PARAMETER IDENTIFICATION IN LINEAR AND NONLINEAR PARABOLIC PARTIAL DIFFERENTIAL EQUATIONS

by

Lan Zhang

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Mathematics

Approved:

J. A. Burns, Chairman

M. D. Gunzburger

T. L. Herdman

R. A. McCoy

J. S. Peterson

August, 1995
Blacksburg, Virginia

PARAMETER IDENTIFICATION IN LINEAR AND NONLINEAR PARABOLIC PARTIAL DIFFERENTIAL EQUATIONS

by

Lan Zhang

Committee Chairman: John A. Burns

Abstract

The research presented in this dissertation is carried out in two parts; the first, which is the main work of this dissertation, involves development of continuous differentiability of the solution with respect to the unknown parameters. For linear parabolic partial differential equations, only mild conditions are assumed on the admissible parameter space. The nonlinear partial differential equation we consider is a generalized Burgers' equation, for which we establish the well-posedness and the smoothness properties of the solution with respect to the parameters.

In the second part, we consider parameter identification problems for these two parameter dependent systems. The identification scheme which we use here is the quasilinearization method. Based on the results in the first part of this work, we obtain existence and local convergence of the algorithm. We also present some numerical examples which demonstrate the performance of the quasilinearization scheme.

ACKNOWLEDGEMENT

First, I would like to express my sincere appreciation to my advisor, Professor John A. Burns, for his help, support and guidance, to whom I am deeply in debt for leading me into new areas and ideas.

I also wish to thank Professors Max D. Gunzburger, Terry L. Herdman, Robert A. McCoy and Janet S. Peterson for serving on my committee and their encouragement throughout the years I am in Blacksburg. Thanks are also due to Professor Peter A. Linnell for his help and encouragement in the beginning of my study at Virginia Tech; also, to Professor Vector I. Shubov in Texas Tech University for valuable discussions.

Finally, and most importantly, special thanks to my husband, Zhijian Qiu, for his love and support. Without his influence, I certainly could not have gone this far.

Contents

Chapter 1

Introduction

During recent years considerable effort has been devoted to the problem of estimating unknown parameters in distributed parameter systems. Many parameter estimation problems are best formulated in an infinite dimensional state space where one must determine the parameter from some admissible parameter set that minimizes an appropriate cost function.

There are two basic classes of approach for optimization based parameter estimation. The first, an indirect approach proceeds by initially approximating the dynamic equations and then using optimization algorithms on the finite dimensional problem. This type of approach, which is typified by the papers [8] - [11], is usually easy to implement and successful. A disadvantage of this approach is that only subsequential convergence of the sequence of generated parameter estimates has been established.

The second more direct approach is based on direct application of an optimization algorithm and employs numerical approximations at each step of the algorithm to compute the necessary solutions of the dynamic equations. This approach is used in [7], [15], [26], [27] and [31]. Direct methods are often limited by the fact that the dependence on unknown parameters of the solution to the infinite dimensional dynamical equations may not be smooth enough to establish convergence of the algorithm. Indeed, some algorithms may not be properly defined without this necessary smoothness. When the direct methods can be applied, however, it is sometimes possible to establish not only sequential convergence but also rates of convergence.

The work presented in this thesis is motivated by the use of a direct method such as quasilinearization to solve parameter estimation problems involving different types of partial differential equations. The crucial part in the proof of the convergence of the algorithm is to establish the smoothness properties with respect to the parameters. In our work, we discuss two kinds of parameter dependent systems. The first one is a linear parabolic system with variable coefficients. This study is presented in Chapter 2. The second one is a generalized Burgers' equation with variable coefficients; it is presented in Chapter 3.

There is a considerable body of research on the problem of estimating coefficients in parabolic equations. In 1985, H. T. Banks and P. D. Lamm ([10]) developed an indirect approach for estimating coefficients in parabolic distributed systems. In 1990, P. W. Hammer ([26]) developed a quasilinearization algorithm for a parameter estimation problem involving a parabolic partial differential equation. Both of these papers assumed strong smoothness assumptions on the admissible parameter space. In 1988, H. T. Banks and K. Ito ([8]) presented a general convergence/stability framework for using indirect methods to treat parameter identification problems involving distributed parameter systems. This framework permits one to give convergence and stability arguments in inverse problems under extremely weak compactness assumptions on the admissible parameter spaces. In Chapter 2, we use this idea in a direct approach. Actually, by modifying the framework, we successfully prove that the quasilinearization method is convergent under mild assumptions on the admissible parameter space.

The next problem we consider is the application of quasilinearization to parameter identification in nonlinear partial differential equations. In [26], P. W. Hammer presented some numerical results for Burgers' equation with constant parameters. In [4], M. G. Armentano presented some numerical results for Burgers' equation with spatially varying parameters. The numerical results presented in these papers were successful. However, neither of these papers contain a convergence proof for the nonlinear equations. The nonlinear partial differential equation we consider in this work is a generalized Burgers' equation on a finite interval. In Chapter 3, we establish the well-posedness and the smoothness properties of the solution with respect to the parameters. These properties are necessary to establish the convergence of the quasilinearization algorithm.

In Chapter 4, we formulate the parameter estimation problems as optimal control problems and prove the existence of solutions to the problems in the first section. In the second section, we prove that with the smoothness properties established in Chapter 2 and Chapter 3, the quasilinearization algorithm converges.

Finally, in Chapter 5, we present numerical examples which demonstrate the performance of the quasilinearization scheme.

Chapter 2

Parameter Dependence in Parabolic Partial Differential Equations

In this chapter, we consider the dependence on an unknown parameter of the solution of parabolic partial differential equations. We prove the smoothness properties under weak assumptions on the parameter space.

2.1 The General Setting

We consider a separable real Hilbert space H and another separable Hilbert space V, which is continuously and densely imbedded in H. We identify H with its own dual space; the dual of V is denoted by V^*. Thus we have $V \subset H \subset V^*$ with continuous and dense imbeddings. We shall use the same notation $(\cdot, \cdot)$ for the inner product in H and for the pairing between V^* and V.

Let T be a positive integer. The space $C([0,T],H)$ has the norm

$$\|u\| = sup_{t\in[0,T]}\|u(t)\|_H,$$

and the space $W_2^1(0,T;V,H) = \{u \in L_2(0,T;V) : u' \in L_2(0,T;V^*)\}$ has the norm

$$\|u\| = (\int_0^T \|u(t)\|_V^2 dt)^{1/2} + (\int_0^T \|u'(t)\|_{V^*}^2 dt)^{1/2}.$$

Notice that, in some books $W_2^1(0,T;V,H)$ is also denoted by $W(0,T)$ (see [43]) or $L^2((0,T),V) \cap H^1((0,T),V^*)$ (see [39]). Here we use the notation in [44].

We consider the following first order system dependent on a parameter q. Suppose $f \in L^2((0,T),V^*)$ and $u_0 \in H$. We investigate a function $u \in W_2^1(0,T;V,H)$ with

$$\begin{cases} \frac{du(t)}{dt} = A(q)u(t) + f(t), \\ u(0) = u_0, \end{cases} \tag{2.1}$$

where $A(q) \in L(V,V^*)$ for each q.

The parameter q belongs to the space Q, which is assumed to be a subset of a separable Hilbert space with norm $\|\cdot\|_Q$. We assume that observations y_i for the solution $u(t_i,q^*)$ of (2.1) at discrete times t_i, $i = 1,2,\cdots,m$ are given. The goal is to find $\hat{q} \in Q$ from this data. In particular, we solve the following inverse problem:

Find $\hat{q} \in Q$ that minimizes the functional

$$J(q) = \sum_i \|Cu(t_i; q) - y_i\|_Y^2,$$

where C is a bounded linear mapping from the state space H to the observation space Y. Clearly, if $y_i = Cu(t_i; q^*)$, then $\hat{q} = q^*$ is the solution to the problem.

With $A(q)$, we can associate the sesquilinear forms

$$\sigma(q)(\phi, \varphi) = (A(q)\phi, \varphi),$$

which is defined from $V \times V$ to R. We assume that $\sigma(q)(\cdot, \cdot)$ has the following properties:

$(A1)$ **Boundedness:** there exists a positive number M_1 such that for all $q \in Q$ and for $\phi, \varphi \in V$, the following holds

$$|\sigma(q)(\phi, \varphi)| \leq M_1 \|\phi\|_V \|\varphi\|_V.$$

$(A2)$ **Coercivity:** there exists a positive number α and a real number λ_0 such that for $q \in Q$, $\phi \in V$ we have

$$\sigma(q)(\phi, \phi) + \lambda_0 \|\phi\|_H^2 \geq \alpha \|\phi\|_V^2.$$

The following two results are well-known, see [39], [43], [44].

Proposition 2.1 *Suppose (A1) and (A2) hold. Then for each $q \in Q$, the equation (2.1) has exactly one solution $u \in W_2^1(0,T;V,H)$.*

Proposition 2.2 *The space $W_2^1(0,T;V,H)$ is continuously embedded in $C([0,T],H)$.*

In order to solve the parameter estimation problem, using direct methods, we require some knowledge on the derivative of the state with respect to the unknown parameter. To obtain differentiability, we need the following additional assumptions:

$(A3)$ **F-differentiable of $\sigma(q)$:** For $q \in Q$ and for $\phi, \varphi \in V$, there exists a linear functional $\partial_q \sigma(q)(\phi, \varphi)$ defined from $Q \to \mathcal{R}$ such that for any $h \in Q$

$$|\sigma(q+h)(\phi,\varphi) - \sigma(q)(\phi,\varphi) - \partial_q\sigma(q)(\phi,\varphi) \cdot h| \leq o(\|h\|_Q)\|\phi\|_V\|\varphi\|_V.$$

$(A4)$ **Boundedness of $\partial_q \sigma(q)(\phi, \varphi)$:** For an element q in Q, there exists $M_2 > 0$ such that for any h in Q and ϕ, φ in V

$$|\partial_q\sigma(q)(\phi,\varphi) \cdot h| \leq M_2\|h\|_Q\|\phi\|_V\|\varphi\|_V.$$

Remark 2.1 *From $(A3)$ and $(A4)$, we see that σ has the continuity property. That is, there exists $\delta = \delta(q) > 0$ such that if $\|h\|_Q < \delta$, then there is $M_3 > 0$ so that the* *following holds*

$$\|\sigma(q+h)(\phi,\varphi)-\sigma(q)(\phi,\varphi)\| \le M_3\|h\|_Q\|\phi\|_V\|\varphi\|_V.$$

2.2 Parameter Dependence

In this section we deduce smoothness properties for the solution $u(q)$ of (2.1) with respect to the parameter q. First, let us recall the definition of the Fréchet derivative ([39]).

Definition 2.1 *Let X and Y be Banach spaces and let x_0 be a point in X. Let F be a mapping from a neighborhood of x_0 into Y. Then F is called Fréchet differentiable at x_0 if there exists a bounded linear operator $A \in L(X,Y)$ such that*

$$\lim_{x\to 0}\frac{\|F(x_0+x)-F(x_0)-Ax\|_Y}{\|x\|_X}=0.$$

If such an A exists, we call it the Fréchet derivative of F at x_0, denoted by $DF(x_0)$.

Theorem 2.1 *Suppose $(A1)-(A4)$ hold. Let $u(t;q) \in C([0,T],H)$ be the solution of (2.1). Then for each $t \in [0,T]$, $u(t;q)$ is Fréchet differentiable with respect to q at every $q \in Q$. Moreover, for each $h \in Q$, $v(t) = D_q u(t;\hat{q})h$ is the unique solution of the weak sensitivity equation*

$$\begin{cases} (v_t,\varphi)+\sigma(\hat{q})(v,\varphi)+\partial_q\sigma(\hat{q})(u(\hat{q}),\varphi)\cdot h = 0, & for \quad \forall\varphi\in V \\ v(0)=0. \end{cases} \tag{2.2}$$

Proof. By Proposition 2.1 and Proposition 2.2, it is clear that both (2.1) and (2.2) have unique solutions in $C([0,T],H)$ for each $q,h \in Q$. For the proof of the remaining part, as in [43] and [44], it suffices to prove the result for the case that $\lambda_0 = 0$ in condition $(A2)$.

Let $h \in Q$, $\|h\|_Q < \delta$, where δ is defined as in Remark 2.1. Let $v(t)$ denote the solution of (2.2) corresponding to h. It is clear that for each fixed $t \in [0,T]$, the mapping from $h \in Q$ to $v(t) \in H$ is linear and continuous. From Definition 2.1, we only need to show that for each $t \in [0,T]$,

$$\frac{\|u(t;\hat{q}+h) - u(t;\hat{q}) - v(t)\|_H}{\|h\|_Q} \to 0,$$

as $\|h\|_Q \to 0$.

Let $l(t) = u(t;q+h) - u(t;q)$. Then $l(t)$ satisfies

$$\begin{cases} (l_t,\varphi) + \sigma(q+h)(u(q+h),\varphi) - \sigma(q)(u(q),\varphi) = 0, & \text{for} \quad \forall\varphi \in V, \\ l(0) = 0. \end{cases}$$

Thus,

$$(l_t,\varphi) + \sigma(q+h)(l,\varphi) + \sigma(q+h)(u(q),\varphi) - \sigma(q)(u(q),\varphi) = 0.$$

Select $\varphi = l(t)$, then we have

$$(l_t,l) + \sigma(q+h)(l,l) \quad = \quad -\sigma(q+h)(u(q),l) + \sigma(q)(u(q),l)$$

$$\begin{aligned} &\leq M_3\|h\|_Q\|u(q)\|_V\|l\|_V \\ &\leq \frac{1}{2\alpha}M_3^2\|h\|_Q^2\|u(q)\|_V^2 + \frac{\alpha}{2}\|l\|_V^2, \end{aligned}$$

here we have used Remark 2.1. Integrating in time and compute the left hand side, we obtain

$$\begin{aligned} \int_0^T [(l_t, l) + \sigma(q+h)(l,l)]dt &= \frac{1}{2}\|l(T)\|_H^2 - 0 + \int_0^T \sigma(q+h)(l,l)dt \\ &\geq \int_0^T \alpha\|l\|_V^2 dt. \end{aligned}$$

Therefore, we have

$$\int_0^T \|l\|_V^2 dt \leq c\|h\|_Q^2 \int_0^T \|u(q)\|_V^2 dt. \tag{2.3}$$

Now, set $w(t) = u(t;\hat{q}+h) - u(t;\hat{q}) - v(t)$, where $\|h\|_Q < \delta$. Then $w(t)$ satisfies the following equation:

$$\begin{cases} (w_t,\varphi) + \sigma(\hat{q}+h)(u(\hat{q}+h),\varphi) - \sigma(\hat{q})(u(\hat{q}),\varphi) - \sigma(\hat{q})(v,\varphi) - \partial_q\sigma(\hat{q})(u(\hat{q}),\varphi)\cdot h = 0 \\ w(0) = 0 \end{cases}$$

for each φ in V.

Therefore, we have

$$\begin{aligned} (w_t,\varphi) &+ \sigma(\hat{q})(w,\varphi) \\ &+ [\sigma(\hat{q}+h)(u(\hat{q}+h),\varphi) - \sigma(\hat{q})(u(\hat{q}+h),\varphi) - \partial_q\sigma(\hat{q})(u(\hat{q}+h),\varphi)\cdot h] \end{aligned}$$

$$+ \quad [\partial_q\sigma(\hat{q})(u(\hat{q}+h),\varphi)\cdot h - \partial_q\sigma(\hat{q})(u(\hat{q}),\varphi)\cdot h] = 0.$$

Since $\sigma(q)(\cdot,\cdot)$ is sesquilinear, $\partial_q\sigma(\hat{q})(\phi,\varphi)$ is linear with respect to ϕ. This implies

$$\begin{aligned}
(w_t,\varphi) \quad &+ \quad \sigma(\hat{q})(w,\varphi) \\
&+ \quad [\sigma(\hat{q}+h)(u(\hat{q}+h),\varphi) - \sigma(\hat{q})(u(\hat{q}+h),\varphi) - \partial_q\sigma(\hat{q})(u(\hat{q}+h),\varphi)\cdot h] \\
&+ \quad \partial_q\sigma(\hat{q})(u(\hat{q}+h)-u(\hat{q}),\varphi)\cdot h = 0.
\end{aligned} \tag{2.4}$$

If we select $\varphi = w(t)$, then (2.4) reduces to

$$\begin{aligned}
\frac{1}{2}\cdot\frac{d}{dt}\|w\|_H^2 \quad &= \quad (w_t,w) \\
&\leq \quad o(\|h\|_Q)\cdot\|u(\hat{q}+h)\|_V\|w\|_V + M_2\|h\|_Q\|u(\hat{q}+h)-u(\hat{q})\|_V\|w\|_V \\
&- \quad \sigma(\hat{q})(w,w \qquad\qquad\quad) \\
&\leq \quad o(\|h\|_Q)\cdot\|u(\hat{q})\|_V\|w\|_V + \hat{c}\|h\|_Q\|u(\hat{q}+h)-u(\hat{q})\|_V\|w\|_V - \alpha\|w\|_V^2 \\
&\leq \quad o(\|h\|_Q^2)\cdot\|u(\hat{q})\|_V^2 + \hat{c}\|u(\hat{q}+h)-u(\hat{q})\|_V^2\|h\|_Q^2.
\end{aligned}$$

Integrating from 0 to t and using the initial condition $w(0) = 0$, we obtain from (2.3)

$$\|w\|_H^2 \leq \hat{c}\|h\|_Q^4 + o(\|h\|_Q^2).$$

Thus, it follows that

$$\|w/\|h\|_Q\|_H \to 0 \qquad \text{as} \qquad \|h\|_Q \to 0,$$

and this completes the proof. $\square$

Theorem 2.2 *Suppose $(A1)-(A4)$ hold. In addition, assume that*

$(A5)$ **Lipschitz continuous of $\partial_q\sigma(q)(\phi,\varphi)$ with respect to q:** *For any ϕ,φ in V, there exists $M_4>0$ such that*

$$\|\partial_q\sigma(q+\Delta q)(\phi,\varphi)-\partial_q\sigma(q)(\phi,\varphi)\| \le M_4\|\Delta q\|_Q\|\phi\|_V\|\varphi\|_V.$$

Then for each $t\in[0,T]$, $D_qu(t;q)$ is locally Lipschitz continuous with respect to q.

Proof. As in the proof of the previous theorem, we only need to consider the case $\lambda_0=0$ in condition $(A2)$. Pick a point h in Q so that $\|h\|_Q=1$. Suppose $v(q)$ and $v(q+\Delta q)$ are the solutions of (2.2) with $\hat{q}=q$ and $\hat{q}=q+\Delta q$, respectively. For small Δq, let $r(t)=v(q+\Delta q)-v(q)$. Then $r(t)$ satisfies

$$\begin{cases} (r_t,\varphi)-\sigma(q)(v(q),\varphi)+\sigma(q+\Delta q)(v(q+\Delta q),\varphi)-\partial_q\sigma(q)(u(q),\varphi)\cdot h \\ \qquad\qquad +\partial_q\sigma(q+\Delta q)(u(q+\Delta q),\varphi)\cdot h=0 \quad \text{for } \forall\varphi\in V \\ \\ r(0)=0. \end{cases}$$

Select $\varphi=r(t)$. Since $\sigma(q+\Delta q)(r,r)\ge\alpha\|r\|_V^2$, it follows that

$$\frac{1}{2}\cdot\frac{d}{dt}\|r\|_H^2 \;=\; (r_t,r)$$

$$
\begin{aligned}
&= [\sigma(q)(v(q), r) - \sigma(q+\Delta q)(v(q), r)] \\
&+ [\partial_q \sigma(q)(u(q), r) \cdot h - \partial_q \sigma(q+\Delta q)(u(q), r) \cdot h] \\
&+ \partial_q \sigma(q+\Delta q)(u(q) - u(q+\Delta q), r) \cdot h - \alpha \|r\|_V^2 \\
&\leq M_3 \|\Delta q\|_Q \|v(q)\|_V \|r\|_V + M_4 \|\Delta q\|_Q \|u(q)\|_V \|r\|_V \|h\|_Q \\
&+ M_2 \|h\|_Q \|u(q) - u(q+\Delta q)\|_V \|r\|_V - \alpha \|r\|_V^2 \\
&\leq C \|\Delta q\|_Q^2 (\|v(q)\|_V^2 + \|u(q)\|_V^2) + C \|u(q) - u(q+\Delta q)\|_V^2;
\end{aligned}
$$

here we have used Remark 2.1, assumptions $(A5)$ and $(A4)$. Integrating from 0 to t, noting that $r(0) = 0$ and employing (2.3), we conclude that

$$
\|r\|_H^2 \leq \hat{C}(t) \|\Delta q\|_Q^2.
$$

□

2.3 An Example

To conclude this chapter, we give an example which shows that the usual parabolic systems, which include the equations discussed in [26], can be treated with the theory we have just discussed. Notice that our assumption on the admissible parameter space is very weak. For other applications, the readers may consult [8].

Consider the following Dirichlet boundary value problem for a one-dimensional parabolic problem:

$$\begin{cases} \frac{du(t)}{dt} = (q_1 u_x)_x + (q_2 u)_x + q_3 u + f(t), \quad \text{on } (0,1), \\ u(0) = u_0, \end{cases} \tag{2.5}$$

where $q = (q_1, q_2, q_3)$ is the parameter. Let the state spaces $H = L^2(0,1)$, $V = H_0^1(0,1)$. The weak form of the equation is given by

$$(\frac{du}{dt}, \varphi) + \sigma(q)(u, \varphi) = (f, \varphi), \quad \varphi \in V,$$

where the sesquilinear form is defined as

$$\sigma(q)(\phi, \varphi) = (q_1 \phi_x, \varphi_x) + (q_2 \phi, \varphi_x) - (q_3 \phi, \varphi). \tag{2.6}$$

Proposition 2.3 *Let $Q = \{(q_1, q_2, q_3) \in C[0,1] \times C[0,1] \times C[0,1] | 0 < m \leq q_1\}$ for some constant m. Then if $q \in Q$, $\sigma(q)(\cdot,\cdot)$ as defined by (2.6) satisfies $(A1) - (A5)$.*

Proof. Let $q \in Q$ and $\phi, \varphi \in H_0^1$. Then by using Hölder's inequality and Poincaré's inequality, we have

$$\begin{aligned} |\sigma(q)(\phi, \varphi)| &\leq |(q_1 \phi_x, \varphi_x)| + |(q_2 \phi, \varphi_x)| + |(q_3 \phi_x, \varphi)| \\ &\leq \|q_1\|_C \int_0^1 |\phi_x||\varphi_x| dx + \|q_2\|_C \int_0^1 |\phi||\varphi_x| dx \\ &+ \|q_3\|_C \int_0^1 |\phi||\varphi| dx \\ &\leq M_1 \|q\|_Q \|\phi\|_{H^1} \|\varphi\|_{H^1} \end{aligned} \tag{2.7}$$

and

$$
\begin{aligned}
\sigma(q)(\phi,\phi) &= (q_1\phi_x,\phi_x)+(q_2\phi,\phi_x)+(q_3\phi,\phi) \\
&\geq \min\{q_1\}\int_0^1 |\phi_x||\phi_x|dx - \|q_2\|_C\int_0^1 |\phi||\phi_x|dx \\
&- \|q_3\|_C\int_0^1 |\phi||\phi|dx \\
&\geq \frac{1}{2}m\int_0^1 |\phi_x||\phi_x|dx - C\|\phi\|_{H^0}^2 \\
&\geq \alpha\|\phi\|_{H^1}^2 - \lambda_0\|\phi\|_{H^0}^2 \qquad (2.8)
\end{aligned}
$$

where M_1, α are positive constants and C, λ_0 are real numbers. From (2.7) and (2.8), we can see that (A1) and (A2) hold.

For (A3) - (A5), notice that $\sigma(q)(\cdot,\cdot)$ is linear with respect to q, which implies that

$$\partial_q\sigma(q)(\phi,\varphi)\cdot h = \sigma(h)(\phi,\varphi), \qquad (2.9)$$

for each $h \in Q$. Combining (2.7), (2.8) and (2.9), we obtain (A3) - (A5). □

The theoretical framework presented in [8], where they used an indirect approach to identify unknown coefficient functions, can be used to treat many types of systems including the problems in which the underlying semigroup is not analytic or the problems involving functional partial differential equations ([6]). With appropriate modifications, the framework we present here should also be able to treat some other types of problems.

Chapter 3

A Generalized Burgers' Equation

3.1 Preliminary

Consider the initial value problem

$$\begin{cases} u_t(t) + Au(t) = f(t, u(t)), & t > 0 \\ u(0) = u_0 \end{cases} \tag{3.1}$$

on a Hilbert space X, here we assume A is a sectorial operator (i.e. $-A$ generates an analytic semigroup) and that the spectrum of A lies entirely in the (open) right half-plane. In this case, the fractional powers of A are well defined, and the space $X_\alpha = D(A^\alpha)$ with the graph norm $\|u\|_\alpha = \|A^\alpha u\|$ is defined for each $\alpha \geq 0$.

For our assumption concerning the function f in (3.1), we will use the following definition:

Definition 3.1 *Let U be an open subset of $R \times X^\alpha$, where α is between $0 \leq \alpha < 1$. We say that f is locally Hölder continuous in t and locally Lipschitz in x on U if for every $(t, u) \in U$, there exists a neighborhood V of (t, u) so that for $(t_1, u_1), (t_2, u_2) \in$*

V, there are $L > 0$ and $\theta > 0$ such taht

$$\|f(t_1, u_1) - f(t_2, u_2)\|_X \leq L(|t_1 - t_2|^\theta + \|u_1 - u_2\|_\alpha).$$

The following local existence theorem for the solution of equation (3.1) can be found in [37], [28] and [18].

Theorem 3.1 *Let A be as before and let f be locally Hölder continuous in t and locally Lipschitz continuous in x in an open set $U \subset R \times X^\alpha$. Then for every initial data $(t_0, u_0) \in U$, there exists $T = T(t_0, u_0) > 0$ such that the initial value problem (3.1) has a unique local solution $u(t)$ on $(t_0, t_0 + T)$ with initial value $u(t_0) = u_0$.*

In this chapter, the equation we consider is a special case of (3.1). In Section 3.2, we will show that the equation satisfies the conditions in Theorem 3.1 and we will establish regularity properties. In Section 3.3 we establish a Maximum Principle for the equation; and in Section 3.4 we prove the global existence and differentiability of the solution with respect to the parameters.

3.2 Local Existence and Regularity Properties

We consider a generalized Burgers' equation which has the Dirichlet boundary condition and is defined on a finite interval $[0,1]$ by

$$\begin{cases} u_t = \epsilon u_{xx} - q(x)uu_x, & 0 < x < 1, \ \ t > 0, \\ u(t,0) = u(t,1) = 0, & t > 0, \\ u(0,x) = u_0(x), & 0 < x < 1, \end{cases} \tag{3.2}$$

where $\epsilon > 0$. We assume $q(x) \in C[0,1]$, and $u_0(x) \in H_0^1(0,1)$.

Let $X = L^2(0,1)$, define an operator A in X by $Au = -\epsilon u_{xx}$ with $\mathcal{D}(A) = H^2(0,1) \cap H_0^1(0,1)$. It is well-known that $-A$ generates an analytic semigroup, and the spectrum $\sigma(A)$ of A consists of all eigenvalues $\epsilon n^2\pi^2, n = 1, 2, \cdots$. Thus we have $Re(\sigma(A)) > 0$. Therefore, we can define fractional powers of A and $\mathcal{D}(A^{1/2})$ is $H_0^1(0,1)$. Let $X^{1/2} = \mathcal{D}(A^{1/2}) = H_0^1(0,1)$.

The function $f(u) = -q(x)uu_x$ is defined from $X^{1/2}$ to X, and (3.2) can be written as the initial-value problem

$$\begin{cases} u_t + Au = f(u), & t > 0 \\ u(0) = u_0. \end{cases}$$

Lemma 3.1 *$f : X^{1/2} \to X$ is locally Lipschitz. In particular, if $u \in X^{1/2}$, then there exists a neighborhood V of u and a constant $C > 0$ such that for $v, w \in V$,*

$$\|f(v) - f(w)\|_X \leq C\|v - w\|_{X^{1/2}}.$$

Proof. For any $v, w \in H_0^1(0,1)$, since $H^1(0,1)$ is continuously embedded in $C_b[0,1]$, it follows from Poincaré's inequality that

$$\begin{aligned}
\|f(v) - f(w)\|_{L^2} &= \|-q(x)vv_x + q(x)ww_x\|_{L^2} \\
&\leq \|q(x)\|_{C_b}\|vv_x - ww_x\|_{L^2} \\
&\leq \|q(x)\|_{C_b}(\|v(v_x - w_x)\|_{L^2} + \|w_x(v-w)\|_{L^2}) \\
&\leq \|q(x)\|_{C_b}(\|v\|_{C_b} \cdot \|v_x - w_x\|_{L^2} + \|w_x\|_{L^2} \cdot \|v - w\|_{C_b}) \\
&\leq \|q(x)\|_{C_b} \cdot C(\|v\|_{H^1} \cdot \|v_x - w_x\|_{L^2} + \|w_x\|_{L^2} \cdot \|v - w\|_{H^1}) \\
&\leq C\|q(x)\|_{C_b} \cdot (\|v\|_{H^1} + \|w\|_{H^1}) \cdot \|v - w\|_{H^1}
\end{aligned}$$

Hence, f is locally Lipschitz. □

The following theorem establishes the existence and smoothness properties of local solutions to (3.2). This result is based on Lemma 1, Definition 3.3.1, Theorem 3.3.3 and Theorem 3.5.2 from Chapter 3 of [28]. Part of the result can also be derived from [37].

Theorem 3.2 *Suppose $\epsilon > 0$ and $q(x) \in C[0,1]$. Then for every $u_0 \in X^{1/2} = H_0^1(0,1)$, there exists $T > 0$ such that (3.2) has a unique solution $u(t)$ on $[0,T)$, where $u(t) \in C([0,T), H_0^1(0,1)) \cap C^1((0,T), H_0^1(0,1))$, and $u(t) \in \mathcal{D}(A)$ for each $t \in (0,T)$.*

For our system (3.2), this solution possess, the following regularity property.

Theorem 3.3 *If $u(t) = u(t,x)$ is the local solution given in Theorem 3.2, then $u(t,x)$ is continuous on $[0,T) \times [0,1]$. Moreover, $u(t,x)$ is a classical solution of (3.2) on $[0,T) \times [0,1]$.*

Proof. First, from Theorem 3.2 we know that

$$u(t) \in C([0,T); H_0^1(0,1)) \cap C^1((0,T); H_0^1(0,1)).$$

Also, since $H_0^1(0,1)$ is continuously imbedded in $C[0,1]$, we see that $u(t,x)$ is continuous on $[0,T) \times [0,1]$ and is continuously differentiable in $t \in (0,T)$ for each $x \in (0,1)$. Since $u(t) \in \mathcal{D}(A) = H^2(0,1) \cap H_0^1(0,1)$, it follows that $u(t,x)$ is continuously differentiable for each $t \in (0,T)$. Clearly, u_{xx} belongs to $L^2(0,1)$.

Notice that every term, except ϵu_{xx}, in the equation is continuous. Thus, u_{xx} is also continuous. Consequently, we see that, for $t \in (0,T)$ and $x \in (0,1)$, $u(t,x)$ is continuously differentiable in t and twice continuously differentiable in x. Hence it is a classical solution, completing the proof. $\square$

3.3 The Maximum Principle

Theorem 3.4 *Let $u(t,x)$ be the solution of (3.2) on $[0,T) \times [0,1]$. Under our assumptions, the maximum absolute value of $u(t,x)$ is reached on $\{0\} \times [0,1]$. That* *is,*

$$|u(t,x)| \leq \max_{0\leq x\leq 1} |u(0,x)|$$

for $t \in [0,T)$, $x \in [0,1]$ *and* $\|u(t,\cdot)\|_C = \max_{0\leq x\leq 1} |u(t,x)|$ *is decreasing on* $t \in [0,T)$.

Proof. Let $v(t,x) = e^{-\frac{\lambda}{2}t}u(t,x)$ for some $\lambda > 0$ and let $\hat{T}$ be any number between 0 and T. From Theorem 3.2, we know that $u(t,x)$ is continuous on $[0,\hat{T}]\times[0,1]$, hence $v^2(t,x)$ is also continuous on $[0,\hat{T}]\times[0,1]$ and thus $v^2(t,x)$ reaches its maximum value on $[0,\hat{T}]\times[0,1]$. Suppose the maximum value of $v^2(t,x)$ is reached at $(t_1,x_1) \in (0,\hat{T}]\times(0,1)$. Without loss of generality, we may assume that $v^2(t_1,x_1) > 0$. By Theorem 3.2, we see that for each $x \in [0,1]$, $v^2(t,x)$ is continuous differentiable in $t \in (0,T) \supseteq (0,\hat{T}]$, and for each $t \in [0,\hat{T}]$, $v^2(t,x)$ is twice continuous differentiable in $x \in (0,1)$. So the first derivative of v^2 in x vanishes at the point (t_1,x_1); that is,

$$(v^2)_x(t_1,x_1) = 0. \tag{3.3}$$

The first derivative of v^2 in t vanishs at the point (t_1,x_1) where $t_1 \in (0,\hat{T})$; and when $t_1 = \hat{T}$, it is greater than or equal to 0, hence we have

$$(v^2)_t(t_1,x_1) \geq 0. \tag{3.4}$$

The second derivative of v^2 in x is less than or equal to 0,

$$(v^2)_{xx}(t_1, x_1) \leq 0. \tag{3.5}$$

Since $(v^2)_x(t_1, x_1) = 2v(t_1, x_1)v_x(t_1, x_1) = 0$ and $v(t_1, x_1) \neq 0$, we have

$$v_x(t_1, x_1) = 0. \tag{3.6}$$

Now lets go back to look at the equation (3.2) and we will get a contradiction. Multiply both sides of the first equation of (3.2) by $u(t, x)$, we get

$$\frac{1}{2}(u^2)_t - \epsilon u_{xx}u + qu^2u_x = 0,$$

Since $(u^2)_{xx} = 2uu_{xx} + 2u_x^2$, it follows that

$$\frac{1}{2}(u^2)_t - \frac{\epsilon}{2}(u^2)_{xx} + \epsilon u_x^2 + \frac{q}{2}u(u^2)_x = 0.$$

Replace u by $e^{\frac{\lambda}{2}t}v$ one obtains

$$(v^2)_t + \lambda v^2 - \epsilon(v^2)_{xx} + 2\epsilon e^{\lambda t}v_x^2 + qe^{\frac{\lambda}{2}t}v(v^2)_x = 0. \tag{3.7}$$

It follows from (3.3), (3.4), (3.5), and (3.6) that at the point (t_1, x_1) we have

$$(v^2)_t \geq 0, \quad \lambda v^2 > 0, \quad -\epsilon(v^2)_{xx} \geq 0,$$

and the last two terms of (3.7) are zero. This means that the equation (3.2) can not hold at (t_1, x_1). Thus, $v^2(t,x)$ does not achieve its maximum value at (t_1, x_1), which is an arbitrary point in $(0, \hat{T}] \times (0,1)$. Therefore, the maximum value of $v^2(t,x)$ is achieved on $\{0\} \times [0,1]$, that is

$$|v^2(t,x)| \le \max_{0 \le x \le 1} |v^2(0,x)| \quad \text{for } \; t \in [0,T), \;\; x \in [0,1],$$

or equivalently,

$$|e^{-\lambda t} u^2(t,x)| \le \max_{0 \le x \le 1} |u_0(x)| \quad \text{for } \; t \in [0,T), \;\; x \in [0,1].$$

Hence, for all $\lambda > 0$, it follows that

$$|u(t,x)| \le e^{\frac{\lambda}{2} t} \max_{0 \le x \le 1} |u_0(x)|.$$

Now letting $\lambda \to 0$, we obtain

$$|u(t,x)| \le \max_{0 \le x \le 1} |u_0(x)| \quad \text{for } \; t \in [0,T), \;\; x \in [0,1],$$

and this completes the proof. □

3.4 Global Existence and Differentiable Dependence Properties

Theorem 3.5 *Assume ϵ, $b(x)$ and $u_0(x)$ satisfy the hypotheses of Theorem 3.4.*

Then the system has a unique global classic solution.

Proof. For the function $f(u) = -q(x)uu_x$, we have

$$\|f(u)\|_{L^2} = \| - q(x)uu_x\|_{L^2} \leq \|q\|_{C_b}\|u\|_{C_b}\|u_x\|_{L^2}.$$

Since $u \in H_0^1(0,1)$, it follows from the fact that $H^1(0,1)$ is continuously embedded in $C[0,1]$ and Poincaré 's inequality that

$$\|f(u)\|_{L^2} \leq C\|u\|_{H^1}^2$$

for some $C > 0$, where C depends only on $q(x)$.

This shows that f maps bounded sets in $X^{1/2} = H_0^1(0,1)$ to bounded sets in $X = L^2(0,1)$. Thus if $u(t)$ is a solution of (3.2) on $[0,T)$ and T is maximal, then either $T = \infty$ or there exists a sequence $t_n \to T$ as $n \to \infty$ such that $\|u(t_n)\|_{X^{1/2}} \to \infty$. (See Theorem 3.3.4. in Chapter 3 of [28].)

Next we prove $\|u(t)\|_{X^{1/2}}$ is bounded on any finite interval.

Let $u(t)$ be a solution of (3.2) on $[0,T)$ with $T < \infty$. Taking the $L^2(0,1)$ inner product of both sides of the first equation of (3.2) with $-u_{xx}$, we have

$$\frac{1}{2}\frac{d}{dt}\|u_x\|^2 = -\epsilon\|u_{xx}\|_{L^2}^2 + \int_0^1 q(x)u_x u u_{xx} dx. \tag{3.8}$$

For the last integral on the right side of (3.8), we have the estimate

$$\begin{aligned} \left|\int_0^1 q(x) u_x u u_{xx} dx\right| &\leq \|u_0\|_{C_b} \|q\|_{C_b} \|u_x\|_{L^2} \|u_{xx}\|_{L^2} \\ &\leq \delta \|u_{xx}\|_{L^2}^2 + \frac{\|u_0\|_{C_b}^2 \|q\|_{C_b}^2}{4\delta} \|u_x\|_{L^2}^2. \end{aligned}$$

Here, we used Theorem 3.3 and Hölder's inequality on the first step and the inequality

$$ab \leq \delta a^2 + \frac{1}{4\delta} b^2$$

on the second step. If we select $\delta = \epsilon/2$, then (3.8) yields

$$\begin{aligned} \frac{d}{dt} \|u_x\|_{L^2}^2 + \epsilon \|u_{xx}\|_{L^2}^2 &\leq \frac{\|u_0\|_{C_b}^2 \|q\|_{C_b}^2}{\epsilon} \|u_x\|_{L^2}^2 \\ &= \bar{C} \|u_x\|_{L^2}^2, \end{aligned}$$

where $\bar{C}$ is a fixed constant for $\epsilon, q(x), u_0(x)$ are given.

Now let us consider the following inequality (which is the previous one without the second term on the left):

$$\frac{d}{dt} \|u_x\|_{L^2}^2 \leq \bar{C} \|u_x\|_{L^2}^2. \tag{3.9}$$

From Theorem 3.1, we know that $u(t) \in C^1((0,T), H_0^1(0,1))$. Thus, for small $\rho \in$

$(0, T)$, it follows from (3.9) that

$$\|u_x(t)\|_{L^2}^2 \leq e^{\ln(\|u_x(\rho)\|_{L^2}^2) - \bar{C}\rho} \cdot e^{\bar{C}t}. \tag{3.10}$$

From (3.10), it is clear that $\|u_x(t)\|_{L^2}$ is bounded on $[0, T)$. So the Poincaré's inequality implies that $\|u(t)\|_{H^1}$ is bounded on $[0, T)$. Now it is clear from the argument at the beginning of the proof that (3.2) has a global solution. □

By Theorem 3.4.4. and Corollary 3.4.5. in [28], we have the following theorem which establishes differentiability of the solution $u(t)$ of (3.2) w.r.t. the parameters.

Theorem 3.6 *Let $u(t)$ be the solution of (3.2) on $[0, \infty)$. Then the mapping $(\epsilon, q, u_o) \rightarrow u(t; \epsilon, q, u_o)$ is infinitely often differentiable from $R^+ \times C_b[0,1] \times H_0^1(0,1)$ into $H_0^1(0,1)$ for $t \in (0, \infty)$. Moreover, the derivatives $w(t) = D_\epsilon u(t)$, $v(t) = D_{q(x)} u(t)$, $z(t) = D_{u_0} u(t)$ are the solutions of the following equations*

$$\begin{aligned}
\frac{dw}{dt} &= \epsilon w_{xx} - q(x) u w_x - q(x) u_x w + u_{xx}, \ w(0) = 0; \\
\frac{dv}{dt} &= \epsilon v_{xx} - q(x) u v_x - q(x) u_x v - u u_x, \ v(0) = 0; \\
\frac{dz}{dt} &= \epsilon z_{xx} - q(x) u z_x - q(x) u_x z, \ z(0) = 1.
\end{aligned}$$

Chapter 4

Parameter Estimation by Quasilinearization

In this chapter we formulate the parameter estimation problems as an optimal control problem in which the parameters are the control variables. We discuss two issues, the first one is the existence of a solution to the resulting optimal control problem, the second one is the convergence result of the quasilinearization method.

4.1 Formulation of the Parameter Estimation Problem

The parameter estimation problem for (2.1) or (3.2) can be formulate as follows. We assume we are given observations or data $y_i \in Y$ at discrete times t_i, $i = 1, 2, \cdots, m$, where the observation space Y is a Hilbert space. The state space H is a Hilbert space, and we assume the admissible parameter set Q is a subset of a Hilbert space. We wish to determine q so that some observed part, $Cu(t_i; q)$, of the state u depending on q best approximates y_i. In other words, we seek to identify

$q^*(x)$ in Q, that minimizes

$$J(q) = \sum_i \|Cu(t_i;q) - y_i\|_Y^2 + \beta\|q\|_{H^1}^2, \tag{4.1}$$

where C is a bounded linear mapping from the state space H to Y and $\beta \geq 0$. When $\beta > 0$, the second term is a 'cost term' and serves as a regularization.

To establish the existence of a solution to our inverse problem, we need the following definition:

Definition 4.1 *Let X and Y be Banach spaces. We say that a mapping $F : X \to Y$ is sequentially weakly continuous if whenever*

$$x_n \rightharpoonup \bar{x} \quad \textit{weakly in } X,$$

we have

$$F(x_n) \to F(\bar{x}) \quad \textit{in } Y.$$

We have the following two solution existence theorems; the first one is for $\beta > 0$ and the second one for $\beta \geq 0$.

Theorem 4.1 *Suppose $q \to Cu(t;q)$ is sequentially weakly continuous in Q for each t. Then there exists a solution $q^*(x) \in Q$ for the inverse problem with $\beta > 0$.*

Proof. It is clear that $\{J(q) : q \in Q\}$ is bounded below by zero, so it has a biggest lower bound. Let y be the biggest lower bound for $\{J(q : q \in Q\}$. That is,

$$y = inf\{J(q : q \in Q\}.$$

Then there is a minimizing sequence $\{q_k\}$ such that

$$J(q_k(x)) \to y, \quad \text{as} \;\; k \to \infty.$$

Since

$$\|q\|_Q \leq \frac{1}{\beta} J(q), \tag{4.2}$$

$\{q_k\}$ is bounded in Q. Hence, the sequence q_k has a weakly convergent subsequence $\{q_{k_j}\}$. Let q^* be such that

$$q_{k_j} \rightharpoonup q^* \quad \text{weakly in} \;\; Q.$$

Then, it follows by our hypothesis that

$$Cu(q_{k_j}) \to Cu(q^*) \quad \text{in} \;\; Y.$$

By the continuity property of our Hilbert space norm, we conclude that

$$lim_{j \to \infty} J(q_{k_j}) = J(q^*).$$

Consequently, we have $J(q^*) = y$. Hence, q^* is a solution of the inverse problem. □

Theorem 4.2 *Suppose $q \to Cu(t;q)$ is sequentially weakly continuous in Q for each t. Moreover, assume Q is a bounded. Then there exists a solution $q^*(x) \in Q$ for the inverse problem.*

Proof. The result can be got in the same way as in the proof of Theorem 4.1 except using the hypotheses instead of (4.2) to get $\{q_k\}$ is bounded. □

As an example, we apply Theorem 4.1 and Theorem 4.2 to a parameter estimation problem governed by the generalized Burgers' equation, which was presented in previous chapter.

The parameter dependent system we are considering is the following

$$\begin{cases} \frac{d}{dt}u(t) = \epsilon u_{xx} - q(x)uu_x, & 0 < x < 1, t > 0, \\ u(t,0) = u(t,1) = 0, & t > 0, \\ u(0,x) = u_0(x), & 0 < x < 1. \end{cases} \tag{4.3}$$

Let the state space H be $H_0^1(0,1)$. When we consider the regularity term (i.e. $\beta > 0$), we let Q be $H^1(0,1)$, otherwise let $Q \in H^1(0,1)$ and bounded in $C[0,1]$. Let Y be any Hilbert space and C be a bounded linear operator from $H_0^1(0,1)$ to Y. In the next Chapter, we will use $Y = R^l \times R^m$, $l, m \in Z^+$, and C to be the projection from $H_0^1(0,1)$ to $R^l \times R^m$ in our numerical examples. Combining the results in Chapter 3 with Theorem 4.1, we have the following result.

Corollary 4.1 *There exists a solution $q^*(x)$ of the parameter estimation problem governed by (4.3).*

Proof. From Theorem 4.1 and Theorem 4.2, we only need to prove that the mapping $q \to Cu(t;q)$, from $H^1(0,1)$ to Y, is sequentially weakly continuous for each t.

Let $q_k \rightharpoonup q$ weakly in $H^1(0,1)$. Since $H^1(0,1)$ is compactly embedded in $C_b(0,1)$, it follows that q_k converges to q in $C_b(0,1)$. It follows from Theorem 3.6 that for each $t \in [0,T)$ $u(t;q_k(x)) \to u(t;q)$ in $H_0^1(0,1)$.

Finally, since C is a continuous mapping from $H_0^1(0,1)$ to Y, we have that $Cu(t;q_k)$ converges to $Cu(t;q)$ in Y for each t. Hence, $Cu(t;q)$ is sequentially weakly continuous with respect to q and the proof is complete. □

4.2 The Quasilinearization Algorithm

In this section we discuss the convergence property of the quasilinearization algorithm for the parameter estimation problem. We consider the parameter estimation problem formulated as in the previous section. For the sake of convenience, we rewrite (4.1) in a simpler form as follows: Let $\hat{Y} = Y^m$ (with the product norm), $y = (y_1, y_2, \cdots, y_m)$ and $U(q) = (Cu(t_1;q), Cu(t_2;q, \cdots, Cu(t_m;q))$. Also, we let $Z = \hat{Y} \times Q$ (with the product norm) and set

$$F(q) = \begin{pmatrix} U(q) - y \\ \sqrt{\beta} q \end{pmatrix}.$$

With this setting, (4.1) becomes

$$J(q) = \|F(q)\|_Z^2. \tag{4.4}$$

The method we use to solve this parameter identification problem in this section is quasilinearization. The concept of quasilinearization was introduced in [12]. Since then, there have been various extensions and variations (see [7], [15], [16], [26], [4]). Quasilinearization is a recursive Newton's method type of algorithm, which can be defined as follows:

Given an initial guess $q_0 \in Q$, define

$$\begin{aligned} q^{k+1} &= q^k - D(q_k)^{-1} M^*(q_k) F(q_k), \\ &= G(q^k) \quad k = 0, 1, 2, 3, \cdots, \end{aligned} \tag{4.5}$$

where

$M(q) = D_q F(q)$,

$D(q) = M^*(q) M(q)$,

and $M^*(q)$ is the adjoint operator of $M(q)$.

The following results are straightforward and can be found in [26]:

Lemma 4.1 *Suppose $u(t; q)$ is Fréchet differentiable with respect to q and the mapping $q \to D_q u(t; q)$ is locally Lipschitz continuous in Q for each t. Then both $M(q)$ and $M^*(q)$ are locally Lipschitz continuous in Q, $D(q)$ is continuous in Q.*

Under some assumptions, we can obtain superlinear convergence when there is an exact fit to data and linear convergence when there is 'small' error in data. These results are typical in the quasilinearization methods, and the proofs given here are in the same idea as those in [7], [15], and [26].

Theorem 4.3 *Suppose the hypotheses of Lemma 4.1 are satisfied. Assume $J(q^*) = 0$ and $D(q^*)^{-1}$ exists. Then for every $\epsilon > 0$, there exists a constant $\delta > 0$ such that*

$$\|G(q) - q^*\| < \hat{C}\|q - q^*\|^2 + \epsilon\|q - q^*\| \qquad \textit{whenever} \qquad \|q - q^*\| < \delta,$$

where $\hat{C}$ is a constant which depends on q^*. In particular, q^* is a point of attraction of the iterative scheme $q^{k+1} = G(q^k)$.

Proof. First, we observe that

$$\begin{aligned} G(q) - q^* &= D(q)^{-1}[D(q)(q - q^*) - M^*(q)F(q)] \\ &= D(q)^{-1}[M^*(q)[M(q)(q - q^*) - F(q)]] \\ &= D(q)^{-1}[M^*(q)[M(q) - M(q^*)](q - q^*)] \\ &- D(q)^{-1}[M^*(q)[F(q) - F(q^*) - M(q^*)(q - q^*)]] \\ &- D(q)^{-1}[M^*(q)F(q^*)]. \end{aligned} \tag{4.6}$$

The last term on the right hand side of (4.6) is equal to zero since, by assumption, $J(q^*) = 0$. From Lemma 4.1, we know $D(q)$ is continuous at q^*. Since by assumption

$D(q^*)^{-1}$ exists, there exists constants δ_0 and K so that for $\|q - q^*\| < \delta_0$, $D(q)^{-1}$ exists and $\|D(q)^{-1}\| < K$. By Lemma 4.1 $M^*(q)$ is continuous at q^*, so there exist constants δ_1 and B, such that if $\|q - q^*\| < \delta_1$, then $\|M^*(q)\| < B$. Therefore, we have

$$\begin{aligned} \|G(q) - q^*\| &\leq KB\|[M(q) - M(q^*)](q - q^*)\| \\ &+ KB\|F(q) - F(q^*) - M(q^*)(q - q^*)\|. \end{aligned} \tag{4.7}$$

Since $M(q)$ is locally Lipschitz continuous at each point q in Q, there exists a constant L such that

$$\|[M(q) - M(q^*)](q - q^*)\| \leq L\|q - q^*\|^2. \tag{4.8}$$

For the second term of (4.7), by the definition of the Fréchet derivative, for any $\epsilon > 0$, there exists a constant $\delta_2 > 0$ so that if $\|q - q^*\| < \delta_2$, then

$$\|F(q) - F(q^*) - M(q^*)(q - q^*)\| < \epsilon\|q - q^*\|. \tag{4.9}$$

Combining (4.8) and (4.9) with the inequality (4.7), we see that

$$\|G(q) - q^*\| < KBm[L\|q - q^*\|^2 + \epsilon\|q - q^*\|] \tag{4.10}$$

whenever $\|q - q^*\| < \delta = \min\{\delta_0, \delta_1, \delta_2\}$. This completes the proof. $\square$

The next theorem does not require an exact fit to data, but it does require the data 'error' is small. Note that if $M^*(q)$ is locally Lipschitz continuous at q^*, then there exists $\delta^* > 0$ such that for $0 < \delta < \delta^*$ there exists a Lipschitz constant $L(\delta)$ if $\|q - q^*\| \leq \delta$. Let $\hat{K} = \liminf_{\delta \to 0} L(\delta)$ and let K be defined as in the proof of Theorem 4.3. Then we have the following.

Theorem 4.4 *Suppose the hypotheses of Lemma 4.1 are satisfied. Assume $D(q^*)^{-1}$ exists and q^* is the fixed point of G. Let K and $\hat{K}$ be the constants as above. If*

$$\|F(q^*)\| < \frac{1}{3K\hat{K}}, \tag{4.11}$$

then q^* is a point of attraction of the iterative scheme $q^{k+1} = G(q^k)$, where G is defined as in (4.5).

Proof. We proceed as in the proof of Theorem 4.2. First, observe that in (4.6) the last term on the right hand is no longer equal to zero. So, for $\|q - q^*\| < \delta$, (4.10) becomes

$$\begin{aligned} \|G(q) - q^*\| &\leq KBm[L\|q - q^*\|^2 + \epsilon\|q - q^*\|] \\ &+ \|D(q)^{-1}[M^*(q)F(q^*)]\|. \end{aligned} \tag{4.12}$$

By our assumption, $q^* = G(q^*)$ and it follows that

$$M^*(q^*)F(q^*) = 0. \tag{4.13}$$

From how we define $\hat{K}$, there exists a constant $\hat{\delta}$ such that for $\|q - q^*\| \leq \hat{\delta}$

$$\|M^*(q) - M^*(q^*)\| \leq 2\hat{K}\|q - q^*\|. \tag{4.14}$$

Now, combining (4.13) and (4.13) with (4.12), we have

$$\begin{aligned} \|G(q) - q^*\| &\leq KBm[L\|q - q^*\|^2 + \epsilon\|q - q^*\|] \\ &+ K\|[M^*(q) - M^*(q^*)]F(q^*)\| \\ &\leq KBm[L\|q - q^*\|^2 + \epsilon\|q - q^*\|] \\ &+ K \cdot 2\hat{K}\|q - q^*\|\|F(q^*)\|. \end{aligned}$$

Hence, it follows by (4.11) that

$$\|G(q) - q^*\| < C_1\|q - q^*\|^2 + C_2\epsilon\|q - q^*\| + \frac{2}{3}\|q - q^*\|$$

whenever $\|q - q^*\| < \min\{\hat{\delta}, \delta\}$, where C_1 and C_2 are constants. Therefore, q^* is a point of attraction of the iterative scheme $q^{k+1} = G(q^k)$. $\square$

Chapter 5

Numerical Results

In this chapter we present numerical results based on the quasilinearization algorithm established in [15], [26] and [4]. In [15], J. A. Burns, E. M. Cliff and D. W. Brewer developed the algorithm for delay differential equation where the unknown parameters include constant coefficients and the delay. In [26], P. W. Hammer developed the algorithm for a parabolic partial differential equation with a spatially varying parameter. P. W. Hammer also presented some numerical examples for a nonlinear parabolic equation (Burgers' equation) with a constant parameter. In [4], M. G. Armentano presented some results for a nonlinear parabolic equation with spatially varying parameter.

The goal of our numerical effort was to present some concrete numerical solutions which demonstrate the problems we discussed. Since many numerical examples have been done for the parameter estimation problems governed by linear parabolic problems (see [10] and [26]), we only consider for the nonlinear equation case. In the first example, we give some numerical solutions for the type of nonlinear equations we discussed in Chapter 3 by using the finite element method. In Examples 2 - 6, we consider parameter estimation problem with spatially varying parameter in a nonlinear parabolic partial differential equation. However, we use linear splines instead of the polynomials used in [4] to approximate the parameters. Finally, in Examples 7 - 9, we present numerical solutions for parameter estimation in another nonlinear partial differential equation. The computer codes are written in Matlab and carried out on either a DEC 3000 machine, housed at ICAM, or a Sun SPARCstation 2 (with a Weitek 80 MHz CPU upgrade) at mathematics department.

In each of the examples we use linear splines to approximate the parameter $q(x) \in Q$. Specifically, we choose

$$\hat{Q} = \{\sum_{i=0}^{N_q} \alpha_i g_i(x) | (\alpha_0, \alpha_1, \cdots, \alpha_{N_q})^T \in R^{N_q+1}\}$$

to approximate the parameter space Q, where $g_i(x)$ denotes the standard hat function on the interval $[\frac{i-1}{N_q}, \frac{i+1}{N_q}]$ for $i = 0, 1, 2, \cdots, N_q$. The functions are defined by

$$g_0(x) = \begin{cases} -N_q x + 1, & x \in [0, \frac{1}{N_q}] \\ 0, & \text{elsewhere} \end{cases}$$

$$g_i(x) = \begin{cases} N_q x - (i-1), & x \in [\frac{i-1}{N_q}, \frac{i}{N_q}] \\ -N_q x + (i+1), & x \in [\frac{i}{N_q}, \frac{i+1}{N_q}] \\ 0, & \text{elsewhere} \end{cases}$$

for $i = 1, 2, \cdots, N_q - 1$ and

$$g_{N_q}(x) = \begin{cases} N_q x - (N_q - 1), & x \in [\frac{N_q-1}{N_q}, 1] \\ 0, & \text{elsewhere.} \end{cases}$$

The finite element method is used to solve the partial differential equations. Linear splines are used to discretize the equation in space. For the state equation, a *linearized backward Euler method* and a *linearized Crank-Nicolson method* ([40]) are used to discretize the equation in time. For the sensitive equation, we use a *backward Euler method* to discretize the equation in time. A *3-point Guassina quadrature formula* is used for the calculation of the the integrals.

Suppose we have the following variational formulation of our problem: Find $u(t) \in V = H_0^1(0,1)$, $t \in [0,T]$, such that

$$\begin{cases} (u_t, v) + (a(u)u_x, v) = (f(u), v)) \ \forall v \in V, \\ u(0) = u_0. \end{cases} \tag{5.1}$$

Let V_h be a finite-dimensional subspace of V with basis $\{h_1, h_2, \cdots, h_{N_x}\}$, where $h_i(x)$ are the standard linear spline functions on the interval $[0,1]$ satisfying zero boundary conditions. The functions are defined by

$$h_i(x) = \begin{cases} (N_x+1)x-(i-1), & x \in [\frac{i-1}{N_x+1}, \frac{i}{N_x+1}] \\ -(N_x+1)x+(i+1), & x \in [\frac{i}{N_x+1}, \frac{i+1}{N_x+1}] \\ 0, & \text{elsewhere} \end{cases}$$

for $i = 1, 2, \cdots, N_x$. Replacing V by the finite-dimensional subspace V_h we get the following semi-discrete analogue of (5.1): Find $u_h(t) \in V_h$, $t \in [0, T]$, such that

$$\begin{cases} (u_{h_t}, v) + (a(u_h)u_{h_x}, v) = (f(u_{h_x}), v) & \forall v \in V_h, \\ (u_h(0), v) = (u_0, v) & \forall v \in V_h. \end{cases} \tag{5.2}$$

In order to solve this initial value problem for a system of ordinary differential equations, we will implement a two-step method, a *linearized Crank-Nicolson method*, supplying by a single-step method, a *linearized backward Euler method*, with fixed step size h_t ([40]). These two methods are carefully discussed and used to solve linear and nonlinear parabolic problems in [30] and [40]. In the *linearized backward Euler method* for the semi-discrete problem (5.2) we seek approximations $u_h^n \in V_h$ of $u(\cdot, t_n)$, $n = 0, 1, \cdots, N_t$, satisfying

$$\begin{cases} (\frac{u_h^n - u_h^{n-1}}{h_t}, v) + (a(u_h^{n-1})u_{h_x}^n, v_x) + = (f(u_h^{n-1}), v) \\ (u_h^0, v) = (u_0, v) \end{cases} \tag{5.3}$$

for $\forall v \in V_h$, $n = 1, 2, \cdots, N_t$. In the *linearized Crank-Nicolson method* for (5.2) we seek $u_h^n \in V_h$, $n = 0, 1, \cdots, N_t$ such that

$$\begin{cases} (\frac{u_h^n - u_h^{n-1}}{h_t}, v) + (a(\frac{3}{2}u_h^{n-1} - \frac{1}{2}u_h^{n-2})\frac{u_{h_x}^n + u_{h_x}^{n-1}}{2}, v) = (f(\frac{3}{2}u_h^{n-1} - \frac{1}{2}u_h^{n-2}), v) \\ (u_h^0, v) = (u_0, v) \end{cases} \tag{5.4}$$

for $\forall v \in V_h$, $n = 1, 2, \cdots, N_t$. In our examples we generate u_h^1 from u_h^0 by (5.3) and then revert to (5.4) for the computation of the subsequent steps.

Example 5.1 In this example, we consider the equation

$$\begin{cases} u_t = \epsilon u_{xx} - q(x)uu_x, & 0 < x < 1, \quad t > 0, \\ u(t,0) = u(t,1) = 0, & t > 0, \\ u(0,x) = u_0(x), & 0 < x < 1. \end{cases}$$

Figure 5.1(a) and Figure 5.1(b) show the solution in case for $\epsilon = \frac{1}{100}$, $q(x) = 1$ and $u_0(x) = \sin(\pi x)$. This numerical solution has also been computed by Kang in [32] and we notice that the result here is the same as the one in [32] (Figure 4.3.19.). Figure 5.1(c) and Figure 5.1(d) show the solution for $\epsilon = \frac{1}{60}$, $q(x) = x\sin(3x)$ and $u_0(x) = \sin(2\pi x)$. In each case, we draw a $2D$ graph and a $3D$ graph.

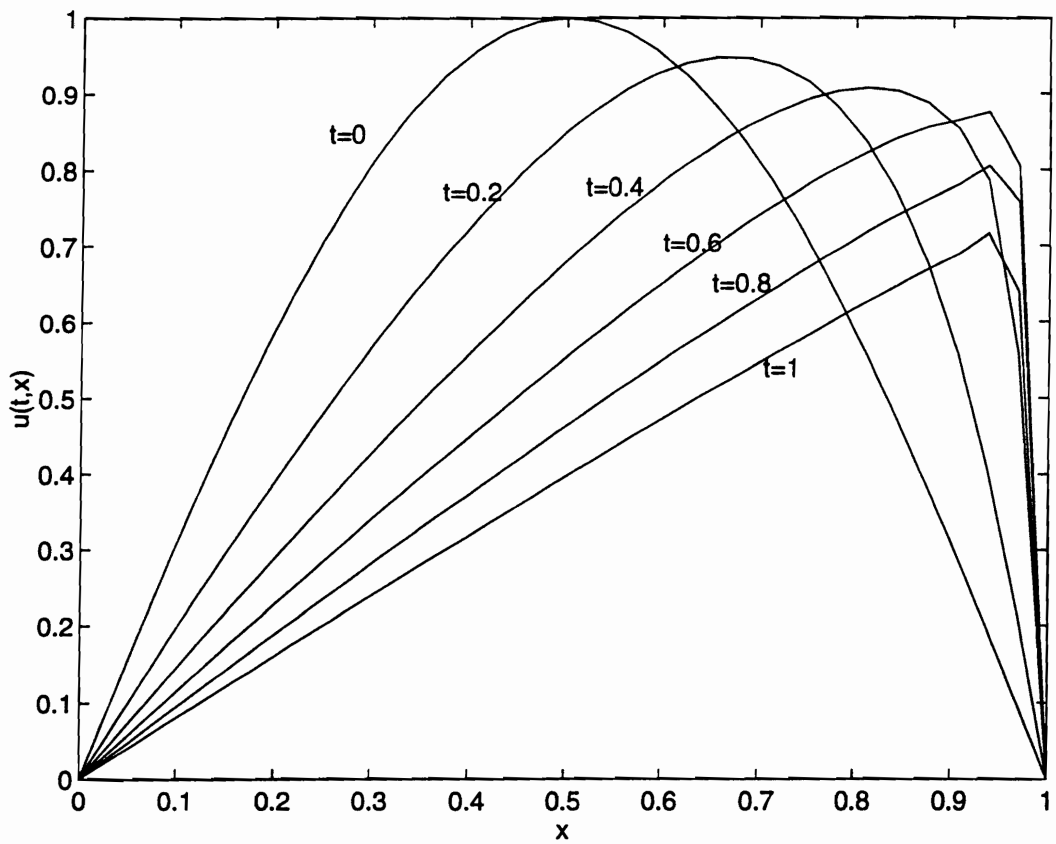

Figure 5.1(a) Solution of the Equation (5.5), $\epsilon = \frac{1}{100}$, $q(x) = 1$

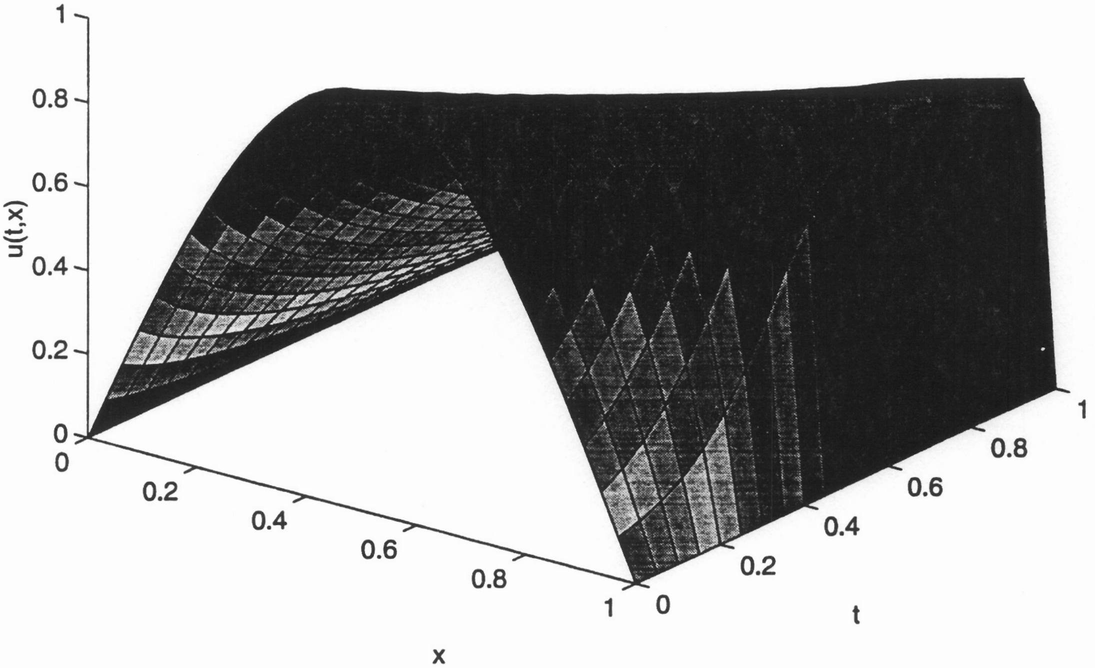

Figure 5.1(b) Solution of the Equation (5.5), $\epsilon = \frac{1}{100}$, $q(x) = 1$

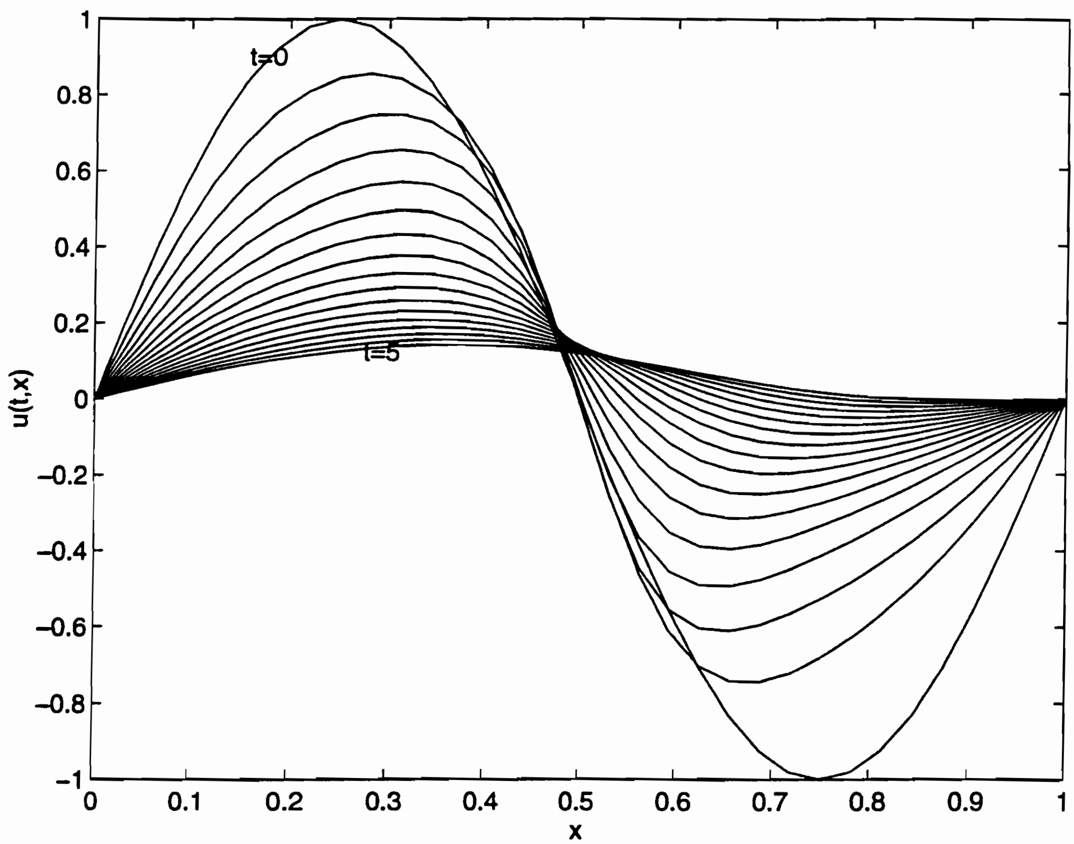

Figure 5.1(c) Solution of the Equation (5.5), $\epsilon = \frac{1}{60}$, $q(x) = x\sin(3x)$

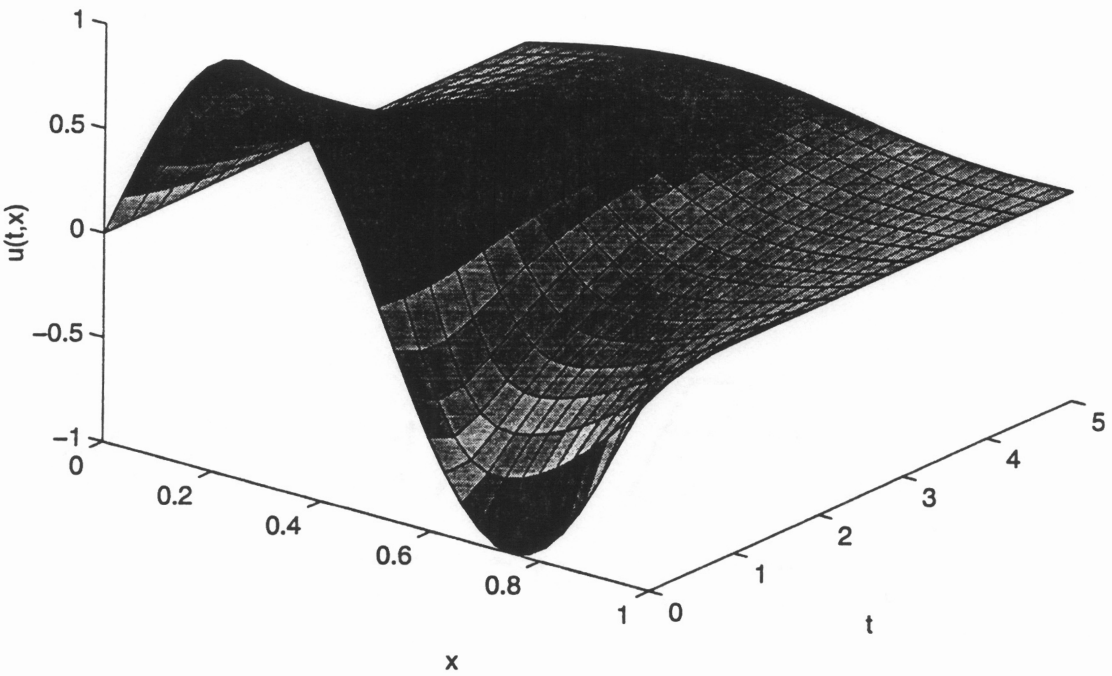

Figure 5.1(d) Solution of the Equation (5.5), $\epsilon = \frac{1}{60}$, $q(x) = x\sin(3x)$

We next consider the following parameter estimation problems:

Given observations $y_{ij} \in R^l \times R^m$, $1 \le i \le m, 1 \le j \le l$, determine the parameter q^* in the admissible space Q that minimizes the functional

$$J(q) = \frac{1}{2} \sum_{i,j} |u(t_i, x_j; q) - y_{ij}|^2,$$

where $u(t;q)$ is the solution to equation (5.1).

For all of the examples considered here, we choose a true parameter and use the finite element method to solve for 'true' solution $u^*(t,x)$. We then add random noise to obtain the noisy observation data $y_{ij}, i = 1, 2, \cdots, m$, $j = 1, 2, \cdots, l$ (or no errors for exact data). In the next four examples, we consider the following state space model discussed in Chapter 3:

$$\begin{cases} u_t = \epsilon u_{xx} - q(x) u u_x, & 0 < x < 1, \quad t > 0, \\ u(t,0) = u(t,1) = 0, & t > 0, \\ u(0,x) = u_0(x), & 0 < x < 1, \end{cases}$$

where $\epsilon = 1/10$ and $u_0(x) = \sin(\pi x)$. The parameter to be estimated is the spatially varying parameter $q(x)$ and the tolarence is 0.001.

Example 5.2 Here $q^*(x) = x \sin(3x)$. The data $y_{ij} = u^*(t_i, x_j)$ is given for $t_i = 0.25, 0.5, 0.75, 1$ and $x_j = 0.25, 0.4, 0.5, 0.6, 0.75, 1$. The iterative scheme is started with an initial estimate $q^0(x) = 0$. The iterative results for $q^k(x)$, $J(q^k)$ and $||u^k - u^*||_{L^2}$ are given in Table 5.2 and Figure 5.2(a). Typical fit to data curves are presented in Figure 5.2(b). We use $N_x = 9$, $N_t = 10$ and $N_q = 4$ in this example. Since we use exact data, we get pretty accurate results.

Table 5.2 Estimate $q(x)$, Exact Data

$q^*(x) = x\sin(3x)$				
k	$J(q^k)$	$\|u^k - u^*\|$	q_1^k	q_2^k
0	1.8712913e-02	4.6062553e-02	0.0000000e+00	0.0000000e+00
1	2.2764263e-03	2.5770944e-02	-9.1502666e-01	3.5941530e-01
2	1.0895802e-03	1.1270525e-02	-1.6995680e-01	2.1389020e-01
3	1.2286964e-06	3.5277521e-04	-4.0855091e-02	1.8009448e-01
4	1.1096643e-09	2.6446144e-05	5.0378634e-03	1.6926737e-01
5	1.2393077e-11	2.5630600e-06	-4.9815535e-04	1.7051592e-01
6	7.0249148e-14	2.1779503e-07	4.2385976e-05	1.7040128e-01
k	q_3^k	q_4^k	q_5^k	
0	0.0000000e+00	0.0000000e+00	0.0000000e+00	
1	6.2534803e-01	6.2063491e-02	2.9961586e+00	
2	3.9903206e-01	6.3178694e-01	5.4593767e-01	
3	4.9162250e-01	5.9223543e-01	1.2845363e-01	
4	4.9957664e-01	5.8367194e-01	1.3957243e-01	
5	4.9867194e-01	5.8353864e-01	1.4122597e-01	
6	4.9875312e-01	5.8355549e-01	1.4111616e-01	

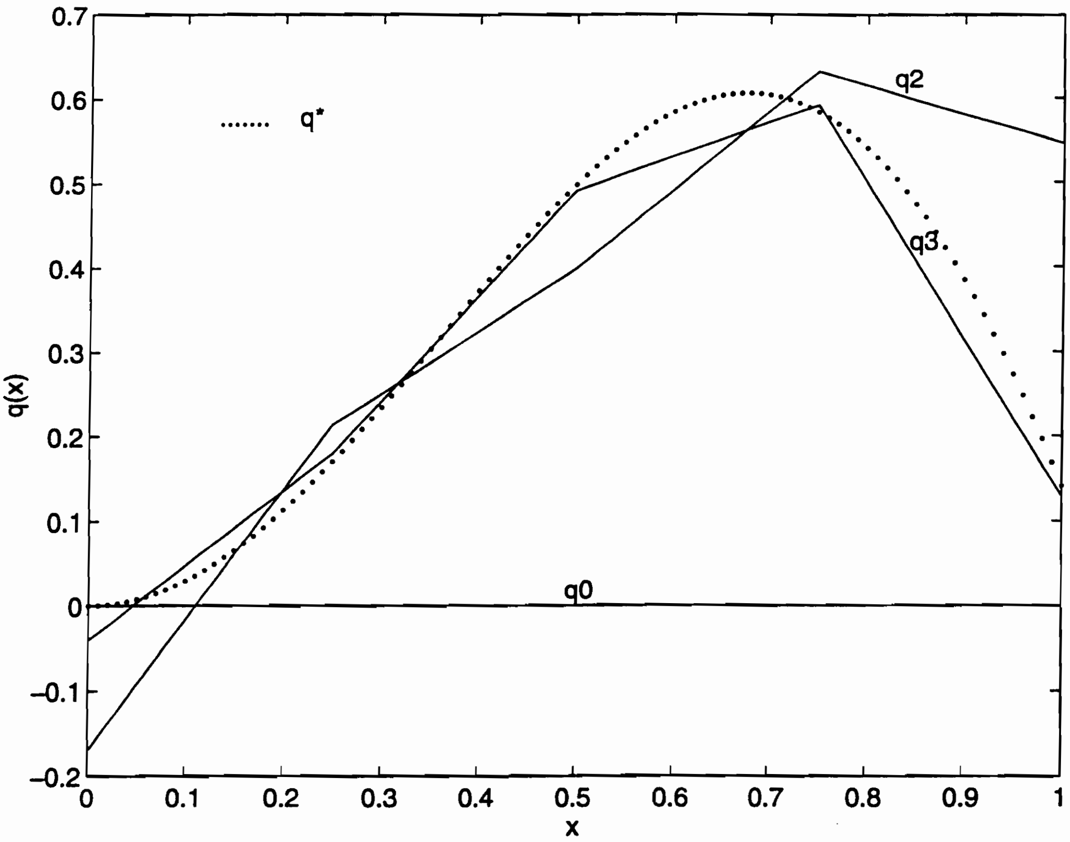

Figure 5.2(a) $q^*(x) = x\sin(3x)$, Exact Data

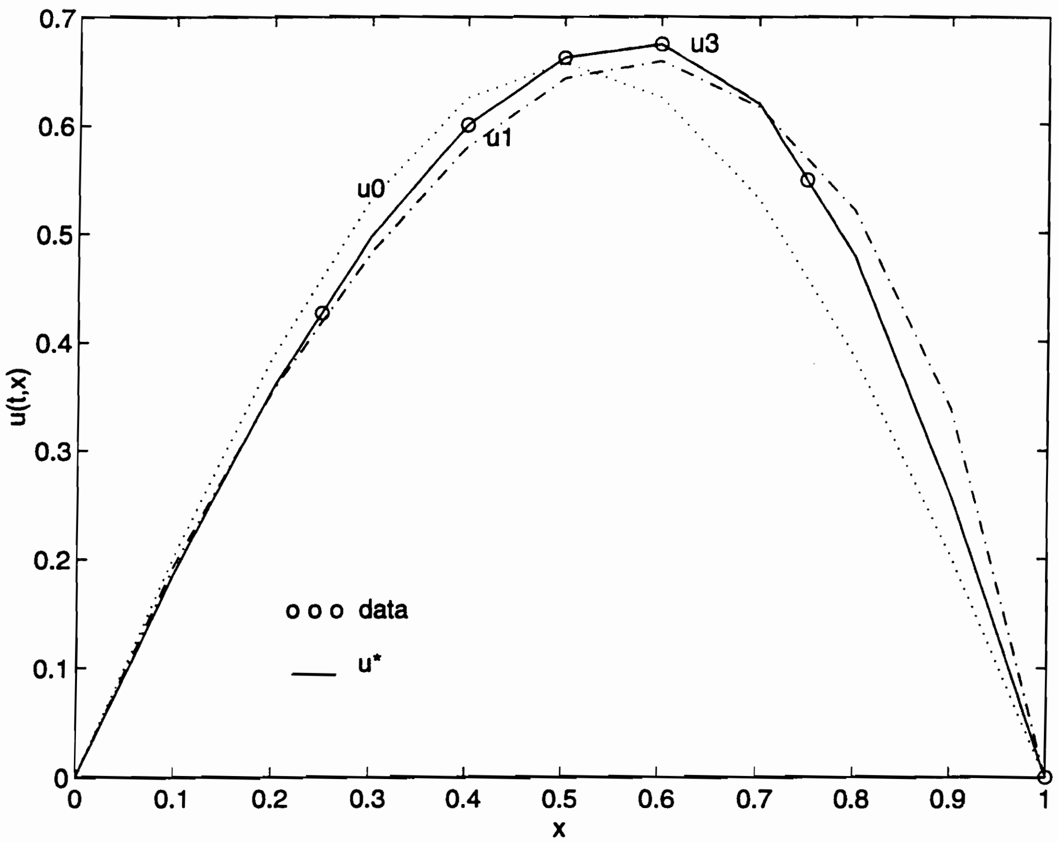

Figure 5.2(b) Fit to Data Curve at $t = 0.5$, Exact Data

Example 5.3 We choose $q^*(x) = e^{-x/2}$ and generate exact data $y_{ij} = u^*(t_i, x_j)$ for $t_i = 0.2, 0.4, 0.6$ and $x_j = 1/10, 2/10, \cdots, 9/10$. We start the iterative estimate scheme with an initial estimate $q^0(x) = 0$. Again, since we use exact data we expect accurate results. The iterative results and typical 2-dimensional fit to data curves are presented in Table 5.3, Figure 5.3(a) Figure 5.3(b) and Figure 5.3(c). Figure 5.3(d) shows the data and the 3-dimensional fit to data curves for the first 3 steps. Since there are a lot of data covered by the surface graph, in Figure 5.3(e) we turn $u^6(t,x)$ (with data) around to see how the data under the surface fits. In this example we use $N_x = 9$, $N_t = 10$ and $N_q = 4$.

Table 5.3 Estimate $q(x)$, Exact Data

$q^*(x) = e^{-x/2}$				
k	$J(q^k)$	$\|u^k - u^*\|$	q_1^k	q_2^k
0	1.1330576e-01	9.0298462e-02	0.0000000e+00	0.0000000e+00
1	4.7498209e-03	1.7681642e-02	-1.1414594e+00	1.3163610e+00
2	2.6762964e-04	4.1283736e-03	1.0771134e+00	8.3685265e-01
3	3.1411724e-07	1.4879095e-04	9.9529337e-01	8.8581729e-01
4	7.6939258e-09	2.4862086e-05	1.0007309e+00	8.8206590e-01
5	1.2812167e-10	3.2522006e-06	9.9990065e-01	8.8254883e-01
6	1.9723594e-12	4.0697630e-07	1.0000134e+00	8.8249072e-01
k	q_3^k	q_4^k	q_5^k	
0	0.0000000e+00	0.0000000e+00	0.0000000e+00	
1	5.3869092e-01	8.2116370e-02	2.8775978e+00	
2	8.1969002e-01	7.0492437e-01	7.7781557e-01	
3	7.7850505e-01	6.8866191e-01	6.0760792e-01	
4	7.7876278e-01	6.8742413e-01	6.0621997e-01	
5	7.7881317e-01	6.8727354e-01	6.0656510e-01	
6	7.7879866e-01	6.8729130e-01	6.0652626e-01	

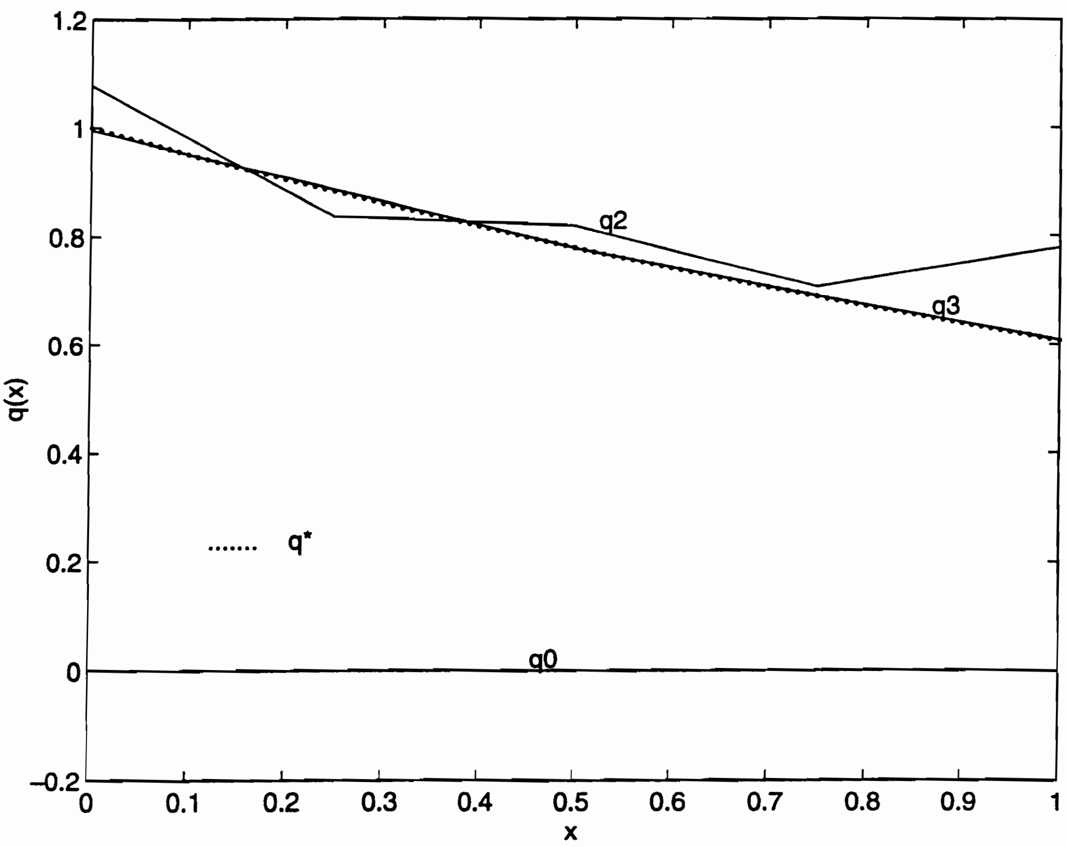

Figure 5.3(a) $q^*(x) = e^{-x/2}$, Exact Data

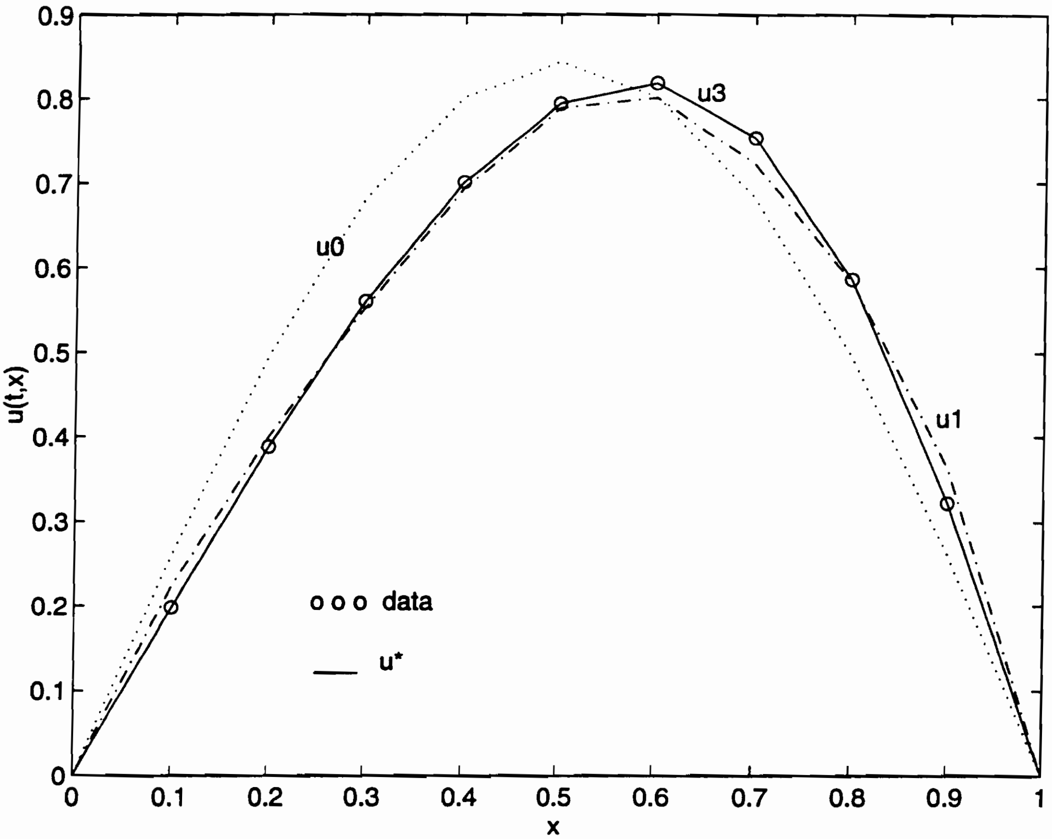

Figure 5.3(b) Fit to Data Curve at $t = 0.2$, Exact Data

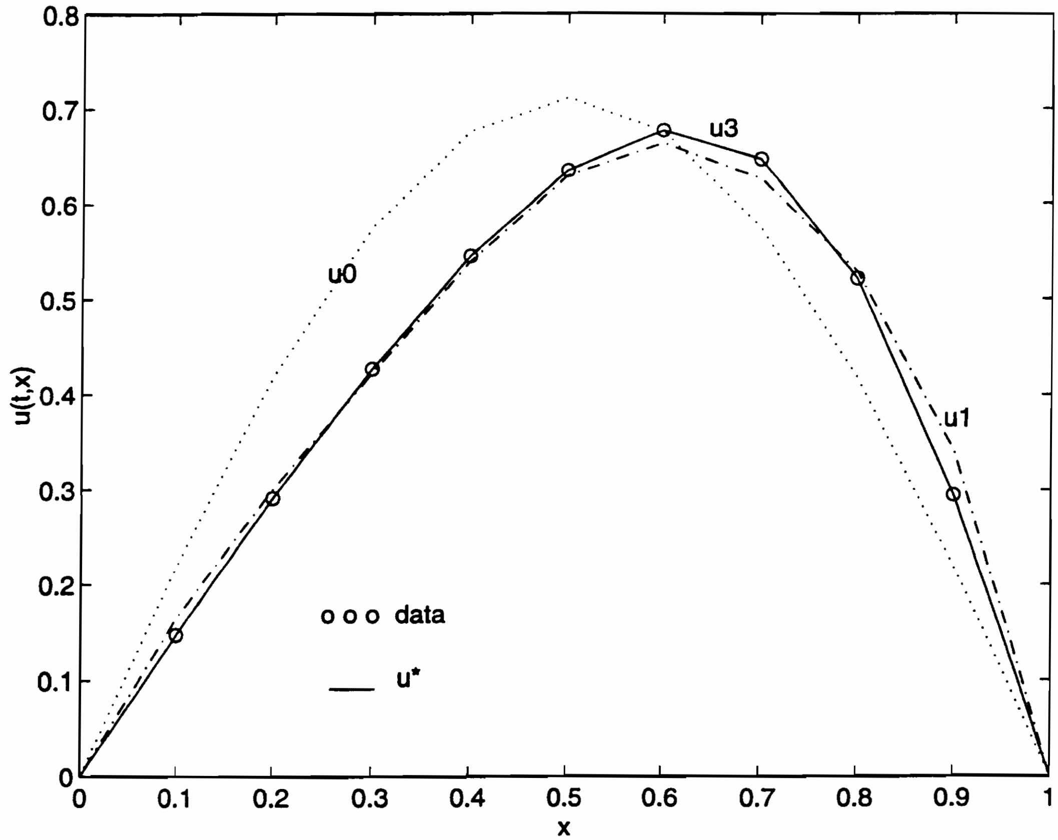

Figure 5.3(c) Fit to Data Curve at $t = 0.4$, Exact Data

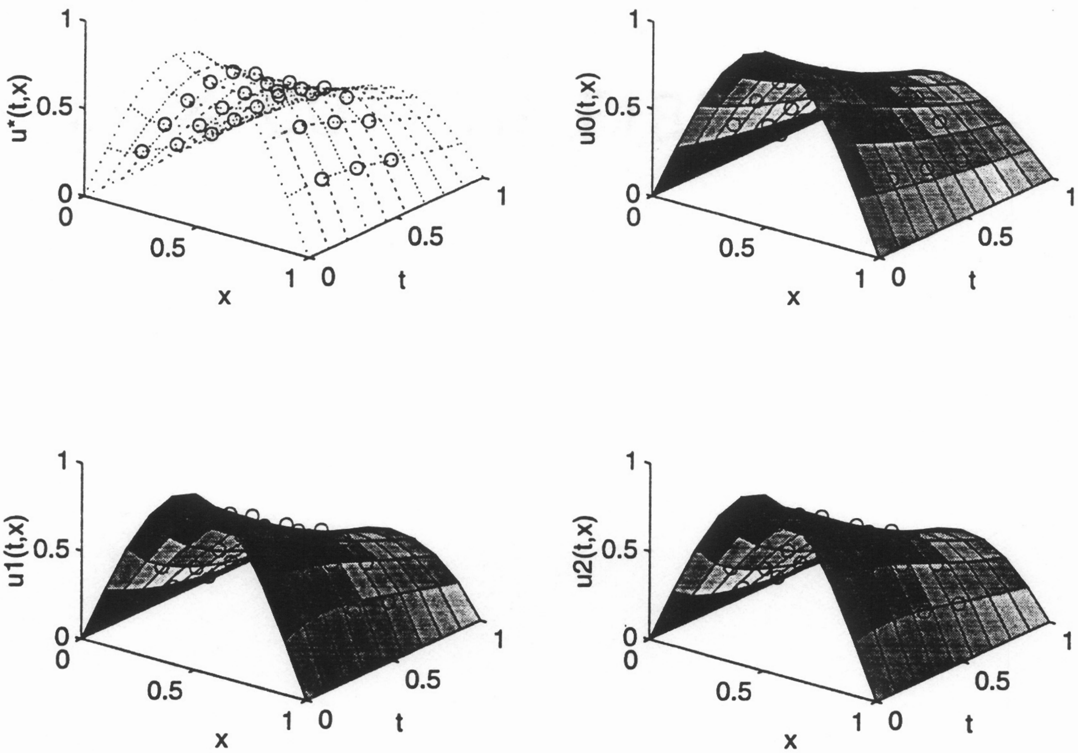

Figure 5.3(d) Fit to Data, (o o o – data)

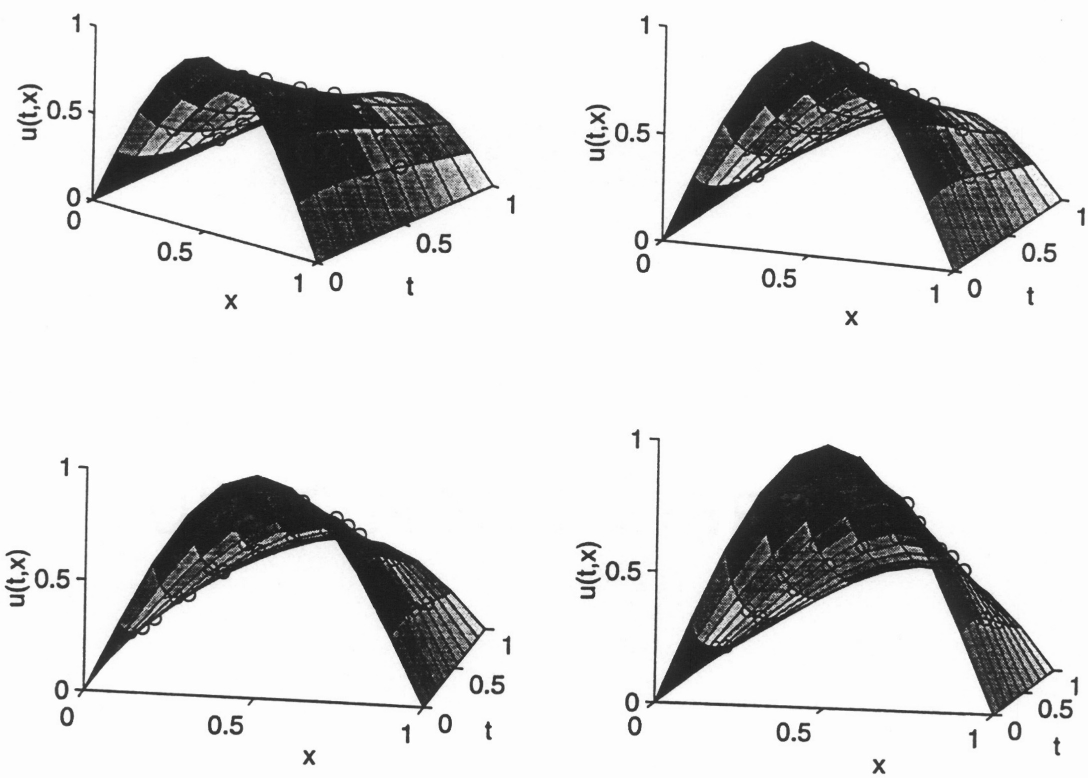

Figure 5.3(e) $u^6(t,x)$ with Data

Example 5.4 In this example, we use $N_x = 31$ to get $u^*(t,x)$ and generate data $y_{ij} = u^*(t_i, x_j) + \delta_{ij}$ for $t_i = 0.25, 0.5, 0.75, 1$, $x_j = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.75, 0.9$, where δ_{ij} are random numbers with uniform distribution that fall in the range $[-0.005, 0.005]$. Specifically, we have

$$\delta =$$

$$\begin{bmatrix} 0.0026 & 0.0048 & -0.0043 & -0.0006 & -0.0023 & 0.0040 & 0.0000 & -0.0001 \\ 0.0049 & 0.0022 & 0.0013 & 0.0027 & -0.0014 & 0.0041 & 0.0002 & -0.0023 \\ -0.0013 & 0.0025 & 0.0038 & -0.0002 & -0.0033 & -0.0044 & -0.0018 & -0.0041 \\ -0.0025 & 0.0015 & -0.0023 & -0.0026 & -0.0001 & 0.0040 & 0.0049 & 0.0045 \end{bmatrix}.$$

We choose $q^*(x) = e^{-x/2}$, $q^0(x) = 0$ and when we solve the equation, we use $N_x = 15$. Since we choose different N_x for $u^*(t,x)$ and $u^k(t,x)$, and add noise to the observation data, we expect less accurate results. The computational findings for q^k and $J(q^k)$ are given in Table 5.4 with corresponding $q^k(x)$ presented in Figure 5.4(a). A typical fit to data curve is shown in Figure 5.4(b). ($N_q = 4$ and $N_t = 10$.)

Table 5.4 Estimate $q(x)$, Noisy Data

$q^*(x) = e^{-x/2}$				
k	$J(q^k)$	q_1^k	q_2^k	q_3^k
0	1.2440204e-01	0.0000000e+00	0.0000000e+00	0.0000000e+00
1	2.4117856e-03	1.3660263e+00	1.6486209e-01	1.6370505e+00
2	3.0208156e-04	3.9453418e-01	1.0050367e+00	7.2752623e-01
3	1.1906665e-04	3.7636845e-01	9.7706711e-01	7.9361691e-01
4	1.1845728e-04	3.6727683e-01	9.8335473e-01	7.8651285e-01
5	1.1845772e-04	3.6806724e-01	9.8274927e-01	7.8700973e-01
6	1.1845782e-04	3.6805193e-01	9.8279161e-01	7.8699552e-01

k	q_4^k	q_5^k	q_6^k
0	0.0000000e+00	0.0000000e+00	0.0000000e+00
1	-4.7859232e-01	1.2422202e+00	3.3994208e-01
2	8.2225610e-01	7.3461032e-01	4.0614400e-01
3	8.4835639e-01	6.9903955e-01	4.8555917e-01
4	8.5241722e-01	6.9886332e-01	4.8388759e-01
5	8.5205323e-01	6.9887111e-01	4.8406660e-01
6	8.5205183e-01	6.9887892e-01	4.8403862e-01

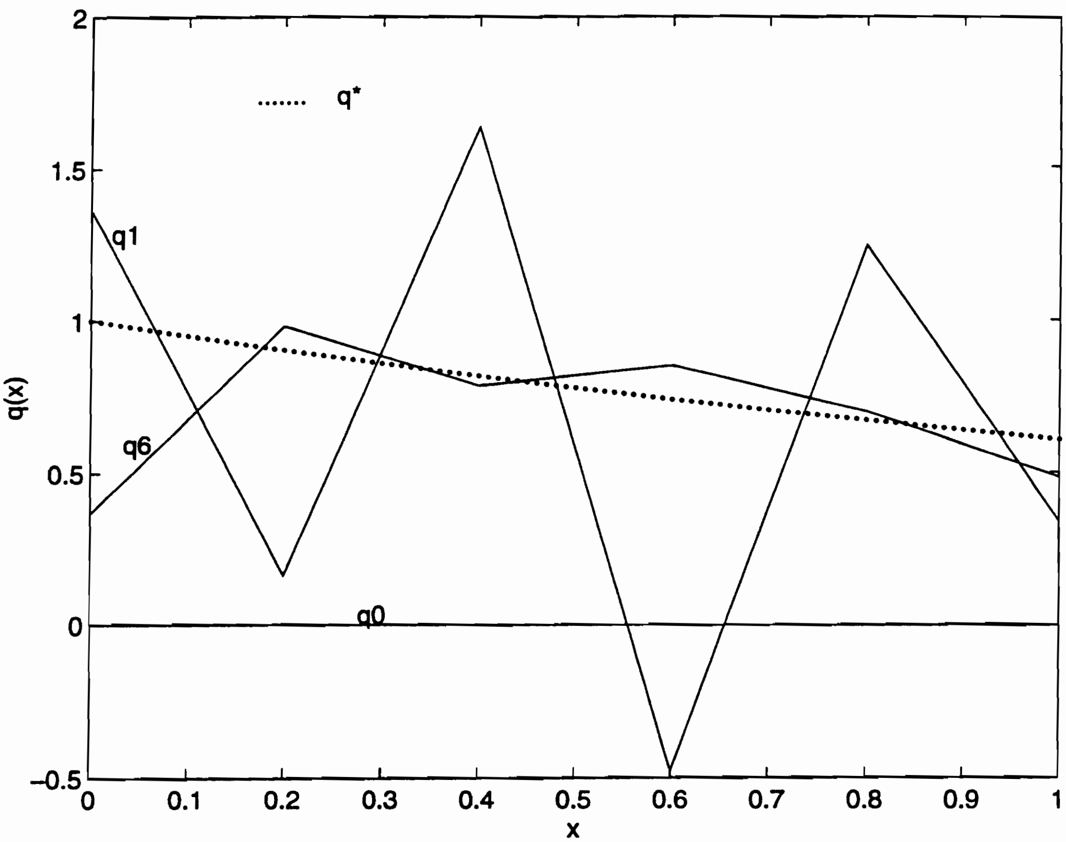

Figure 5.4(a) $q^*(x) = e^{-x/2}$, Noisy Data

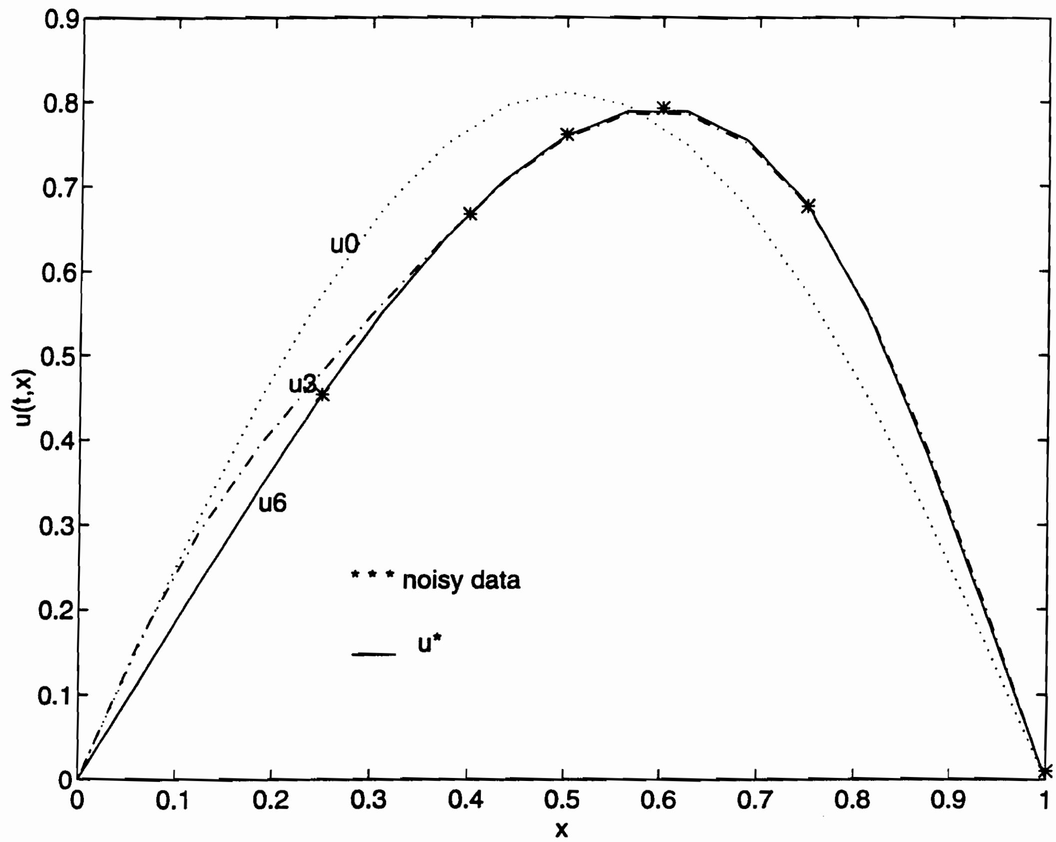

Figure 5.4(b) Fit to Data Curve at $t = 0.5$, Noisy Data

Example 5.5 This example is the same as the last one except we choose $q^*(x) = x\sin(3x)$ and $y_{ij} = u^*(t_i, x_j) + \hat{\delta_{ij}}$ with $\hat{\delta_{ij}} = \frac{1}{10}\delta_{ij}$, where δ_{ij} is the one in Example 5.4. Again, since we choose different N_x for $u^*(t,x)$ and $u^k(t,x)$, and add noise to the observation data. Since the noise is smaller than in Example 5.4, we expect to see better parameter estimations. The computational findings for q^k are given in Table 5.5 with corresponding $q^k(x)$ presented in Figure 5.5(a). A typical fit to data curve is shown in Figure 5.5(b). ($N_q = 4$ and $N_t = 10$.)

Table 5.5 Estimate $q(x)$, Noisy Data

$q^*(x) = x\sin(3x)$				
k	$J(q^k)$	q_1^k	q_2^k	q_3^k
0	2.8878106e-02	0.0000000e+00	0.0000000e+00	0.0000000e+00
1	1.5488178e-03	2.7066786e-01	-2.9726219e-02	6.4296945e-01
2	9.2831812e-05	-1.6274533e-01	1.5052944e-01	3.1909187e-01
3	6.5148763e-06	-8.8031936e-02	1.2146274e-01	3.6544739e-01
4	6.3061199e-06	-9.2101078e-02	1.2309801e-01	3.6286997e-01
5	6.3058897e-06	-9.1890142e-02	1.2300720e-01	3.6302443e-01
6	6.3058588e-06	-9.1900312e-02	1.2301205e-01	3.6301526e-01
k	q_4^k	q_5^k	q_6^k	
0	0.0000000e+00	0.0000000e+00	0.0000000e+00	
1	6.3607771e-02	9.3402033e-01	-9.5367131e-02	
2	7.0583765e-01	5.2690846e-01	1.4345777e-01	
3	6.4829558e-01	5.3825280e-01	1.3405970e-01	
4	6.5231982e-01	5.3673258e-01	1.3635384e-01	
5	6.5203429e-01	5.3685475e-01	1.3614861e-01	
6	6.5205493e-01	5.3684543e-01	1.3616524e-01	

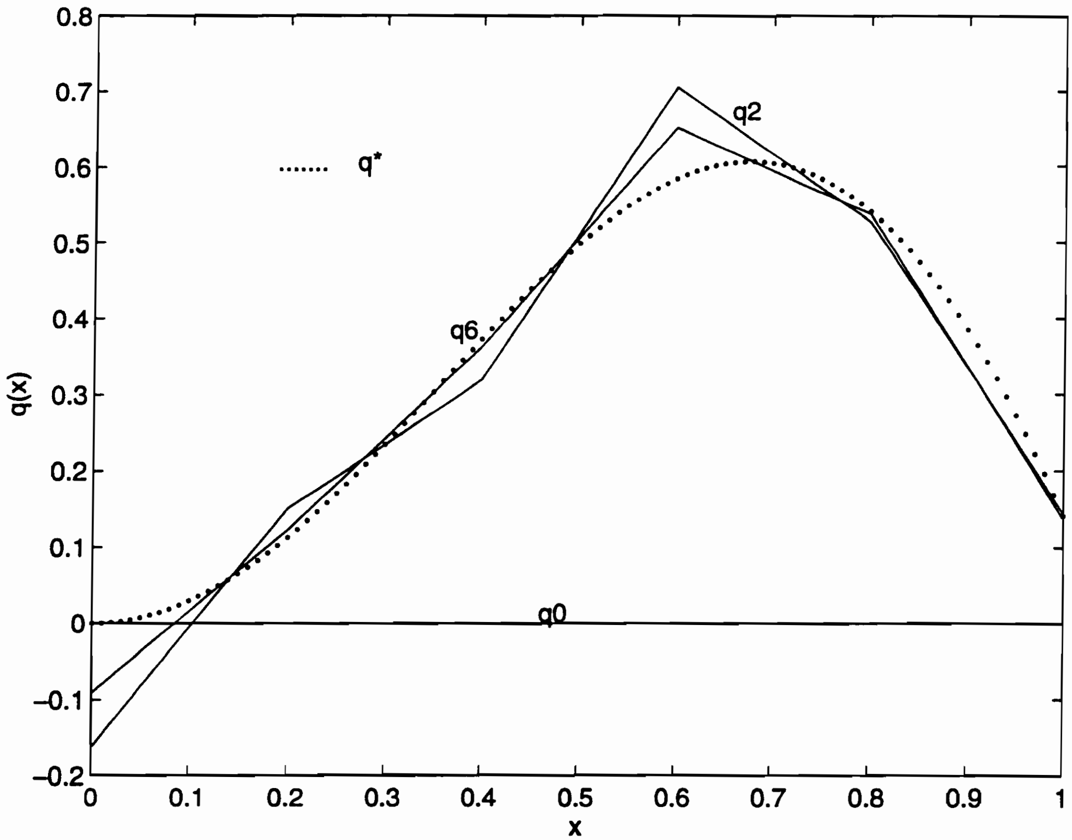

Figure 5.5(a) $q^*(x) = x \sin(3x)$, Noisy Data

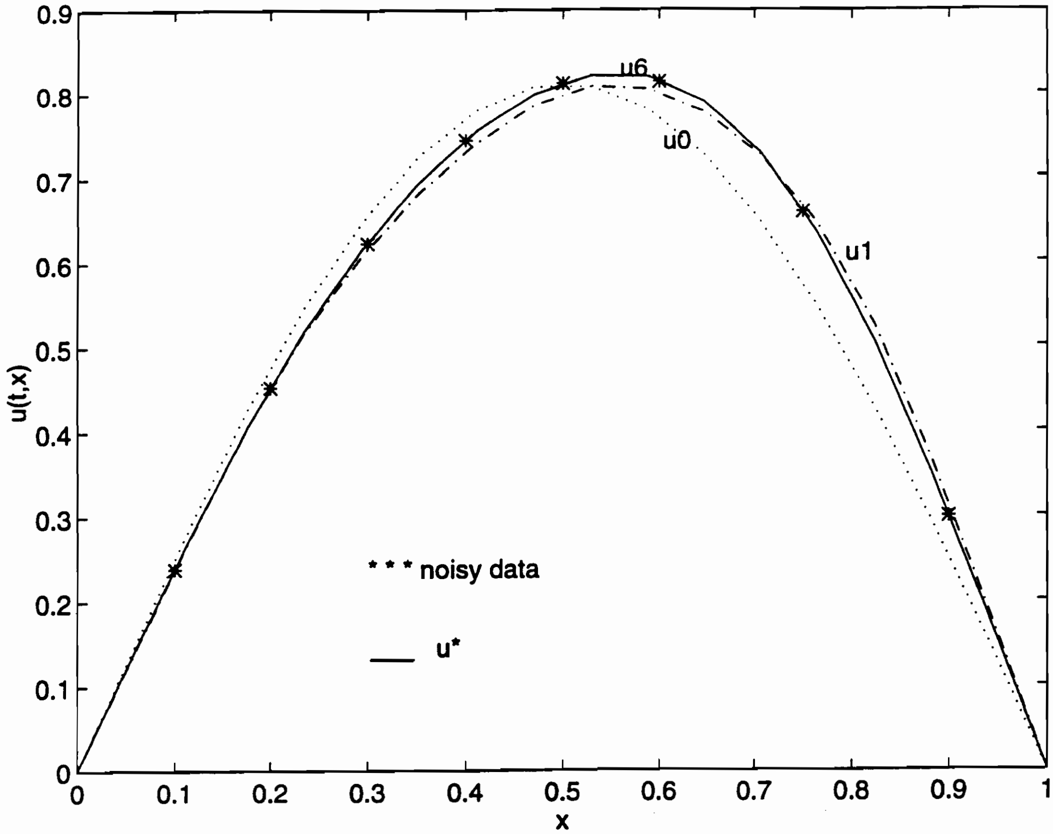

Figure 5.5(b) Fit to Data Curve at $t = 0.25$, Noisy Data

Example 5.6 In this example, we increase the range of the noise to be $[-0.05, 0.05]$. We choose $q^* = x\sin(3x)$ and generate data $y_{ij} = u^*(t_i, x_j) + \delta_{ij}$ for $t_i = 0.2, 0.4, 0.6$ and $x_j = 1/10, 1/20, \cdots, 9/10$ where δ_{ij} are random noise. We start the iterative estimate scheme with initial estimate $q^0(x) = 0$. Since the noise added to the observation data is large in this example, we expect less accurate results. The iterative results and 2-dimensional fit to data curves are presented in Table 5.6, Figure 5.6(a), Figure 5.6(b) and Figure 5.6(c). ($N_x = 9$, $N_q = 4$ and $N_t = 10$.)

Table 5.6 Estimate $q(x)$, Noisy Data

$q^*(x) = x\sin(3x)$				
k	$J(q^k)$	q_1^k	q_2^k	q_3^k
0	5.1567995e-02	0.0000000e+00	0.0000000e+00	0.0000000e+00
1	2.9961498e-02	1.0232796e+00	-3.1943739e-01	2.3895762e+00
2	1.1812766e-02	2.5700674e+00	-7.7042121e-01	1.8359208e+00
3	8.9351825e-03	2.1263864e+00	-5.9692588e-01	1.8031708e+00
4	8.9265192e-03	2.2147449e+00	-6.2714665e-01	1.8412713e+00
5	8.9266759e-03	2.1931268e+00	-6.1938236e-01	1.8283935e+00
6	8.9265966e-03	2.1974871e+00	-6.2095068e-01	1.8312534e+00
k	q_4^k	q_5^k		
0	0.0000000e+00	0.0000000e+00		
1	-3.5344787e-01	3.3322282e+00		
2	6.8538232e-01	5.7157491e-01		
3	5.3262591e-01	6.0790314e-01		
4	5.6394563e-01	5.0665733e-01		
5	5.6051397e-01	5.1449227e-01		
6	5.6147943e-01	5.1184326e-01		

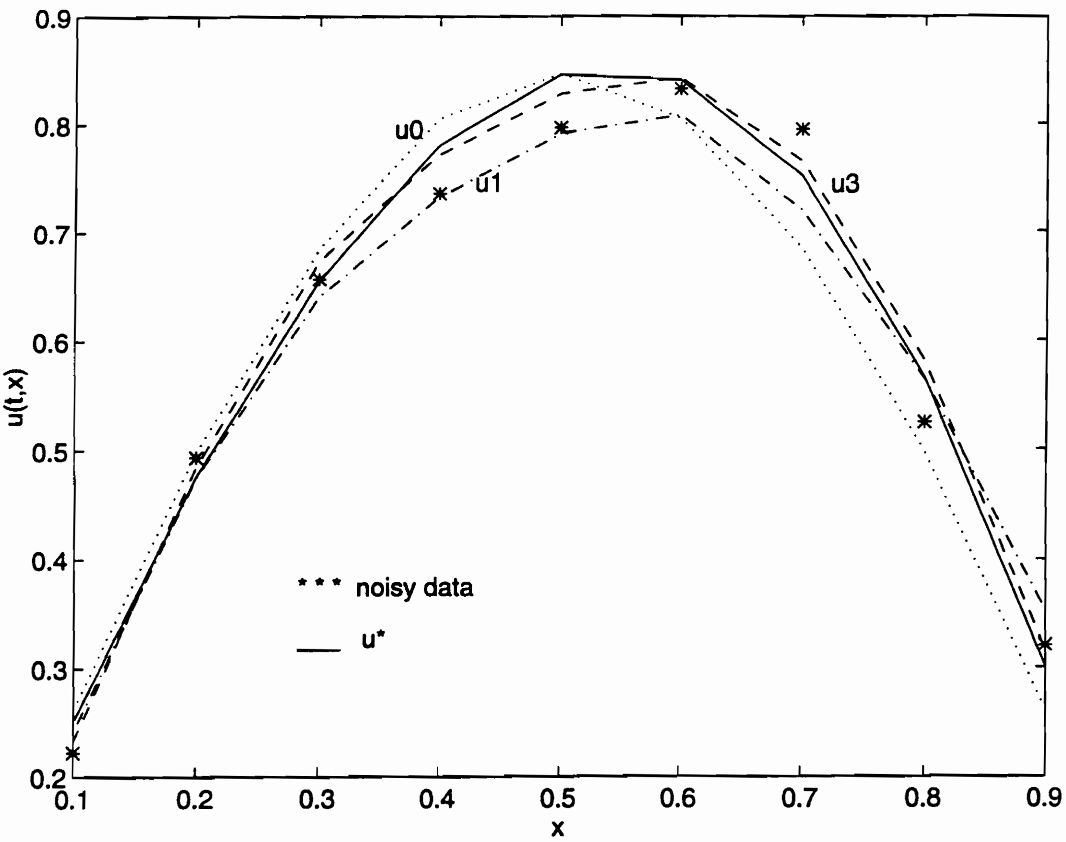

Figure 5.6(a) Fit to Data Curve at $t = 0.2$, Noisy Data

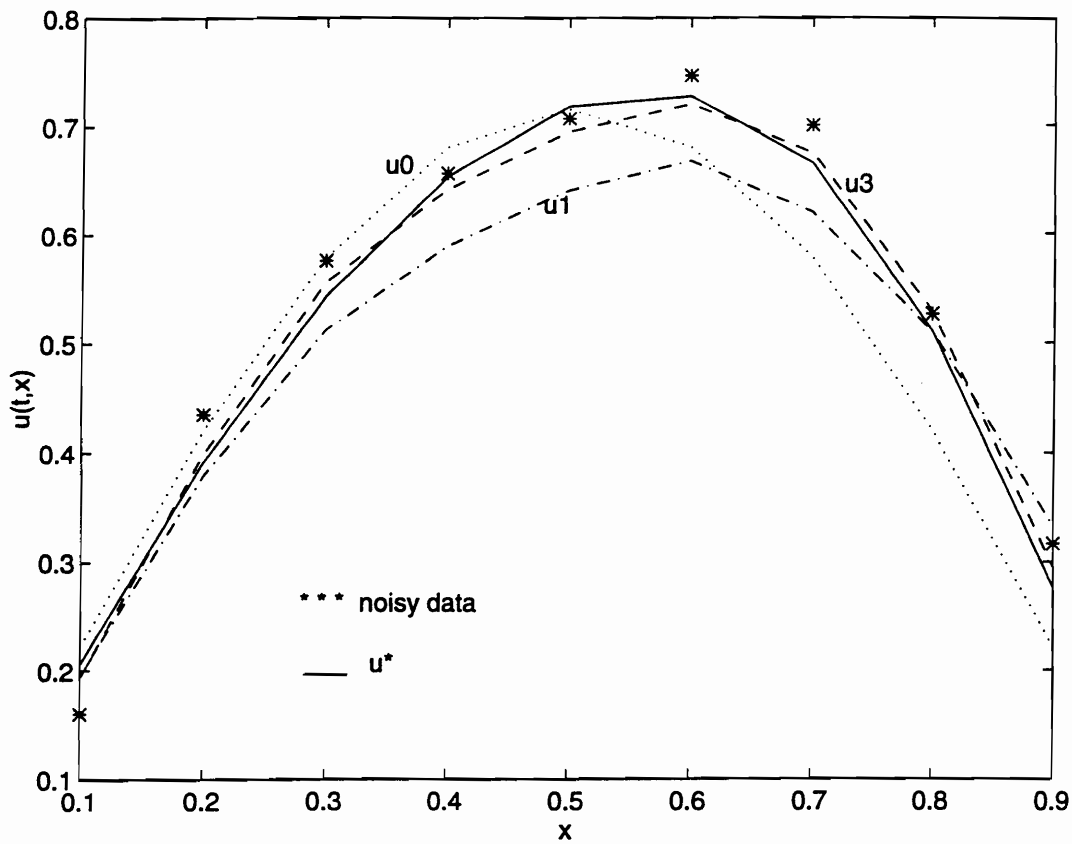

Figure 5.6(b) Fit to Data Curve at $t = 0.4$, Noisy Data

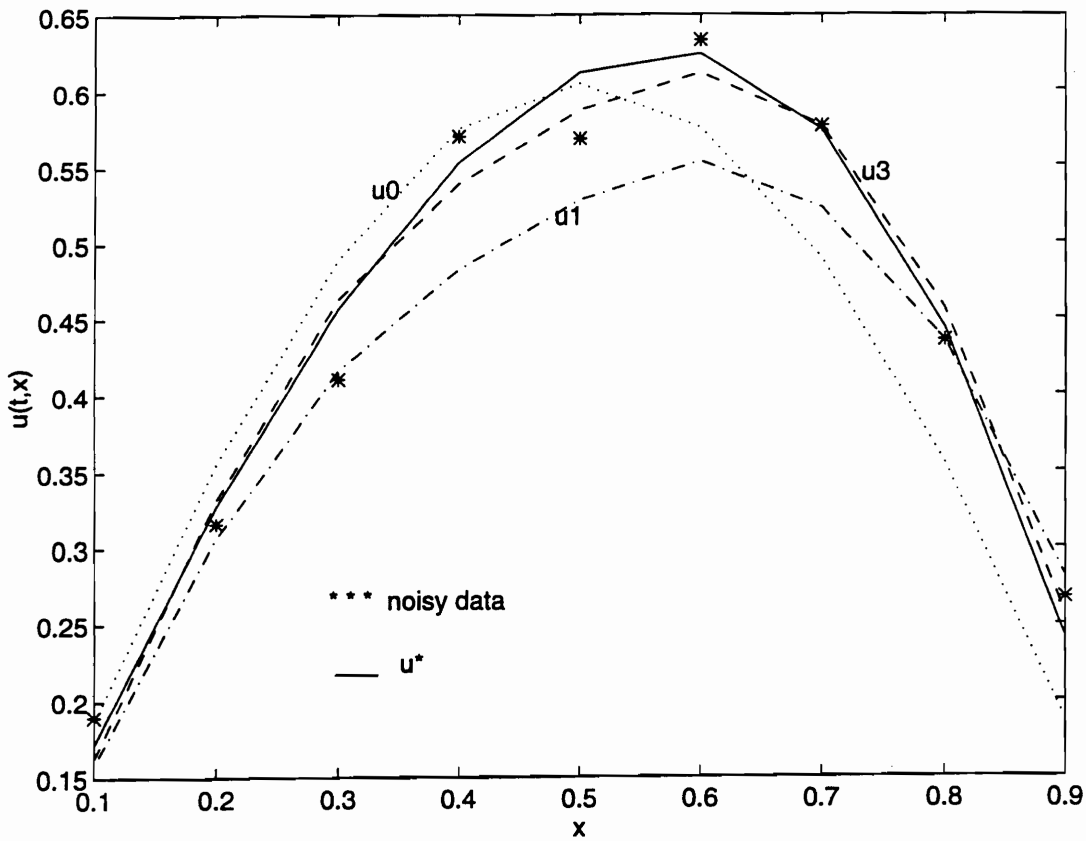

Figure 5.6(c) Fit to Data Curve at $t = 0.6$, Noisy Data

The last three examples are for parameter estimation problems governed by the following equation:

$$\begin{cases} u_t - quu_x - ru^2u_x = \epsilon u_{xx}, & t \in [0,1],\ x \in [0,1] \\ u(t,0) = u(t,1) = 0 \\ u(0,x) = u_0(x) \end{cases}$$

where $\epsilon > 0$. We do not have any theoretical results for this type of equation. Here we only present numerical findings. In all these examples, we generate data $y_{ij} = u^*(t_i, x_j) + \delta_{ij}$ for $t_i = 0.25, 0.5, 0.75, 1$ and $x_j = 0.25, 0.4, 0.5, 0.6, 0.75, 1$, where $\{\delta_{ij}\}$ represent observation errors.

Example 5.7 In this example, we set $r = 1$ and $q = 1$, estimate ϵ. We choose $\epsilon^* = \frac{1}{60}$ and start the iterative scheme with an initial estimate $\epsilon^0 = 0.1$. The computational findings for exact data are given in Table 5.7(a) with corresponding fit to data graph presented in Figure 5.7(a). We then add random noise to the data. In this case, $y_{ij} = u^*(t_i, x_j) + \delta_{ij}$ where δ_{ij} are random numbers with uniform distribution that fall in range $[-0.01, 0.01]$. For exact data we expect accurate results; for noisy data, we expect some error. The numerical findings for noisy data are given in Table 5.7(b) with corresponding fit to data graph presented in Figure 5.7(a). ($N_x = 31$, $N_t = 10$.)

Table 5.7(a) Estimate ϵ, Exact Data

$\epsilon^* = 1/60$			
k	ϵ^k	$J(\epsilon^k)$	$\|u^k - u^*\|$
0	1.0000000e-01	4.7974567e-02	1.6563412e-01
1	4.7854675e-02	2.5962191e-03	8.3804947e-02
2	2.6791189e-02	2.3161879e-04	3.3208839e-02
3	1.6754350e-02	1.7650008e-08	3.2928204e-04
4	1.6665608e-02	2.5715208e-12	3.9798413e-06

Table 5.7(b) Estimate ϵ, Noisy Data

$\epsilon^* = 1/60$			
k	ϵ^k	$J(\epsilon^k)$	$\|u^k - u^*\|$
0	1.0000000e-01	4.7816859e-02	1.6563412e-01
1	4.8101391e-02	2.8012559e-03	8.4304765e-02
2	2.7704784e-02	6.0298699e-04	3.5827992e-02
3	1.8741741e-02	4.1417794e-04	7.5566486e-03
4	1.8366698e-02	4.1393780e-04	6.2256280e-03
5	1.8356829e-02	4.1393984e-04	6.1904070e-03

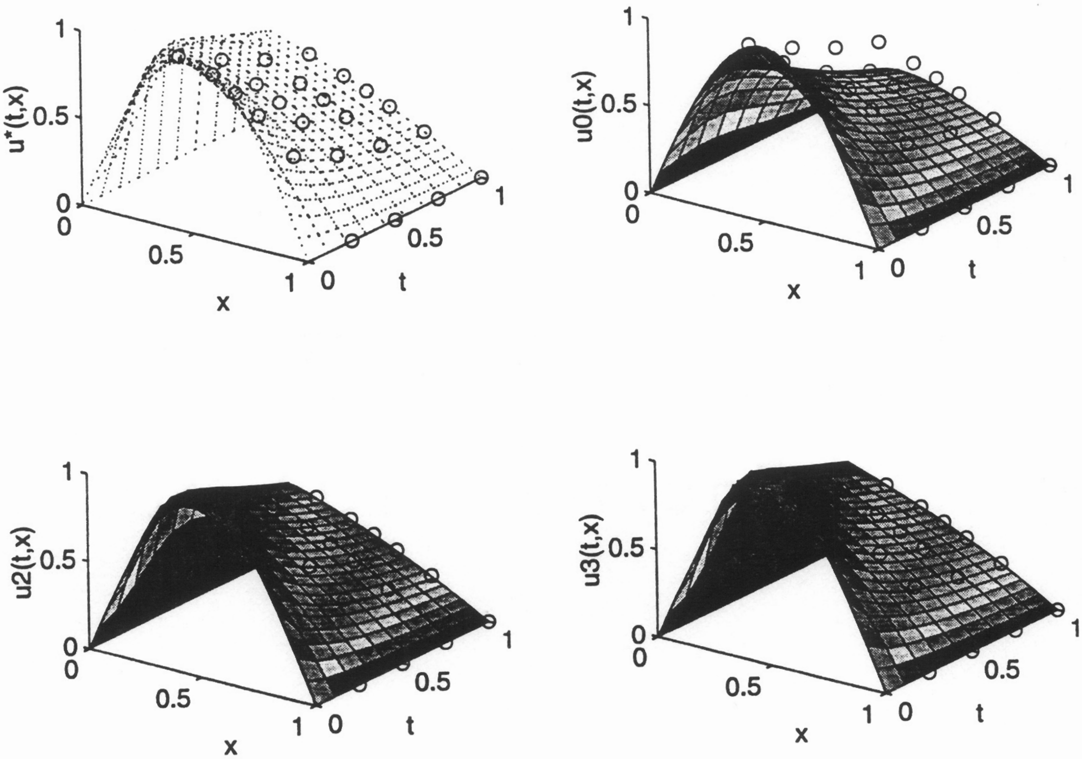

Figure 5.7(a) Fit to Data, Exact Data, (o o o – data)

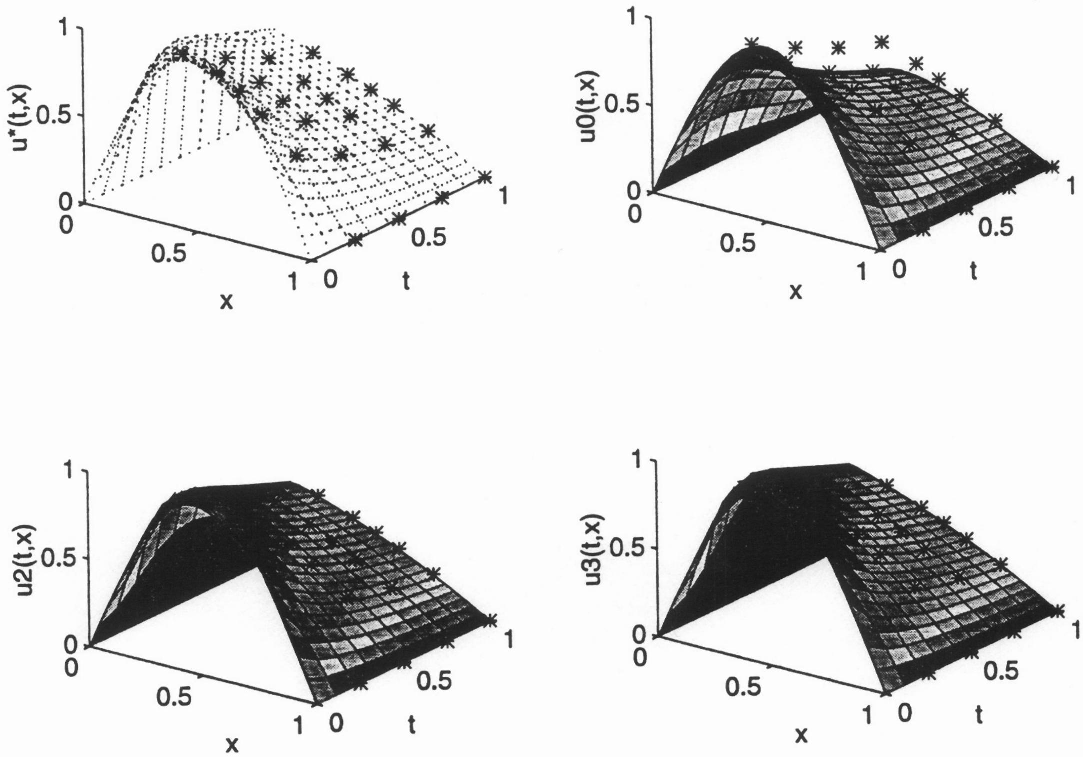

Figure 5.7(b) Fit to Data, Noisy Data, ($* * *$ – noisy data)

Example 5.8 In this example, we set $\epsilon = 1/10$ and $q = 1$, estimate r. We choose $r^* = 2$ and start the iterative scheme with an initial estimate $r^0 = 0.5$. Again, since we use exact data we expect accurate results. The computational findings for r^k are given in Table 5.8 with corresponding fit to data graph presented in Figure 5.8. ($N_x = 9$, $N_t = 10$.)

Table 5.8 Estimate r, Exact Data

$r^* = 2$			
k	r^k	$J(r^k)$	$\|u^k - u^*\|$
0	5.0000000e-01	4.6216406e-02	6.4954838e-02
1	2.0109018e+00	1.5740903e-06	4.3779437e-04
2	1.9960245e+00	2.1018210e-07	1.5979406e-04
3	2.0014358e+00	2.7376383e-08	5.7693945e-05
4	1.9994796e+00	3.5982257e-09	2.0913252e-05
5	2.0001884e+00	4.7139010e-10	7.5699165e-06
6	1.9999318e+00	6.1828294e-11	2.7414857e-06
7	2.0000247e+00	8.1060212e-12	9.9265722e-07

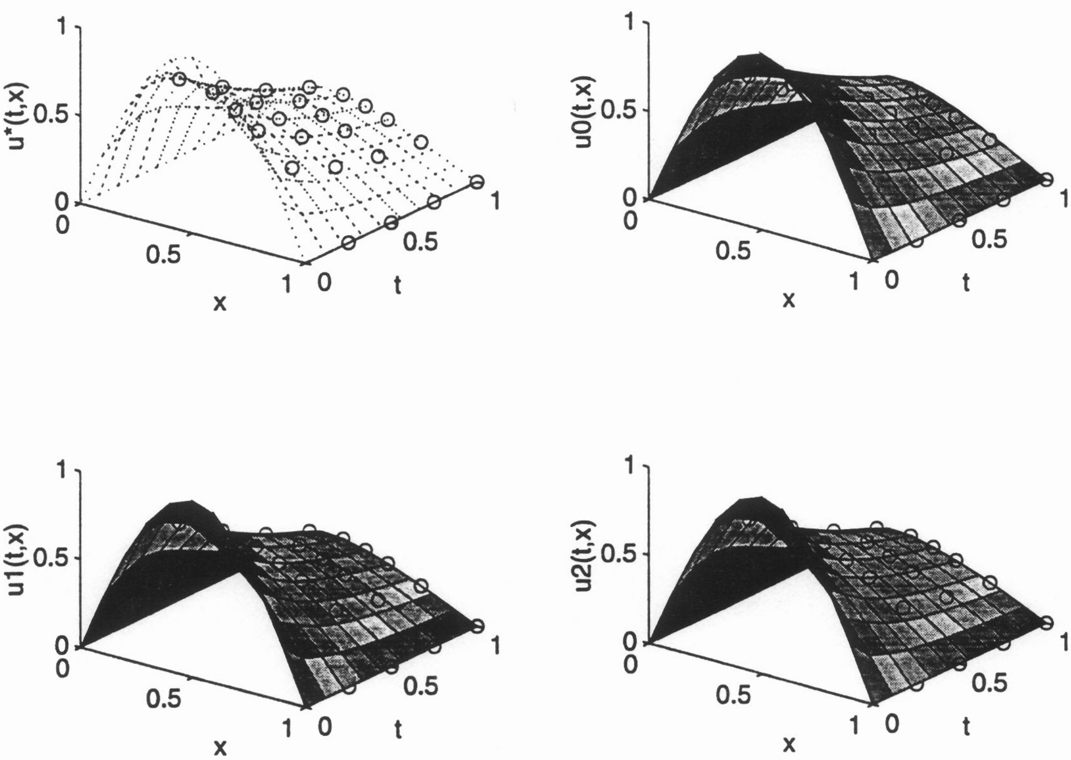

Figure 5.8 Fit to Data, (o o o – data)

Example 5.9 In this example, we set $\epsilon = 1/10$ and $r = 1$, estimate q. We choose $q^* = 5$ and start the iterative scheme with an initial estimate $q^0 = 0$. Again, since we use exact data we expect accurate results. The computational findings for r^k are given in Table 5.9 with corresponding fit to data graph presented in Figure 5.9. ($N_x = 9$, $N_t = 10$.)

Table 5.9 Estimate q, Exact Data

$q^* = 5$			
k	q^k	$J(q^k)$	$\|\|u^k - u^*\|\|$
0	0.0000000e+00	8.3479196e-01	2.6030279e-01
1	1.8637211e+00	1.7727140e-01	1.3697375e-01
2	4.0335507e+00	7.9322887e-03	3.6058255e-02
3	4.9720370e+00	5.0518150e-06	9.7488322e-04
4	5.0023797e+00	3.6281802e-08	8.2785252e-05
5	4.9997889e+00	2.8573531e-10	7.3454024e-06
6	5.0000187e+00	2.2329548e-12	6.4935210e-07
7	4.9999983e+00	1.7462012e-14	5.7423127e-08

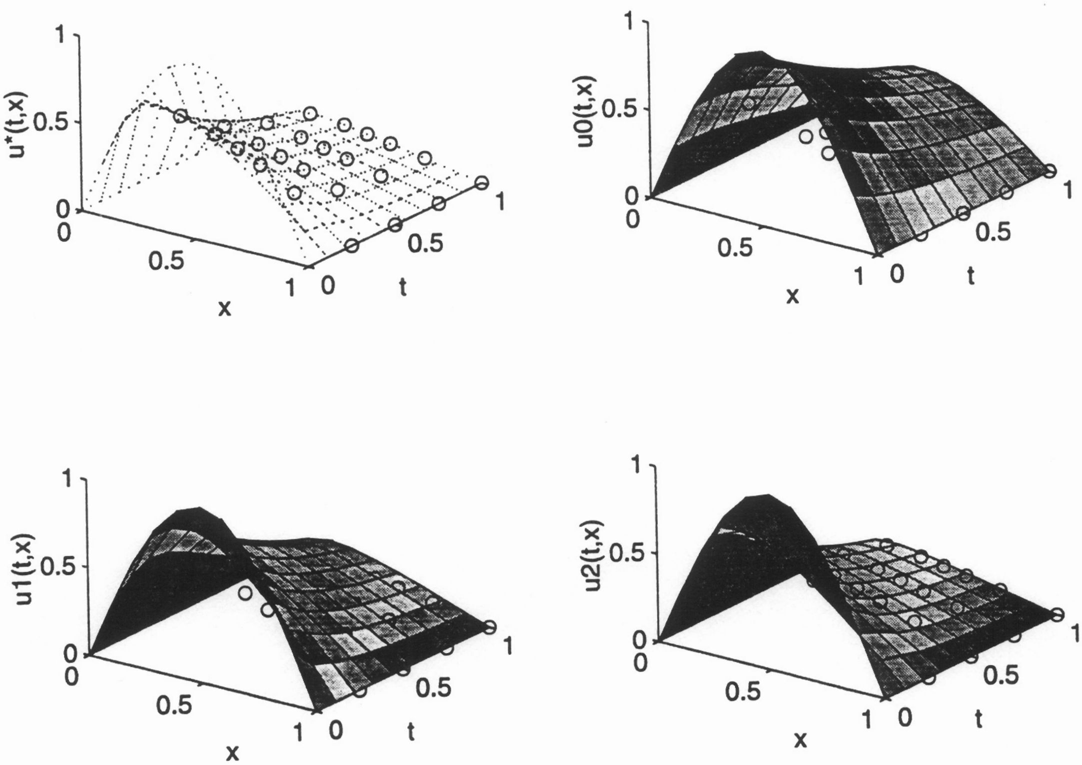

Figure 5.9 Fit to Data, (o o o – data)

Chapter 6

Concluding Remarks

In this thesis, we studied parameter identification problems for linear and nonlinear parabolic problems. The conditions are investigated under which the gradient of the state with respect to a parameter possesses smoothness properties which lead to local convergence of an estimation algorithm based on quasilinearization. For linear parabolic partial differential equations, we presented a framework which provides smoothness and convergence arguments under weak assumptions on the admissible parameter spaces. We established the maximum principle, the well-posedness and the smoothness properties of the solutions of a generalized Burgers' equation. Convergence of the quasilinearization algorithm is considered.

Numerical examples based on this algorithm were used to test the method. The numerical effort demonstrated that the method has potential, but also indicated that much work remains to be done before a complete theory can be developed. In particular, the algorithm seemed to work for problems with exact data and not for noisy data. Moreover, the numerical method was not analized for convergence. These issues need to be addressed in future work.

Bibliography

[1] R. A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.

[2] E. L. Allgower, and K. Böhmer, *A Mesh-Independence Principle for Operator Equations and Their Discretizations*, Arbeitspapiere der GMD 129, Bonn, January 1985.

[3] E. L. Allgower, K. Böhmer, F. A. Potra, and W. C. Rheinboldt, *A Mesh-Independence Principle for Operator Equations and Their Discretizations*, SIAM J. NUMER. ANAL., VOL. 23, No. 1, pp.160-169, February 1986.

[4] M. G. Armentano, *Parameter Identification in Burgers' Equation*, Preprint.

[5] H. T. Banks, J. A. Burns and E. M. Cliff, *Parameter Estimation and Identification for System with Delays*, SIAM J. Control and Optimization, 19, pp.791-828, 1981.

[6] H. T. Banks, R. Fabiano and Y. Wang, *Estimation of Boltzmann damping coefficients in beam models*, to appear.

[7] H. T. Banks and G. M. Groome, Jr., *Convergence Theorems for Parameter Estimation by Quasilinearization*, J. Math. Anal. Appl., 42, pp.91-109, 1973.

[8] H. T. Banks and K. Ito, *A Unified Framework for Approximation in Inverse Problems for Distributed Parameter Systems*, Control-Theory and Advanced Technology, 4, pp.73-90, 1988.

[9] H. T. Banks and K. Kunish, *Estimation Techniques for Distributed Parameter Systems*, Birkhauser, New York, 1989.

[10] H. T. Banks and P. D. Lamm, *Estimation of Variable Coefficients in Parabolic Distributed Systems*, IEEE Trans. Automat. Control, 30, pp.386-398, 1985.

[11] H. T. Banks and D. A. Rebnord, *Analytic Semigroups: Applications to Inverse Problems for Flexible Structures*, ICASE Report No. 90-36, May 1990.

[12] R. Bellman and R. Kalaba, *Quasilinearization and Nonlinear Boundary Value Problems*, American Elsevier, New York, 1965.

[13] J. Borggard, J. A. Burns, E. Cliff, and M. Gunzburger, *Sensitivity Calculations for a 2D, Inviscid, Supersonic Forebody Problem*, ICASE Report No. 93-13, March 1993.

[14] D. W. Brewer, *The Differentiability with Respect to a Parameter of the Solution of a Linear Abstract Cauchy Problem*, SIAM J. Math. Anal. 13, 607-620 (1982).

[15] D. W. Brewer, J. A. Burns and E. M. Cliff, *Parameter Identification for an Abstract Cauchy Problem by Quasilinearization*, Quart. of Appl. Math., 1(1993), pp.1-22.

[16] J. A. Burns and E. M. Cliff, *An Abstract Quasilinearization Algorithm for Estimating Parameters in Hereditary Systems*, IEEE Trans. Automat. Contr., 25(1980), pp.126-129.

[17] J. A. Burns, E. M. Cliff and M. D. Gunzburger, *An Optimization Problem Involving a Nonlinear Two-Point Boundary Value Problem*, ICAM Report, 91-07-03.

[18] J. A. Burns and S. Kang, *A Control Problem for Burgers' Equation with Bounded Input/Output*, Nonlinear Dynamics 2: pp.235-262, 1991.

[19] J. A. Burns and Y. Yan, *Dynamics of Nonhomogeneous Boundary Value of Burgers' Equation*, Preprint.

[20] C. I. Byrnes, D. S. Gilliam and V. I. Shubov, *On the Zero and Pole Dynamics of a Nonlinear Distributed Parameter System*, Preprint, 1994.

[21] P. G. Ciarlet, *The Finite Element Methods for Elliptic Problems*, North-Holland, Amsterdam, 1978.

[22] J. D. Cole, *On a Quasi-Linear Parabolic Equation Occurring in Aerodynamics*, Quart. Appl. Math., Vol. IX, No. 3, pp. 225-236, 1951.

[23] R. F. Curtain and A. J. Pritchard, *Infinite Dimensional Linear Systems Theory*, Springer-Verlag, 1978.

[24] J. E. Dennis, Jr. and R. B. Schnabel, *Numerical Methods for Nonlinear Equations and Unconstrained Optimization*, Prentice Hall, Englewood Cliffs, NJ, 1983.

[25] C. A. J. Fletcher, *Burgers' Equation: A Model for All Reasons*, Numerical Solution of Partial Differential Equations, J. Noye ed., North-Holland Publ. Co., pp.139-225, 1982.

[26] P. W. Hammer, *Parameter Identification in Parabolic Partial Differential Equations Using Quasilinearization*, Ph.D. dissertation, Virginia Tech, Blacksburg, Virginia, July, 1990.

[27] M. Heinkenschloss, *Mesh Independence for Nonlinear Least Squares Problems with Norm Constraints*, SIAM J. Optimization, Vol. 3, No. 1, pp. 81-117, 1993.

[28] D. Henry, *Geometric Theory of Semilinear Parabolic Equations*, Springer-Verlag, New York, 1981.

[29] E. Hopf, *The Partial Differential Equation* $u_t + uu_x = \mu u_{xx}$, Comm. Pure and Appl. Math., Vol. 3, pp.201-230, 1950.

[30] C. Johnson, *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Cambridge University Press, 1992.

[31] D. M. Hwang and C. T. Kelley, *Sequential Quadratic Programming for Parameter Identification Problems*, Parameter Systems, Perpignan, France, 1989.

[32] S. Kang, *A Control Problem for Burgers' Equation*, Ph.D. dissertation, Virginia Tech, Blacksburg, Virginia, April 1990.

[33] C. T. Kelley and E. W. Sachs, *Approximate Quasi-Newton Methods*, Mathematical Programming 48, pp.41-70, 1990.

[34] C. T. Kelley and E. W. Sachs, *Pointwise Broyden Methods*, SIAM J. Optimization, Vol. 3, No. 2, pp. 423-441, May 1993.

[35] T. Lin and B. Zhang, *A Finite Control Problem for an Initial Value Inverse Problem with Overspecified Boundary Data*, Preprint.

[36] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, Orlando, Florida, 1970.

[37] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983.

[38] M. H. Protter and H. Weinberger, *Maximum Principles in Differential Equations*, Printice-Hall, Englwood Cliffs, New Jersey, 1967.

[39] M. Renardy and R. C. Rogers, *An Introduction to Partial Differential Equations*, Springer-Verlag, New York, 1992.

[40] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Springer-Verlag, 1984.

[41] L. W. White, *A Study of Uniqueness for the Initialization Problem for Burgers' Equation*, Journal of Mathematical Analysis and Applications 172: pp.412-431, 1993.

[42] G. B. Whitham, *Linear and Nonlinear Waves*, John Wiley & Sons, 1974.

[43] J. Wloka, *Partial Differential Equations*, Cambridge University Press, New York, 1987.

[44] E. Zeidler, *Nonlinear Functional Analysis and its Applications II/A*, Springer-Verlag, New York, 1990.

VITA

Lan Zhang was born in Sichuan, China on March 25, 1968, daughter of Anxun Zhang and Suzhen Zhu. She received the B.S. degree in Mathematics from Chengdu University of Science and Technology in 1989. She came to the United States in 1989 and received her Ph.D. in mathematics from Virginia Polytechnic Institute and State University in 1995.

Lan Zhang