JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 53, 521–538 (1976)

On the Identification of Inhomogeneous Parameters in Dynamic Linear Partial Differential Equations

S. YAKOWITZ

Department of Systems and Industrial Engineering, University of Arizona, Tucson, Arizona 85721

AND

P. NOREN

Department of Mathematics, University of Arizona, Tucson, Arizona 85721

Submitted by E. Stanley Lee

The task of identifying inhomogeneous (position-dependent) coefficients of linear dynamic partial differential equations on the basis of a finite collection of points of the solution has practical importance and is the subject of many published analyses, some of which are described herein. The purpose of the present paper is to present new developments on a simple yet appealing method due to the hydrologist B. Sagar. The technique exploits the viewpoint that the coefficient values of the partial differential equation at any point x are uniquely determined by the solution values in a small neighborhood of x. The identification algorithm which results from these considerations is extremely simple, and yet, in view of technical considerations and experimental evidence set forth here, it seems effective. In particular, we have been able to derive error bounds, which the authors believe is a new feature in the literature of identification of partial differential systems.

1. INTRODUCTION

The problem of parameter identification is of considerable interest and importance in engineering and hydrology and, in the case of ordinary differential equations and linear systems, a large body of literature (reviewed by Eykhoff [1] and Mendel [2]) is already available. In the case of partial differential equations, engineering needs exist; Phillipson [3] cites applications relevant to the problem of vibrations in a slender airframe and to seismic signals wherein the propagation of waves is used to detect anomalous underground objects. Bellman and Kalaba [4] study a problem of finding parameters in a radiative transfer phenomenon. The largest body of literature to come to

our attention on the subject of identification of parameters of partial differential equations of an aquifer on the basis of piezometric head measurements (e.g., [5–8]).

The purpose of the present paper is to set forth an inclusive method based on an idea of Sagar (see Sagar [9] or Sagar *et al.* [10]) for performing parameteric identification. The chief advantages of the Sagar method, in comparison to its published alternatives, is that it is extremely easy to implement, the computational burden is relatively light, and error bounds are available.

As background and by way of contrast, we provide here a brief review of some of the main avenues of attack on the PDE parameter identification problem. The first avenue we discuss is inspired by the technique called, "quasilinearization," developed mainly by Bellman and Kabala (see especially [4]). Basically, quasilinearization is a successive approximation technique for solving nonlinear differential equations by means of iterative solution of simpler linear differential equations. But, as proposed in [4, Chapter 6], quasilinearization may be made to serve as an important step in parametric identification, as follows: Let us suppose that a finite set $\{u(x_i, t_i)\}$ of observations of the solution u of a PDE parameterized by an unknown vector K is available. A criterion function $J(u, u_{K'})$ is chosen to measure, in some sense, how well the function $u_{K'}$ agrees with the known values of u. $u_{K'}$ is a solution (or approximation thereof) of the PDE having parameter K'. Now variational techniques are used to convert the minimization problem to a PDE problem. Quasilinearization is then employed to simultaneously obtain a new estimate, K'', of K, and a solution of a linearized version of the PDE with parameter K''. K'' now plays the role of K', and the above steps are repeated. Further illustrations of this technique include Bellman *et al.* [11] and, for the aquifer problem, Yeh and Tauxe [5]. A technique that is closely related to the preceding is one wherein the gradient of J with respect to K' is computed and K'' is chosen by a steepest descent technique. Seinfeld [12] applies this method to the problem of finding a parameter of a parabolic PDE describing a chemical reaction. Quasilinearization (and, to a lesser extent, the steepest descent method) has the desirable property that convergence is fast in a neighborhood of the unknown parameter, but in general, convergence is not guaranteed at all. Bellman and Kalaba assert [4, p. 156]: "The principal difficulty in the application of the theory of quasilinearization lies in the restrictive conditions required to ensure convergence of the sequence of successive approximations."

An additional drawback in the application of the above methods to the aquifer identification problem is that we are interested in the case of perhaps several hundred coordinates in the unknown parameter. In this circumstance, the quasilinearization and gradient methods become unwieldy, and computa-

tions which we have found reported have been applied only to the case in which the parameter is homogeneous in space and of low dimension.

Filtering theory inspires the techniques suggested by Phillipson [3, Appendix A1] and Chen and Seinfeld [13]. Let $u(x, t)$ denote the state variable of the PDE and $K(x)$ denote the unknown parameter. In the above references, one augments $K(x)$ to $u(x, t)$ to form the new state variable, $z(x, t) = (u(x, t), K(x))$. It is clear how to convert the original linear differential equation and boundary and initial conditions to the new nonlinear PDE system for the variable z. An additional constraint is that the partial with respect to t of that portion of z corresponding to the K, must be 0. With the new differential equation system before us, the original parameter identification problem has been converted to a state identification problem with quadratic cost and a nonlinear system. Phillipson [3] recommends linearizing the resultant PDE and making successive approximations (and this avenue would lead to something akin to the quasilinearization approach), whereas Chen and Seinfeld [13] derive an admittedly complicated solution to the nonlinear quadratic-cost filtering problem.

Let $\mathscr{A}$ denote a finite set of nonnegative integer n-tuples. D^α is the operator $\partial^{\alpha_1+\cdots+\alpha_n}/\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}$ where $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathscr{A}$. D_t is the operator $\partial/\partial t$.

The idea (inspired by [9]) motivating the present parameter identification study is simple. Let x be fixed and observe that if in the PDE

$$\Sigma_{\alpha \in \mathscr{A}} K_\alpha(x)\, D^\alpha u(x, t) = \sum_{v=0}^{V} C_v D_t^v u(x, t) + q(x, t), \tag{1}$$

the partials $D^\alpha u$ are known for J different times t_j, $1 \leqslant j \leqslant J$, where J is greater than or equal to the number of unknown parameters K_α, $\alpha \in \mathscr{A}$, then one can solve the resulting J simultaneous linear equations for the K_α values (unless the system happens to be degenerate—a situation which one would not anticipate in physical applications). In the overdetermined case, a least-squares solution would probably be indicated, but other alternatives may be preferred in certain physical situations. For example (as in [10]), one can find the least-squares fit among parameter values eonstrained to satisfy some property (e.g., nonnegativity of the coefficients K_α and ellipticity of the operator on the left-hand side of (1)) which is known to hold in light of physical considerations.

In general, one would anticipate that u is known at only finitely many space-time points $\{(x_i, t_j),\ 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J\}$. Splines and/or numerical differentiation may be used to compute estimates $D^\alpha u'(x, t)$ of the requisite partials of Eq. (1). Elementary methods are available for computing a bound on the differences $| K_\alpha - K_\alpha' |$ between the exact and computed parameter values on the basis of bound on the differences $| D^\alpha(u(x, t_j) - u'(x, t_j))|$

between the actual and interpolated values of u and its derivatives evaluated at x and t_j. The method is described more fully and the above notions are substantiated in Section 2.

The appeal of the Sagar method is that the computations for finding $K_\alpha(x)$, $\alpha \in \mathscr{A}$, x fixed, requires solving a relatively loworder linear equation, and, therefore, calculations are inexpensive and do not require much computer memory. The fact that the identification scheme requires solution of a relatively low-order system and (when differencing is used, at least) only local samples of u near x accords with reason in that intuitively the parameter value $K_\alpha(x)$ is uniquely determined by values of u in the neighborhood of x. There should be no need to solve global equations to find the values $K_\alpha(x)$, as must be done in the other approaches cited. In fact, it is not clear that values of u at some distance from x are of value for determining K_α at x.

In summary, we feel that the two important features of the Sagar method are: (i) it reduces the identification problem for $K_\alpha(x)$, x fixed, to one of relatively small degree; and (ii) it admits computation of meaningful error bounds. The advance in this paper over Sagar [9] is that we have generalized the approach from the unnecessarily restrictive aquifer setting and have provided error bounds.

2. The Method

Let $\mathscr{A}$ and D^α be as described in the Introduction. The general format of an inhomogeneous (i.e., spacially variable), time-invariant linear PDE is

$$\Sigma_{\alpha\in\mathscr{A}} K_\alpha(x)\, D^\alpha u(x, t) = q(x, t) + \Sigma_v C_v D_t^v u(x, t). \tag{1a}$$

The parameter identification problem of concern to us is the following: Given the function u (or an approximation thereof) the constants C_v, $1 \leqslant v \leqslant V$, and the forcing function $q(x, t)$, compute $K_\alpha(x)$, all $\alpha \in \mathscr{A}$, for any given x.

Let M denote the number of unknown parameters (i.e., the cardinality of $\mathscr{A}$). Our generalization of Sagar's method requires first computing $D^\alpha u(x, t_j)$ and $D_t^v u(x, t_j)$ (or approximations denoted by $D^\alpha u'(x, t_j)$ and $D_t^v u'(x, t_j)$ at J distinct time points t_j, $1 \leqslant j \leqslant J$, where $J \geqslant M$.

From these values, we obtain the simultaneous linear equations

$$\Sigma_{\alpha\in\mathscr{A}} K_\alpha(x)\, D^\alpha u(x, t_j) = q(x, t_j) + \Sigma C_v D_t^v u(x, t_j), \qquad 1 \leqslant j \leqslant J. \tag{1b}$$

which for notational convenience is restated in matrix form:

$$\mathbf{U}_x \mathbf{K} = \mathbf{Q} + \mathbf{U}_t .$$

Here $\mathbf{K}$ is a tuple of the $K_\alpha(x)$ terms, $\alpha \in \mathscr{A}$, and $\mathbf{U}_x$ is the matrix whose coordinate at the jth row and column corresponding to α is $D^\alpha u(x, t_j)$. The columns are ordered to correspond with the correct rows of $\mathbf{K}$. $\mathbf{Q}$ is the vector $(q(x, t_1),..., q(x, t_j))^T$ and $\mathbf{U}_t$ is the vector whose jth coordinate equals $\Sigma_v C_v D_t^v u(x, t_j)$.

It is evident that if $\mathbf{U}_x$ and $\mathbf{U}_t$ are determined by exact values of u then one may as well set $J = M$ (M being the cardinality of $\mathscr{A}$), and provided $\mathbf{U}_x$ is nonsingular, we have

$$\mathbf{K} = \mathbf{U}_x^{-1}(\mathbf{Q} + \mathbf{U}_t). \tag{3}$$

Of course, in all but rare instances it will be necessary to work with approximations u' of u. (Presently, we will review plausible candidates for constructing these requisite approximations from finite sets of physical measurements.) Under this circumstance, the matrix $\mathbf{U}_x'$ and vector $\mathbf{U}_t'$ constructed from the approximating function u' differs from $\mathbf{U}_x$ and $\mathbf{U}_t$. The estimate $\mathbf{K}'$ of the parameter $\mathbf{K}$ got by solving the linear equation

$$\mathbf{U}_x'\mathbf{K}' = \mathbf{Q} + \mathbf{U}_t' \tag{4}$$

is, consequently, in error. The discussion to follow applies to any vector norm. Elementary numerical analysis considerations lead to a bound on the norm of $\mathbf{K}' - \mathbf{K}$ in terms of a bound for $\| \mathbf{U}_t' - \mathbf{U}_t \|$ and a bound for $\| \mathbf{U}_x' - \mathbf{U}_x \|$, the matrix norm being the norm generated by the vector norm. It is clear how bounds are readily established from the following result.

PROPOSITION 1. *Let* $\delta\mathbf{U} \equiv \mathbf{U}_x - \mathbf{U}_x'$ *satisfy* $\|(\mathbf{U}_x')^{-1}\| \| \delta\mathbf{U} \| < 1$. *Then*

$$\| \mathbf{K}' - \mathbf{K} \| \leqslant \|(\mathbf{U}_x')^{-1}\|^2 (1 - \|(\mathbf{U}_x')^{-1}\| \| \delta\mathbf{U} \|)^{-1} \| \delta\mathbf{U} \| (\| \mathbf{Q} + \mathbf{U}_t' \| + \| \mathbf{U}_t' - \mathbf{U}_t \|) + \|(\mathbf{U}_x')^{-1}\| \| \mathbf{U}_t' - \mathbf{U}_t \|.$$

Proof. For brevity of exposition, for this proof only we define $A = \mathbf{U}_x'$, $B = \mathbf{U}_x'$, $y = \mathbf{Q} + \mathbf{U}_t$, $y' = \mathbf{Q} + \mathbf{U}_t'$, and $\Delta = \mathbf{A} - \mathbf{B}$. Then

$$\begin{aligned} \| \mathbf{K} - \mathbf{K}' \| &= \| B^{-1}y' - A^{-1}y \| \\ &\leqslant \| B^{-1} - A^{-1} \| \| y' \| + \| A^{-1} \| \| y - y' \| \\ &\leqslant \| B^{-1} - A^{-1} \| \| y' \| + (\| B^{-1} \| + \| B^{-1} - A^{-1} \|) \| y - y' \|. \end{aligned}$$

But

$$\begin{aligned} \| B^{-1} - A^{-1} \| &= \| B^{-1} - (B + \Delta)^{-1} \| \leqslant \| B^{-1} \| \| I - (I + \Delta B^{-1})^{-1} \| \\ &= \| B^{-1} \| \| (I + \Delta B^{-1})^{-1} \Delta B^{-1} \| \\ &\leqslant \| B^{-1} \| \| \Delta B^{-1} \| \| (I + \Delta B^{-1})^{-1} \|. \end{aligned}$$

If $\| \Delta \| \| B^{-1} \| < 1$, as hypothesized,

$$\|(I + \Delta B^{-1})^{-1}\| \leqslant (1 - \| \Delta B^{-1} \|)^{-1} \leqslant (1 - \| \Delta \| \| B^{-1} \|)^{-1}.$$

Thus, in summary,

$$\| \mathbf{K} - \mathbf{K}' \| \leqslant \| B^{-1} \|^2 \| \Delta \| (1 - \| \Delta \| \| B^{-1} \|)^{-1} (\| y' \| + \| y - y' \|) + \| B^{-1} \| \| y - y' \|,$$

which yields the proposition when the correct substitutions are made for Δ, B, y, etc.

In the usual case of inexact values u', it would seem a sensible procedure to use as large a value J as possible, consistent with reasonable cost of measurement and computation. Whenever $J > M$, the error bounding problem can be cast into a regression analysis framework. For $1 \leqslant j \leqslant J$, let us define

$$\eta_j = \Sigma C_v D_t^v (u(x, t_j) - u'(x, t_j)) - \Sigma K_\alpha(x) (D^\alpha (u(x, t_j) - u'(x, t_j)). \tag{4}$$

If it is supposed that the above differences η_j are independent random variables which are bounded by σ^2, or more generally, have a known variance σ^2, then error bounds on $\| \mathbf{K}' - \mathbf{K} \|$ are afforded by the Gauss–Markov theorem, which we restate in the terminology of our problem.

PROPOSITION 2. *Suppose* $E[\eta_j] = 0$ *and the variance (or bound) of* η_j *is* σ^2. *Then provided the columns of* U_x *are linearly independent, the vector*

$$\mathbf{K}' = [(\mathbf{U}_x')^T \mathbf{U}_x']^{-1} \mathbf{U}_x'^T (\mathbf{Q} + \mathbf{U}_t')$$

is the minimum variance linear unbiased estimator of $\mathbf{K}$, *and the covariance matrix of* $\mathbf{K}' - \mathbf{K}$ *is* $[(\mathbf{U}_x')^T \mathbf{U}_x']^{-1} \sigma^2$. *In particular, for the Euclidean norm,*

$$E[\| \mathbf{K}' - \mathbf{K} \|^2] = \operatorname{trace}(((\mathbf{U}_x')^T \mathbf{U}_x')^{-1}) \sigma^2.$$

$\mathbf{K}'$ *is also the least square estimator of* $\mathbf{K}$.

Proof. See Wilks [14, p. 283].

If by virtue of physical reasons or experience, one has some bound for the vector $\mathbf{K}$ as well as bounds for the interpolation error $[u(x, t_j) - u'(x, t_j)]$ and its derivatives, then Eq. (4) yields directly a bound for η_j, and hence, from the proposition, a bound on the estimation error $\| \mathbf{K}' - \mathbf{K} \|$. On the other hand, if no bound on $\mathbf{K}$ is available, the estimate itself can be used in (4) to approximate the requisite bound of η_j. The size of the resultant error bound will suggest the accuracy of the approximate "bound" for η_j.

In most engineering applications, the solution u is not available, but must be approximated by a function u' computed from a finite collection of observations $\{u(x_i, t_j)\}$, $1 \leqslant i \leqslant I$, $1 \leqslant j \leqslant J$. In order for the identification procedure to be effective, it is necessary not only that u' approximate u, but also that $D^\alpha u'$ and $D_t^v u'$ approximate $D^\alpha u$ and $D_t^v u$, respectively. Additionally, in order to avail ourselves of the error bounds of Propositions 1 and 2, it is necessary to be able to bound the differences $|D^\alpha(u - u')|$, $|D_t^v(u - u')|$.

We catalog below the techniques known to us for achieving convergence of the needed derivatives, with error bounds. At this point, we note that in the aquifer problem as well as other applications, it may be possible to obtain measurements of the time derivatives of $u(x, t)$ at given locations x, and consequently, in some cases, interpolation with respect to time can be avoided.

Spline approximations were recommended and employed by Sagar [9]. Splines have appeal over numerical differentiation in that under reasonable regularity conditions the rate convergence is faster and the approximation takes advantage of points other than those in the immediate vicinity of the point in question. However, a serious drawback of the spline approach to our interpolation problem is that (with the exception of a preliminary study by Shah [5]) the properties of multivariate interpolation splines on non rectangular meshes have not been derived. Thus, to achieve bounds, it is necessary to make the measurements at regularly spaced points so that for fixed coordinates $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ one has a sequence of points in the x_i direction from which to construct the one-dimensional spline $S_{x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n}(x)$. From this family of splines indexed by $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$, $1 \leqslant i \leqslant N$, the one-dimensional spline theory yields error bounds on partial derivatives of u' with respect to x_i, the partials being of all orders up to one less than the degree of the spline polynomials. To evaluate mixed partials or partials at points not along one of the spline lines, it is necessary to use numerical differentiation. But error bounds are available for this operation also. The text by Nilson, Ahlberg, and Walsh [16] has been useful for our studies, and, in particular, from [16, Theorem 4.4.1] we have the asymptotic result that if the spline is composed of polynomials of order $2v - 1$, under lengthy but reasonable conditions on the function u to be approximated, uniformly

$$\frac{\partial^p}{\partial x_i^p}\left(u(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) - S_{x_1,\ldots,x_{i-1},x_{i+1},\ldots x_n}(x)\right) = O(\|\Delta_i\|^{2v-2-p})$$

where $\|\Delta_i\|$ is a bound on the partition length and p ranges from 0 to $2v - 2$.

The following general result [16, p. 142] holds.

PROPOSITION 3. *If an interval* $[a, b]$ *is partitioned by a mesh* Δ, *and* S_Δ *is a* $2n - 1$ *degree spline interpolation of* $f \in \mathscr{C}^{2n-2}[a, b]$, *then for* $p \leqslant 2n - 2$,

$$| S_\Delta^{(p)}(x) - f^{(p)}(x)| \leqslant \mu[f^{(p)}; p/2 \,\|\, \Delta \,\|] + \mu[S_\Delta^{(p)}; p/2 \,\|\, \Delta \,\|].$$

In the above, $\|\, \Delta \,\|$ *refers to the maximum distance between gridpoints and for* g *a real function on* $[a, b]$ *and* $c > 0$,

$$\mu(g; c) \equiv \sup_{|x-y|<c} (g(x) - g(y)).$$

To determine a spline interpolation uniquely, it is necessary to prescribe additional "end conditions," which are not always available from the data of a given physical problem. Theorem 2.9.3 of [16] shows, however, that even if incorrect end conditions are prescribed, the second derivatives of a cubic spline will converge uniformly in closed subintervals.

Some numerical differentiation theory is available for multivariate functions, but the situation is far less satisfactory than the one-dimensional theory. For the existing multivariate theory, there is no constraint on the measurement locations other than that the difference vectors about a point of interest span the space. Issacson and Keller [17, p. 298] outline the theory and error bounds for multivariate numerical differentiation.

In many engineering applications it may be anticipated that the measurements $\{u(x_i, t_i)\}$ are corrupted by additive noise. One of the present authors has devoted much consideration to the problem of interpolation from noisy measurements and on the basis of both study of available literature and computer simulation, has concluded that at present the method of potential functions (see [18] and references therein) appear to be the strongest weapon. Briefly, the method of potential functions yields an iterative sequence of approximations $\{u^N\}_N$. u^N is based on the noisy sample pairs

$$\{((x_i, t_i), u(x_i, t_i) + \eta_i): 1 \leqslant i \leqslant N\}$$

where the sequence $\{\eta_i\}$ denotes an independent sequence of random variables whose variance is bounded but whose distribution is unknown. The approximation u^N of u turns out to be of the form

$$u^N(x, t) = \sum_{i=1}^{M} C_i^N \phi_i(x, t) \qquad (M \text{ possibly infinite}),$$

where the ϕ_i's comprise a basis for some Hilbert space known to contain u. It has long been known that if M is finite, an error bound can be achieved. Recently, Fisher and Yakowitz [19] have derived an error bound for $M = \infty$.

The problem of bounding the errors of the derivatives has not been investigated for $M = \infty$, but for $M < \infty$, the bound is obvious.

3. Illustrative Computations Using the Identification Algorithm

To test the effectiveness of the methodology proposed in the previous section, we chose to study a parabolic equation in one space variable of the form

$$K_1(x)\, u_{xx} + K_2(x)\, u_x + K_3(x) = u_t\,. \tag{5}$$

The inhomogeneous parameters K_1, K_2, K_3 were selected in such a manner that one is able to derive a simple closed-form solution and thus avoid errors resulting from numerical solution of (5). Toward that end, it was decided that u should have the form

$$u(x, t) = (x + 1/3x^3)\, t^2 + (1/3x^3 + 7/30x^5 + 1/42x^7)\, t + 3/10x^5 + 1/2x^3. \tag{6}$$

One may verify that the associated functions $K_j(x)$ are as below

$$\begin{aligned} K_1(x) &= (1 + x^2)^{-1} \\ K_2(x) &= d/dx(K_1(x)) \\ K_3(x) &= 1/3x^3 + 7/30x^5. \end{aligned} \tag{7}$$

In this section, several identification studies of the preceding PDE will be reported in which different interpolation schemes are employed and in which the interval length and gridpoint spacings are varied. To begin with, however, we will relate carefully the details and output of one particular program so that the reader may gain some acquaintance with how closely interpolation methods approximate the function u and additionally, the manner in which the estimates track the actual values of the functions K_1, K_2, and K_3 as the space variable changes.

For the program about which we speak in detail, we chose as our set of positions the interval $[-5, 5]$. The sample points were evenly spaced with a separation of 0.5 (i.e., there were 21 gridpoints). The samples $u(x_i, t_j)$ were obtained from each gridpoint x_i at each of the times $t_j = 0, 1, 2$. For purposes of this particular run, (cubic) spline interpolation was employed. The end conditions were (incorrectly) taken to satisfy

$$\begin{aligned} u_1'' &= \frac{1}{2}\left[\frac{6(((u_3 - u_2)/(x_3 - x_2)) - ((u_2 - u_1)/(x_2 - x_1)))}{(x_3 - x_1)} - u_2''\right] \\ u_{21}'' &= \frac{1}{2}\left[\frac{6(((u_{21} - u_{20})/(x_{21} - x_{20})) - ((u_{20} - u_{19})/(x_{20} - x_{19})))}{(x_{21} - x_{19})} - u_{20}''\right] \end{aligned} \tag{8}$$

TABLE I

Tabulation of a Function, Its Derivatives, and Its Spline Approximation

		u			$\partial u/\partial x$			$\partial^2 u/\partial x^2$		
		Exact	By spline	Relative error	Exact	By spline	Relative error	Exact	By spline	Relative error
$x = -5.$,	$t = 0$	−1000			975	942	0.033	−765	−536	0.298
(Gridpoint)	$t = 1$	−3677			4359	4109	0.057	−4493	−2705	0.398
	$t = 2$	−6448			7795	7327	0.060	−8241	−4890	0.406
$x = -4.7500$,	$t = 0$	−779	−781	−0.0033	797	803	−0.007	−657	−576	0.122
	$t = 1$	−2718	−2737	−0.0071	3351	3397	−0.013	−3594	−2986	0.169
	$t = 2$	−4738	−4775	−0.0077	5953	6039	−0.014	−6550	−5414	0.173
$x = -4.5000$,	$t = 0$	−599			645	654	−0.013	−560	−616	−0.101
(Gridpoint)	$t = 1$	−1984			2549	2615	−0.026	−2848	−3268	−0.147
	$t = 2$	−3439			4495	4620	−0.027	−5155	−5938	−0.151
$x = -4.2500$,	$t = 0$	−454	−453	0.0016	516	514	0.003	−473	−496	−0.049
	$t = 1$	−1429	−1424	0.0039	1916	1904	0.006	−2235	−2423	−0.084
	$t = 2$	−2464	−2454	0.0043	3354	3331	0.006	−4013	−4367	−0.088

$x = -4.0000$,	$t = 0$	-339			408	405	0.005	-396	-376	0.048
(Gridpoint)	$t = 1$	-1014			1422	1404	0.012	-1734	-1578	0.090
	$t = 2$	-1741			2470	2436	0.013	-3089	-2796	0.094
$x = -3.7500$,	$t = 0$	-248	-249	-0.0007	317	318	-0.001	-327	-323	0.013
	$t = 1$	-709	-710	-0.0017	1041	1044	-0.003	-1330	-1298	0.024
	$t = 2$	-1211	-1214	-0.0019	1794	1800	-0.003	-2347	-2288	0.025
$x = -3.5000$,	$t = 0$	-179			243	244	-0.002	-267	-269	-0.006
(Gridpoint)	$t = 1$	-486			750	754	-0.006	-1007	-1017	-0.010
	$t = 2$	-830			1283	1292	-0.006	-1760	-1780	-0.011
$x = -3.2500$,	$t = 0$	-125	-125	0.0005	183	183	0.000	-215	-218	-0.013
	$t = 1$	-327	-327	0.0016	531	531	0.001	-751	-772	-0.028
	$t = 2$	-559	-558	0.0017	903	902	0.001	-1300	-1340	-0.030
$x = -3.0000$,	$t = 0$	-86			135	134	0.001	-171	-167	0.019
(Gridpoint)	$t = 1$	-216			370	368	0.004	-552	-527	0.043
	$t = 2$	-369			625	622	0.004	-945	-900	0.047
$x = -2.7500$,	$t = 0$	-57	-57	0.0001	97	97	-0.000	-133	-133	-0.004
	$t = 1$	-139	-139	-0.0001	252	252	-0.001	-398	-401	-0.007
	$t = 2$	-240	-240	-0.0001	424	424	-0.001	-674	-697	0.007

where $x_1, \ldots, x_{21}$ are the gridpoints of the space variable x, and for fixed t,

$$u_j = u(x_j, t), \qquad 1 \leqslant j \leqslant 21.$$

u_j'' denotes the approximation of $(\partial^2/\partial x^2)\, u(x_j, t)$. Equations (8) and the condition

$$\frac{(x_i - x_{i-1})}{(x_{i+1} - x_i)} u_{i-1}'' + \frac{2(x_{i+1} - x_{i-1})}{(x_{i+1} - x_i)} u_i'' + u_{i+1}''$$
$$= \frac{6}{x_{i+1} - x_i}\left[\frac{u_{i+1} - u_i}{x_{i+1} - x_i} - \frac{u_i - u_{i-1}}{x_i - x_{i-1}}\right],$$

$i = 2, \ldots, 20$, give a complete set of estimates $\{u_i''\}_{i=1}^{21}$ which, in turn, uniquely determine the cubic spline.

Table I compares the spline interpolated and estimated values of u and its derivatives with the exact value (Eq. (6)) over a portion of the space interval. We omitted inclusion of the spline values at the gridpoints because, of course, they are exact, as a consequence of the defining properties of splines. The approximations tabulated here are not exemplary; our experience is that Lagrange interpolation performs significantly better for the problem at hand.

Using the approximate values of u given by the spline interpolation as described above, Eq. (3) was solved for K' at a series of position values both at the gridpoints and between gridpoints. The results of these computations are given in their entirety in Table II.

Notice that the parameter values are changing reasonably briskly with change in position. The parameter value at the position 0 cannot be calculated because the matrix $\mathbf{U}_x'$ in (3) is singular at this point. The error in estimation at the ends of the space interval (i.e., -5 and 5) are to be anticipated because we used incorrect end conditions in evaluating the spline.

We tested spline interpolation with grid spacing 0.1 and 0.05 also. All these results will be summarized, along with error behavior arising in other interpolation schemes we invesitagted. First, however, we describe the details of these alternative interpolation methods, all of which surpassed the spline method in accuracy for the differential equation identification problem studied in this section. For future reference, we designate identification using spline interpolation as *Method A*.

Method B (*Finite difference*). This is the simplest possible technique; no interpolation is performed at all, and consequently, this technique can be used only for identification at measurement points. Let x_i, $1 \leqslant i \leqslant 21$, denote the positions at which measurements are made. Then the standard

TABLE II

Comparison of Exact and Estimated (Method A) Parameters[a]

	K_1			K_2			K_3		
	Exact	Estimate	Relative error	Exact	Estimate	Relative error	Exact	Estimate	Relative error
$x = -4.50$*	0.04	0.03	0.17	0.01	0.01	0.14	−1337.0	−1337.0	−0.0005
$x = -4.25$	0.05	0.04	0.12	0.02	0.02	0.11	−932.0	−933.0	−0.0007
$x = -4.00$*	0.05	0.06	−0.16	0.02	0.03	−0.15	−638.0	−637.0	0.0014
$x = -3.75$	0.06	0.06	−0.04	0.03	0.03	−0.04	−427.0	−427.0	0.0005
$x = -3.50$*	0.07	0.07	0.01	0.03	0.03	0.01	−279.0	−297.0	−0.0002
$x = -3.25$	0.08	0.08	0.05	0.04	0.04	0.05	−177.0	−177.0	−0.0016
$x = -3.00$*	0.10	0.10	−0.08	0.06	0.06	−0.08	−108.0	−108.0	0.0042
$x = -2.75$	0.11	0.11	0.01	0.07	0.07	0.01	−63.0	−63.0	−0.0005
$x = -2.50$*	0.13	0.14	−0.05	0.09	0.10	−0.05	−35.0	−34.0	0.0054
$x = -2.25$	0.16	0.15	0.04	0.12	0.11	0.04	−17.45	−17.62	−0.0100
$x = -2.00$*	0.20	0.22	−0.11	0.16	0.17	−0.11	−7.18	−6.82	0.0489
$x = -1.75$	0.24	0.23	0.04	0.21	0.20	0.04	−1.56	−1.64	−0.0541
$x = -1.50$*	0.30	0.35	−0.13	0.28	0.32	−0.13	1.19	1.43	−0.1990
$x = -1.25$	0.39	0.36	0.07	0.38	0.35	0.07	2.27	2.17	0.0450
$x = -1.00$*	0.50	0.61	−0.22	0.50	0.61	−0.22	2.40	2.63	−0.0927
$x = -0.75$	0.64	0.58	0.08	0.61	0.56	0.08	2.05	2.02	0.0134
$x = -0.50$*	0.80	1.04	−0.30	0.64	0.83	−0.30	1.45	1.52	−0.0508
$x = -0.25$	0.94	0.83	0.11	0.44	0.39	0.11	0.74	0.73	0.0115
$x = 0.00$*	1.00	−999.00	1000.00	−0.00	−999.00	—	0.00	−999.00	—

[a] *Note.* Parameters and their estimates are symmetric about 0.

* Gridpoints.

equation for numerical differentiation (e.g., [20, p. 191]) is employed (assuming t is fixed and, therefore, suppressed)

$$\begin{aligned}(\partial/\partial x)\, u_i &\cong [u_{i+1} - u_{i-1}]/2(x_{i+1} - x_{i-1}) \\ (\partial^2/\partial x^2)\, u_i'' &\cong (u_{i+1} - 2u_i + u_{i-1}))/(x_{i+1} - x_{i-1})^2.\end{aligned} \tag{9}$$

Method C (*Lagrange interpolation*). A Lagrange interpolation polynomial is passed through the 21 measurement values and, consequently, the degree of the polynomial is 20. Derivatives are calculated as in Method B to find the parameter at measurement points and to compute the values of the parameters $K_j(x)$, $j = 1, 2, 3$, for x not a measurement point, derivatives of u are obtained by Eq. (9) using the Lagrange interpolation polynomial. A sensible variation, which we did not investigate, is to use the fact that the interpolation function is a polynomial to obtain derivatives "synthetically."

Method D (*Lagrange interpolation of the function and its derivatives*). First, a Lagrange interpolation polynomial P_1 is passed through the measured values $u(x_i)$, $1 \leqslant i \leqslant 21$. Then, using the values $u(x_i) = u_i$, for x_i's being the measurement (or grid) positions, numerical differentiation is performed according to Eq. (9) to obtain derivatives $\partial/\partial x\, u_i$ at the interior gridpoints x_i, $2 \leqslant i \leqslant 20$. Then an eighteenth degree polynomial P_2 is used to interpolate these values. An interpolation polynomial P_3 is passed through the values $\partial^2/\partial x^2\, u(x_i)$ also obtained by numerical differentiation. Then for any x, the values $u(x)$, $\partial/\partial x\, u(x)$ and $\partial^2/\partial x^2\, u(x)$ needed for identification are evaluated directly from the rule

$$\begin{aligned}u(x) &= P_1(x) \\ \partial/\partial x\, u(x) &= P_2(x) \\ \partial^2/\partial x^2\, u(x) &= P_3(x).\end{aligned}$$

Table III tabulates the parameter estimates obtained using Method C for the same problem as reported in Table II. Table IV compares the accuracy of all the methods on various intervals and with different grid spacings.

4. Closing Comments

The identification algorithm we have studied here is enormously less sophisticated than its alternatives in the published literature, and it is correspondingly easier to program and less demanding of computation and memory. Yet, it has some appealing properties (beyond its simplicity):

TABLE III

Comparison of Exact and Estimated (Method C) Parameters[a]

	K_1			K_2			K_3		
	Exact	Estimate	Relative error	Exact	Estimate	Relative error	Exact	Estimate	Relative error
$x = -4.50$*	0.0471	0.0470	0.0017	0.0199	0.0198	0.0056	−1337.1298	−1337.2474	−0.0001
$x = -4.25$	0.0525	0.0524	0.0018	0.0234	0.0232	0.0062	−932.6852	−932.8074	−0.0001
$x = -4.00$*	0.0588	0.0587	0.0020	0.0277	0.0275	0.0069	−638.3619	−638.4889	−0.0002
$x = -3.75$	0.0664	0.0662	0.0023	0.0331	0.0328	0.0077	−427.6587	−427.7907	−0.0003
$x = -3.50$*	0.0755	0.0753	0.0025	0.0399	0.0395	0.0088	−297.5315	−279.6687	−0.0005
$x = -3.25$	0.0865	0.0862	0.0029	0.0486	0.0481	0.0100	−177.4845	−177.6268	−0.0008
$x = -3.00$*	0.1000	0.0997	0.0033	0.0600	0.0593	0.0115	−108.7714	−108.9187	−0.0014
$x = -2.75$	0.1168	0.1163	0.0038	0.0750	0.0740	0.0134	−63.6992	−63.8512	−0.0024
$x = -2.50$*	0.1379	0.1373	0.0046	0.0951	0.0936	0.0159	−35.0270	−35.1828	−0.0044
$x = -2.25$	0.1649	0.1640	0.0056	0.1224	0.1201	0.0191	−17.4527	−17.6112	−0.0091
$x = -2.00$*	0.2000	0.1985	0.0073	0.1600	0.1562	0.0236	−7.1810	−7.3401	−0.0222
$x = -1.75$	0.2462	0.2437	0.0101	0.2121	0.2057	0.0300	−1.5630	−1.7197	−0.1003
$x = -1.50$*	0.3077	0.3030	0.0151	0.2840	0.2727	0.0398	1.1963	1.0484	0.1253
$x = -1.25$	0.3902	0.3807	0.0244	0.3807	0.3598	0.0551	2.2733	2.1359	0.0605
$x = -1.00$*	0.5000	0.4793	0.0413	0.5000	0.4602	0.0797	2.4095	2.2913	0.0490
$x = -0.75$	0.6400	0.5949	0.0704	0.6144	0.5422	0.1175	2.0508	1.9584	0.0451
$x = -0.50$*	0.8000	0.7100	0.1125	0.6400	0.5325	0.1680	1.4509	1.3885	0.0430
$x = -0.25$	0.9412	0.7948	0.1555	0.4429	0.3468	0.2169	0.7446	0.7136	0.0416
$x = 0.00$*	1.0000	−999.0000	1000.0000	−0.0000	−999.0000	—	0.0000	−999.0000	—

[a] *Note.* Parameters and their derivatives are symmetric about 0.

* Gridpoints.

TABLE IV

Comparison of Methods (Average Absolute Relative Error, Excluding Singular Point $x = 0$)

Method / Δx	A			B			C			D			Interval
	K_1	K_2	K_3	K_1	K_2	K_3	K_1	K_2	K_3	K_1	K_2	K_3	
0.05	0.0244	0.0116	0.0019	0.0018	0.0025	0.0004	0.0011	0.0015	0.0003	0.0017	0.0025	0.0004	[−0.5, 0.5]
0.1	0.0494	0.0352	0.0077	0.0050	0.0076	0.0018	0.0030	0.0046	0.0011	0.0051	0.0076	0.0018	[−1, 1]
0.5	0.1735	0.1552	0.0271	0.0211	0.0400	0.0237	0.0140	0.0258	0.0165	0.0253	0.0452	0.0274	[−5, 5]

(a) Error bounds are available (although in the example, they turned out to be overly inflated).

(b) No initial "guesses" are required (in contrast to the gradient and quasilinearization methods, for example.)

(c) The technique utilizes the notion that the effect of the parameter value at a point x is uniquely determined by values of u is a neighborhood of x.

Suppose we had wished to estimate the coefficients at 41 positions, as in Tables II and III, using the quasilinearization or gradient methods. It would have then been necessary to augment 123 coordinates to the 3-coordinate state variable and then repetitively solve boundary value differential equations of dimension 126.

ACKNOWLEDGMENTS

The authors are extremely grateful for the able programming assistance of Mr. Pak Chan. National Science Foundation Grant GK-35791 supported, in part, the efforts of both authors.

REFERENCES

1. P. EYKHOFF, Process parameter and state estimation, *Automatica* **4** (1968), 205–233.
2. J. MENDEL, "Discrete Techniques of Parameter Estimation," Dekker, New York, 1973.
3. G. PHILLIPSON, "Identification of Distributed Systems," Elsevier, New York, 1971.
4. R. BELLMAN AND R. KALABA, "Quasilinearization and Nonlinear Boundary Value Problems," Elsevier, New York, 1965.
5. W. G. YEH AND G. W. TAUXE, Quasilinearization and the identification of aquifer parameters, *Water Resources Res.* **7** (1971), 375–381.
6. V. VEMURI AND W. J. KARPLUS, Identification of nonlinear parameters of groundwater basins by hybrid computations, *Water Resources Res.* **5** (1969), 172–185.
7. Y. EMSELLEM AND G. DE MARSILY, An automatic solution for the inverse problem, *Water Resources Res.* **7** (1971), 1264–1283.
8. D. KLEINECKE, Use of linear programming for estimating geohydrologic parameters of groundwater basins, *Water Resources Res.* **7** (1971), 367–381.
9. B. SAGAR, "Calibration and Validation of Aquifer Models," Technical Report Number 17, Hydrology and Water Resources Department, University of Arizona, 1973.
10. B. SAGAR, L. DUCKSTEIN, AND S. YAKOWITZ, A direct method to solve the inverse problem of saturated flow in aquifers, *Water Resources Res.* **11** (1975), 563–570.
11. R. BELLMAN, H. KAGIWADA, AND R. KALABA, Quasilinearization, boundary value problems and linear programming, *IEEE Trans. Automatic Contr.* **AC-10** (1965), 199.
12. J. SEINFELD, Identification of parameters in partial differential equations, *Chem. Engrg. Sci.* **24** (1969), 65–74.

13. W. H. CHEN AND J. H. SEINFELD, Estimation of spatially varying parameters in partial differential equations, *Int. J. Contr.* **15** (1972), 487–495.
14. S. WILKS, "Mathematical Statistics," Wiley, New York, 1962.
15. J. SHAH, Two dimensional polynomial splines, *Numer. Math.* **15** (1970), 1–14.
16. J. AHLBERG, E. NILSON, AND J. WALSH, "The Theory of Splines and Their Applications," Academic Press, New York, 1967.
17. E. ISAACSON AND H. KELLER, "Analysis of Numerical Methods," Wiley, New York, 1966.
18. M. AIZERMAN, E. BRAVERMAN, AND L. ROZONOER, Extrapolative problems in automatic control and the method of potential functions, *Amer. Math. Soc. Transl.* **87** (1970), 281–303.
19. L. FISHER AND S. YAKOWITZ, Uniform convergence of the potential function algorithm, *SIAM Journal on Control*, to appear.
20. C. FROBERG, "Introduction to Numerical Analysis," 2nd ed., Addison Wesley, Reading, Mass., 1969.