_{31} x_3 - a_{11} x_4 - a_{21} x_5 - a_{31} x_6 = 0,$$
$$x_8 + a_{12} x_1 + a_{22} x_2 + a_{32} x_3 - a_{12} x_4 - a_{22} x_5 - a_{32} x_6 = 0,$$
$$x_9 + x_1 + x_3 + x_4 + x_5 + x_6 = 1,$$
$$x_j \geq 0; \quad j = 1, \cdots 9.$$

Implementation

Suppose, for the given problem, we have

$a_{11} = 1$; $a_{12} = 1$; $b_1 = 1.1$;
$a_{21} = 2$; $a_{22} = 1$; $b_2 = 1.6$;
$a_{31} = 1$; $a_{32} = 2$; $b_3 = 1.6$.

These numerical values are substituted in the standard model.

Computation

The Simplex algorithm applied to the above fives

$$u_1 = 0.54; \quad u_2 = 0.54; \quad u_3 = 0.02.$$

This computation was performed on the IBM 704 and required less than one minute of machine time, and much of that was consumed in reading in control and data cards.

Interpretation

The best Chebyshev fit for the three equations indicates that the unknowns have values, 0.54 and 0.54, with the error on measurement at 0.02,

Example 4. Problem on a Linear Second Order Ordinary Differential Equation

An unknown function $u(s)$ satisfies the differential equation,

$$u'' + u = s,$$

on the interval $S: 0 \leq s \leq 0.2,$

and satisfies the prescribed conditions,

$$u(0) + u'(0) = 2.0,$$

$$u(0.2) = 1.18007.$$

We are required to find $u(0.08)$.

Formulation and Implementation

We consider three mesh points: $s_1 = 0$, $s_2 = 0.08$, and $s_3 = 0.2$, and use as approximation formulas Taylor expansions relating values of u, u' and u'' at adjacent mesh points.

Variables:

u_k $u(s_k)$; $k = 1(1)3$,

u'_k $u'(s_k)$; $k = 1(1)3$,

u''_k $u''(s_k)$; $k = 1(1)3$,

E_{jk}, F_{jk} approximation errors $j = 1, 2$; $k = 1, 2$,

u_0 $\max(|E_{jk}|, |F_{jk}|)$.

Constants:

c^1	2.0,	h_1	0.08,
c^2	1.18008,	h_2	0.12,
c_1	0,	h_1^2	0.0064,
c_2	0.08,	h_2^2	0.0144.
c_3	0.2,		

Relations:

$$\left.\begin{aligned} u_1 + u_1' &= 2.0 \\ u_3 &= 1.18007 \end{aligned}\right\} \text{ prescribed conditions.}$$

$$\left.\begin{aligned} u_1 + u_1'' &= 0. \\ u_2 + u_2'' &= 0.08 \\ u_3 + u_3'' &= 0.2 \end{aligned}\right\} \text{ differential equation discretized.}$$

$$\left.\begin{aligned} u_1 + 0.08u_1' + 0.0032u_1'' - u_2 &= E_{11} \\ u_2 + 0.12u_2' + 0.0072u_2'' - u_3 &= E_{12} \\ 0.08u_1' + 0.0064u_1'' - 0.08u_2' &= E_{21} \\ 0.12u_2' + 0.0144u_2'' - 0.12u_3' &= E_{22} \end{aligned}\right\} \text{ Taylor expansions (forward)}$$

$$\left.\begin{aligned} u_2 - 0.08u_2' + 0.0032u_2'' - u_1 &= F_{11} \\ u_3 - 0.12u_3' + 0.0072u_3'' - u_2 &= F_{12} \\ -0.08u_2' + 0.0064u_2'' + 0.08u_1' &= F_{21} \\ -0.12u_3' + 0.0144u_3'' + 0.12u_2' &= F_{22} \end{aligned}\right\} \text{ Taylor expansions (backward)}$$

Linear model:

minimize $w = u_0$,

subject to conditions:

$$
\begin{aligned}
(x_1)\quad & u_1 + u_1' \geq 2.0,\\
(x_2)\quad & -u_1 - u_1' \geq -2.0,\\
(x_3)\quad & u_3 \geq 1.18007,\\
(x_4)\quad & -u_3 \geq -1.18007,\\
(x_5)\quad & u_1 + u_1'' \geq 0.0,\\
(x_6)\quad & -u_1 - u_1'' \geq 0.0,\\
(x_7)\quad & u_2 + u_2'' \geq 0.08,\\
(x_8)\quad & -u_2 - u_2'' \geq -.08,\\
(x_9)\quad & u_3 + u_3'' \geq 0.2,\\
(x_{10})\quad & -u_3 - u_3'' \geq -.2,\\
(x_{11})\quad & u_0 + u_1 + 0.08u_1' + 0.0032u_1'' - u_2 \geq 0,\\
(x_{12})\quad & u_0 - u_1 - 0.08u_1' - 0.0032u_1'' + u_2 \geq 0,\\
(x_{13})\quad & u_0 + u_2 + 0.12u_2' + 0.0072u_2'' - u_3 \geq 0,\\
(x_{14})\quad & u_0 - u_2 - 0.12u_2' - 0.0072u_2'' + u_3 \geq 0,\\
(x_{15})\quad & u_0 + 0.08u_1' + 0.0064u_1'' - 0.08u_2' \geq 0,\\
(x_{16})\quad & u_0 - 0.08u_1' - 0.0064u_1'' + 0.08u_2' \geq 0,\\
(x_{17})\quad & u_0 + 0.12u_2' + 0.0144u_2'' - 0.12u_3' \geq 0,\\
(x_{18})\quad & u_0 - 0.12u_2' - 0.0144u_2'' + 0.12u_3' \geq 0,\\
(x_{19})\quad & u_0 + u_2 - 0.08u_2' + 0.0032u_2'' - u_1 \geq 0,\\
(x_{20})\quad & u_0 - u_2 + 0.08u_2' - 0.0032u_2'' + u_1 \geq 0,\\
(x_{21})\quad & u_0 + u_3 - 0.12u_3' + 0.0072u_3'' - u_2 \geq 0,\\
(x_{22})\quad & u_0 - u_3 + 0.12u_3' - 0.0072u_3'' + u_2 \geq 0,\\
(x_{23})\quad & u_0 - 0.08u_2' + 0.0064u_2'' + 0.08u_1' \geq 0,\\
(x_{24})\quad & u_0 + 0.08u_2' - 0.0064u_2'' - 0.08u_1' \geq 0,\\
(x_{25})\quad & u_0 - 0.12u_3' + 0.0144u_3'' + 0.12u_2' \geq 0,\\
(x_{26})\quad & u_0 + 0.12u_3' - 0.0144u_3'' - 0.12u_2' \geq 0,
\end{aligned}
$$

The above linear model is the dual of a linear program whose variables are x_n, $n = 1(1)26$.

Linear program:

maximize the linear function

$$z = 2x_1 - 2x_2 + 1.18007x_3 - 1.18007x_4 + 0.08x_7 - 0.08x_8 + 0.2x_9 - 0.2x_{10},$$

subject to restraints:

$$(u_0)\quad x_9 + x_{10} + x_{11} + x_{12} + x_{13} + x_{14} + x_{15} + x_{16} + x_{17} + x_{18} + x_{19} + x_{19} + x_{20} + x_{21} + x_{22} + x_{23} + x_{24} = 1.0,$$

$$(u_1)\quad x_1 - x_2 + x_5 - x_6 + x_{11} - x_{12} - x_{19} + x_{20} = 0.0,$$

$$(u_1')\quad x_1 - x_2 + 0.08x_{11} - 0.08x_{12} + 0.08x_{15} - 0.08x_{16} + 0.08x_{23} - 0.08x_{24} = 0.0,$$

$$(u_1'')\quad x_5 - x_6 + 0.003x_{11} - 0.0032x_{12} + 0.0064x_{15} - 0.0064x_{16} = 0.0,$$

$$(u_2)\quad x_7 - x_8 - x_{11} + x_{12} + x_{13} - x_{14} + x_{19} - x_{20} - x_{21} + x_{22} = 0.0,$$

$$(u_2')\quad 0.12x_{13} - 0.12x_{14} - 0.08x_{15} + 0.08x_{16} + 0.12x_{17} - 0.12x_{18} - 0.08x_{19} + 0.08x_{20} - 0.08x_{23} + 0.08x_{24} + 0.12x_{25} - 0.12x_{26} = 0.0,$$

$$(u_2'')\quad x_7 - x_8 + 0.0072x_{13} - 0.0072x_{14} + 0.0144x_{17} - 0.0144x_{18} - + 0.0032x_{10} - 0.0032x_{20} + 0.0064x_{23} - 0.0064x_{24} = 0.0,$$

$$(u_3)\quad x_3 - x_4 + x_9 - x_{10} - x_{13} + x_{14} + x_{21} - x_{22} = 0.0,$$

$$(u_3')\quad -.12x_{17} + 0.12x_{18} - 0.12x_{21} + 0.12x_{22} - 0.12x_{25} + 0.12x_{26} = 0.0,$$

$$(u_3'')\quad x_9 - x_{10} + 0.0072x_{21} - 0.0072x_{22} + 0.0144x_{25} - 0.0144x_{26} = 0.0,$$

(The above formulation appears very tedious, but it is given in all detail for reasons of clarity. Both the linear model and the linear program can be presented in a single tableau that, columnwise, depicts original model and, rowwise, the program. Even this detail can be by-passed by computer code that will take minimum problem information and prepare an input data tape for the SCROL linear program code.)

Computation

Computation of the above program by the IBM 704 code, SCROL, gave the following results:

u_0 = 0.00004 (maximum approximation error),
u(0) = 0.99998, u(0.08) = 1.07682, u(0.2) = 1.18007,
u'(0) = 1.00002, u'(0.08)= 0.92052, u'(0.2) = 0.80124,
u''(0) = - .99998, u''(0.08)= -.99682, u''(0.2)= -.98007.

Interpretation

The required solution value, u(0.08), is approximately 1.07682 with a possible error of 0.00004.

Comment

The analytic function $u = s + \cos s$ satisfies the differential equation and prescribed conditions. The given differential system can be transformed into a system of first-order differential equations with continuous "coefficients" and continuous right-hand sides satisfying Lepschitz conditions. Two independent conditions are prescribed, hence according to Ince,[21] the differential problem admits of a unique continuous solution u and u' having continuous first derivatives. The functions u and u' are of this kind, hence are unique among functions of this kind.

Comparison of values for u, u', and u'' with the computed values u, u', u'' at the mesh-points reveals a close agreement particularly for u and u''. The discrepancy in u', although larger, still permits a very good approximation to u for any point s in S by means of truncated expansions about the mesh points using computed values for u, u', and u''.

The finite-difference method described by Fox (Ref. 26) applied to this problem gives u(0.08) = 1.07681, which is a slightly better approximation to $\mu(0.08 = 1.07680$ than we obtained. However, this method requires additional mesh points (to provide a uniform mesh), the first of which is outside the specified domain and does not give approximate values for the first and second derivatives.

Example 5. Problem on a Linear Ordinary Differential System of Two Second-Order Equations

The unknown functions u(s) and v(s) satisfy the differential equations,

$$\left.\begin{aligned} u + v + 2v' + v'' &= \cos s \\ u + u'' - v - v' &= 0 \end{aligned}\right\} \quad \text{on the interval } S\text{: } 0.4 \leq s \leq 0.7,$$

and satisfy the prescribed conditions,

$$u'(0.4) = -.38942,$$
$$u'(0.7) = -.64422,$$
$$v'(0.5) = -.60653,$$
$$v'(0.6) = -.54881.$$

We are required to find the values of u and v at s = 0.4, 0.5, 0.6, 0.7.

Formulation and Implementation

We consider the four mesh points $s_1 = 0.4$, $s_2 = 0.5$, $s_3 = 0.6$, and $s_4 = 0.7$, and for approximation we use Taylor expansions relating u, u' and u'' and relating v, v' and v'' at adjacent mesh points. These linear relations, together with the differential equations localized at the mesh points and the prescribed conditions, provide us (as in the previous example) with a system of linear inequalities involving as variables u_0, the absolute approximation error, and values of u and v and their first two derivatives at the mesh points. With the objective of minimizing u_0, we have a linear model whose dual is a linear program in unknowns, x_n, n = 1(1)72.

Computation and Interpretation

Computation of the above linear program by the Simplex algorithm using the SCROL code gives these results:

$$u_0 = 0.0003,$$
$$u(0.4) = 0.92460,\ u(0.5) = 0.88072,\ u(0.6) = 0.82810,\ u(0.7) = 0.76719,$$
$$v(0.4) = 0.67081,\ v(0.5) = 0.60683,\ v(0.6) = 0.54891,\ v(0.7) = 0.49707,$$

along with first and second derivative values for the same points.

The values given above approximate the required solution. The error variable u_0 applies to a single mesh step. Such errors can accumulate, hence the error in any one of the solution values may be much greater.

Comment

The analytic functions $\mu = \cos s$ and $\nu = e^{-s}$ satisfy the differential equations and prescribed conditions. According to the same argument given for Example 4, these solution functions are unique among functions having continuous first and second derivatives.

Comparison of values for μ and ν with the computed values u and v given above reveals a maximum difference between u and μ of .00354 at the mesh point s = 0.4 and a maximum difference between v and ν of .00049 at s = 0.4 and at s = 0.8. The derivative value u', v' have maximum departures from μ' of .00032 at s = 0.5, and from ν' of .00029 at s = 0.4. For u'' and v'' the maximum departures are .00061 at s = 0.4 and .00305 at s = 0.5 respectively. The close agreement on the first derivatives is due to the fact all the prescribed conditions assign numerical values to first derivatives.

Example 6. Problem on a First-Order Partial Differential Equation in Two Independent Variables

The unknown function u(s, t) must satisfy the partial differential equation,

$$2u - su_s + u_t = s^2 \text{ on } R: 0.6 \leq s \leq 1.0;\ 0 \leq t \leq 0.2,$$

and the prescribed conditions,

$$u(s, 0) = 0,$$
$$u_s(0.6, t) = 1.2t.$$

We are required to find u for s = 0.6, 0.8, 1.0, and t = 0.1, 0.2.

Formulation and Implementation

We consider the grid,

$$s_1 = 0.6, \quad t_1 = 0.0, \quad h_\ell = 0.2, \quad \ell = 1, 2,$$
$$s_2 = 0.8, \quad t_2 = 0.1 \quad h_k = 0.1, \quad k = 1, 2,$$
$$s_3 = 1.0, \quad t_3 = 0.2,$$

and construct a linear model whose unknowns are an error variable u_0 and values of u, u_s and u_t at the lattice points (s_ℓ, t_k), $\ell = 1(1)3$, and $k = 1(1)3$. We seek to minimize u_0 subject to linear inequality restraints as follows:

$$\left.\begin{aligned} &\pm(2u^{1k} - 0.6u_s^{1k} + u_t^{1k}) \geq \pm .36 \\ &\pm(2u^{2k} - 0.8u_s^{2k} + u_t^{2k}) \geq \pm .64 \\ &\pm(2u^{3k} - 1.0u_s^{3k} + u_t^{3k}) \geq \pm 1.00 \end{aligned}\right\} \quad k = 1(1)3, \quad (+,+;\ -,-),$$

$$\begin{aligned} &\pm u^{\ell 1} \geq 0, && \ell = 1(1)3, \\ &\pm u_s^{11} \geq 0, \\ &\pm u_s^{21} \geq \pm .12, && (+,+;\ -,-), \\ &\pm u_s^{31} \geq \pm .24, && (+,+;\ -,-), \end{aligned}$$

$$\left.\begin{aligned} &u_0 \pm (u^{\ell k} + 0.2u_s^{\ell k} - u^{\ell+1,\,k}) \geq 0 \\ &u_0 \pm (u^{\ell k} - u^{\ell+1,\,k} + 0.2u_s^{\ell+1,\,k}) \geq 0 \end{aligned}\right\} \quad \ell = 1, 2, \quad k = 1(1)3,$$

$$\left.\begin{aligned} &u_0 \pm (u^{\ell k} + 0.1u_t^{\ell k} - u^{\ell,\,k+1}) \geq 0 \\ &u_0 \pm (u^{\ell\ k} - u^{\ell,\,k+1} + 0.1u_t^{\ell,\,k+1}) \geq 0 \end{aligned}\right\} \quad \ell = 1(1)3, \quad k = 1, 2.$$

The above restraints are obtained from the localization of the differential equation and prescribed conditions and from Taylor expansions. The associated linear program has 28 (equality) restraints in 78 nonnegative (formal) unknowns x_n.

Computation and Interpretation

Computation of the linear program gives the following results:

$u = 0.0032$

$u(0.6, 0.1) = 0.0387, \quad u(0.6, 0.2) = 0.0769,$

$u(0.8, 0.1) = 0.0659, \quad u(0.8, 0.2) = 0.1282,$

$u(1.0, 0.1) = 0.0951, \quad u(1.0, 0.2) = 0.1859,$

along with approximation for first partial derivative values at the nine lattice points.

The above values approximate the required solution. As shown by the comparative magnitude of u_0, the approximation is admittedly crude. If greater accuracy is desired, any one or more of the following modifications can be employed:

a. refinement of the lattice;

b. three-point approximation;

c. term-by-term differentiation of the original differential equation with respect to s and to t, to provide two second-order partial differential equations which must be satisfied by u.

Comment

The analytic function $\mu = s^2t$ satisfies the partial differential equation and the prescribed conditions, and is the only such function that does (extension of Theorem 6, see Ref. 23).

Comparison of the computed values for u above with μ shows a maximum difference of .0141 or about 7% at (1., 0.2) The difference at (0.8, 0.1) is only .0019 or about 3%, which shows the effect on the departure from the theoretical solution which occurs when the solution is extended to points farther from points of prescription. Test runs on simpler problems have indicated that the departure

can definitely be reduced by use of finer mesh; the results are not conclusive enough to project a definitive statement as to the degree of improvement attainable. Other tests indicate that much more significant improvement can be made by following modification c above.

Example 7. Problems on a Parabolic Partial Differential Equation

A. The unknown function $u(s, t)$ must satisfy the linear second-order differential equation,

$$u + u_s/t + u_t - u_{ss} = e^{-t}/t, \quad \text{for } 0.5 < s < 1.5; \; t > 0,$$

and the conditions,

$$u(1.0, \; 0) = 1 \quad \text{(exact)}$$

$$u(0.5, \; 0.05) = 0.50595 \quad \text{(measured)},$$

$$u(1.5, \; 0.05) = 1.43801 \quad \text{(measured)}.$$

Values for u and its derivatives at $(1.0, \; 0.05)$ are desired.

Formulation and Implementation

Let u^*, u_s^*, u_t^*, u_{ss}^* denote the required values at $(1.0, \; 0.05)$.

Localizing the differential equation at this point yields.

$$u^* + 20.u_s^* + u_t^* - u_{ss}^* = 19.0246.$$

Central difference approximation gives

$$u_s^* \sim (1.43801 - 0.50595)/2(0.5),$$

$$u_{ss}^* \sim (1.43801 - 2u^* - 0.50595/0.25.$$

Linear approximation in t gives

$$u^* - 0.05u_t^* \sim 1.$$

Because we have four linear relations in four unknowns, the problem is barely conditioned. There is no opportunity for minimization of approximation errors. The approximations are treated as equations producing a linear algebraic system to be solved.

Computation and Interpretation

The linear algebraic system is satisfied by

$$u^* = 0.9699,$$

$$u_s^* = 0.9321,$$

$$u_s^* = -.602,$$

$$u_{ss}^* = 0.0156.$$

The approximations for u^* and u_t^* are suspect. It is suggested that more data be obtained.

Comment

Since the conditions are prescribed in discrete form (not as functions) we are not able to predicate a unique functional solution. Additional discrete data might completely alter the numerical result. We are confronted here with a problem analogous to curve-fitting to discrete data--we simply use all the data available to obtain the best answer possible.

However, the analytic function $u = se^{-t} + te^{-s}$ does satisfy the partial differential equation and prescribed conditions. At (1.0, 0.05) we have

$$u = .9693, \quad u_s = .9328, \quad u_t = -.5834, \quad u_{ss} = .0184$$

to compare with the values for u* and its derivatives above. It seems we have a surprisingly good approximation despite the sparse data.

B. The differential equation,

$$u + u_s/t + u_t - u_{ss} = e^{-t}/t,$$

holds on the closed domain: $0.5 \leq s \leq 1.5$; $0.05 \leq t \leq 1.0$, and the following values of u are furnished:

$$u(0.5, 0) = 0.5, \quad u(1.0, 0) = 1.0 \quad u(1.5, 0) = 1.5,$$

which are exact, and

$u(0.5, 0.05) = 0.50595$, $u(1, 0.05) = 0.970$, $u(1.5, 0.05) = 1.43801$,
$u(0.5, 0.1) = 0.51307$, $u(1.5, 0.1) = 1.37957$,
$u(0.5, 0.2) = 0.53066$, $u(1.5, 0.2) = 1.27272$,

which are measured with a possible error of 5 in the last digit.

Approximate values for u at $(1, 0.05)$, $1, 0.1)$ and $(1, 0.2)$ are required. Values of u_s, u_t and u_{ss} at these points and at the measured prescription points would be very useful.

Formulation and Implementation

We consider the grid,

$s_1 = 0.5$, $h_1 = 0.5$, $t_1 = 0.05$, $h_1 = 0.05$,
$s_2 = 1.0$, $h_2 = 0.5$, $t_2 = 0.10$, $h_2 = 0.10$,
$s_3 = 1.5$, $t_3 = 0.20$,

and the unknowns,

$u^{\ell k}$, $u_s^{\ell k}$, $u_t^{\ell k}$, and $u_{ss}^{\ell k}$; $\ell = 1, 2, 3$, $k = 1, 2, 3$.

We allow for measurement errors, and write the prescribed conditions as inequalities:

$u^{11} \geq 0.50590$	$u^{21} \geq 0.965$	$u^{31} \geq 1.43796$
$-u^{11} \geq -.50600$	$-u^{21} \geq -.975$	$-u^{31} \geq -1.43806$
$u^{12} \geq 0.51302$		$u^{32} \geq 1.37952$
$-u^{12} \geq -.51312$		$-u^{32} \geq 1.37962$
$u^{13} \geq 0.53061$		$u^{33} \geq 1.27267$
$-u^{13} \geq -.53071$		$-u^{33} \geq 1.27277$

Localization of the differential equation, with allowance for rounding errors in evaluating e^{-t}/t, provides

$$
\left.
\begin{aligned}
u^{\ell 1} + 20u_s^{\ell 1} + u_t^{\ell 1} - u_{ss}^{\ell 1} &\geq 19.02455 \\
-u^{\ell 1} - 20u_s^{\ell 1} - u_t^{\ell 1} + u_{ss}^{\ell 1} &\geq -19.02465 \\
u^{\ell 2} + 10u_s^{\ell 2} + u_t^{\ell 2} - u_{ss}^{\ell 2} &\geq 9.04835 \\
-u^{\ell 2} - 10u_s^{\ell 2} - u_t^{\ell 2} + u_{ss}^{\ell 2} &\geq -9.04845 \\
u^{\ell 3} - 5u_s^{\ell 3} + u_t^{\ell 3} - u_{ss}^{\ell 3} &\geq 4.09360 \\
-u^{\ell 3} - 5u_s^{\ell 3} - u_t^{\ell 3} + u_{ss}^{\ell 3} &\geq -4.09370
\end{aligned}
\right\} \ell = 1, 2, 3.
$$

Central difference formulas are assumed to hold for $s = 1$.

$$
\left.
\begin{aligned}
&\pm(u_s^{2k} - u^{1k} - u^{3k}) \geq 0 \qquad \text{(implies equality)} \\
&\pm(u_{ss}^{2k} - 4u^{1k} + 8u^{2k} - 4u^{3k}) \geq 0 \quad \text{(implies equality)}
\end{aligned}
\right\} k = 1, 2, 3,
$$

and for $t = 0.05$ and 0.1,

$$
\begin{aligned}
&\pm(10u^{12} - u_t^{11}) \geq \pm\ 5.0, \\
&\pm(10u^{22} - u_t^{21}) \geq \pm\ 10.0, \\
&\pm(10u^{32} - u_t^{31}) \geq \pm\ 15.0, \\
&\pm(5u^{13} - u_t^{12}) \geq \pm\ 2.5, \\
&\pm(5u^{23} - u_t^{22}) \geq \pm\ 5.0, \\
&\pm(5u^{33} - u_t^{32}) \geq \pm\ 7.5.
\end{aligned}
$$

Simpson's Rule is assumed to hold in s;

$$
\pm[u^{1k} + (0.5u_s^{1k} + 2u_s^{2k} + 0.5u_s^{3k})/3 - u^{3k}] \geq 0, \quad k = 1, 2, 3.
$$

Taylor expansions, with errors to be minimized, applied at the s-boundary points give

$$
\left.
\begin{aligned}
u_0 \pm (u^{1k} + 0.5u_s^{1k} + u_{ss}^{1k} - u0^{2k}) &\geq 0 \\
u_0 \pm (u^{3k} - 0.5u_s^{3k} + u_{ss}^{3k} - u0^{2k}) &\geq 0
\end{aligned}
\right\} \quad k = 1, 2, 3.
$$

The trapezoid rule, with errors to be minimized, applied at the upper t points yields

$$u_0 \pm (u^{2} + 0.05u_t^{2} - u^{k} + 0.05u_t^{3}) \geqslant 0, \qquad = 1, 2, 3.$$

The linear model consisting of the above inequalities with the objective of minimizing u_0 is the dual of a standard linear program.

Computation and Interpretation

Computation of the linear program was performed by the IBM 709 code, SCROL. The problem required six minutes of computer time.

The following results were obtained:

$u(0.5, 0.05) = 0.5060$,	$u(1, 0.05) = 0.9709$,	$u(1.5, 0.05) = 1.4380$,
$u_s = 0.9218$,	$u_s = 0.9320$,	$u_s = 0.9420$,
$u_t = 0.1312$,	$u_t = -.5770$,	$u_t = -1.2048$,
$u_{ss} = 0.0493$,	$u_{ss} = 0.0183$,	$u_{ss} = 0.0091$;
$u(0.5, 0.10) = 0.5131$,	$u(1, 0.10) = 0.9423$,	$u(1.5, 0.10) = 1.3795$,
$u_s = 0.8452$,	$u_s = 0.8664$,	$u_s = 0.8875$,
$u_t = 0.1530$,	$u_t = -.5258$,	$u_t = -1.1366$,
$u_{ss} = 0.0699$,	$u_{ss} = 0.0321$	$u_{ss} = 0.0197$;
$u(0.5, 0.20) = 0.5306$,	$u(1, 0.20) = 0.8948$,	$u(1.5, 0.20) = 1.2727$,
$u_s = 0.7045$,	$u_s = 0.7421$,	$u_s = 0.7796$,
$u_t = 0.1535$,	$u_t = -.4570$,	$u_t = -.9641$,
$u_{ss} = 0.1130$,	$u_{ss} = 0.0545$,	$u_{ss} = 0.1128$.

The value for the error term u_0 was 0.0022. The values given for $t = 0.10$ may be in error by this amount. The values at $t = 0.2$, particularly for u_{ss} and u_t, are not very reliable because of lower-order approximations that had to be used for these points. More data for additional (increasing) t values along the s boundaries will make

extension of the tabulation in t possible, with greater accuracy, except that the values at the greatest t will always be more or less suspect.

Comment

Again as in A, discrete conditions do not predicate unicity. This discrete problem is also satisfied by

$u = se^{-t} + te^{-s}$. At (1.0, 0.2) we have

$u = .8923,\ u_s = .7452,\ u_t = -.4508,\ u_{ss} = .0736,$

which compares very favorably with the computed values for this point given above.

A technique for measurement-error allowance is illustrated in this example.

Example 8. Problems on a Hyperbolic Partial Differential Equation

A. An unknown function u(s, t) must satisfy the linear second-order partial differential equation,

$$u - u_s/s + u_{ss} - u_{tt} = t^2 - 1, \quad \text{for } s > 0, \quad t > 0,$$

and the conditions,

$u(0.2, 0) = 1.04$ $\qquad$ $u_t(0.2, 0) = 0.04,$

$u(0.5, 0) = 1.25,$ $\qquad$ $u_t(0.5, 0) = 0.25,$

$u(1.0, 0) = 2.00,$ $\qquad$ $u_t(1.0, 0) = 1.00.$

Values for u, u_s, u_t and u_{ss} are desired at the points (0.2, 0.1), (0.5, 0.1), and (1.0, 0.1).

Formulation and Implementation

The points (0.2, 0.1) and (1.0, 0.1) lie outside the domain determined by characteristics through (0.2, 0) and (1.0, 0). Values at these points cannot be determined from the given data (Ref. 27, p. 16).

We consider the grid points,

$s_0 = 0.2,\ h_0 = 0.3,$ $\qquad$ $t_0 = 0,$ $\qquad$ $h_0 = 0.1,$

$s_1 = 0.5,\ h_1 = 0.5,$ $\qquad$ $t_1 = 0.1$

$s_2 = 1.0.$

Localization of the differential equation yields

$$u^{11} - 2u_s^{11} + u_{ss}^{11} - u_{tt}^{11} = -.99.$$

Taylor expansions at (s_1, t_0) give

$$1.25 - 0.3u_s^{10} + 0.45u_{ss}^{10} \sim 1.04,$$

$$1.25 + 0.5u_s^{10} + 0.125u_{ss}^{10} \sim 2.00,$$

$$0.25 - 0.3u^{10}_{st} + 0.045u^{10}_{sst} \sim 0.04,$$
$$0.25 + 0.5u_{st} + 0.125u^{10}_{sst} \sim 1.00,$$
$$u^{10}_{s} + 0.1u_{st} \sim u^{11}_{s},$$
$$u^{10}_{ss} + 0.1u^{10}_{sst} \sim u^{11}_{ss}.$$

Expansions at (s_1, t_1) yield

$$u^{11} - 0.1u^{11}_{t} + 0.005u^{11}_{tt} \sim 1.25,$$
$$u^{11} - 0.1u^{11}_{tt} \sim 0.25.$$

Altogether there are nine linear relations in nine unknowns. The problem is barely conditioned. The approximations must be treated as equalities.

Computation and Interpretation

The linear algebraic system is satisfied by

$$u^{11} = 1.2714,$$
$$u^{11}_{s} = 1.1000,$$
$$u^{11}_{t} = 0.4788,$$
$$u^{tt}_{ss} = 2.2000,$$
$$u^{11}_{tt} = 2.2878.$$

Linear approximation had to be used for u_t and u_{tt}, hence these values are suspect. More data along the line $t = 0$ would be useful to improve accuracy and extend the scope of the problem.

Comment

Remarks on discrete conditions given in comment on Example 7A also apply here. An analytic function that satisfies the discrete problem is $u = s^2e^t + t^2 + 1$.

At (0.5, 0.1) we have

$$u = 1.2638,\ u_s = 1.1052,\ u_t = 0.4538,\ u_{ss} = 2.2104,\ u_{tt} = 2.2538,$$

which compares favorably with computed values for u and its derivatives given above.

B. The function $u(s,t)$ must satisfy

$$u - u_s/s + u_{ss} - u_{tt} = t^2 - 1, \quad \text{for } s > 0, \quad t > 0,$$

and the conditions,

$$u(0.20, 0) = 1.0400, \qquad u_t(0.20, 0) = 0.0400,$$
$$u(0.35, 0) = 1.1225, \qquad u_t(0.35, 0) = 0.1225,$$
$$u(0.50, 0) = 1.2500, \qquad u_t(0.50, 0) = 0.2500,$$
$$u(0.75, 0) = 1.5625, \qquad u_t(0.75, 0) = 0.5625,$$
$$u(1.00, 0) = 2.0000, \qquad u_t(1.00, 0) = 1.0000.$$

Values for u, u_s, u_t, u_{ss}, and u_{tt} are desired at the points,

$$(0.35,\ 0.1), \qquad (0.50,\ 0.1), \qquad (0.75,\ 0.1),$$
$$(0.50,\ 0.2).$$

Formulation and Implementation

The characteristics for the equation are

$$s \pm t = \text{constant}.$$

The requirement points along $t = 0.1$ all lie within the domain bounded by the characteristics,

$$s \pm t = 0.20, \qquad s \pm t = 1.00;$$

hence the values at these points can be determined from the data given along the segment $t = 0$, $0.20 \leq s \leq 1.00$. The point $(0.50, 0.2)$ lies in the domain bounded by the characteristics,

$$s \pm t = 0.45, \qquad s \pm t = 0.85;$$

hence the values here can be determined once the values along the segment $t = 0.1$, $0.35 \leq s \leq 0.75$ are known. This is true because the segments are nowhere parallel to the characteristics.

We consider the grid,

$$s_0 = 0.20, \quad h_0 = 0.15, \qquad t_0 = 0, \quad h_0 = 0.1,$$
$$s_1 = 0.35, \quad h_1 = 0.15, \qquad t_1 = 0.1, \quad h_1 = 0.1.$$
$$s_2 = 0.50, \quad h_2 = 0.25, \qquad t_2 = 0.2,$$
$$s_3 = 0.75, \quad h_3 = 0.25,$$
$$s_4 = 1.00.$$

At every lattice point, grid steps in t are less than those in s with u, and u_t are given along a line t = constant, namely t = 0; therefore there should be no stability difficulty(Ref. 27, p. 24).

Localization of the differential equation yields

$$\pm(u^{11} - 2.85714u_s^{11} + u_{ss}^{11} - u_{tt}^{11}) \gtrless .99,$$
$$\pm(u^{21} - 2.00000u_s^{21} + u_{ss}^{21} - u_{tt}^{21}) \gtrless .99,$$
$$\pm(u^{31} - 1.33333u_s^{31} + u_{ss}^{31} - u_{tt}^{31}) \gtrless .99,$$
$$\pm(u^{22} - 2.00000u_s^{22} + u_{ss}^{22} - u_{tt}^{22}) \gtrless .96.$$

The following approximations are used with no error admitted where known numerical values appear. (This concession is forced by the fact that the differential equation is not assumed to hold along the line t = 0.)

Second-order expansions in s along t = 0 yield

$$\pm(1.1225 - 0.15u_s^{10} + 0.01125u_{ss}^{10}) \gtrless 1.0400,$$
$$\pm(1.1225 + 0.15u_s^{10} + 0.01125u_{ss}^{10}) \gtrless 1.2500,$$
$$\pm(1.2500 - 0.15u_s^{20} + 0.01125u_{ss}^{20}) \gtrless 1.1225,$$
$$\pm(1.2500 + 0.25u_s^{20} + 0.03125u_{ss}^{20}) \gtrless 1.5625,$$
$$\pm(1.5625 - 0.25u_s^{30} + 0.03125u_{ss}^{30}) \gtrless 1.2500,$$
$$\pm(1.5625 + 0.25u_s^{30} + 0.03125u_{ss}^{30}) \gtrless 2.0000,$$
$$\pm(0.1225 - 0.15u_{st}^{10} + 0.01125u_{sst}^{10}) \gtrless 0.0400,$$
$$\pm(0.1225 + 0.15u_{st}^{10} + 0.01125u_{sst}^{10}) \gtrless 0.2500,$$
$$\pm(0.2500 - 0.15u_{st}^{20} + 0.01125u_{sst}^{20}) \gtrless 0.1225,$$
$$\pm(0.2500 + 0.25u_{st}^{20} + 0.03125u_{sst}^{20}) \gtrless 0.5626,$$
$$\pm(0.5625 + 0.25u_{st}^{30} + 0.03125u_{sst}^{30}) \gtrless 1.0000.$$

Second-order expansion in s along t = 0.1 yield

$$u_0 \pm (u^{11} + 0.15u_s^{11} + 0.01125u_{ss}^{11} - u^{21}) \geq 0,$$
$$u_0 \pm (u^{21} + 0.25u_s^{21} + 0.03125u_{ss}^{21} - u^{31}) \geq 0,$$
$$u_0 \pm (u^{11} - u^{21} + 0.15u_s^{21} - 0.01125u_{ss}^{21}) \geq 0,$$
$$u_0 \pm (u^{21} - u^{31} + 0.25u_s^{31} - 0.03125u_{ss}^{31}) \geq 0,$$
$$u_0 \pm (u_t^{21} - 0.15u_{st}^{21} + 0.01125u_{sst}^{21} - u_t^{11}) \geq 0,$$
$$u_0 \pm (u_t^{21} + 0.25u_{st}^{21} + 0.03125u_{sst}^{21} - u_t^{31}) \geq 0.$$

Expansions in t give

$$\pm (u^{11} - 0.1u_t^{11} + 0.005u_{tt}^{11} - 1.1225) \geq 0,$$
$$\pm (u^{21} - 0.1u_t^{21} + 0.005u_{tt}^{21} - 1.2500) \geq 0,$$
$$\pm (u^{31} - 0.1u_t^{31} + 0.005u_{tt}^{31} - 1.5625) \geq 0,$$
$$u_0 \pm (u^{21} + 0.1u_t^{21} + 0.005u_{tt}^{21} - u^{22}) \geq 0,$$
$$u_0 \pm (u^{22} - 0.1u_t^{22} + 0.005u_{tt}^{22} - u^{21}) \geq 0,$$
$$\pm (u_t^{11} - 0.1u_{tt}^{11} - 0.1225) \geq 0,$$
$$\pm (u_t^{21} - 0.1u_{tt}^{21} - 0.2500) \geq 0,$$
$$\pm (u_t^{31} - 0.1u_{tt}^{31} - 0.5625) \geq 0.$$
$$u_0 \pm (u_t^{21} + 0.1u_{tt}^{21} - u_t^{31}) \geq 0,$$
$$u_0 \pm (u_t^{31} - 0.1u_{tt}^{31} - u_t^{21}) \geq 0,$$
$$u_0 \pm (u_s^{10} + 0.1u_{st}^{10} - u_s^{11}) \geq 0,$$
$$u_0 \pm (u_s^{20} + 0.1u_{st}^{20} - u_s^{21}) \geq 0,$$
$$u_0 \pm (u_s^{30} + 0.1u_{st}^{30} - u_s^{31}) \geq 0,$$
$$u_0 \pm (u_s^{21} + 0.1u_{st}^{21} - u_s^{22}) \geq 0,$$
$$u_0 \pm (u_s^{22} + 0.1u_{st}^{22} - u_s^{21}) \geq 0,$$
$$u_0 \pm (u_s^{10} + 0.1u_{sst}^{10} - u_{ss}^{11}) \geq 0,$$

$$u_0 \pm (u^{20}_{ss} + 0.1u^{20}_{sst} - u^{21}_{ss}) \geq 0,$$
$$u_0 \pm (u^{20}_{ss} + 0.1u^{20}_{sst} - u^{21}_{ss}) \geq 0,$$
$$u_0 \pm (u^{20}_{ss} + 0.1u^{30}_{sst} - u^{31}_{ss}) \geq 0,$$
$$u_0 \pm (u^{21}_{ss} + 0.1u^{21}_{sst} - u^{22}_{ss}) \geq 0,$$
$$u_0 \pm (u^{22}_{ss} - 0.1u^{22}_{sst} - u^{21}_{ss}) \geq 0.$$

The above linear restraints with the objective of minimizing u_0 constitute a model which is the dual of a standard linear program.

Computation and Interpretation

Computation of linear program by the SCROL code gives

$u_0 = 0.0013,$

$u(0.35, 0.1) = 1.145,\ u_s = 0.771,\ u_t = 0.336,\ u_{ss} = 2.201,\ u_{tt} = 2.133,$

$u(0.50, 0.1) = 1.286,\ u_s = 1.101,\ u_t = 0.478,\ u_{tt} = 2.201,\ u_{tt} = 2.275,$

$u(0.75, 0.1) = 1.632,\ u_s = 1.651,\ u_t = 0.824,\ u_{ss} = 2.199,\ u_{tt} = 2.619,$

$u(0.50, 0.2) = 1.3\ 4,\ u_s = 1.212,\ u_t = 0.706,\ u_{ss} = 2.421,\ u_{tt} = 2.302.$

The approximation error along $t = 0.1$ is no greater than u_0. For increased accuracy, data at intermediate points along $t = 0$ is needed. A smaller grid step in t can then be used.

Comment

The differential equation and prescribed conditions are satisfied by $\mu = s^2 e^t + t^2 + 1$. At the point $(0.5, 0.2)$ we have

$$\mu = 1.345,\ \mu_s = 1.221,\ \mu_t = 705,\ \mu_{ss} = 2.443,\ \mu_{tt} = 2.305.$$

Despite the fact that this comparison point is four times as far from the prescription, we have essentially as good comparison of u and μ as in A. This result shows the effect of the mesh refinement in s.

Example 9. Problems on an Elliptic Partial Differential Equation

A. An unknown function $u(s, t)$ must satisfy Laplace's partial differential equation for two independent variables,

$$u_{ss} + u_{tt} = 0,$$

on an open two-dimensional domain consisting of those points (s, t) that are in the circle, $s^2 + t^2 < 1.0$, but not in the square,
$-.6 \leq s \leq 0.6,\ -.6 \leq t \leq 0.6.$

The potential u can be measured on the inner and outer boundaries. The following measurements have been made:

$$u(0.6,\ 0) = 0.564, \qquad u(1.0,\ 0) = 0.841,$$
$$u(0.8,\ 0.6) = 1.307, \qquad u(0.8,\ -.6) = 0.394.$$

Values for u and its derivatives at $(0.8,\ 0)$ are desired.

Formulation and Implementation

We denote the required values at $(0.8,\ 0)$ by u^*, u^*_s, u^*_t, u^*_{ss}, and u^*_{tt} and use the following central difference approximations:

$$u^*_s \sim (0.841 - 0.564)/0.4,$$
$$u^*_{ss} \sim (0.841 - 2u^* + 0.564)/0.04,$$
$$u^*_t \sim (1.307 - 0.394)/1.2,$$
$$u^*_{tt} \sim (1.307 - 2u^* + 0.394)/0.36.$$

Localization of the differential equation yields

$$u^*_{ss} + u^*_{tt} = 0.$$

As there are only five linear relations in five unknowns, the problem is barely conditioned. The approximations must be treated as equalities.

Computation and Interpretation

The linear algebraic system is satisfied by

$$u\,(0.8,\ 0) = 0.717,$$
$$u_s(0.8,\ 0) = 0.692,$$
$$u_t(0.8,\ 0) = 0.761,$$
$$u_{ss}(0.8,\ 0) = -.740,$$
$$u_{tt}(0.8,\ 0) = 0.740.$$

The above results are the best obtainable with the given data. For greater accuracy, it is suggested that additional measurements be made, to four decimal places if possible, at the following boundary points (s, t):

(a) The points already measured;

(b) Those points at which $t = 0.3$ and $t = -.3$.

Comment

This example may be thought of as a discretization of a plane Dirichlet problem of the form

$$u_{ss} + u_{tt} = 0 \quad \text{on } D, \text{ a bounded region in the } sxt \text{ plane}$$

with

$$u = f(s, t) \text{ on } \Gamma, \text{ the boundary of } D,$$

in which a solution u is sought which is analytic on D and continuous on $D + \Gamma$. This problem has been the subject of considerable mathematical investigation, the question usually being on sufficient conditions of $f(s, t)$ as to differentiability in order to insure the existence and uniqueness for a more-or-less specific domain D of a solution u as required.

The discrete problem differs from the above in that (1) the prescribed condition is

$$u = F(P_\gamma),$$

where F is not a function of continuous variables s, t along Γ but rather a function of some finite set of discrete points P_γ of Γ, and (2) the only solution attainable is likewise not the function u of continuous variables s, t but rather a function U of a finite set of discrete points (s_k, t_k) in D.

Forsythe and Wasow (Ref. 27, p. 177) list some eight questions relative to discretization of elliptic partial differential problems, and preface this listing with the statement, "The answers to these problems are generally unknown, and their tentative answers already fill a considerable literature which records some of the current state of the art."

When confronted with the problem of finding potential values at some points within a specific domain D when values are known (presumably measured) at some points on the boundary Γ of D, one is forced to an experimental attitude.

Two questions arise:

(1) How to use the known information (boundary values and differential equation) to obtain approximations for the required potential values?

(2) How can the approximation be tested and improved?

On (1) we have the reassurance that there is a polynomial P (s, t) which satisfies the discrete problem, that is the polynomial whose values at the (discrete) boundary points (P_γ) coincide with those of F and whose second derivatives at all the discrete requirement points (s_ℓ, t_k) satisfy the localized differential equation. Usually some other function U is constructed, since the determination of P may be indeed a cumbrous process.

In (2) we are concerned with what can happen if more data are available. Presumably it is impossible or impractical to measure u at all the requirement points (s_ℓ, t_k), but spot measurement would certainly provide information on the agreement of u and U. It may be possible that values for u may be obtained at additional points of Γ. The question then becomes: For a sequence of discrete problems approaching the analytic problem, do the discrete solutions approach the analytic solution? This question requires considerable investigation, and it is hoped the linear program method may provide some assistance.

There is an analytic function, $u = e^t \sin s$ which satisfies the differential equation and prescribed conditions. The computed values for u and its derivatives at (0.8, 0) compare quite well with those of u.

B. The potential u for part A above has been remeasured to obtain

u(0.8, 0.6) = 1.3072,

u(0.954, 0.3) = 1.1011, u(0.6, 0.3) = 0.7622,

u(1.0, 0) = 0.8415, u(0.6, 0) = 0.5646,

u(0.954, -.3) = 0.6043 u(0.6, -.3) = 0.4182.

u(0.8, -.6) = 0.3937,

Values for u and its derivatives at u and its derivatives at the interior points (0.8, 0.3), (0.8, 0) and (0.8, -.3) are desired.

Formulation and Implementation

We consider the grid

$t_1 = -.6$, $h_1 = 0.3$, For t_3: For t_2, t_4:

$t_2 = -.3$, $h_2 = 0.3$, $s_1 = 0.6$, $h_1 = 0.2$, $s_1 = 0.6$, $h_1 = 0.2$

$t_3 = 0$, $h_3 = 0.3$, $s_2 = 0.8$, $h_2 = 0.2$, $s_2 = 0.8$, $h_2 = 0.154$,

$t_4 = 0.3$, $h_4 = 0.3$, $s_3 = 1.0$; $s_3 = 0.954$,

$t_5 = 0.6$;

and seek values for u, u_s, u_t, u_{ss} at the points (s_2, t_k), $k = 2, 3, 4$.

Localization of the differential equation at these points gives

$$\pm (u_{ss} + u_{tt}) \geq 0; \qquad k = 2, 3, 4.$$

Taylor expansions in t along the line $s = s_2$ give

$$u_0 \pm (u^{2k} + 0.3u_t^{2k} + 0.045u_{tt}^{2k} - u^{2,k+1}) \geq 0, \quad k = 2, 3,$$

$$u_0 \pm (u^{24} + 0.3u_t^{24} + 0.045u_{tt}^{24}) \geq \pm 1.3072,$$

$$u_0 \pm (u^{22} - 0.3u_t^{22} + 0.045u_{tt}^{22}) \geq \pm 0.3937,$$

$$u_0 \pm (u^{2k} - u^{2,k+1} + 0.3u_t^{2,k+1} - 0.045u_{tt}^{2,k+1}) \geq 0, \quad k = 2, 3.$$

Taylor expansions in s give

$$u_0 \pm (u^{22} - 0.2u_s^{22} + 0.02u_{ss}^{22}) \geq \pm .4182,$$
$$u_0 \pm (u^{22} + 0.154u_s^{22} + 0.01186u_{ss}^{22}) \geq \pm .6043,$$
$$u_0 \pm (u^{23} - 0.2u_s^{23} + 0.02u_{ss}^{23}) \geq \pm .5646,$$
$$u_0 \pm (u^{23} + 0.2u_s^{23} + 0.02u_{ss}^{23}) \geq \pm .8415,$$
$$u_0 \pm (u^{24} - 0.2u_s^{24} + 0.02u_{ss}^{24}) \geq \pm .7622,$$
$$u_0 \pm (u^{24} + 0.154u_s^{24} + 0.01186u_{ss}^{24}) \geq \pm .1011$$

Since there are uniform steps in t, Simpson's Rule can be applied to yield:

$$\pm (u_t^{22} + 0.1u_{tt}^{22} + 0.4u_{tt}^{23} - u_t^{24} + 0.1u_{tt}^{24}) \geq 0.$$

(The above is assumed to hold without admitted error because it is a higher-order approximation.)

With the objective of minimizing u_0, the above linear model is the dual of a standard linear program.

Computation and Interpretation

Computation of the above linear program by the SCROL code required four minutes on the IBM 704 computer. The following results were obtained:

$u_0 = 0.00015;$

$u(0.8, -.3) = 0.53168,$ $u_s = 0.51348,$ $u_{ss} = -.53161,$
$u_t = 0.54018,$ $u_{tt} = 0.53161;$

$u(0.8, 0) = 0.71750,$ $u_s = 0.69225,$ $u_{ss} = -.73021,$
$u_t = 0.72844,$ $u_{tt} = 0.73021;$

$u(0.8, 0.3) - 0.96874,$ $u_s = 0.93506,$ $u_{ss} = -.96905,$
$u_t = 0.98232,$ $u_{tt} = 0.96905.$

The accuracy of the approximation for u is very nearly the same as that of the measured values. Additional values for u at nearby points

in either the s or t direction may be obtained by using a truncated Taylor expansion with the given derivative values. Increased accuracy in derivative values could be obtained by more measurements.

Comment

General comments on 9A also apply here. The agreement between computed values for u and its derivatives and $u = e^t \sin s$ (which satisfies the problem) and its derivatives is quite good.

Example 10. Problem on a System of First-Order Partial Differential Equations

The unknown functions u(s, t) and v(s, t) must satisfy the partial differential equations:

$$u_s + v_t = e^s + e^t,$$
$$su_t - tv_s = 0, \qquad \text{for } s > 0, \quad t > 0.$$

Along the line t = 1.0, measurements of u and v give

u(0.6, 1.0) = 2.822,	v(0.6, 1.0) = 3.078,
u(0.8, 1.0) = 3.225,	v(0.8, 1.0) = 3.358,
u(1.0, 1.0) = 3.718,	v(1.0, 1.0) = 3.718,
u(1.2, 1.0) = 4.320,	v(1.2, 1.0) = 4.158,
u(1.4, 1.0) = 5.055,	v(1.4, 1.0) = 4.678,

with possible errors of ±.0005.

Values for the first partial derivatives at the above points and for the solution functions and these derivatives at the points (0.8, 1.2), (1.0, 1.2), and (1.2, 1.2) are desired.

Formulation and Implementation

Required values are determinable from the given data. Stability is ensured for the following grid

$s_1 = 0.6$, $h_1 = 0.2$, $t_1 = 1.0$, $h_1 = 0.2$.
$s_2 = 0.8$, $h_2 = 0.2$, $t_2 = 1.2$,
$s_3 = 1.0$, $h_3 = 0.2$,
$s_4 = 1.2$, $h_4 = 0.2$,
$s_5 = 1.4$,

Localization of the differential equations yields

$$\pm (u_s^{11} + v_t^{11}) \geqslant \pm (4.540 \mp 0.0005),$$
$$\pm (u_s^{11} + v_t^{21}) \geqslant \pm (4.944 \mp 0.0005),$$
$$\pm (u_s^{31} + v_t^{31}) \geqslant \pm (5.437 \mp 0.0005),$$
$$\pm (u_s^{41} + v_t^{41}) \geqslant \pm (6.038 \mp 0.0005),$$
$$\pm (u_s^{51} + v_t^{51}) \geqslant \pm (6.773 \mp 0.0005),$$
$$\pm (u_s^{22} + v_t^{22}) \geqslant \pm (5.546 \mp 0.0005),$$
$$\pm (u_s^{32} + v_t^{32}) \geqslant \pm (6.038 \mp 0.0005),$$
$$\pm (u_s^{42} + v_t^{42}) \geqslant \pm (6.640 \mp 0.0005),$$

with allowance for rounding error in the numerical evaluation of $e^s + e^t$, and

$$\pm (0.6u_t^{11} - 1.0v_s^{11}) \geqslant 0,$$
$$\pm (0.8u_t^{21} - 1.0v_s^{21}) \geqslant 0,$$
$$\pm (1.0u_t^{31} - 1.0v_s^{31}) \geqslant 0,$$
$$\pm (1.2u_t^{41} - 1.0v_s^{41}) \geqslant 0,$$
$$\pm (1.4u_t^{51} - 1.0v_s^{51}) \geqslant 0,$$
$$\pm (0.8u_t^{22} - 1.2v_s^{22}) \geqslant 0,$$
$$\pm (0.0u_t^{32} - 1.2v_s^{32}) \geqslant 0,$$
$$\pm (1.2u_t^{42} - 1.2v_s^{42}) \geqslant 0.$$

Prescribed conditions yield

$$\pm u^{11} \geqslant \pm (2.822 \mp 0.0005), \qquad \pm v^{11} \geqslant \pm (3.078 \mp 0.0005),$$
$$\pm u^{21} \geqslant \pm (3.225 \mp 0.0005), \qquad \pm v^{21} \geqslant \pm (3.358 \mp 0.0005),$$
$$\pm u^{31} \geqslant \pm (3.718 \mp 0.0005), \qquad \pm v^{31} \geqslant \pm (3.718 \mp 0.0005),$$
$$\pm u^{41} \geqslant \pm (4.320 \mp 0.0005), \qquad \pm v^{41} \geqslant \pm (4.158 \mp 0.0005),$$
$$\pm u^{51} \geqslant \pm (5.055 \mp 0.0005), \qquad \pm v^{51} \geqslant \pm (4.678 \mp 0.0005).$$

Taylor expansions in s yield

$$\left.\begin{aligned} u_0 &\pm (u^{k1} + 0.2u_s^{k1} - u^{k+1,1}) \geq 0 \\ u_0 &\pm (v^{k1} + 0.2v_s^{k1} - v^{k+1,1}) \geq 0 \end{aligned}\right\} k = 1(1)4,$$

$$\left.\begin{aligned} u_0 &\pm (u^{k2} + 0.2u_s^{k2} - u^{k+1,2}) \geq 0 \\ u_0 &\pm (v^{k2} + 0.2v_s^{k2} - v^{k+1,2}) \geq 0 \end{aligned}\right\} k = 2, 3$$

$$\left.\begin{aligned} u_0 &\pm (u^{k-1,1} - u^{k1} + 0.2u_s^{k1}) \geq 0 \\ u_0 &\pm (v^{k-1,1} - v^{k1} + 0.2v_s^{k1}) \geq 0 \end{aligned}\right\} k = 2(1)5,$$

$$\left.\begin{aligned} u_0 &\pm (u^{k-1,2} - u^{k2} + 0.2u_s^{k2}) \geq 0 \\ u_0 &\pm (v^{k-1,2} - v^{k2} + 0.2v_s^{k2}) \geq 0 \end{aligned}\right\} k = 3, 4.$$

Taylor expansions in t yield

$$\left.\begin{aligned} u_0 &\pm (u^{k1} + 0.2u_t^{k1} - u^{k2}) \geq 0 \\ u_0 &\pm (v^{k1} + 0.2v_t^{k1} - v^{k2}) \geq 0 \\ u_0 &\pm (u^{k1} - u^{k2} + 0.2u_t^{k2}) \geq 0 \\ u_0 &\pm (v^{k1} - v^{k2} + 0.2v_t^{k2}) \geq 0 \end{aligned}\right\} k = 2, 3, 4.$$

Computation and Interpretation

Computation of the linear program, minimizing u_0, gives

$u_0 = 0.0655,$

$u(0.6, 1.0) = 2.822, \quad u_s = 1.692, \quad u_t = 1.796,$

$u(0.8, 1.0) = 3.225, \quad u_s = 2.132, \quad u_t = 2.167,$

$u(1.0, 1.0) = 3.718, \quad u_s = 2.788, \quad u_t = 2.122,$

$u(1.2, 1.0) = 4.320, \quad u_s = 3.342, \quad u_t = 1.623,$

$u(1.4, 1.0) = 5.055, \quad u_s = 3.342, \quad u_t = 1.623,$

$u(0.8, 1.2) = 3.643, \quad u_s = 2.734, \quad u_t = 2.417,$

$u(1.0, 1.2) = 4.207, \quad u_s = 3.148, \quad u_t = 2.270,$

$u(1.2, 1.2) = 4.807, \quad u_s = 3.326, \quad u_t = 2.267;$

$v(0.6, 1.0) = 3.078, \quad v_s = 1.078, \quad v_t = 2.847,$
$v(0.8, 1.0) = 3.358, \quad v_s = 1.732, \quad v_t = 2.812,$
$v(1.0, 1.0) = 3.717, \quad v_s = 2.122, \quad v_t = 2.649,$
$v(1.2, 1.0) = 4.157, \quad v_s = 2.528, \quad v_t = 2.695,$
$v(1.4, 1.0) = 4.677, \quad v_s = 2.272, \quad v_t = 3.430,$
$v(0.8, 1.0) = 3.855, \quad v_s = 1.612, \quad v_t = 2.812,$
$v(1.0, 1.0) = 4.243, \quad v_s = 2.266, \quad v_t = 2.891,$
$v(1.2, 1.0) = 4.762, \quad v_s = 2.266, \quad v_t = 3.313.$

The relatively low accuracy as reflected in the size of the error term u_0 is caused by the use of linear approximations on a rather coarse grid. Greater accuracy can be attained if values for u and v are prescribed for a finer grid in s along $t = 1.0$. It is preferable that the grid be uniform. The grid step in t can then be reduced correspondingly.

Comment

The differential equations and prescribed conditions are satisfied by

$$\mathcal{u} = e^s + t^2, \qquad \mathcal{v} = e^t + s^2.$$

Comparison of the computed values for u, u_s, u_t, v, v_s, and v_t with their counterparts in $\mathcal{u}$ and $\mathcal{v}$ at the mesh points shows agreement to two-decimal-place accuracy.

REFERENCES

1. V. Riley and S. I. Gass, Linear Programming and Associated Techniques (Johns Hopkins Press, Baltimore, 1958).
2. J. von Neumann, Zur Theorie der Gesellshaftspiele, Mathematische Annalen 100, 295 (1928).
3. H. Weyl, Elementaire Theorie der Konvexen Polyeder, Commentarii Mathematici Helvetici 7, 290 (1935).
4. J. von Neumann and O. Morgenstern, Theory of Games and Economic Behaviour (Princeton University Press, Princeton, N. J., 1944).
5. F. L. Hitchcock, The Distribution of a Product from Several Sources to Numerous Localities, 20, No. 3, 224 (1941).
6. G. J. Stigler, The Cost of Subsistence, 27, No. 2, 303 (1945).
7. Directorate of Management Analysis, Scientific Planning of Military Problems, Project SCOOP No. 4-PU (U. S. Air Force, Washington, 1948).
8. G. B. Dantzig, The Simplex Method, RAND Report P-891 (The Rand Corporation, Santa Monica, Calif., 1956).
9. T. C. Koopmans, Activity Analysis of Production and Allocation, Cowles Commission Monograph No. 13 (John Wiley & Sons, Inc., New York, 1951).
10. A. Orden and L. Goldstein (eds.), Symposium on Linear Inequalities and Programming (U. S. Air Force, Washington, 1952).
11. H. Antosiewicz, Proceedings of the Second Symposium in Linear Programming, Vol. 1, 2 (Bureau of Standards, Washington, 1955).
12. A. Charnes, W. W. Cooper, and A. Henderson, An Introduction to Linear Programming (John Wiley & Sons, Inc., New York, 1953).
13. W. Orchard-Hays, Background, Development and Extensions of the Revised Simplex Method, RAND Report RM-1433 (The Rand Corporation, Santa Monica, Calif. 1954).

14. W. Orchard-Hays, Evolution of Computer Codes for Linear Programming, RAND Report P-810 (The Rand Corporation, Santa Monica, Calif., 1956).

15. W. Orchard-Hays, Evolution of Linear Programming Computing Techniques, RAND Report P-900 (The Rand Corporation, Santa Monica, Calif. 1956).

16. S. I. Gass, Linear Programming (McGraw-Hill, Book Co., Inc., New York, 1958).

17. K. Eisemann, Linear Programming, Quarterly of Applied Mathematics 13, No. 3, 209 (1955).

18. W. Orchard-Hays, Matrices, Elimination and the Simplex Method CEIR, Inc., Washington, 1957).

19. G. B. Dantzig and A. Orden, Duality Theorems, RAND Report RM-1265 (The Rand Corporation, Santa Monica, Calif. 1953).

20. L. R. Ford, Differential Equations (McGraw-Hill Book Co., Inc., New York, 1933).

21. E. L. Ince, Ordinary Differential Equations (Dover Publications, Inc., New York, 1926).

22. E. Kamke, Differentialgleichungen reeler Funktionen (Chelsea Publishing Company, New York, 1947).

23. D. L. Bernstein, Existence Theorems in Partial Differential Equations (Princeton University Press, Princeton, N. J., 1950).

24. I. N. Sneddon, Elements of Partial Differential Equations (McGraw-Hill Book Co., Inc., New York, 1957).

25. W. E. Milne, Numerical Solution of Differential Equations (John Wiley and Sons, Inc., New York, 1953).

26. L. Fox, The Numerical Solution of Two-Point Boundary Problems in Ordinary Differential Equations (Oxford University Press, London, 1957).

27. G. E. Forsythe and W. R. Wasow, Finite-Difference Methods for Partial Differential Equations (John Wiley & Sons, Inc., New York, 1960).

28. H. Levy and E. A. Baggott, Numerical Solutions of Differential Equations (Dover Publications, Inc., New York, 1950).

29. E. Goursat and E. R. Hedrick, A Course in Mathematical Analysis (Ginn and Co., Boston, 1904).

30. W. Orchard-Hays, SCROL, A Comprehensive Operating System for Linear Programming on the IBM 704 (CEIR, Inc., Washington, 1959).