J. Austral. Math. Soc. Ser. B 40(1998), 1–22

A QUASI-NEWTON APPROACH TO IDENTIFICATION OF A PARABOLIC SYSTEM

WENHUAN YU[1]

(Received 3 January 1996; revised 20 January 1997)

Abstract

A quasi-Newton method (QNM) in infinite-dimensional spaces for identifying parameters involved in distributed parameter systems is presented in this paper. Next, the linear convergence of a sequence generated by the QNM algorithm is also proved. We apply the QNM algorithm to an identification problem for a nonlinear parabolic partial differential equation to illustrate the efficiency of the QNM algorithm.

1. Introduction

Quasi-Newton methods play an important role in numerically solving optimization problems on the Euclidean spaces. But few papers discuss these methods in identification of infinite-dimensional systems.

Formulating parameter estimation problems as constrained, regularized optimization problems, Kunisch *et al.* [13] investigated the reduced SQP (Sequential Quadratic Programming) methods with BFGS (Broyden-Flecher-Goldfarb-Shanno) update for the identification of an elliptic system.

In this paper we formulate an identification problem as an unconstrained optimization one. We suggest a Quasi-Newton Method (QNM) to solve an unconstrained optimization problem in Section 2. Following Broyden *et al.* [3] and using the Hilbert-Schmidt class defined in [7], we prove that the approximate sequence generated by the QNM procedure converges to the optimal element of the optimization problem if the latter exists. In Section 3 we apply the QNM algorithm to estimating a coefficient appearing in a nonlinear parabolic partial differential equation and we prove that the assumptions, which ensure the convergence of the approximate sequence obtained by the QNM algorithm, are satisfied. Finally, we illustrate a numerical example to show the efficiency of the QNM algorithm.

---

[1]Department of Mathematics, Tianjin University, Tianjin 300072, People's Republic of China.

https://doi.org/10.1017/S0334270000012339 Published online by Cambridge University Press

There are many papers dealing with the parameter identification problem for distributed parameter systems. Methods of solving some of those problems are listed in the following:

(1) the gradient or conjugate gradient methods, for example, Chavent *et al.* [4], Seinfeld *et al.* [17, 16], and Yu [18], which need to compute the derivative maps of the operators described by partial differential equations;
(2) the generalized pulse spectrum technique (GPST), for example, Chen *et al.* [5, 22];
(3) the finite-dimensional approximate modal methods, for example, Banks *et al.* [1, 2];
(4) the regularization methods, for example, Yu *et al.* [19, 20, 21];
(5) the sequential quadratic programming (SQP) methods, for example, Kunisch and Sachs [13] and Huang *et al.* [9];
(6) the Lagrangian method, for example, Ito and Kunisch [10, 11];
(7) Quasi-Newton methods for solving unconstrained optimal control problems, for example, Kelley and Sachs [13].

Finally, it should be pointed out that proving a superlinear rate of convergence for quasi-Newton methods in infinite-dimensional spaces is not trivial as it is in finite-dimensional spaces. The Q-superlinear convergence for the above-mentioned sequence can be obtained under an additional condition assumed by Griewank [8].

The QNM algorithm presented in this paper can also be applied to identification problems of other PDS's.

2. A quasi-Newton method in Hilbert spaces

We consider the following unconstrained optimization problem (UOP):

$$\text{minimize } f(x), \tag{2.1}$$

where $f : H \to \mathbb{R}$, and H is a Hilbert space. A point x^* is called optimal for UOP if f attains a local minimum at x^*.

It is well-known that the necessary condition for x^* being optimal is

$$f'(x^*) = 0, \tag{2.2}$$

where $f'(x^*) \in \mathscr{L}(H; \mathbb{R}) \equiv H'$ is the Fréchet derivative of f at x^*, $\mathscr{L}(X; Y)$ denotes the space of bounded linear operators from a Banach space X to a Banach space Y with the operator norm and H' is the adjoint space of H.

If $f : \mathbb{R}^n \to \mathbb{R}$, one frequently uses quasi-Newton methods for solving UOP because of their high efficacy. So, we use a quasi-Newton method for solving UOP as an iterative scheme which generates the sequences $\{x_k\}$ and $\{A_k\}$ from the formulas

$$A_k s_k = -f'(x_k), \tag{2.3}$$

$$x_{k+1} = x_k + s_k, \tag{2.4}$$

$$y_k = f'(x_{k+1}) - f'(x_k), \tag{2.5}$$

$$A_{k+1} = A_k + y_k \langle y_k, \cdot\rangle / \langle y_k, s_k\rangle - A_k s_k \langle A_k s_k, \cdot\rangle / \langle A_k s_k, s_k\rangle, \tag{2.6}$$

where x_0 and A_0 are given, $\langle\cdot,\cdot\rangle$ is the dual product between H' and H, and for any $y \in H'$ the operator $\langle y, \cdot\rangle : H \to \mathbb{R}$ is defined by

$$\langle y, \cdot\rangle x \equiv \langle y, x\rangle, \qquad \forall x \in H.$$

Obviously, $A_k \in \mathscr{L}(H, H')$ and $y_k, f'(x_k) \in H'$.

The above algorithm is just a BFGS formula in a Hilbert space. If set $B_k \equiv A_k^{-1}$, then by the Sherman-Morrison-Woodbury formula, we obtain

$$B_{k+1} = B_k + \frac{(s_k - B_k y_k)\langle\cdot, s_k\rangle + s_k\langle\cdot, s_k - B_k y_k\rangle}{\langle y_k, s_k\rangle} - \frac{\langle y_k, s_k - B_k y_k\rangle}{\langle y_k, s_k\rangle^2} s_k \langle\cdot, s_k\rangle, \tag{2.7}$$

and $B_k \in \mathscr{L}(H'; H)$.

In this paper $K : H \to H'$ is the canonical isometry, that is, for any $x \in H$ $Kx \in H'$ and

$$\langle Kx, s\rangle = (x, s), \qquad \forall s \in H,$$

$(\cdot,\cdot)$ and $((\cdot,\cdot))$ are the inner products of H and H', respectively.

The following definition can be found in [7].

DEFINITION 2.1. Let $\mathscr{B}_0(H; H')$ be the class of all compact operators on H and H'. For any $T \in \mathscr{B}_0(H; H')$, define

$$\|T\|_2 \equiv \left(\sum_{k=1}^{\infty} \|T\phi_k\|^2\right)^{1/2}, \tag{2.8}$$

where $\|\cdot\|$ is the norm of H' and $\{\phi_k\}$ is a complete orthonormal family in H. If the series in the right-hand side does not converge, set $\|T\|_2 = +\infty$. Moreover, $\|T\|_2$ is independent of the choice of the complete orthonormal family $\{\phi_k\}$ in (2.8). $\|T\|_2$ is called the Schmidt norm of T. The subset of $\mathscr{B}_0(H; H')$ consisting of all T with $\|T\|_2 < +\infty$ is called the Hilbert-Schmidt class, which is denoted by $\mathscr{B}_2(H; H')$.

$\mathscr{B}_2(H; H')$ is a Banach space with the norm $\|\cdot\|_2$.

DEFINITION 2.2. For any $T \in \mathscr{B}_2(H'; H)$, define the norm

$$\|T\|_M \equiv \|KMTKM\|_2, \tag{2.9}$$

where $M : H \to H$ is positive and self-adjoint.

$\mathscr{B}_2(H'; H)$ is a Banach space with the norm $\|\cdot\|_M$.

Obviously, we have the following properties.

LEMMA 2.1. *If* $T \in \mathscr{B}_2(H; H')$, $S_1 \in \mathscr{L}(H)$, *and* $S_2 \in \mathscr{L}(H')$, *where* $\mathscr{L}(X)$ *is the space of bounded linear operators with the operator norm from* X *to* X, *then* S_2T, $TS_1 \in \mathscr{B}_2(H; H')$ *and*

$$\|TS_1\|_2 \le \|S_1\|\|T\|_2, \qquad \|S_2T\|_2 \le \|S_2\|\|T\|_2. \tag{2.10}$$

Moreover, there are positive constants η_1 *and* η_2 *with* $\eta_2 \ge 1$ *such that* $\forall T \in \mathscr{B}_2(H; H')$

$$\eta_1\|T\|_M \le \|T\| \le \eta_2\|T\|_M. \tag{2.11}$$

In this paper we always suppose that the following assumptions are satisfied:

H1 $f : H \to \mathbb{R}$ is twice continuously Fréchet differentiable in $D_0 \subset H$, where D_0 is a convex and open set.

H2 There exists an $x^* \in D_0$ such that $f'(x^*) = 0$, $\|f''(x^*)\| \le \beta$, and that

$$\|f''(x) - f''(x^*)\| \le L\|x - x^*\|, \qquad \forall x \in D_0, \tag{2.12}$$

where L is a constant.

H3 $f''(x^*)$ is selfadjoint and strictly positive in the sense that $f''(x^*)h^2 \ge \lambda\|h\|^2$, $\forall h \in H$, where $\lambda > 0$, hence $f''(x^*)$ is invertible, $[f''(x^*)]^{-1} \equiv \Lambda \in \mathscr{L}(H'; H)$ and $\|\Lambda\| \le \theta$.

LEMMA 2.2. *Let the assumptions* **H1–H3** *be true. In addition, assume that there are non-negative constants* α_1 *and* α_2 *such that the operator sequence* $\{B_k\}$ *defined by* (2.7) *satisfies*

$$\|B_{n+1} - \Lambda\|_M \le (1 + \alpha_1\sigma_n)\|B_n - \Lambda\|_M + \alpha_2\sigma_n, \tag{2.13}$$

where $\sigma_n \equiv \max\{\|x_n - x^*\|, \|x_{n+1} - x^*\|\}$. *Then for each* $\gamma \in (0, 1)$ *there exist* $\epsilon = \epsilon(\gamma)$ *and* $\delta = \delta(\gamma)$, *such that if* B_0 *and* x_0 *satisfy*

$$\|x_0 - x^*\| \le \epsilon, \qquad \|B_0 - \Lambda\|_M \le \delta, \tag{2.14}$$

then the sequence $\{x_n\}$ *defined by the* QNM *algorithm is well-defined, converges to* x^*, *and satisfies*

$$\|x_{n+1} - x^*\| \le \gamma\|x_n - x^*\|, \qquad n = 0, 1, \ldots. \tag{2.15}$$

Furthermore, B_n^{-1} *exists and the sequences* $\{\|B_n\|\}$ *and* $\{\|B_n^{-1}\|\}$ *are uniformly bounded.*

https://doi.org/10.1017/S0334270000012339 Published online by Cambridge University Press

PROOF. By the assumption **H2** for any $\gamma \in (0, 1)$ we can choose $\delta = \delta(\gamma) > 0$ and $\epsilon = \epsilon(\gamma) > 0$ such that

$$6\beta(1+\gamma)\delta\eta_2 \leq \gamma, \tag{2.16}$$

$$(2\alpha_1\delta + \alpha_2)\epsilon/(1-\gamma) \leq \delta, \qquad \epsilon < \epsilon_0, \tag{2.17}$$

$$(\theta + 4\eta_2\delta)[L\epsilon/2 + 2(1+\gamma)^2\beta^2\eta_2\delta] \leq \gamma, \tag{2.18}$$

where ϵ_0 is so small that

$$B(x^*, \epsilon_0) \equiv \{x \in H; \|x - x^*\| < \epsilon_0\} \subset D_0.$$

It follows from (2.11) that

$$\|B_0 - \Lambda\| \leq \eta_2 \|B_0 - \Lambda\|_M \leq \eta_2\delta \tag{2.19}$$

and

$$\|B_0\| \leq \|\Lambda\| + \|B_0 - \Lambda\| \leq \theta + \eta_2\delta. \tag{2.20}$$

Because $B_0 = \Lambda + (B_0 - \Lambda)$ and $\|B_0 - \Lambda\|_M \leq \delta < 1/\beta \leq \|\Lambda^{-1}\|^{-1}$, by the Banach inverse theorem, we deduce that B_0 is invertible and that

$$\|A_0\| = \|B_0^{-1}\| \leq \|\Lambda^{-1}\|/(1 - \|\Lambda^{-1}\|\|B_0 - \Lambda\|) \leq \beta/(1 - \beta\eta_2\delta) < \beta/(1 - 6\beta\eta_2\delta).$$

But by (2.16)

$$1 - 6\beta\eta_2\delta > 1 - \gamma/(1+\gamma) = 1/(1+\gamma),$$

so

$$\|A_0\| = \|B_0^{-1}\| < (1+\gamma)\beta. \tag{2.21}$$

Furthermore,

$$\begin{aligned} \|A_0 - f''(x^*)\| &= \|A_0(\Lambda - B_0)f''(x^*)\| \leq \|A_0\|\|B_0 - \Lambda\|\|f''(x^*)\| \\ &\leq (1+\gamma)\beta^2\eta_2\delta. \end{aligned} \tag{2.22}$$

It follows from the mean-value theorem that

$$\begin{aligned} \|x_1 - x^*\| &= \|(x_1 - x_0) + (x_0 - x^*)\| = \| - B_0^{-1} f'(x_0) + (x_0 - x^*)\| \\ &= \|B_0^{-1}\{-[f'(x_0) - f'(x^*) - f''(x^*)(x_0 - x^*)] \\ &\qquad + [A_0 - f''(x^*)](x_0 - x^*)\}\| \\ &\leq \|B_0\| \left\{ \left\| \int_0^1 [f''(x^* + t(x_0 - x^*)) - f''(x^*)](x_0 - x^*)\, dt \right\| \right. \\ &\qquad \left. + \|A_0 - f''(x^*)\|\|x_0 - x^*\| \right\} \\ &\leq (\theta + \eta_2\delta)[L\epsilon/2 + (1+\gamma)\beta^2\eta_2\delta]\|x_0 - x^*\| \leq \gamma\|x_0 - x^*\|. \end{aligned} \tag{2.23}$$

https://doi.org/10.1017/S0334270000012339 Published online by Cambridge University Press

Next, from (2.13), (2.14), and (2.17) we have

$$\|B_1 - \Lambda\|_M \le (1+\alpha_1\epsilon)\|B_0 - \Lambda\|_M + \alpha_2\epsilon \le \delta + (\alpha_1\delta + \alpha_2)\epsilon < 2d\delta. \quad (2.24)$$

Using induction, we prove

$$\|B_k - \Lambda\| \le 2\delta \quad \text{and} \quad \|x_{k+1} - x^*\| \le \gamma\|x_k - x^*\|. \quad (2.25)$$

In fact, suppose that (2.25) are true for $k \le m-1$. By (2.13)

$$\|B_{k+1} - \Lambda\|_M \le (1+\alpha_1\epsilon\gamma^k)\|B_k - \Lambda\|_M + \alpha_2\epsilon\gamma^k,$$

that is,

$$\|B_{k+1} - \Lambda\|_M - \|B_k - \Lambda\|_M \le 2\alpha_1\epsilon\gamma^k\delta + \alpha_2\epsilon\gamma^k = (2\alpha_1\delta + \alpha_2)\epsilon\gamma^k. \quad (2.26)$$

Adding (2.26) from $k = 0$ to $m-1$, one gets

$$\begin{aligned}\|B_m - \Lambda\|_M &\le \|B_0 - \Lambda\|_M + (2\alpha_1\delta + \alpha_2)\epsilon\sum_{k=0}^{m-1}\gamma^k \\ &< \delta + (2\alpha_1\delta + \alpha_2)\epsilon/(1-\gamma) \le 2\delta.\end{aligned} \quad (2.27)$$

Therefore

$$\|B_m - B_0\|_M \le \|B_m - \Lambda\|_M + \|B_0 - \Lambda\|_M < 3\delta \qquad \forall m. \quad (2.28)$$

In addition, by (2.11), (2.16), (2.21) and the above

$$\|I - B_0^{-1}B_m\| \le \|B_0^{-1}\|\|B_m - B_0\| < 3(1+\gamma)\beta\eta_2\delta < \gamma < 1,$$

where $I \in \mathscr{L}(H')$ is the identity operator. By the Banach theorem $(B_0^{-1}B_m)^{-1}$ exists, hence B_m is invertible. Furthermore,

$$\begin{aligned}\|A_m\| = \|B_m^{-1}\| &= \|[B_0 + (B_M - B_0)]^{-1}\| \le \|B_0^{-1}\|\sum\|B_0^{-1}\|^k\|B_m - B_0\|^k \\ &\le (1+\gamma)\beta\sum[3(1+\gamma)\beta\eta_2\delta]^k = (1+\gamma)\beta/[1-3(1+\gamma)\beta\eta_2\delta] \\ &< (1+\gamma)\beta/(1-6\beta\eta_2\delta) < (1+\gamma)^2\beta, \qquad \forall m,\end{aligned} \quad (2.29)$$

that is, for any $m \in \mathbb{N}$, B_m is invertible and $\{\|B_m^{-1}\|\}$ is uniformly bounded as well. Moreover,

$$\|B_m\| \le \|B_0\| + \|B_m - B_0\| \le (\theta + \eta_2\delta) + 3\eta_2\delta = \theta + 4\eta_2\delta, \quad (2.30)$$

$$\begin{aligned}\|A_m - f''(x^*)\| &= \|A_m(B_m - \Lambda)f''(x^*)\| \le \|A_m\|\|B_m - \Lambda\|\|f''(x^*)\| \\ &\le (1+\gamma)^2\beta\eta_2 2\delta\beta = 2(1+\gamma)^2\beta^2\eta_2\delta.\end{aligned} \quad (2.31)$$

https://doi.org/10.1017/S0334270000012339 Published online by Cambridge University Press

By the inductive assumption one has $\|x_m - x^*\| < \epsilon$ and so

$$\begin{aligned} \|x_{m+1} - x^*\| &= \|(x_{m+1} - x_m) + (x_m - x^*)\| = \| - B_m^{-1} f'(x_m) + (x_m - x^*)\| \\ &= \|B_m^{-1}\{-[f'(x_m) - f'(x^*) - f''(x^*)(x_m - x^*)] \\ &\qquad + [A_m - f''(x^*)](x_m - x^*)\}\| \\ &\le \|B_m^{-1}\| \left\{ \left\| \int_0^1 [f''(x^* + t(x_m - x^*)) - f''(x^*)](x_m - x^*)\, dt \right\| \right. \\ &\qquad \left. + \|A_m - f''(x^*)\| \|x_m - x^*\| \right\} \\ &\le (\theta + 4\eta_2\delta)[L\epsilon/2 + 2(1+\gamma)^2\beta^2\eta_2\delta]\|x_m - x^*\| \le \gamma\|x_m - x^*\|. \end{aligned} \tag{2.32}$$

Furthermore,

$$\|x_{m+1} - x^*\| \le \gamma\|x_m - x^*\| \le \cdots \le \gamma^{m+1}\|x_0 - x^*\| < \gamma^{m+1}\epsilon < \epsilon_0. \tag{2.33}$$

Thus $\{x_m\} \subset D_0$, $x_m \to x^*$ in H, and $\{\|B_m\|\}$ and $\{\|B_m^{-1}\|\}$ are uniformly bounded.

LEMMA 2.3. *Let $C, B \in \mathscr{L}(H'; H)$ be selfadjoint, $y \in H'$, $s \in H$ with $\langle y, s\rangle \ne 0$, and set*

$$\bar{B} = B + \frac{(s - By)\langle \cdot, s\rangle + s\langle \cdot, s - By\rangle}{\langle y, s\rangle} - \frac{\langle y, s - By\rangle}{\langle y, s\rangle^2} s\langle \cdot, s\rangle. \tag{2.34}$$

If $M \in \mathscr{L}(H)$ is invertible and selfadjoint, then

$$\bar{E} = P^*EP + \frac{KM(s - Cy)}{\langle y, s\rangle}\langle KMs, \cdot\rangle + \frac{KMs}{\langle y, s\rangle}\langle P^*KM(s - Cy), \cdot\rangle, \tag{2.35}$$

where $E = KM(B - C)KM$, $\bar{E} = KM(\bar{B} - C)KM, P = I - M^{-1}K^{-1}y\langle KMs, \cdot\rangle/\langle y, s\rangle$, $I \in \mathscr{L}(H)$ is the identity operator, and P^* is the adjoint operator of P.

PROOF. Premultiplying and postmultiplying both sides of (2.34) by KM and subtract-

https://doi.org/10.1017/S0334270000012339 Published online by Cambridge University Press

ing $KMCKM$, we have the simple calculation

$$\begin{aligned}
\bar{E} &= E + \{KM[(s - Cy) - (B - C)y]\langle KMs, \cdot\rangle + KMs\langle KM[(s - Cy) \\
&\qquad - (B - C)y], \cdot\rangle\}/\langle y, s\rangle - KMs\langle y, (s - Cy) - (B - C)y\rangle\langle KMs, \cdot\rangle/\langle y, s\rangle^2 \\
&= E - EM^{-1}K^{-1}y\langle KMs, \cdot\rangle/\langle y, s\rangle - KMs\langle EM^{-1}K^{-1}y, \cdot\rangle/\langle y, s\rangle \\
&\qquad + \{KM(s - Cy)\langle KMs, \cdot\rangle + KMs\langle KM(s - Cy), \cdot\rangle\}/\langle y, s\rangle \\
&\qquad + \langle EM^{-1}K^{-1}y, M^{-1}K^{-1}y\rangle KMs\langle KMs, \cdot\rangle/\langle y, s\rangle^2 \\
&\qquad - \langle y, s - Cy\rangle KMs\langle KMs, \cdot\rangle/\langle y, s\rangle^2 \\
&= E[I - M^{-1}K^{-1}y\langle KMs, \cdot\rangle/\langle y, s\rangle] + KMs\{\langle EM^{-1}K^{-1}y, M^{-1}y\rangle\langle KMs, \cdot\rangle/\langle y, s\rangle \\
&\qquad - \langle EM^{-1}K^{-1}y, \cdot\rangle\}/\langle y, s\rangle + KM(s - Cy)\langle KMs, \cdot\rangle/\langle y, s\rangle \\
&\qquad + KMs\{\langle KM(s - Cy), \cdot\rangle - \langle y, s - Cy\rangle\langle KMs, \cdot\rangle/\langle y, s\rangle\}/\langle y, s\rangle.
\end{aligned} \tag{2.36}$$

It is obvious that E and $\bar{E}$ are selfadjoint and that

$$P^* = I - KMs\langle \cdot, M^{-1}K^{-1}y\rangle/\langle y, s\rangle \in \mathscr{L}(H'). \tag{2.37}$$

Considering the above results, from (2.36) we have

$$\begin{aligned}
\bar{E} &= EP - KMs\langle EM^{-1}K^{-1}y - KMs\langle M^{-1}K^{-1}y, M^{-1}K^{-1}y\rangle/\langle y, s\rangle, \cdot\rangle/\langle y, s\rangle \\
&\qquad + KM(s - Cy)\langle KMs, \cdot\rangle/\langle y, s\rangle + KMs\langle KM(s - Cy) \\
&\qquad - KMs\langle KM(s - Cy), M^{-1}K^{-1}y\rangle/\langle y, s\rangle, \cdot\rangle/\langle y, s\rangle \\
&= EP - KMs\langle [I - KMs\langle \cdot, M^{-1}K^{-1}y\rangle/\langle y, s\rangle]EM^{-1}K^{-1}y, \cdot\rangle/\langle y, s\rangle \\
&\qquad + KM(s - Cy)\langle KMs, \cdot\rangle/\langle y, s\rangle \\
&\qquad + KMs\langle [I - KMs\langle \cdot, M^{-1}K^{-1}y\rangle/\langle y, s\rangle]KM(s - Cy), \cdot\rangle/\langle y, s\rangle \\
&= EP - KMs\langle P^*EM^{-1}K^{-1}y, \cdot\rangle/\langle y, s\rangle + KM(s - Cy)\langle KMs, \cdot\rangle/\langle y, s\rangle \\
&\qquad + KMs\langle P^*KM(s - Cy), \cdot\rangle/\langle y, s\rangle \\
&= [I - KMs\langle M^{-1}K^{-1}y, \cdot\rangle/\langle y, s\rangle]EP \\
&\qquad KM(s - Cy)\langle KMs, \cdot\rangle/\langle y, s\rangle + KMs\langle P^*KM(s - Cy), \cdot\rangle/\langle y, s\rangle \\
&= P^*EP + KM(s - Cy)\langle KMs, \cdot\rangle/\langle y, s\rangle + KMs\langle P^*KM(s - Cy), \cdot\rangle/\langle y, s\rangle.
\end{aligned}$$

LEMMA 2.4. *Let $M \in \mathscr{L}(H)$ be a non-singular selfadjoint operator such that*

$$\|Ms - M^{-1}K^{-1}y\| \leq \rho\|M^{-1}K^{-1}y\|, \tag{2.38}$$

where $\rho \in (0, 1/3)$, $s \in H$ and $y \in H'$ with $y \neq 0$. Then

$$(1 - \rho)\|M^{-1}K^{-1}y\|^2 \leq \langle y, s\rangle \leq (1 + \rho)\|M^{-1}K^{-1}y\|^2 \tag{2.39}$$

https://doi.org/10.1017/S0334270000012339 Published online by Cambridge University Press

and for each $E \in \mathscr{B}_2(H, H')$,

$$\|E[I - M^{-1}K^{-1}y\langle KM^{-1}K^{-1}y, \cdot\rangle/\langle y, s\rangle]\|_2 \le (1 - \mu\nu^2)^{1/2}\|E\|_2, \tag{2.40}$$

$$\|E[I - M^{-1}K^{-1}y\langle KMs, \cdot\rangle/\langle y, s\rangle]\|_2 \le \{(1-\mu\nu^2)^{1/2} + \|Ms - M^{-1}K^{-1}y\|/[(1-\rho)\|M^{-1}K^{-1}y\|]\}\|E\|_2, \tag{2.41}$$

and $\forall y \in H'$, $A \in \mathscr{L}(H, H')$, $s \in H$:

$$\|(y - As)\langle KMs, \cdot\rangle/\langle y, s\rangle\|_2 \le 2\|y - As\|/\|M^{-1}K^{-1}y\|, \tag{2.42}$$

where

$$\mu = (1 - 2\rho)/(1 - \rho^2) \in (3/8, 1)$$

and

$$\nu = \|EM^{-1}K^{-1}y\|/(\|E\|_2\|M^{-1}K^{-1}y\|) \in (0, 1).$$

PROOF. We have (2.39) from

$$\begin{aligned}\langle y, s\rangle &= (K^{-1}y, s) = (MM^{-1}K^{-1}y, s) = (M^{-1}K^{-1}y, Ms - M^{-1}K^{-1}y + M^{-1}K^{-1}y)\\ &= \|M^{-1}K^{-1}y\|^2 + (M^{-1}K^{-1}y, Ms - M^{-1}K^{-1}y).\end{aligned}$$

Observing that $\forall u, v \in H$,

$$\begin{aligned}&\|E[I - u(v, \cdot)]\|_2^2\\ &\quad= \sum_k \|[E - Eu(v, \cdot)]\phi_k\|^2\\ &\quad= \sum_k ((E\phi_k - (v, \phi_k)Eu, E\phi_k - (v, \phi_k)Eu))\\ &\quad= \sum_k ((E\phi_k, E\phi_k)) - 2\sum_k (v, \phi_k)((Eu, E\phi_k)) + \sum_k (v, \phi_k)^2((Eu, Eu))\\ &\quad= \|E\|_2^2 - 2((Eu, E\sum_k (v, \phi_k)\phi_k)) + \|v\|^2\|Eu\|^2\\ &\quad= \|E\|_2^2 - 2((Eu, Ev)) + \|Eu\|^2\|v\|^2,\end{aligned} \tag{2.43}$$

and taking $u = M^{-1}K^{-1}y/\langle y, s\rangle$ and $v = M^{-1}K^{-1}y$, we immediately obtain

$$\begin{aligned}&\|E[I - M^{-1}K^{-1}y(M^{-1}K^{-1}y, \cdot)/\langle y, s\rangle]\|_2^2\\ &\quad= \|E\|_2^2 - 2(EM^{-1}K^{-1}y, EM^{-1}K^{-1}y)/\langle y, s\rangle\\ &\qquad\quad + \|EM^{-1}K^{-1}y\|^2\|M^{-1}K^{-1}y\|^2/\langle y, s\rangle^2\\ &\quad= \|E\|_2^2 + \{-2\langle y, s\rangle + \|M^{-1}K^{-1}y\|^2\}\|EM^{-1}K^{-1}y\|^2/\langle y, s\rangle^2\\ &\quad\le \|E\|_2^2 - (1 - 2\rho)\|EM^{-1}K^{-1}y\|^2/[(1 - \rho)\|M^{-1}K^{-1}y\|^2] = (1 - \mu\nu^2)\|E\|_2^2,\end{aligned}$$

which reduces to (2.40).

Owing to (2.40), in order to establish (2.41) we need only prove

$$\begin{aligned}\|EM^{-1}K^{-1}y[\langle KMs,\cdot\rangle - (M^{-1}K^{-1}y,\cdot)]/\langle y,s\rangle\|_2 &= \|EM^{-1}K^{-1}y(M^{-1}K^{-1}y - Ms,\cdot)/\langle y,s\rangle\|_2\\ &= \|M^{-1}K^{-1}y\|\|M^{-1}K^{-1}y - Ms\|/\langle y,s\rangle\|\|E\|_2\\ &\le \|M^{-1}K^{-1}y - Ms\|/[(1-\rho)\|M^{-1}K^{-1}y\|]\|E\|_2.\end{aligned}$$

Inequality (2.42) can be reduced in a similar fashion. In fact,

$$\begin{aligned}\|(y-As)\langle KMs,\cdot\rangle/\langle y,s\rangle\|_2^2 &= \sum_k \|(y-As)(Ms,\phi_k)/\langle y,s\rangle\|^2\\ &= \sum_k \|y-As\|^2(Ms,\phi_k)^2/\langle y,s\rangle^2 = \|y-As\|^2\|Ms\|^2/\langle y,s\rangle^2,\end{aligned}$$

so

$$\begin{aligned}\|(y-As)\langle KMs,\cdot\rangle/\langle y,s\rangle\|_2 &= \|y-As\|\|Ms\|/\langle y,s\rangle\\ &\le \|y-As\|/\langle y,s\rangle\{\|M^{-1}K^{-1}y - Ms\| + \|M^{-1}K^{-1}y\|\}\\ &\le \|y-As\|/[(1-\rho)\|M^{-1}K^{-1}y\|^2](1+\rho)\|M^{-1}K^{-1}y\|\\ &\le 2\|y-As\|/\|M^{-1}K^{-1}y\|.\end{aligned}$$

LEMMA 2.5. *Let M satisfy the conditions in Lemma 2.4 and let $B_0 - \Lambda \in \mathscr{B}_2(H)$. Then $\langle s_k, y_k\rangle \ne 0$, and B_{k+1} is well-defined and satisfies*

$$\begin{aligned}\|B_{k+1}-\Lambda\|_M &\le 2(1+\rho)\|KM\|\|s_k - \Lambda y_k\|/[(1-\rho)^2\|M^{-1}K^{-1}y_k\|]\\ &+ \{\sqrt{1-\mu\nu^2} + 5/2\|Ms_k - M^{-1}K^{-1}y_k\|/[(1-\rho)\|M^{-1}K^{-1}y_k\|]\}\|B_k - \Lambda\|_M,\end{aligned} \tag{2.44}$$

where $\mu = (1-2\rho)/(1-\rho^2) \in [3/8, 1]$ and $\nu = \|KM(B_k-\Lambda)y_k\|/(\|B_k - \Lambda\|_M\|M^{-1}K^{-1}y_k\|)$.

PROOF. It follows by the QNM algorithm and (2.39) that $y_k \ne 0$ and $\langle s_k, y_k\rangle \ne 0$. So B_{k+1} is well-defined by (2.7).

According to Lemma 2.3 we have

$$\begin{aligned}E_{k+1} = P^*E_kP &+ KM(s_k - \Lambda y_k)\langle KMs_k,\cdot\rangle/\langle y_k,s_k\rangle\\ &+ Ms_k\langle P^*KM(s_k-\Lambda y_k),\cdot\rangle/\langle y_k,s_k\rangle,\end{aligned} \tag{2.45}$$

where $E_{k+1} = KM(B_{k+1}-\Lambda)KM$, $E_k = KM(B_k - \Lambda)KM$, and $P = I - M^{-1}K^{-1}y_k\langle KMs_k,\cdot\rangle/\langle y_k,s_k\rangle$.

Applying Lemma 2.4, we have

$$\begin{aligned}\|P^*E_kP\|_2 &\le \left\{\sqrt{1-\mu\nu^2} + \|Ms_k - M^{-1}K^{-1}y_k\|/[(1-\rho)\|M^{-1}K^{-1}y_k\|]\right\}\|P^*E_k\|_2 \\ &\le \{1+\|Ms_k - M^{-1}K^{-1}y_k\|/[(1-\rho)\|M^{-1}K^{-1}y_k\|]\}\|E_kP\|_2.\end{aligned}$$

Hence,

$$\begin{aligned}\|P^*EP\|_2 &\le \{1+\|Ms_k - M^{-1}K^{-1}y_k\|/[(1-\rho)\|M^{-1}K^{-1}y_k\|]\} \\ &\quad\times\left\{\sqrt{1-\mu\nu^2} + \|Ms_k - M^{-1}K^{-1}y_k\|/[(1-\rho)\|M^{-1}K^{-1}y_k\|]\right\}\|E_k\|_2 \\ &\le \left\{\sqrt{1-\mu\nu^2} + 5\|Ms_k - M^{-1}K^{-1}y_k\|/[2(1-\rho)\|M^{-1}K^{-1}y_k\|]\right\}\|E_k\|_2.\end{aligned} \tag{2.46}$$

Next, we estimate the other two terms of (2.45). Since

$$\begin{aligned}\|KM(s_k-\Lambda y_k)\langle KMs_k,\cdot\rangle/\langle y_k,s_k\rangle\|_2^2 &= \sum_n \|(Ms_k,\phi_n)KM(s_k-\Lambda y_k)\|^2/\langle y_k,s_k\rangle^2 \\ &= \sum_n (Ms_k,\phi_n)^2\|KM(s_k-\Lambda y_k)\|^2/\langle y_k,s_k\rangle^2 \\ &= \|Ms_k\|^2\|KM(s_k-\Lambda y_k)\|^2/\langle y_k,s_k\rangle^2\end{aligned}$$

and

$$\|Ms_k\| \le \|M^{-1}K^{-1}y_k\| + \|Ms_k - M^{-1}K^{-1}y_k\| \le (1+\rho)\|M^{-1}K^{-1}y_k\|,$$

we have

$$\begin{aligned}\|KM(s_k-\Lambda y_k)\langle KMs_k,\cdot\rangle/\langle y_k,s_k\rangle\|_2 &= \|Ms_k\|\|KM(s_k-\Lambda y_k)\|/\langle y_k,s_k\rangle \\ &\le (1+\rho)\|M^{-1}K^{-1}y_k\|\|KM\|\|s_k-\Lambda y_k\|/[(1-\rho)\|M^{-1}K^{-1}y_k\|^2] \\ &= (1+\rho)\|KM\|\|s_k-\Lambda y_k\|/[(1-\rho)\|M^{-1}K^{-1}y_k\|].\end{aligned} \tag{2.47}$$

On the other hand,

$$\begin{aligned}\|P^*\| = \|P\| &\le \|I\| + \|M^{-1}K^{-1}y_k\langle KMs_k,\cdot\rangle/\langle y_k,s_k\rangle\| \\ &\le 1+(1+\rho)/(1-\rho) = 2/(1-\rho),\end{aligned} \tag{2.48}$$

and hence

$$\begin{aligned}&\|KMs_k\langle P^*KM(s_k-\Lambda y_k),\cdot\rangle/\langle y_k,s_k\rangle\|_2 \\ &\quad= \|P^*KM(s_k-\Lambda y_k)\|\|Ms_k\|/\langle y_k,s_k\rangle \\ &\quad\le \|P^*\|\|KM\|\|s_k-\Lambda y_k\|(1+\rho)\|M^{-1}K^{-1}y_k\|/[(1-\rho)\|M^{-1}K^{-1}y_k\|^2] \\ &\quad\le 2(1+\rho)\|KM\|\|s_k-\Lambda y_k\|/[(1-\rho)\|M^{-1}K^{-1}y_k\|].\end{aligned} \tag{2.49}$$

Summing up (2.46), (2.47) and (2.49) and considering $\|B_{k+1} - \Lambda\|_M = \|KM(B_{k+1} - \Lambda)KM\|_2 = \|E_{k+1}\|_2$, we immediately obtain the estimate (2.44) from (2.45).

THEOREM 2.6. *Let assumptions* **H1**, **H2**, *and* **H3** *be satisfied. Then the sequence* $\{x_n\}$ *generated by the* QNM *algorithm is well-defined and converges to* x^* *provided that the initial guesses* x_0 *and* B_0 *satisfy the conditions of Lemma* 2.2.

PROOF. Because $f''(x^*)[\cdot, \cdot]$ is a bounded symmetric bilinear form on $H \times H$, there is a selfadjoint operator $T \in \mathscr{L}(H)$ such that

$$f''(x^*)[s, t] = (Ts, t), \qquad \forall s, t \in H.$$

Moreover, by assumption **H3**, there exists a selfadjoint positive operator $M \in \mathscr{L}(H)$ such that $M^2 = T$. Hence

$$f''(x^*)[s, t] = (M^2 s, t) = (Ms, Mt), \qquad \forall s, t \in H.$$

For any $y \in H'$ we have $K^{-1}y \in H$ and

$$y - f''(x^*)s(\cdot) = (K^{-1}y - M^2 s, \cdot) = K[M(M^{-1}K^{-1}y - Ms)](\cdot). \tag{2.50}$$

But, by assumption **H2** and the QNM algorithm one has

$$\begin{aligned} \|y_k - f''(x^*)s_k\| &= \|f'(x_{k+1}) - f'(x_k) - f''(x^*)s_k\| \\ &= \left\| \int_0^1 [f''(x^* + t(x_{k+1} - x_k) - f''(x^*)](x_{k+1} - x_k)\, dt \right\| \\ &\leq L\|x_{k+1} - x_k\|^2/2 = L\|s_k\|^2/2. \end{aligned} \tag{2.51}$$

So

$$\begin{aligned} \|Ms_k - M^{-1}K^{-1}y_k\| &= \|M^{-1}K^{-1}[y_k - f''(x^*)s_k]\| \\ &\leq \|M^{-1}K^{-1}\|L\|s_k\|^2/2. \end{aligned} \tag{2.52}$$

Moreover, by assumption **H3** there is a $\kappa > 0$ such that

$$\|s_k\|/\kappa \leq \|M^{-1}K^{-1}y_k\| \leq \kappa\|s_k\|. \tag{2.53}$$

Summarizing (2.50), (2.51) and (2.53), we have

$$\|Ms_k - M^{-1}y_k\| \leq \rho\|M^{-1}K^{-1}y_k\|, \tag{2.54}$$

https://doi.org/10.1017/S0334270000012339 Published online by Cambridge University Press

where $\rho \in (0, 1/3)$. It follows by Lemma 2.1 that $B_k - \Lambda \in \mathscr{B}_2(H', H)$. Therefore by Lemma 2.5

$$\|B_{k+1} - \Lambda\|_M \le \left\{\sqrt{1-\mu\nu^2} + 5\|Ms_k - M^{-1}K^{-1}y_k\|/[2(1-\rho)\|M^{-1}K^{-1}y_k\|]\right\} \times \|B_k - \Lambda\|_M + 2(1+\rho)\|KM\|\|s_k - \Lambda y_k\|/[(1-\rho)^2\|M^{-1}K^{-1}y_k\|]. \tag{2.55}$$

Considering (2.52) and (2.53), we have

$$\begin{aligned}\|s_k - \Lambda y_k\| = \|\Lambda(y_k - f''(x^*)s_k\| &\le \|\Lambda\|L\|s_k\|^2/2 \\ &\le \theta L\kappa\|M^{-1}K^{-1}y_k\|/2\|s_k\|.\end{aligned} \tag{2.56}$$

If we set $\alpha_1 = 5L\kappa\|M^{-1}K^{-1}/[4(1-\rho)]$ and $\alpha_2 = (1+\rho)L\kappa\|KM\|/(1-\rho)^2$, then from (2.55) we have

$$\|B_{k+1} - \Lambda\|_M \le (1 + \alpha_1\sigma_k)\|B_k - \Lambda\|_M + \alpha_2\sigma_k,$$

where $\sigma_k \equiv \max(\|x_{k+1} - x^*\|, \|x_k - x^*\|)$.

It follows by Lemma 2.2 that the conclusions of Theorem 2.6 are true.

From [8] we quote the following result.

THEOREM 2.7. *Assume that the requirements in Theorem* 2.6 *are satisfied and that* $A_0 - F'(q^*)$ *is compact. Then the sequence* $\{x_n\}$ *generated by* QNM *is* Q-*superlinear convergent, that is,*

$$\lim_{k\to\infty} \|x_{k+1} - x^*\|/\|x_k - x^*\| = 0,$$

provided $\|x_0 - x^*\|$ *is sufficiently small.*

3. Identification of a nonlinear parabolic system

The problem we address here is to identify the parameter q appearing in a parabolic semilinear equation

$$\begin{aligned}\partial_t u - \Delta u + q^2u^2 = 0, &\qquad (x,t) \in D \equiv \Omega \times (0,T), \\ u|_{\partial\Omega} = 0, &\qquad u|_{t=0} = g(x),\end{aligned} \tag{3.1}$$

based on the final measurement of the state u

$$u|_{t=T} = z(x), \qquad x \in \Omega. \tag{3.2}$$

The assumptions we use in this section are as follows:

https://doi.org/10.1017/S0334270000012339 Published online by Cambridge University Press

A1. $\Omega \subset \mathbb{R}^m$ is bounded and its boundary $\partial\Omega \in C^{2+\beta}$, where $\beta \in (0, 1)$;
A2. $g \in C^{2+\beta}(\bar{\Omega})$.

For any $q \in C^{\beta}(\bar{\Omega})$ it was proved in [14] that problem (3.1) has a unique classical solution, $u \in C^{2+\beta,1+\beta/2}(\bar{D})$. So, we denote $u = u(q) = u(x, t; q)$ to show the dependence of u on q.

The problem of identification is stated as an optimization problem

$$I(q) = 1/2\|u(\cdot, T; q) - z\|^2_{L^2(\Omega)} \to \min. \tag{3.3}$$

But the above problem is, usually, ill-posed in the Hadamard sense. Thus, we introduce a regularization term as follows:

$$J(q) = 1/2\|u(\cdot, T; q) - z\|^2_{L^2(\Omega)} + \alpha/2\|q\|^2_H, \tag{3.4}$$

where $\alpha > 0$ is a constant, $H = H^l(\Omega)$, and the order l of the Sobolev space is chosen such that H is compactly embedded in $C^{\beta}(\bar{\Omega})$. For example, $l = 1$ when $m = 1$ and $l = 2$ when $m = 2$ or 3.

It follows by [21] that the optimal parameter for the problem (3.4) converges to the optimal parameter of the problem (3.3) as $\alpha \to 0$.

THEOREM 3.1. *The function $u : H \to V \equiv C^{2+\beta,1+\beta/2}(\bar{D})$ defined by* (3.1) *is infinitely differentiable, that is, $u \in C^N(H; V)$, $\forall N \in \mathbb{N} \cup \{0\}$, where $C^N(H; V)$ denotes the linear space of N-times continuously Fréchet differentiable functions on H to V. Moreover, the first Fréchet derivative $u'(\cdot) : H \to \mathscr{L}(H; V)$ and the second Fréchet derivative $u''(\cdot) : H \to \mathscr{L}(H; \mathscr{L}(H; V))$ of u at q are implicitly determined by $u'(q)h = \dot{u}$ and $u''(q)hk = \ddot{u}$, $\forall h, k \in H$, respectively, where $\dot{u}$ and $\ddot{u}$ are determined by the problems*

$$\begin{aligned} &\partial_t \dot{u} - \Delta \dot{u} + 2q^2 u \dot{u} = -2qu^2 h, \qquad (x, t) \in D \\ &\dot{u}|_{\partial\Omega} = 0, \qquad \dot{u}|_{t=0} = 0, \end{aligned} \tag{3.5}$$

and

$$\begin{aligned} &\partial_t \ddot{u} - \Delta \ddot{u} + 2q^2 u \ddot{u} = -4qu\dot{v}h - 4qu\dot{u}k - 2q^2\dot{u}\dot{v} - 2u^2hk, \qquad (x, t) \in D \\ &\ddot{u}|_{\partial\Omega} = 0, \qquad \ddot{u}|_{t=0} = 0, \end{aligned} \tag{3.6}$$

where $u = u(q)$ is determined by (3.1), *$\dot{u} = u'(q)h$ and $\dot{v} = u'(q)k$ are determined by* (3.5), *respectively.*

PROOF. First, we prove $u \in C^N(H; V)$.

Take, for example, $N = 0$. The proof is similar for any order N. Let $q, \bar{q} \in H$. Then it follows from (3.1) that $\bar{u} = u(\bar{q})$ and $u = u(q)$. Set $h = \bar{q} - q$ and $\delta u = \bar{u} - u$, so δu satisfies

$$\begin{aligned} \partial_t(\delta u) - \Delta(\delta u) + q^2(\bar{u} + u)(\delta u) &= -(\bar{q} + q)\bar{u}^2 h, \qquad (x, t) \in D, \\ (\delta u)|_{\partial\Omega} = 0, &\qquad (\delta u)|_{t=0} = 0. \end{aligned} \tag{3.7}$$

It is obvious by [14] that $\|\delta u\|_V = O(\|h\|_H)$. Hence $u(\cdot) \in C(H; V)$.

Secondly, we prove $u'(q)h = \dot{u}$. For $q, h \in H$, there exists a unique solution $\dot{u} \in V$ to the problem (3.5). Thus, set $\bar{q} = q + h$, $u = u(q)$, $\bar{u} = u(\bar{q})$, and $\tilde{u} = \bar{u} - u - \dot{u} = \delta u - \dot{u}$. It is evident that $\tilde{u}$ satisfies

$$\begin{aligned} \partial_t\tilde{u} - \Delta\tilde{u} + q^2(\bar{u} + u)\tilde{u} &= -q^2\dot{u}\delta u - (\bar{q} + q)(\bar{u} + u)h\delta u - u^2h^2, \qquad (x, t) \in D, \\ \tilde{u}|_{\partial\Omega} = 0, &\qquad \tilde{u}|_{t=0} = 0. \end{aligned}$$

It follows from (3.5) by [14] that $\dot{u} = O(\|h\|_H)$. Moreover, by the above argument we obtain $\delta u = o(1)$. Therefore, $\tilde{u} = o(\|h\|)$. Hence, $u'(q)h = \dot{u}$. It is similar to prove $u''(q)hk = \ddot{u}$.

Next, we give the following result.

THEOREM 3.2. *There exists an optimal element for the optimization problem* (3.4).

PROOF. Suppose that $\{q_n\}$ is a minimizing sequence for the optimization problem (3.4), that is,

$$J(q_n) \to h \equiv \inf_{q \in H} J(q). \tag{3.8}$$

Thus $\{q_n\}$ is bounded in H. Since H is a Hilbert space, there exists a subsequence, which is still denoted $\{q_n\}$, such that $q_n \overset{w}{\to} \bar{q}$, in H.[1] By the property of compact embeddedness of H, it follows that $q_n \overset{s}{\to} \bar{q}$ in $C^\beta(\bar{\Omega})$. Furthermore, by Theorem 3.1 we obtain that $u(q_n) \overset{s}{\to} u(\bar{q})$ in $C^{2+\beta, 1+\beta/2}(\bar{D})$.

We can write $J(q_n)$ as

$$J(q_n) = 1/2\|u(\cdot, T; q_n) - z\|^2_{L^2(\Omega)} + \alpha/2\|q_n\|^2_H \equiv J_1(q_n) + J_2(q_n). \tag{3.9}$$

Obviously, $J_2(q_n) \equiv \alpha/2\|q_n\|^2_H$ is convex and strongly lower semi-continuous, so it is also weakly lower semi-continuous. Letting $n \to \infty$ in (3.9) and considering (3.8), we obtain

$$J(\bar{q}) = 1/2\|u(\cdot, T; \bar{q}) - z\|^2_{L^2(\Omega)} + \alpha/2\|\bar{q}\|^2_H = \inf_{q \in H} J(q).$$

That is, $\bar{q}$ is the optimal element.

[1] "$x_n \overset{s}{\to} \bar{x}$ (or $x_n \overset{w}{\to} \bar{x}$), in X" means that x_n strongly (or weakly) converges to $\bar{x}$ in X.

For calculating $J'(q)$ we have the following results.

THEOREM 3.3. *The functional $J(q)$ defined by* (3.4) *is twice continuously Fréchet differentiable and its first Fréchet differential $J'(q)h$ is determined by the formula*

$$J'(q)h = \langle \tilde{j}(q), h \rangle, \qquad \forall h \in H, \tag{3.10}$$

where $\tilde{j}(q) \in H'$ is defined by

$$\tilde{j}(q) \equiv -2 \int_0^T qp(q)u^2(q)\,dt + \alpha K q, \tag{3.11}$$

$K : H \to H'$ is the canonical isometry, $u(q) = u$ is determined by (3.1), *and $p(q) = p$ is defined by the problem*

$$\begin{aligned} -\partial_t p - \Delta p + 2q^2 up = 0, \qquad (x,t) \in D, \\ p|_{\partial\Omega} = 0, \qquad p|_{t=T} = u(T;q) - z. \end{aligned} \tag{3.12}$$

Moreover, the second Fréchet differential $J''(q)hk$ is determined by the formula

$$\begin{aligned} J''(q)hk = \int_0^T \langle -4qu\dot{u}k - 4qu\dot{v}h - 2q^2\dot{u}\dot{v} - 2u^2hk, p\rangle\,dt \\ + \langle \dot{u}(T), \dot{v}(T)\rangle + \alpha(h,k), \qquad \forall h, k \in H, \end{aligned} \tag{3.13}$$

where $\dot{u} = u'(q)h$ and $\dot{v} = u'(q)k$ are determined by (3.5) *and $\langle\cdot,\cdot\rangle$ and $(\cdot,\cdot)$ denote the inner products in $L^2(\Omega)$ and H, respectively.*

PROOF. It follows from the calculation

$$\begin{aligned} &J(q+h) - J(q) - \{\langle \dot{u}(T;q), u(T;q) - z\rangle + \alpha(q,h)\} \\ &\quad = 1/2\langle u(T;q+h) - u(T;q) - u'(q)h(T), u(T;q+h) - z\rangle \\ &\qquad + 1/2\langle u'(q)h(T), u(T;q+h) - u(T;q)\rangle \\ &\qquad + 1/2\langle u(T;q) - z, u(T;q+h) - u(T;q) - u'(q)h(T)\rangle + \alpha/2\|h\|_H^2 \\ &\quad = o(\|h\|), \end{aligned}$$

that

$$J'(q)h = \langle \dot{u}(T;q), u(T;q) - z\rangle + \alpha(q,h), \qquad \forall h \in H. \tag{3.14}$$

Consideration of $\dot{u}(0;q) = 0$ and $p(T;q) = u(T;q) - z$ and use of Green's formula lead to

$$\begin{aligned}\langle \dot{u}(T;q), u(T;q) - z\rangle &= \langle \dot{u}(T;q), p(T;q)\rangle = \int_0^T \partial_t \langle \dot{u}, p\rangle\, dt \\ &= \int_0^T \{\langle \partial_t \dot{u}, p\rangle + \langle \dot{u}, \partial_t p\rangle\}\, dt = \int_0^T \{\langle \partial_t \dot{u}, p\rangle + \langle \dot{u}, -\Delta p + 2q^2 up\rangle\}\, dt \\ &= \int_0^T \langle \partial_t \dot{u} - \Delta \dot{u} + 2q^2 u\dot{u}, p\rangle\, dt = \left\langle -\int_0^T 2qu^2 p\, dt, h\right\rangle.\end{aligned} \tag{3.15}$$

Since H is a Sobolev space of order l with $l > 0$, H is compactly embedded in $L^2(\Omega)$. If we choose $L^2(\Omega)$ to be a pivot space, that is , $L^2(\Omega) = [L^2(\Omega)]'$, then

$$H \subset L^2(\Omega) \subset H'. \tag{3.16}$$

Because K is the canonical isometry from H onto H',

$$(q, h) = \langle Kq, h\rangle_{L^2(\Omega)}, \qquad \forall q, h \in H. \tag{3.17}$$

Substitution of (3.15) and (3.17) into (3.14) leads to

$$J'(q)h = \left\langle -2\int_0^T qu^2 p\, dt + Kq, h\right\rangle, \qquad \forall h \in H. \tag{3.18}$$

From (3.11) we see that (3.10) holds.

Furthermore, by (3.14) we obtain

$$\begin{aligned}J'(q+k)h &- J'(q)h - \{\langle u''(q)hk(T), u(T;q) - z\rangle + \langle \dot{u}(T), \dot{v}(T)\rangle + \alpha(h,k)\} \\ &= \{\langle u'(q+k)h(T), u(T;q+k) - z\rangle + \alpha(q+k, h)\} \\ &\quad - \{\langle u'(q)h(T), u(T;q) - z\rangle + \alpha(q,h)\} \\ &\quad - \{\langle \ddot{u}(T), u(T;q) - z\rangle + \langle \dot{u}(T), \dot{v}(T)\rangle + \alpha(h,k)\} \\ &= \langle u'(q+k)h(T) - u'(q)h(T) - \ddot{u}(T), u(T;q) - z\rangle \\ &\quad + \langle u'(q+k)h(T), u(T;q+k) - u(T;q) - \dot{v}(T)\rangle \\ &\quad + \langle u'(q+k)h(T) - u'(q)h(T), \dot{v}(T)\rangle = o(\|h\|^2 + \|k\|^2).\end{aligned}$$

Thus, $J(q)$ is twice Fréchet differentiable and its second Fréchet differential is

$$J''(q)hk = \langle \ddot{u}(T), u(T;q) - z\rangle + \langle \dot{u}(T), \dot{v}(T)\rangle + \alpha(h,k) \qquad \forall h, k \in H. \tag{3.19}$$

Using $u(T;q) - z = p(T;q)$ from (3.12) and using an argument similar to (3.15), we have

$$\begin{aligned}\langle \ddot{u}(T), u(T;q) - z\rangle &= \langle \ddot{u}(T), p(T;q)\rangle = \int_0^T \partial_t \langle \ddot{u}, p\rangle\, dt \\ &= \int_0^T \{\langle \partial_t \ddot{u}, p\rangle + \langle \ddot{u}, -\Delta p + 2q^2 up\rangle\}\, dt = \int_0^T \langle \partial_t \ddot{u} - \Delta \ddot{u} + 2q^2 u\ddot{u}, p\rangle\, dt \qquad (3.20) \\ &= \int_0^T \langle -4qu\dot{u}k - 4qu\dot{v}h - 2q^2\dot{u}\dot{v} - 2u^2hk, p\rangle\, dt.\end{aligned}$$

Substituting (3.20) for the first term of (3.19) we obtain immediately (3.13).

The following Lemma is quoted from [7].

LEMMA 3.4. *Let (S, Σ, μ) be a positive measure space. Then an operator A in the Hilbert space $L^2(S, \Sigma, \mu)$ is of Hilbert-Schmidt class if and only if there exists a $\mu \times \mu$ measurable function $A(\cdot, \cdot)$ on $S \times S$ such that*

$$\left\{\int_S \int_S |A(x,t)|^2 \mu(ds)\mu(dt)\right\}^{1/2} < \infty \qquad (3.21)$$

and such that

$$Af(s) \equiv \int_S A(s,t) f(t)\mu(dt), \qquad f \in L^2(S, \Sigma, \mu), \qquad (3.22)$$

for μ-almost all s. Moreover, $\|A\|_2$ is exactly equal to the finite quantity (3.21).

LEMMA 3.5. *The operator $I''(q) : L^2(\Omega) \to L^2(\Omega)$ is of Hilbert-Schmidt class, where $I(q)$ is defined by* (3.3).

PROOF. It is obvious that

$$I''(q)hk = \langle u'(T;q)k, u'(T;q)h\rangle + \langle u''(q)hk(T), u(T;q) - z\rangle \qquad (3.23)$$

and

$$I''(q)h = [u'(T;q)]^* u'(q)h(T) + [u''(T;q)h]^*[u(T;q) - z], \qquad (3.24)$$

where $[u'(T;q)]^*$ and $[u''(T;q)h]^*$ are the dual operators of $u'(T;q)$ and $u''(T;q)h$, respectively. Moreover, from [14] and problem (3.5), it is easy to see that

$$u'(T;q)h(x) = -2\int_0^T \int_\Omega G(x,\xi,T,\tau)q(\xi)u^2(\xi,\tau;q)h(\xi)\, d\xi d\tau, \qquad (3.25)$$

https://doi.org/10.1017/S0334270000012339 Published online by Cambridge University Press

where $G(x, \xi, t, \tau)$ is the Green function of the problem (3.5), and hence

$$[u'(T;q)]^* k(x) = -2q(x) \int_\Omega \int_0^T G(\xi, x, T, \tau) u^2(x, \tau; q)\, d\tau\, k(\xi)\, d\xi \qquad \forall k \in H. \tag{3.26}$$

Substitution of (3.25) into (3.26) and consideration of Lemma 3.4 lead to the conclusion that the operator

$$[u'(T;q)]^* u'(T;q) : L^2(\Omega) \to L^2(\Omega)$$

is of Hilbert-Schmidt class.

We need the following assumption.

A3. Suppose that there exists a $q^* \in H$ such that

$$I(q^*) = \min_{q \in H} I(q). \tag{3.27}$$

It is obvious that assumption **A3** implies $I'(q^*) = 0$ and $I''(q^*) \geq 0$. Thus,

$$J''(q^*)h^2 = I''(q^*)h^2 + \alpha(h, h) \geq \alpha \|h\|_H^2, \qquad \forall h \in H, \tag{3.28}$$

that is, $J''(q^*)$ is strictly positive. Therefore, q^* solves the following operator equation

$$K^{-1} I'(q^*) = 0. \tag{3.29}$$

When α is small, instead of (3.29) we consider the equation

$$f'(q) \equiv K^{-1} J'(q) \equiv K^{-1} I'(q) + \alpha q = 0. \tag{3.30}$$

Using the QNM algorithm stated in Section 2, we obtain an approximate sequence $\{q_n\}$, and its convergence is proved in the following theorem.

THEOREM 3.6. *Let assumptions* **A1**, **A2**, *and* **A3** *be true. The sequence* $\{q_n\}$ *determined by the* QNM *algorithm is superlinearly convergent providing that* q_0 *and* A_0 *satisfy the conditions of Lemma* 2.2 *and that* $A_0 - f''(q^*)$ *is compact.*

PROOF. Under assumptions **A1–A3**, the function $f'(q)$ defined by (3.30) satisfies assumptions **H1–H3** of Section 2. Thus, by Theorem 2.7 the conclusions of Theorem 3.6 follow immediately.

In order to solve (3.30) we used the QNM algorithm for the one-dimensional case, that is, $\Omega = (0, 1)$. Moreover, we take $H = H_0^1(\Omega)$, and hence $K = -\partial^2/\partial x^2$.

Suppose that $g(x) = \sin^2 \pi x$ and that the true parameter in (3.1) is $q_{tr}(x) = x(1-x)$. Then z in (3.2) is $z = u(T; q_{tr})$. We used the Crank-Nicholson implicit finite difference method, for example see [15], to discretize (3.1) and (3.2), that is, we could, for example, replace $\partial_t u(i\Delta x, j\Delta t)$ with

$$\partial_t u(i\Delta x, (j+1/2)\Delta t) \approx 2(u_{i,j+1/2} - u_{i,j})/\Delta t,$$

where $u_{i,j} = u(i\Delta x, j\Delta t)$ and $u_{i,j+1/2} = u(i\Delta x, (j+1/2)\Delta t)$.

Therefore one could obtain the three-level stable approximate equations

$$\frac{1}{\Delta x^2}(u_{i+1,j+1/2} - 2u_{i,j+1/2} + u_{i-1,j+1/2}) = \frac{2}{\Delta t}(u_{i,j+1/2} - u_{i,j}) + q_i^2 u_{i,j}^2$$

and

$$\begin{aligned} u_{i-1,j+1}/\Delta x^2 &- 2(1/\Delta x^2 + 1/\Delta t)u_{i,j+1} + u_{i+1,j+1}/\Delta x^2 \\ &= -u_{i-1,j}/\Delta x^2 + 2(1/\Delta x^2 - 1/\Delta t)u_{i,j} - u_{i+1,j}/\Delta x^2 + q_i^2 u_{i,j+1/2}^2. \end{aligned}$$

We also have similar discrete equations for problem (3.12).

In this paper we took $T = 1$, $\Delta x = 0.1$, $\Delta t = 0.05$, and $N_x = 1/\Delta x$.

The computational results are summarized in Table 1, where

$$\begin{aligned} \Delta q_i &= q(i\Delta x) - q_{tr}(i\Delta x), \\ m_1 &= \textstyle\sum_{i=0}^{N_x} \Delta q_i/(N_x + 1), \\ \sigma_1 &= \sqrt{\textstyle\sum_{i=0}^{N_x}[\Delta q_i - m_1]^2/N_x}, \\ s &= \sqrt{\textstyle\sum_{i=0}^{N_x}[\Delta q_i]^2}, \\ y_i &= u(i\Delta x) - z(i\Delta x) \\ m_2 &= \textstyle\sum_{i=0}^{N_x} y_i/(N_x + 1), \\ \sigma_2 &= \sqrt{\textstyle\sum_{i=0}^{N_x}[y_i - m_2]^2/N_x}, \\ J &= \Delta x \textstyle\sum_{i=0}^{N_x} y_i^2/2 + \alpha \sum_{i=0}^{N_x}[q((i+1)\Delta x) - q(i\Delta x)]^2/2\Delta x. \end{aligned}$$

Acknowledgement

This work has partially been supported by the National Nature Science Foundation of China.

TABLE 1. Convergence for the algorithm QNM

Iteration Times	m_1	σ_1	s	m_2	σ_2	J
1	-8.359	14.087	53.841	$1.25E{-5}$	$1.309E{-5}$	635.135
2	-2.057	4.466	38.353	$1.68E{-6}$	$1.724E{-6}$	78.117
3	0.849	$9.4E{-2}$	$4.37E{-5}$	$5.21E{-7}$	$7.459E{-7}$	$1.084E{-5}$
4	0.85	$9.06E{-2}$	$2.33E{-7}$	$3.22E{-7}$	$3.007E{-7}$	$1.381E{-13}$

References

[1] H. T. Banks and K. Kunisch, *Estimation Techniques for Distributed Parameter Systems* (Birkhäuser, Boston, 1989).

[2] H. T. Banks, S. Reich and I. G. Rosen, "An approximation theory for the identification of nonlinear distributed parameter systems", *SIAM J. Control and Optim.* **28** (1990) 523–569.

[3] C. G. Broyden, J. E. Dennis and J. J. Moore, "On the local and superlinear convergence of quasi-Newton methods", *J. Inst. Math. Appl.* **12** (1973) 223–245.

[4] G. Chavent, "Identification of distributed parameter systems: About the output least squares method, its implementation and identifiability", in: *Identification and System Parameter Estimation* (ed. R. Isermann) (Pergamon, New York, 1982) 85–97.

[5] Y. M. Chen and F. G. Zhang, "Hierarchical multigrid strategy for efficiency improvement of the GPST inversion algorithm", *Applied Numerical Mathematics* **6** (1990) 431–446.

[6] T. F. Coleman and A. R. Conn, "On the local convergence of a quasi-Newton method for the nonlinear programming problem", *SIAM J. Numer. Anal.* **21** (1984) 755–769.

[7] N. Dunford and J. T. Schwartz, *Linear Operators, Part II: Spectral Theory* (Interscience, New York, 1963).

[8] A. Griewank, "The local convergence of Broyden-like methods on Lipchitzian problems in Hilbert spaces", *SIAM J. Numer. Anal.* **24** (1987) 684–705.

[9] D. M. Hwang and C. T. Kelley, "Sequential quadratic programming for parameter identification problems", in: *Fifth Symp. on Control of Distributed Parameter Systems*, (ed. A. E. Jai and M. Amouroux) (1989) 105–109.

[10] K. Ito and K. Kunisch, "The augmented Lagrangian method for equality and inequality constrained problems in Hilbert spaces", *Math. Programming* **46** (1990) 341–360.

[11] K. Ito and K. Kunisch, "The augmented Lagrangian method for parameter estimation in elliptic systems", *SIAM J. Control and Optim.* **28** (1990) 113–136.

[12] C. T. Kelley and E. W. Sachs, "Quasi-Newton methods and unconstrained optimal control problems", *SIAM J. Control and Optim.* **23** (1987) 1503–1516.

[13] C. K. Kunisch and E. W. Sachs, "Reduced SQP methods for parameter identification problems", *SIAM J. Numer. Anal.* **29** (1992) 1793–1820.

[14] O. A. Ladyzenskaja et al., *Linear and Quasi-linear Equations of Parabolic Type*, (AMS, Providence, 1968).

[15] L. Lapidus and G. F. Pinder, *Numerical Solution of Partial Differential Equations in Sciences and Engineering* (John Wiley and Sons, New York, 1982).

[16] Tai-Yong Lee and John H. Seinfeld, "Estimation of two-phase petroleum reservoir properties by regularization", *J. Comput. Phys.* **69** (1987) 397–419.

[17] J. H. Seinfeld and W. H. Chen, "Identification of petroleum reservoir properties", in: *Distributed Parameter Systems: Identification, Estimation and Control* (ed. W. H. Ray amd D. G. Lainiotis)

(Marcel Dekker, New York, 1978).

[18] W. H. Yu, "A Rapid convergent iteration method for identifying distributed parameters in partial differential equations", *Preprints 1988 'IFAC Symp. Identification and System Parameter Estimation* (Beijing, 1988).

[19] W. H. Yu and J. H. Seinfeld, "Identification of distributed paremeter systems by regularization with differential operator", *J. Math. Anal. Appl.* **132** (1988) 365–398.

[20] W. H. Yu and J. H. Seinfeld, "Identification of distributed parameter systems with pointwise constrains on the parameters", *J. Math. Anal. Appl.* **136** (1988) 497–520.

[21] W. H. Yu, "Solving inverse problems for hyperbolic equations via the regularization method", *J. Computational Mathematics* **11** (1993) 142–153.

[22] Jianping Zhu and Yung Ming Chen, "Multilevel grid method for history matching multi-dimensional multi-phase reservoir models", *Applied Numer. Math.* **10** (1992) 159–174.