Azmy S. Ackleh; Robert R. Ferdinand; Simeon Reich

Numerical studies of parameter estimation techniques for nonlinear evolution equations

Kybernetika, Vol. 34 (1998), No. 6, [693]--712

Persistent URL: http://dml.cz/dmlcz/135255

KYBERNETIKA — VOLUME 34 (1998), NUMBER 6, PAGES 693–712

NUMERICAL STUDIES OF PARAMETER ESTIMATION TECHNIQUES FOR NONLINEAR EVOLUTION EQUATIONS

AZMY S. ACKLEH, ROBERT R. FERDINAND AND SIMEON REICH

We briefly discuss an abstract approximation framework and a convergence theory of parameter estimation for a general class of nonautonomous nonlinear evolution equations. A detailed discussion of the above theory has been given earlier by the authors in another paper. The application of this theory together with numerical results indicating the feasibility of this general least squares approach are presented in the context of quasilinear reaction diffusion equations.

1. INTRODUCTION

There is a growing literature concerning the development of theoretical and computational methods for inverse problems involving the identification of nonlinear distributed parameter systems (see [1, 5, 6, 7, 8, 15] and the many references cited therein). In this paper we report on our recent results in this area with the main focus being the numerical studies. However, we shall provide a short discussion of the theoretical basis and convergence results that we obtained in our earlier paper [3]. In [3] we have extended the results in [8] to a more general class of nonlinear evolution systems that arise in many important applications including nonlinear reaction diffusion equations.

Our paper is organized as follows. In Section 2, we discuss the identification problem and review the theoretical results concerning convergence of parameter estimates established via Galerkin approximation. Section 3 is devoted to numerical results concerning the estimation of parameters in nonlinear reaction diffusion equations. Concluding remarks and future research issues are addressed in Section 4. Finally, we present in the Appendix an example indicating the convergence of the Galerkin approximations to the forward problem.

2. OUTLINE OF APPROXIMATION AND CONVERGENCE THEORY

Let D be a metric space with Q (the admissible parameter set) a compact subset of D. Let Z, the observation space, be a normed linear space with norm $|\cdot|_Z$. Let H be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $|\cdot|$, and let V be a reflexive Banach space with norm $\|\cdot\|$ which is densely and continuously embedded in H. This means that there exists a constant $\mu > 0$ such that $|\phi| \leq \mu \|\phi\|$ for all $\phi \in V$. Letting V^* be the space of continuous linear functionals defined on V, we denote the dual space norm on V^* by $\|\cdot\|_*$. Further, H is identified with its dual by the Riesz Representation Theorem and $V \subset H = H^* \subset V^*$ where H^* is the dual of H which is densely and continuously embedded in V^*. Since $\|\phi\|_* \leq \mu |\phi|$ for all $\phi \in H$, we get $\|\phi\|_* \leq \mu^2 \|\phi\|$ for all $\phi \in V$. The duality pairing between $\phi \in V^*$ and $\psi \in V$ is denoted by $\langle \phi, \psi \rangle$. Thus for $\phi \in H$, $\langle \phi, \psi \rangle$ is the usual inner product of ϕ and ψ.

For any fixed $T > 0$, $q \in Q$ and all $t \in [0,T]$, let $A(t;q)$ be a hemicontinuous (see [9] for the definition), in general nonlinear operator from V into V^* whose domain is all of V. In our discussion below, it is assumed that $A(t;q)$ and the nonlinear function $F(t,u;q)$, $t \in [0,T]$, $q \in Q$, satisfy the following conditions.

(A) (Continuity): For each $\phi \in V$, the map $q \rightarrow A(t;q)\,\phi$ is continuous from $Q \subset D$ into V^* for almost every $t \in [0,T]$.

(B) (Equi–V–Monotonicity): There exist constants ω and α, $\alpha > 0$, which don't depend on $q \in Q$ or $t \in [0,T]$ such that

$$\langle A(t;q)\,\phi - A(t;q)\,\psi, \phi - \psi \rangle + \omega\,|\phi - \psi|^2 \geq \alpha\,\|\phi - \psi\|^2$$

for all $\phi, \psi \in V$ and almost every $t \in [0,T]$.

(C) (Equiboundedness): There exists a constant $\beta > 0$, independent of $q \in Q$ or $t \in [0,T]$, such that

$$\|A(t;q)\,\phi\|_* \leq \beta\,(\|\phi\| + 1)$$

for all $\phi \in V$ and almost every $t \in [0,T]$.

(D) (Measurability): For each $q \in Q$, the function $A(t;q)\,u(t) : [0,T] \rightarrow V^*$ is strongly measurable for every $u \in L^2(0,T;V)$.

(E) (Lipschitz Continuity): The function $F : [0,T] \times H \times Q \rightarrow H$ is continuous. Moreover, F is locally Lipschitz continuous in H, uniformly for $q \in Q$ and $t \in [0,T]$.

For each $q \in Q$, let $\xi(q) \in H$ and assume that the mapping $q \rightarrow \xi(q)$ is continuous from $Q \subset D$ into H. For each $z \in Z$, we assume that the mapping $\Phi(\cdot; z)$ is defined on $L^2(0,T;V)$ with the non-negative real numbers containing its range. The functional Φ is assumed to be continuous when its domain is restricted to either one of the spaces $C(0,T;H)$ or $L^2(0,T;V)$ with their respective usual topologies. The following abstract parameter identification problem is considered.

(ID) Given observations $z \in Z$, find parameters $\overline{q} \in Q$ which minimize the performance index

$$J(q) = \Phi(u(\cdot, q); z),$$

where for each $q \in Q$, $u(\cdot\,,q)$ is the solution of the initial value problem

$$\begin{cases} \dot{u}(t) + A(t;q)\,u(t) = F(t,u(t);q) \\ u(0) = \xi(q). \end{cases} \tag{2.1}$$

Note that in this general formulation, the initial value ξ is part of the unknown parameters vector q (as in, for example, [4]). In [3], the following theorem concerning existence and uniqueness of solutions to equation (2.1) has been established using theoretical results on monotone operators discussed in [9].

Theorem 2.1. There exists a $T^* > 0$ such that equation (2.1) has a unique solution in the interval $[0,T^*]$, i.e., a function $u(\cdot;q)$ which is V^*-absolutely continuous on $[0,T^*]$, and satisfies $u(\cdot;q) \in L^2(0,T^*;V) \cap C(0,T^*;H)$, $\dot{u}(\cdot;q) \in L^2(0,T^*;V^*)$ and (2.1) for almost every $t \in [0,T^*]$.

We now present an abstract Galerkin based approximation theory for the problem (ID). For each $N = 1,2,\ldots$, let H^N be a finite dimensional subspace of H with $H^N \subset V$. Denoting by $P^N : H \to H^N$ the orthogonal projection of H onto H^N with respect to the inner product $\langle\cdot,\cdot\rangle$, we impose the following standard condition on our approximation elements.

(F) $$\lim_{N\to\infty} \left\|P^N\phi - \phi\right\| = 0 \text{ for each } \phi \in V.$$

We remark that condition (F) is satisfied by many finite elements and spectral schemes (see [4, 11, 14]). From condition (F) and the fact that $|P^N| = 1$, it follows that $\lim_{N\to\infty} |P^N\phi - \phi| = 0$ for each $\phi \in H$. Thus, the projection operator P^N converges strongly to the identity on H as well. For each $q \in Q$, $N = 1,2,\ldots$, and almost every $t \in [0,T^*]$, the Galerkin approximation $A^N(t;q)$ to $A(t;q)$ is defined in the following fashion: We let the operator $A^N(t;q) : H^N \to H^N$ be the restriction of the operator $A(t;q)$ to H^N with the image in V^* of $\phi^N \in H^N$ being $A(t;q)\,\phi^N \in V^*$, considered to be a linear functional on H^N. Identifying H^N with its dual, for $\phi^N \in H^N$ we have $A^N(t;q)\,\phi^N = \psi^N$, where $\psi^N \in H^N$ satisfies

$$\left\langle A(t;q)\,\phi^N, \chi^N\right\rangle = \left\langle \psi^N, \chi^N\right\rangle \quad \text{for all} \quad \chi^N \in H^N.$$

For each $N = 1,2,\ldots$ and $q \in Q$, $\xi^N(q) \in H^N$ is defined as $\xi^N(q) = P^N\xi(q)$. Thus we consider the following sequence of approximating identification problems:

(ID^{N}) Find parameters $\overline{q}^N \in Q$ which minimize the performance index

$$J^N(q) = \Phi\left(u^N(\cdot;q);z\right)$$

using the given observations $z \in Z$, where $u^N(\cdot,q)$ is the solution of the initial value problem

$$\begin{cases} \dot{u}^N + A^N(t;q)\,u^N(t) = P^N F(t,u^N(t);q) \\ u^N(0) = \xi^N(q) \end{cases} \tag{2.2}$$

corresponding to $q \in Q$.

The problem (2.2) in H^N is the standard Galerkin approximation to problem (2.1). Note that existence-uniqueness of the solution to equation (2.2) can be proved by simple modification of the arguments used to prove the existence-uniqueness of solutions of equation (2.1). We now state the following convergence theorem which has been proved in [3].

Theorem 2.2. If conditions (A) – (F) are satisfied, then

(i) $\lim_{N\to\infty} u^N(\cdot; q^N) = u(\cdot; q)$ in $C(0,T^*;H)$ and $L^2(0,T^*;V)$ whenever $\{q^N\}$ is a sequence in Q with $\lim_{N\to\infty} q^N = q$.

(ii) For each fixed $N = 1,2,\ldots$, $\lim_{K\to\infty} u^N(\cdot; q^K) = u^N(\cdot; q)$ in $C(0,T^*;H)$ and $L^2(0,T^*;V)$ whenever $\{q^K\}$ is a sequence in Q with $\lim_{K\to\infty} q^K = q$.

It follows from Theorem 2.2 above that the inverse problem ($\mathrm{ID^N}$) has a solution. Indeed, since $u^N(\cdot\,;\,q)$ depends continuously on q by (ii) of Theorem 2.2, we know that the approximate cost functional J^N is continuous on Q. Hence, the existence of such a solution $\overline{q}^N \in Q$ follows immediately, since Q is a compact metric space. The infinite dimensional identification problem (ID) also has a solution which is the limit of a subsequence of $\overline{q}^N$. In fact, from the compactness of Q we have that for any sequence $\{\overline{q}^N\} \subset Q$, there exists a convergent subsequence $\{\overline{q}^{N_j}\}$ of $\{\overline{q}^N\}$. Denoting the limit of this subsequence by $\overline{q}$, we obtain

$$\begin{aligned}
J(\overline{q}) &= \Phi(u(\cdot\,;\overline{q});z) = \Phi\left(\lim_{j\to\infty} u^{N_j}(\cdot\,;\overline{q}^{N_j});z\right)\\
&= \lim_{j\to\infty} \Phi\left(u^{N_j}(\cdot\,;\overline{q}^{N_j});z\right) = \lim_{j\to\infty} J^{N_j}(\overline{q}^{N_j})\\
&\le \lim_{j\to\infty} J^{N_j}(q) = \lim_{j\to\infty} \Phi\left(u^{N_j}(\cdot\,;q);z\right)\\
&= \Phi\left(\lim_{j\to\infty} u^{N_j}(\cdot\,;q);z\right) = \Phi(u(\cdot\,;q);z)\\
&= J(q)
\end{aligned}$$

for each $q \in Q$. This shows that $\overline{q}$ is indeed a solution of the problem (ID).

When the parameter set Q is also infinite dimensional, i.e., the parameters to be estimated are elements in a function space, the parameter set Q has to be discretized as well. In this case, for each $M = 1,2,\ldots$ let $I_M : Q \subset D \to D$ be a continuous mapping with its range $Q_M = I_M(Q)$ in a finite dimensional space, such that $\lim_{M\to\infty} I_M(q) = q$, uniformly on Q. In this situation, a doubly indexed sequence of parameter identification problems, ($\mathrm{ID_M^N}$), where for each N and M, ($\mathrm{ID_M^N}$) is the problem ($\mathrm{ID^N}$) with Q replaced by Q_M, is considered. It can be argued that each of these problems has a solution $\overline{q}_M^N \in Q_M$ and the sequence $\{\overline{q}_M^N\}$ has a D-convergent subsequence $\left\{q_{M_i}^{N_j}\right\}$ whose limit is in Q. Thus $\lim_{i,j\to\infty} q_{M_i}^{N_j} = \overline{q} \in Q$, and $\overline{q}$ can be shown to be a solution of problem (ID). Note that the approximating identification problems ($\mathrm{ID_M^N}$) involve only the minimization of functionals over compact subsets of finite dimensional Euclidean spaces.

3. NUMERICAL RESULTS

We apply our theory to the following partial differential equation which arises in studying nonlinear reaction diffusion equations and population models (see [16]):

$$\begin{cases} u_t - (a(t,x,u_x)\, u_x)_x = c(t)\, F(u(t,x)) \\ u_x(t,0) = 0 = u_x(t,1) \\ u(0,x) = u^0(x), \end{cases} \tag{3.1}$$

with $x \in \Omega = [0,1]$, $t \in [0,T^*]$ and $a \in C_B([0,T^*] \times \Omega \times I\!R)$, the space of bounded continuous functions defined on the given domain and possessing the usual supremum metric denoted by d_∞. It is assumed here that $u^0 \in L^2(0,1)$ and $F(u)$ is locally Lipschitz continuous. The logistic form $F(u) = u(\hat{k} - u)$ or the monod form $F(u) = \dfrac{u}{\hat{k} + u}$ are very commonly used in the literature. Letting $D = C_B([0,T^*] \times \Omega \times I\!R) \times C[0,T^*]$ and $q = (a,c) \in Q$, we choose Q to be a compact subset of D with $q \in Q$ if the following conditions hold.

1. The mapping $\theta \to a(t,x,\theta)$ is C^1 for almost every $(t,x) \in [0,T^*] \times \Omega$.
2. There exists a constant $\delta > 0$, independent of $q \in Q$, for which
$$(a(t,x,\theta)\,\theta - a(t,x,\eta)\,\eta) \cdot (\theta - \eta) \geq \delta\, |\theta - \eta|^2$$
for almost every $(t,x) \in [0,T^*] \times \Omega$ and every $\theta, \eta \in I\!R$.
3. $c(t)$ is Lipschitz continuous.

In this example, we take $H = L^2(0,1)$, $V = H^1(0,1)$ and the observation space $Z = C\left(0,T^*; L^2(0,1)\right)$. For each $q \in Q$ and almost every $t \in [0,T^*]$ we define the operator $A(t;q) : V \to V^*$ by

$$\langle A(t;q)\,\phi, \psi\rangle = \int_\Omega a(t,x,\phi_x(x))\,\phi_x(x)\,\psi_x(x)\,\mathrm{d}x \qquad \text{for all } \phi, \psi \in \mathrm{V}.$$

In [8] conditions (A)−(D) have been verified for the operator A and condition (E) easily follows. Hence, all the results in Section 2 of this paper apply to equation (3.1). Now, let $\{t_i\}_{i=1}^{\kappa}$ with $0 \leq t_1 \leq t_2 \leq \ldots \leq t_\kappa \leq T^*$ be given, and for each $z \in Z$ define the least squares cost functional by

$$\Phi(u;z) = \sum_{i=1}^{\kappa} \int_0^1 |u(t_i,x) - z(t_i,x)|^2\, \mathrm{d}x.$$

For $N = 1,2,\ldots$, let $H^N = \operatorname{span}\left\{\phi_j^N\right\}_{j=0}^{N}$ where ϕ_j^N is the jth linear B-spline on $[0,1]$ defined on the uniform mesh $\left\{0, \frac{1}{N}, \frac{2}{N}, \ldots, 1\right\}$. In other words,

$$\phi_j^N(x) = \begin{cases} 0 & 0 \leq x \leq \frac{j-1}{N} \\ Nx - j + 1 & \frac{j-1}{N} \leq x \leq \frac{j}{N} \\ j + 1 - Nx & \frac{j}{N} \leq x \leq \frac{j+1}{N} \\ 0 & \frac{j+1}{N} \leq x \leq 1 \end{cases}$$

where $j = 0, 1, 2, \ldots, N$. Clearly, $H^N \subset V = H^1(0,1)$ for $N = 1, 2, \ldots$. Let $P^N : H \to H^N$ denote the orthogonal projection of $L^2(0,1)$ onto H^N with respect to the usual L^2 inner product. Using standard approximation results for interpolatory splines as in [18] we can argue that condition (F) is satisfied. In our computational efforts we consider the following two parameter identification problems.

3.1. 1–D estimation problem

In this problem, we estimate the function a as a one-dimensional function of θ, independent of t and x. We also estimate c as a one-dimensional function of t. The following parameter values are used by us to generate the observed data $z \in Z$.

$$c(t) = 2 + \sin(100t) \qquad u^0(x) = x(1-x)$$

$$F(u) = u(1-u) \qquad a(t, x, \theta) = 1 - 0.5\exp(-0.1\theta^2)$$

For our least-squares method we generate data from the solution of our forward problem using the above fixed parameters and collect the data $z(t_i, \cdot)$ at points t_i, $i = 0, \ldots, 250$, where $t_i = 0.0002 \cdot i$.

Define $D = C_B(\mathbb{R}) \times C[0, T^*]$. For given fixed values of $\alpha_0, \rho_0, \sigma_0, K_0$ and $\theta_0 > 0$, we choose $Q = A \times \widehat{C}$, where A is the D-closure of

$$\{a \in C_B(\mathbb{R}) \ : \ |a(\theta)| \le \rho_0, \ \left|\frac{\mathrm{d}a}{\mathrm{d}\theta}(\theta)\right| \le \sigma_0,$$

$$\frac{\mathrm{d}a}{\mathrm{d}\theta}(\theta)\,\theta + a(\theta) \ge \alpha_0 \text{ for } \theta \in \mathbf{R} \text{ and } \mathrm{a}(\theta) \text{ is constant}$$

$$\text{for } \theta \le \theta_{\mathrm{a}} \text{ and } \theta \ge \widehat{\theta}_a, \text{where } \theta_{\mathrm{a}}, \widehat{\theta}_{\mathrm{a}} \text{ satisfy } -\theta_0 \le \theta_a \le \widehat{\theta}_a \le \theta_0\}$$

and

$$\widehat{C} = \{c \in C[0, T^*] : |c(t)| \le K_0, |c(t) - c(r)| \le K_0 \, |t - r| \ \forall\, t, r \in [0, T^*]\}.$$

Hence, a straightforward application of the Arzelà–Ascoli theorem shows Q to be a compact subset of D. We approximate the infinite dimensional parameter space as follows: For M_1, a positive integer, and $a \in A$, we set

$$(I_{M_1} a)(\theta) = \sum_{j=0}^{M_1} a\left(\theta_a + j\left(\frac{\widehat{\theta}_a - \theta_a}{M_1}\right)\right) \psi^j_{M_1}\left(\theta\, ; \theta_a, \widehat{\theta}_a\right)$$

where $\theta \in \mathbb{R}$. In this formula, $\psi^j_{M_1}\left(\theta\, ; \theta_a, \widehat{\theta}_a\right)$, $j = 0, \ldots, M_1$, are the linear B-splines defined on the uniform partition $\left\{\theta_a, \theta_a + \left(\frac{\widehat{\theta}_a - \theta_a}{M_1}\right), \ldots, \widehat{\theta}_a\right\}$ of the interval $\left[\theta_a, \widehat{\theta}_a\right]$. We then extend our approximation on the interval $[\,\theta_a, \widehat{\theta}_a]$ to a continuous function on the entire real line via $\psi^j_{M_1}\left(\theta\,;\theta_a,\widehat{\theta}_a\right) = \psi^j_{M_1}\left(\theta_a;\theta_a,\widehat{\theta}_a\right)$ for $\theta \le \theta_a$ and $\psi^j_{M_1}\left(\theta\,;\,\theta_a,\widehat{\theta}_a\right) = \psi^j_{M_1}\left(\widehat{\theta}_a\,;\,\theta_a,\widehat{\theta}_a\right)$ for $\theta \ge \widehat{\theta}_a$. Then, using the Peano Kernel Theorem given in [17] we get

$$\lim_{M_1\to\infty} I_{M_1}a = a \quad \text{in } \mathrm{C_B}(\mathrm{R}),$$

uniformly in a, for $a \in A$. Similarly, for $c \in \widehat{C}$ we have

$$(I_{M_2}c)(t) = \sum_{j=0}^{M_2} c\left(\frac{j}{M_2}T^*\right)\lambda^j_{M_2}(t\,;T^*),$$

where, the functions $\lambda^j_{M_2}(t\,;T^*)$, $j = 0,\ldots,M_2$, represent the linear B-splines defined on the uniform mesh $\left\{0,\frac{T^*}{M_2},\frac{2T^*}{M_2},\ldots,T^*\right\}$ of the interval $[0,T^*]$. Clearly $\lim_{M_2\to\infty} I_{M_2}c = c$, in $C[0,T^*]$, uniformly in c, for $c \in \widehat{C}$. Thus defining $M = (M_1,M_2)$, setting $q_M = (a_{M_1},c_{M_2}) \in Q_M = I_M(Q)$ (where $I_M(Q) = I_{M_1}(A) \times I_{M_2}(\hat{C})$) and approximating $u^N(t,x)$ by

$$u^N(t,x) = \sum_{i=0}^{N} w^N_i(t)\,\phi^N_i(x),$$

we have the following form for the finite dimensional initial value problem(2.2):

$$\begin{cases} \Lambda^N\dot{w}^N(t) + G^N\left(w^N(t)\,;\,a_{M_1}\right) = \Upsilon^N\left(t,w^N(t)\right) \\ \Lambda^N w^N(0) = \left(w^N\right)^0, \end{cases} \tag{3.2}$$

where $t \in [0,T^*]$ and $w^N(t) = \left(w^N_0,w^N_1,\ldots,w^N_N\right) \in I\!R^{N+1}$. The matrix Λ^N is an $(N+1)\times(N+1)$ Gram matrix whose (i,j)th entry is given as $\Lambda^N_{i,j} = \langle\phi^N_i,\phi^N_j\rangle$ and $\left(w^N\right)^0$ is an $(N+1)$ dimensional vector whose ith element is given by $\left(w^N\right)^0_i = \langle\xi,\phi^N_i\rangle$. Further, using the definition of $\phi^N_j(x)$, $j = 0,1,\ldots,N$, the inner product $\langle a,b\rangle = \int_0^1 a\cdot b\,\mathrm{d}x$ and the following integral approximation

$$\int_0^1 f(x)\,\mathrm{d}x = \sum_{i=1}^{N} f\left(\frac{i}{N}\right)\Delta x,$$

where $\Delta x = \frac{1}{N}$, we evaluate $\Upsilon^N\left(t,w^N(t)\right)$, an $(N+1)$-dimensional vector, as follows:

$$\begin{aligned} \Upsilon^N_0\left(t,w^N(t)\right) &= 0 \\ \Upsilon^N_i\left(t,w^N(t)\right) &= \Delta x\, c_{M_2}(t)\,F(w^N_i(t)) \quad \text{for } i = 1,\ldots,N. \end{aligned}$$

Similarly, as described in [5], $G^N(\cdot\,;\,a_{M_1}) : I\!R^{N+1} \to I\!R^{N+1}$ can be expressed as

$$\begin{aligned} G_0^N(\gamma\,;\,a_{M_1}) &= a_{M_1}\left(\frac{\gamma_1-\gamma_0}{\Delta x}\right)\left(\frac{\gamma_0-\gamma_1}{\Delta x}\right) \\ G_N^N(\gamma\,;\,a_{M_1}) &= a_{M_1}\left(\frac{\gamma_N-\gamma_{N-1}}{\Delta x}\right)\left(\frac{\gamma_N-\gamma_{N-1}}{\Delta x}\right) \end{aligned}$$

and

$$\begin{aligned} G_i^N(\gamma\,;\,a_{M_1}) &= a_{M_1}\left(\frac{\gamma_i-\gamma_{i-1}}{\Delta x}\right)\left(\frac{\gamma_i-\gamma_{i-1}}{\Delta x}\right) - a_{M_1}\left(\frac{\gamma_{i+1}-\gamma_i}{\Delta x}\right)\left(\frac{\gamma_{i+1}-\gamma_i}{\Delta x}\right) \\ &\quad \text{for } i = 1,\ldots,N-1, \end{aligned}$$

where $\gamma \in I\!R^{N+1}$. Hence, if $q_M = (a_{M_1}, c_{M_2}) \in Q_M$ is given by

$$a_{M_1}(\theta) = \sum_{j=0}^{M_1} \nu_{M_1}^j \psi_{M_1}^j\left(\theta\,;\theta_{a_{M_1}}, \widehat{\theta}_{a_{M_1}}\right)$$

and

$$c_{M_2}(t) = \sum_{j=0}^{M_2} F_{M_2}^j \lambda_{M_2}^j(t\,;T^*),$$

then solving (ID_M^N) involves the identification of the (M_1+3) coefficients $\left\{\nu_{M_1}^j\right\}_{j=0}^{M_1}$, $\theta_{a_{M_1}}$ and $\widehat{\theta}_{a_{M_1}}$ and the (M_2+1) coefficients $\left\{F_{M_2}^j\right\}_{j=0}^{M_2}$ from a compact subset of $I\!R^{M_1+M_2+4}$ so as to minimize the functional $J^N(q_M) = \Phi\left(u^N(\cdot\,;q_M)\,;z\right)$. Figures 1 and 2 show the estimates of a and c with $N = 9$ and $M = (M_1, M_2) = (5,5)$, respectively. Figures 3 and 4 do the same for $N = 9$ and $M = (7,7)$ while in Figures 5 and 6, $N = 9$ and $M = (9,9)$. Clearly the numerical results obtained in these figures corroborate the convergence results of the previous section. To test our scheme against measurement error, we added noise with mean $\mu = 0$ and standard deviation $\sigma = 0.03$ to our observed data $z \in Z$. We present the corresponding estimates of a and c in Figures 7 and 8 using $N = 9$ and $M = (5,5)$. For convenience, we give in Table 1 the values of $||a_{M_1} - a||_\infty$ and $||c_{M_2} - c||_\infty$ for the numerical results presented in Figures 1–8. To conclude this numerical experiment we tested our inverse method when less data points are collected. That is, we now generate the data $z(t_i,\cdot)$ at points t_i, $i = 0,\ldots,10$, where $t_i = 0.005\cdot i$. Our results, presented in Figures 9 and 10, indicate that our estimates of a and c remain essentially unchanged.

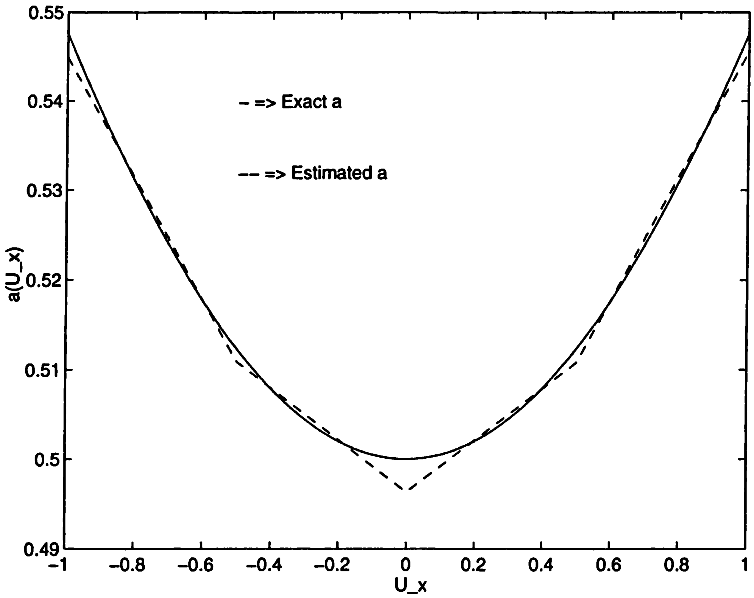

Fig. 1. This figure represents the true versus the estimated function $a(\theta)$ with $N = 9$, $M_1 = 5$ and no noise added to the computationally generated data.

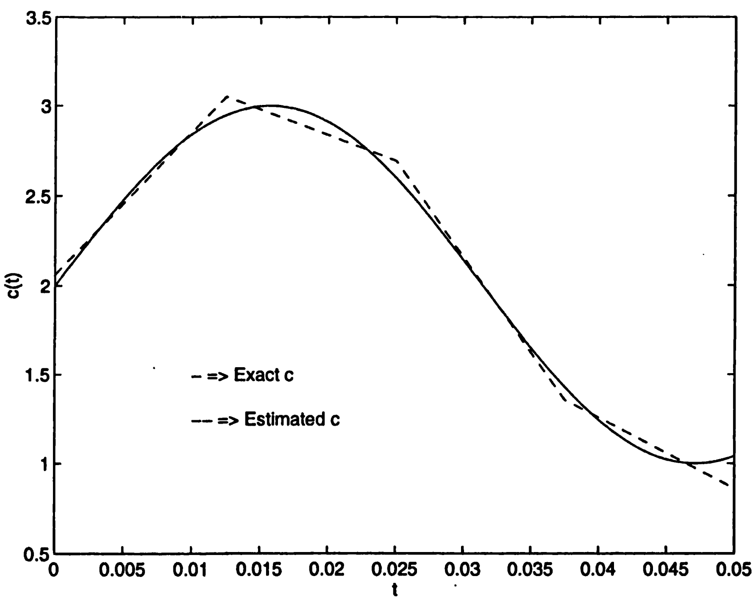

Fig. 2. This figure represents the true versus the estimated function $c(t)$ with $N = 9$, $M_2 = 5$ and no noise added to the computationally generated data.

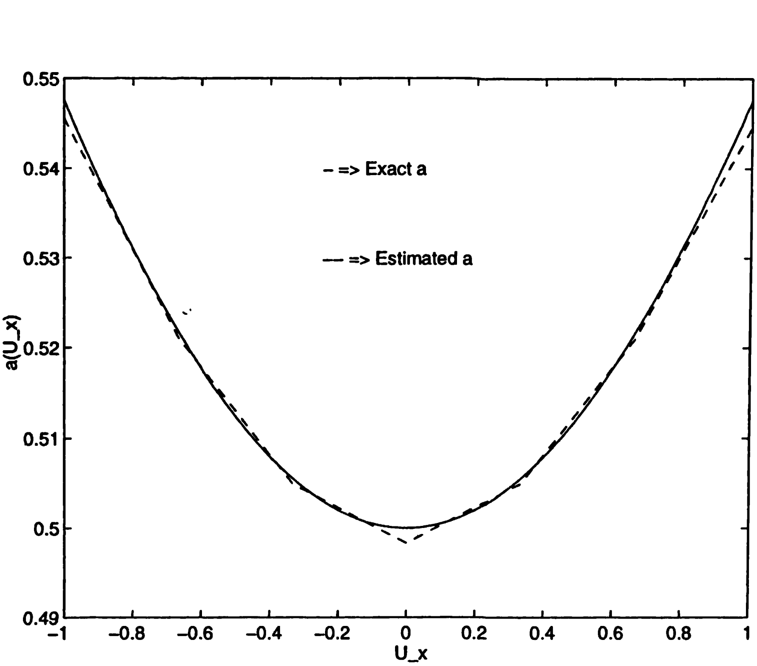

Fig. 3. This figure is the same as Figure 1, with $N = 9$, $M_1 = 7$.

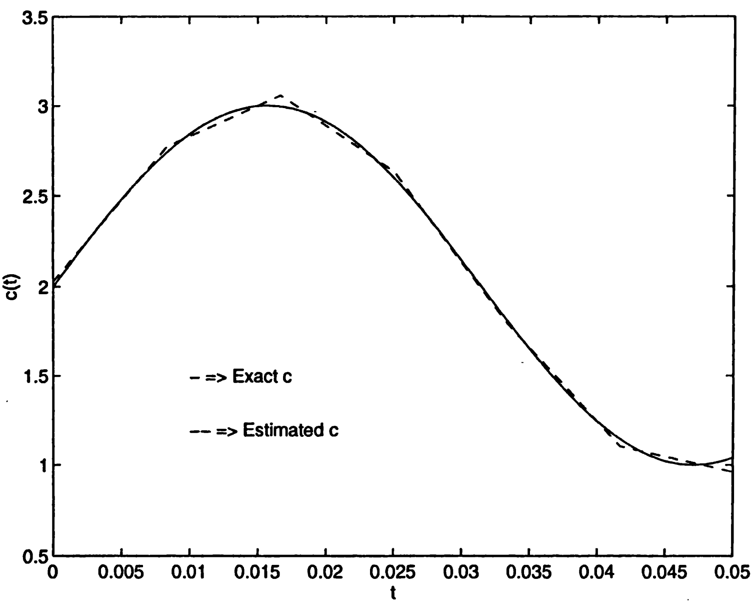

Fig. 4. This figure is the same as Figure 2, with $N = 9$, $M_2 = 7$.

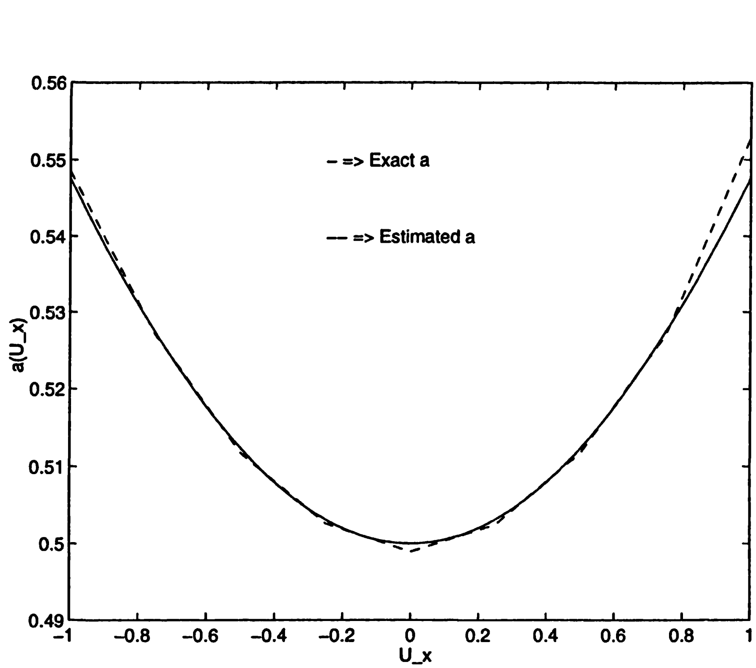

Fig. 5. This figure is the same as Figure 1, with $N = 9$, $M_1 = 9$.

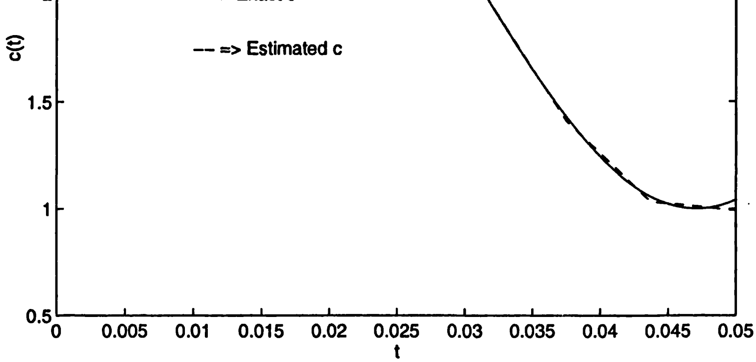

Fig. 6. This figure is the same as Figure 2, with $N = 9$, $M_2 = 9$.

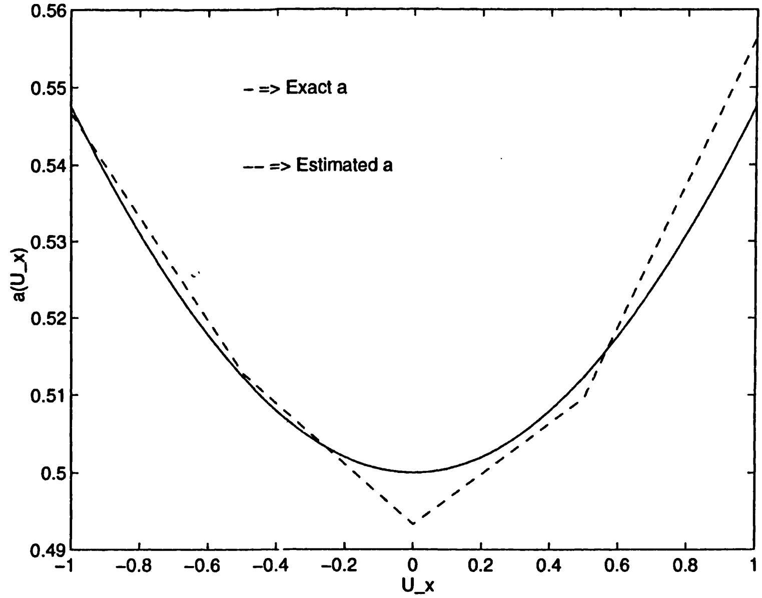

Fig. 7. This figure represents the true versus the estimated function $a(\theta)$ with $N = 9$, $M_1 = 5$ and noise with mean $\mu = 0$ and standard deviation $\sigma = 0.03$ added to the computationally generated data.

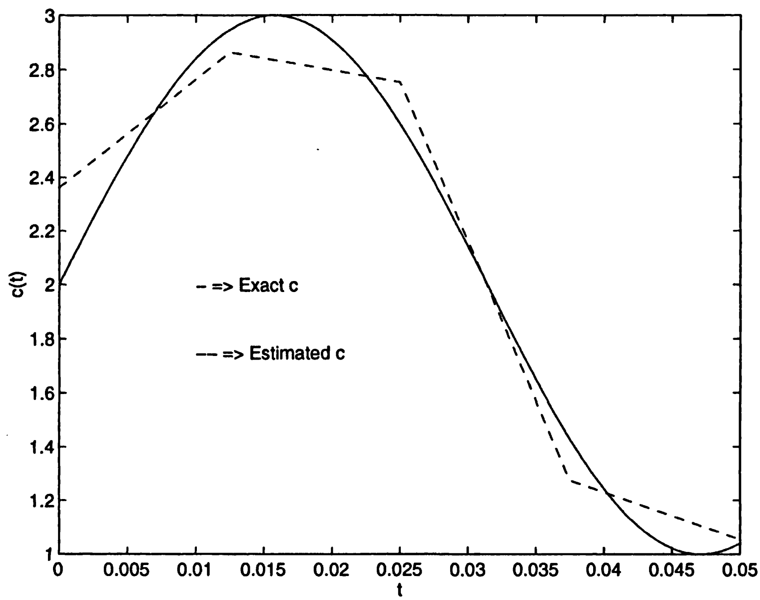

Fig. 8. This figure represents the true versus the estimated function $c(t)$ with $N = 9$, $M_2 = 5$ and noise with mean $\mu = 0$ and standard deviation $\sigma = 0.03$ added to the computationally generated data.

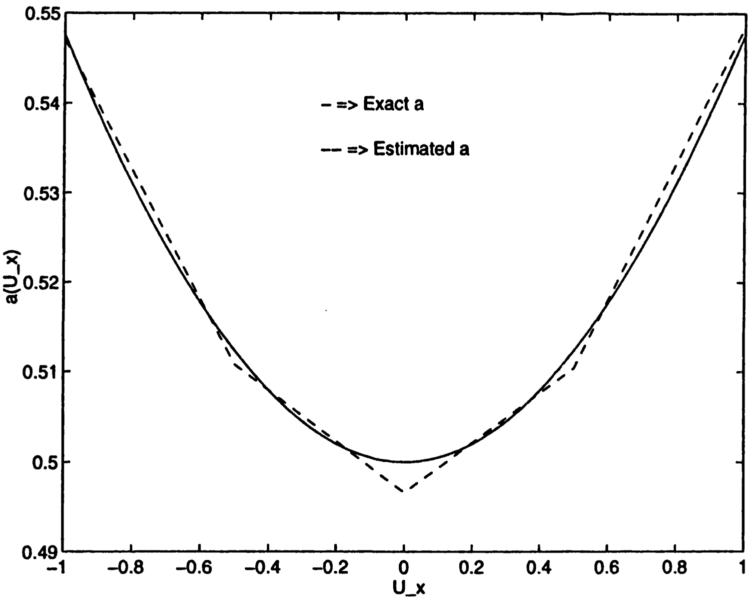

Fig. 9. This figure is the same as Figure 1, with $N = 9$, $M_1 = 5$ and less data points provided.

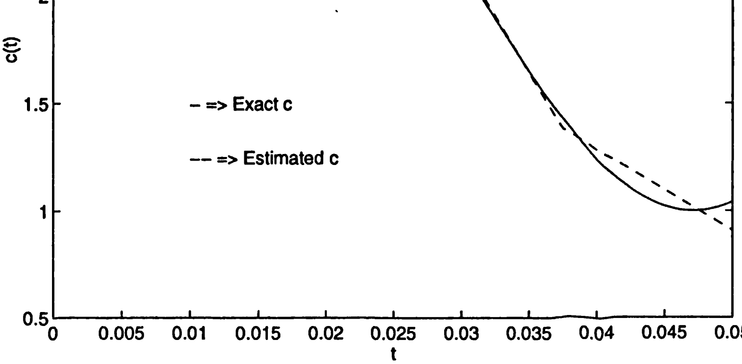

Fig. 10. This figure is the same as Figure 2, with $N = 9$, $M_2 = 5$ and less data points provided.

Table 1. This table gives the values of $\|a_{M_1}-a\|_\infty$ and $\|c_{M_2}-c\|_\infty$ for $M_i=5,7$ and 9, $i=1,2$, with no noise added to the computationally generated data and the same values for $M_i=5$, $i=1,2$, when noise with mean $\mu=0$ and standard deviation $\sigma=0.03$ is added to the computationally generated data.

	$M_i,\ i=1,2$	$\|a_{M_1}-a\|_\infty$	$\|c_{M_2}-c\|_\infty$
Data without noise	5	0.0037	0.1818
Data without noise	7	0.0030	0.0802
Data without noise	9	0.0025	0.0530
Noisy data with $\sigma=0.03$	5	0.0088	0.3618

3.2. 2–D estimation problem

This problem involves the estimation of a as a two-dimensional function of θ and t, while c is regarded as a known function in our computations. We use the same parameter values as in the 1-D problem above with the following exception:

$$a(t,\theta)=1-0.5\exp\left[-0.1(t+0.5)\,\theta^2\right].$$

For our least-squares method we generate data from the solution of our forward problem using the above fixed parameters and collect the data $z(t_i,\cdot)$ at points t_i, $i=0,\dots,50$ where $t_i=0.0002\cdot i$.

For given fixed values of α_0,ρ_0,σ_0 and $\theta_0>0$, we let Q be the D-closure of the following set:

$$\{a\in C_B([0,T^*]\times I\!R)\ :\ |a(t,\theta)|\le\rho_0,\ |a_\theta(t,\theta)|,|a_t(t,\theta)|\le\sigma_0,$$

$$a_\theta(t,\theta)\,\theta+a(t,\theta)\ge\alpha_0\ \text{ for }\ t\in[0,T^*],\ \theta\in I\!R\ \text{ and }\ a(t,\theta)=a(t)$$

$$\text{for }\theta\le\theta_a(t)\,\text{and}\ \theta\ge\widehat{\theta}_a(t)\ \text{where}\ \theta_a,\widehat{\theta}_a\ \text{ satisfies } -\theta_0\le\theta_a(t)\le\widehat{\theta}_a(t)\le\theta_0\}.$$

Again, a straightforward application of the Arzelà–Ascoli theorem shows that Q is a compact subset of D. Here we approximate our parameter set as follows:

$$\begin{aligned}(I_{M_1,M_2}a)(t,\theta)\ &=\ \sum_{i=0}^{M_2}\sum_{j=0}^{M_1}a\left(\frac{iT^*}{M_2},\theta_a(\frac{iT^*}{M_2})+j\left(\frac{\widehat{\theta}_a(\frac{iT^*}{M_2})-\theta_a(\frac{iT^*}{M_2})}{M_1}\right)\right)\\ &\qquad\psi^j_{M_1}\left(\theta\,;\theta_a(\frac{iT^*}{M_2}),\widehat{\theta}_a(\frac{iT^*}{M_2})\right)\lambda^i_{M_2}(t\,;T^*)\end{aligned}$$

where ψ and λ have similar definitions as in the 1-D experiment and again the function $(I_{M_1,M_2}a)(t,\theta)$ is extended to a continuous function over the entire real line in a similar manner as in the 1-D case above. The Peano Kernel Theorem is used here once again to yield

$$\lim_{M_1,M_2\to\infty} I_{M_1,M_2}a = a \quad \text{in } \mathrm{C_B}\left([0,\mathrm{T}^*]\times\mathrm{R}\right)$$

uniformly in a, for $a \in Q$. Using the compact notation $M = (M_1, M_2)$, with $q_M = a_{M_1,M_2} \in Q_M = I_M(Q)$ and approximating $u^N(t,x)$ as in the 1-D case above, we can derive a finite dimensional initial value problem similar to the one given in (3.2).

Once again, if $a_{M_1,M_2} \in Q_M$ is given by

$$a_{M_1,M_2}(t,\theta) = \sum_{i=0}^{M_2}\sum_{j=0}^{M_1} \Theta^{i,j}_{M_1,M_2}\psi^j_{M_1}\left(\theta\,;\theta^i_{M_2},\widehat{\theta}^i_{M_2}\right)\lambda^i_{M_2}(t\,;T^*),$$

then the solution of the identification problem (ID^N_M) involves identifying the (M_1+3) (M_2+1) coefficients $\left\{\Theta^{i,j}_{M_1,M_2},\theta^i_{M_2},\widehat{\theta}^i_{M_2}\right\}^{M_2,M_1}_{i,j=0}$ from a compact subset of $\mathbb{R}^{M_1M_2+M_1+3M_2+3}$ so as to minimize the least squares cost functional $J^N(q_M) = \Phi\left(u^N(\cdot;q_M);z\right)$. Figures 11 and 12 show the exact and estimated two-dimensional plots of a, respectively, while in Figure 13 we present the difference between the exact and estimated plots of a. We use $N = 9$ and $M = (5,5)$ in the calculations presented in Figures 11, 12 and 13.

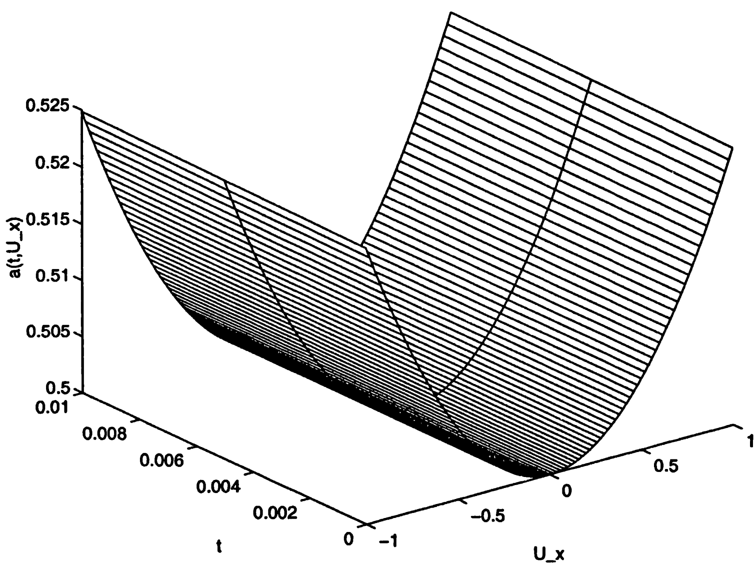

Fig. 11. This figure represents the exact function $a(t,\theta)$.

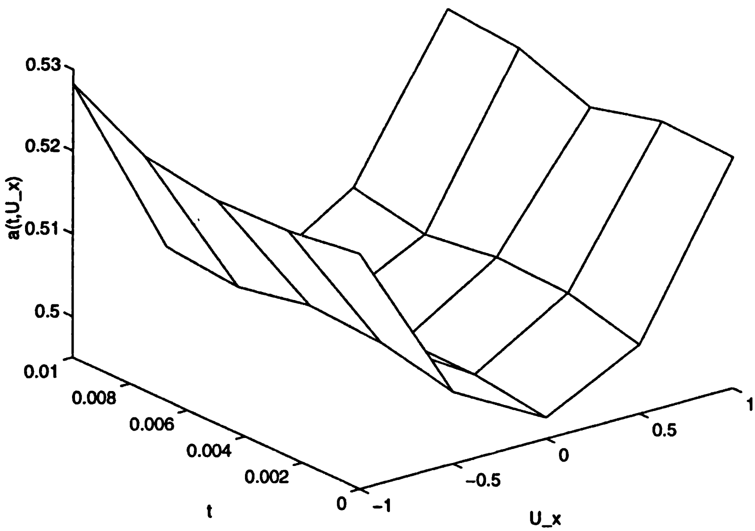

Fig. 12. This figure represents the estimated function $a(t,\theta)$ with $N=9$, $M=(5,5)$ and no noise added to the computationally generated data.

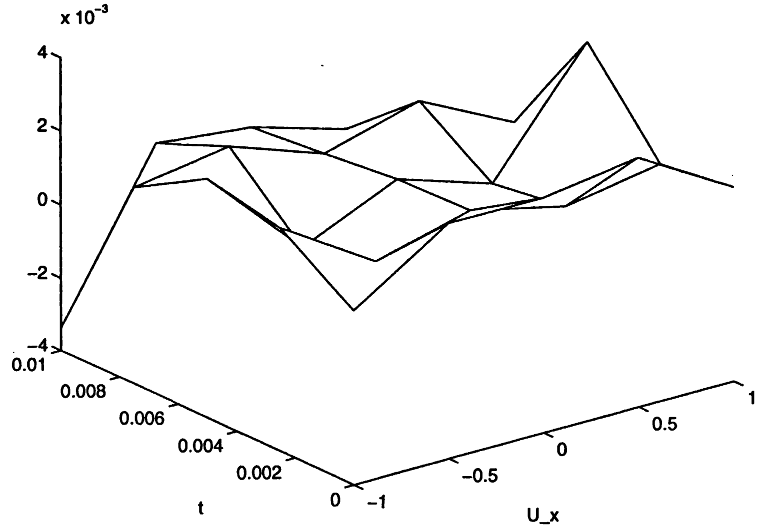

Fig. 13. This figure represents the difference between the exact and estimated function $a(t,\theta)$.

3.3. Comments on our computations

In our computations involving the estimation of parameters we solved the initial value problem (3.2) for a given choice q_M^N of parameters, by using a modification of the Runge–Kutta method. All the integrals were computed using the Riemann sum representation of integrals on the interval $[0,1]$. We used the subroutine *LIMDIF1* from *NETLIB*, an application of the Levenberg–Marquardt algorithm, to solve the finite dimensional nonlinear least squares minimization problem $(\mathrm{ID}_{\mathrm{M}}^{\mathrm{N}})$. We set the positive constant defined in the admissible parameter space $\theta_0 = 1$, and chose the initial guesses for a and c to be 0.5 and 2, respectively. For simplicity and to shorten the computational time required for each experiment, rather than identifying the numbers θ_a and $\widehat{\theta}_a$ we set $\theta_a = -\theta_0$ and $\widehat{\theta}_a = \theta_0$, that is, we assume that every $a \in Q$ is constant outside the interval $[-\theta_0, \theta_0] = [-1,1]$. Hence, in our computations we only identified the coefficients $\left\{\nu_{M_1}^j\right\}_{j=0}^{M_1}$ of $\psi_{M_1}^j(\theta\,; -1, 1)$, $j = 0, \ldots, M_1$, the linear B-splines defined on the uniform partition $\left\{-1, -1 + \left(\frac{2}{M_1}\right), \ldots, 1\right\}$ of the interval $[-1,1]$. Similarly in the 2–D experiments we set $\theta_{M_2}^i = -\theta_0$ and $\widehat{\theta}_{M_2}^i = \theta_0$ and identified only $\left\{\Theta_{M_1,M_2}^{i,j}\right\}_{i,j=0}^{M_2,M_1}$. All the computations for estimating a and c were executed on a *Pentium 166* machine at the University of Southwestern Louisiana Computational Research Laboratory. The two-dimensional estimation of a as a function of t and θ was also carried out on an *Ultra-Sparc 2000* machine at the same location. It was observed that the *Ultra-Sparc 2000* took about 36 % of the time that the *Pentium* took to perform the same calculations. For the one-dimensional case, the CPU times taken by the *Pentium* ranged between 22–47 hours in all of our numerical experiments. The final least squares value at the end of our programs ranged from 10^{-10} for data without noise to 10^{-3} for data with noise. We point out that for the case $M = (5,5)$ without noise, the Levenberg–Marquardt algorithm required the solution of equation (3.1) 725 times for convergence to a minimizer and 921 times for the case $M = (9,9)$ without noise.

4. CONCLUDING REMARKS

In this paper we have presented an abstract approximation framework for parameter estimation in a class of nonautonomous nonlinear evolution equations, and more importantly, have for the first time provided numerical evidence supporting our theory. The numerical results presented in this paper indicate that the general least-squares approach is indeed very promising. To date, all our experiments have dealt with computationally generated data with one dimensional reaction diffusion equations. Our near future efforts will focus on testing these techniques using experimental data as well as efficient implementation in two and three dimensional transport problems arising from modeling bioremediation of contaminated groundwater (see [10, 12, 13]). Furthermore, we hope to extend the current nonlinear theory to allow the identification of discontinuous parameters as was established in [2] for the linear case.

APPENDIX

To test the accuracy of our algorithm which numerically computes the solution of (3.1) we considered the following partial differential equation.

$$\begin{cases} u_t - (a(t,x,u_x)\, u_x)_x = c(t)\; F(u(t,x)) + y(t,x) \\ u_x(t,0) = 0 = u_x(t,1) \\ u(0,x) = u^0(x). \end{cases} \tag{A}$$

Choosing $u(t,x) = \exp(-t)\left(x^2 - \frac{2x^3}{3}\right)$ to be the solution of equation (A) with $t \in [0, 0.05]$, $x \in [0,1]$, $u^0(x) = \left(x^2 - \frac{2x^3}{3}\right)$, $c(t) = 1$, $F(u(t,x)) = u(1-u)$ and $a(t,x,\theta) = 1 - 0.5\exp(-0.1\theta^2)$, one can easily find the forcing term $y(t,x)$ from the differential equation as follows:

$$y(t,x) = u_t - (a(t,x,u_x)\, u_x)_x - u(1-u).$$

Thus, solving (A) numerically, using the Galerkin approximation scheme discussed in this paper, we compared our solution $u^N(t,x)$ to the exact solution u, with $N = 9$, $N = 18$ and $N = 36$ and presented the L^2 norm of the errors in Figure 14. The results indicate a linear convergence rate as expected.

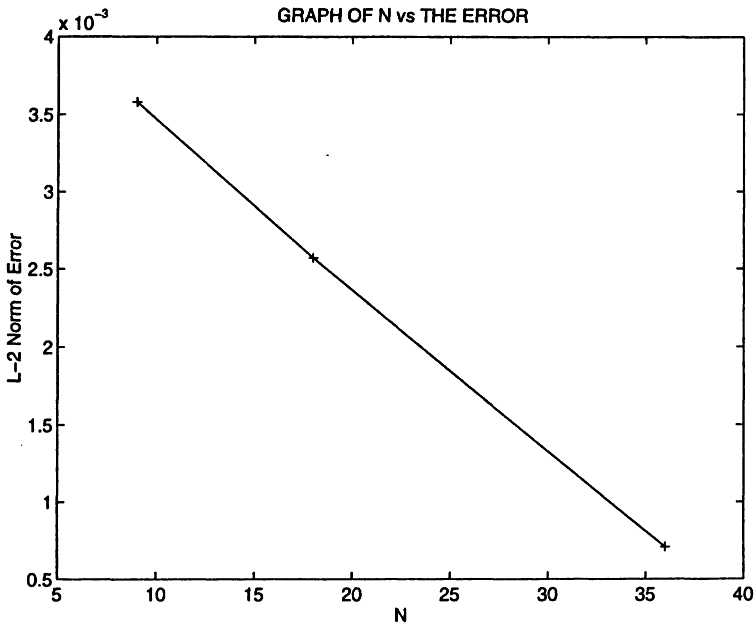

Fig. 14. This figure shows the linear order of convergence of our Galerkin approximation scheme used to solve the forward problem.

ACKNOWLEDGEMENTS

The work of the first two authors was partially supported by the Louisiana Education Quality Support funding under grant LEQSF (1996-1999)-RD-A-36. The work of the third author was partially supported by the Israel Science Foundation founded by the Israel Academy of Sciences and Humanities and by the Fund for the Promotion of Research at the Technion. All the authors thank the two anonymous referees for their helpful comments and suggestions.

(Received April 8, 1998.)

REFERENCES

[1] A. S. Ackleh and B. G. Fitzpatrick: Estimation of time dependent parameters in general parabolic evolution systems. J. Math. Anal. Appl. *203* (1996), 464–480.

[2] A. S. Ackleh and B. G. Fitzpatrick: Estimation of discontinuous parameters in general parabolic systems. Kybernetika *32* (1996), 543–556.

[3] A. S. Ackleh and S. Reich: Parameter Estimation in Nonlinear Evolution Equations. Numer. Funct. Anal. Optim. *19* (1998), 933–947.

[4] H. T. Banks and K. Kunisch: Estimation Techniques for Distributed Parameter Systems. Birkhäuser, Boston – Basel 1989.

[5] H. T. Banks, C. K. Lo, S. Reich and I. G. Rosen: Numerical studies of identification in nonlinear distributed parameter systems. Internat. Ser. Numer. Math. *91* (1989), 1–20.

[6] H. T. Banks, S. Reich and I. G. Rosen: An approximation theory for the identification of nonlinear distributed parameter systems. SIAM J. Control Optim. *28* (1990), 552–569.

[7] H. T. Banks, S. Reich and I. G. Rosen: Estimation of nonlinear damping in second order distributed parameter systems. Control Theory and Advanced Technology *6* (1990), 395–415.

[8] H. T. Banks, S. Reich and I. G. Rosen: Galerkin Approximation for inverse problems for nonautonomous nonlinear distributed systems. Appl. Math. Optim. *24* (1991), 233–256.

[9] V. Barbu: Nonlinear Semigroups and Differential Equations in Banach Spaces. Noordhoff, Leyden 1976.

[10] J. Bear: Dynamics of Fluids in Porous Media. Elsevier, New York 1972.

[11] C. Canuto and A. Quateroni: Approximation results for orthogonal polynomials in Sobolev spaces. Math. Comp. *38* (1982), 67–86.

[12] B. G. Fitzpatrick: Analysis and approximation for inverse problems in contaminant transport and biodegradation models. Numer. Funct. Anal. Optim. *16* (1995), 847–866.

[13] R. A. Freeze and J. Cherry: Groundwater. Prentice-Hall, Englewood Cliffs, N. J. 1979.

[14] C. Johnson: Numerical Solution of Partial Differential Equations by The Finite Element Method. Cambridge Univ. Press, Cambridge 1987.

[15] R. Kluge and H. Langmach: On some Problems of Determination of Functional Parameters in Partial Differential Equations, In: Distributed Parameter Systems: Modeling and Identification (Lecture Notes in Control and Information Sciences 1). Springer-Verlag, Berlin 1978, pp. 298–309.

[16] C. V. Pao: Nonlinear Parabolic and Elliptic Equations. Plenum Press, New York 1992.

[17] M. H. Schultz: Spline Analysis. Prentice–Hall, Englewood Cliffs, N. J. 1973.

[18] B. K. Swartz and R. S. Varga: Error bounds for spline and L-spline interpolation. J. Approx. Theory *6* (1972), 6–49.

Dr. Azmy S. Ackleh, Department of Mathematics, University of Southwestern Louisiana, Lafayette, LA 70504-1010. U. S. A.
e-mail: ackleh@usl.edu

Mr. Robert R. Ferdinand, Department of Mathematics, University of Southwestern Louisiana, Lafayette, LA 70504-1010. U. S. A.
e-mail: robf@usl.edu

Dr. Simeon Reich, Department of Mathematics, The Technion – Israel Institute of Technology, 32000, Haifa. Israel.
e-mail: sreich@techunix.technion.ac.il