Journal of Computational Mathematics, Vol.17, No.3, 1999, 315–326.

A FINITE DIMENSIONAL METHOD FOR SOLVING NONLINEAR ILL-POSED PROBLEMS*1)

Qi-nian Jin
(*Institute of Mathematics, Nanjing University, Nanjing 210008, China*)
Zong-yi Hou
(*Institute of Mathematics, Fudan University, Shanghai 200433, China*)

Abstract

We propose a finite dimensional method to compute the solution of nonlinear ill-posed problems approximately and show that under certain conditions, the convergence can be guaranteed. Moreover, we obtain the rate of convergence of our method provided that the true solution satisfies suitable smoothness condition. Finally, we present two examples from the parameter estimation problems of differential equations and illustrate the applicability of our method.

Key words: Nonlinear ill-posed problems, Finite dimensional method, Convergence and convergence rates.

1. Introduction

In this paper we consider the nonlinear problems of the form

$$F(x) = y_0, \tag{1}$$

where $F : D(F) \subset X \to Y$ is a nonlinear operator between real Hilbert spaces X and Y and $y_0 \in R(F)$. The norms in X and Y will be denoted by $\|\cdot\|_X$ and $\|\cdot\|_Y$ respectively. We are mainly interested in those problems of the form (1) for which the solution does not depend continuously on the right hand side. Such problems are called ill-posed. We refer to [5] for a number of important inverse problems in natural sciences which lead to such ill-posed problems.

Let L be a linear operator

$$L : D(L) \subset X \to Z$$

with Z a Hilbert space (the norm is denoted by $\|\cdot\|_Z$) and $D(F) \cap D(L) \neq \emptyset$. L need not be bounded and allows us to define a seminorm $|\cdot|$ on $D(L)$ by means of $|x| := \|Lx\|_Z$. Let $(x, z)_L := (x, z)_X + (Lx, Lz)_Z$ for each pair $x, z \in D(L)$, then $(\cdot, \cdot)_L$

* Received October 31, 1996.
1)This work is carried out during the first author's stay at Fudan University and is supported by National Natural Science Foundation of China.

is an inner product on $D(L)$, and the induced norm is denoted by $\|z\|_L = \sqrt{(z,z)_L}$ for $z \in D(L)$. If L is closed then $(D(L), \|\cdot\|_L)$ forms a Hilbert space.

Now we choose a concept of "solution" for problem (1). An element $x_0 \in X$ is called an x^*-minimum-seminorm-solution (x^*-MSS) of problem (1) for given $y_0 \in Y$ and $x^* \in D(L)$ if

$$F(x_0) = y_0,$$

and

$$\|Lx_0 - Lx^*\|_Z = \inf\{\|Lx - Lx^*\|_Z \mid F(x) = y_0,\ x \in D(F) \cap D(L)\},$$

where $x^* \in D(L)$ is an *a priori* guess of x_0 and it plays the role of a selection criterion. In the following we always assume the existence of an x^*-MSS x_0 for problem (1). Due to the nonlinearity of F, this solution need not be unique.

Since in practice, we often only know the approximation data y_δ of y_0, $\|y_\delta - y_0\| \le \delta$, regularization technique is required to obtain a reasonable solution to x^*-MSS x_0 due to the ill-posedness of problem (1). Tikhonov regularization is the well known method. In [7], Tikhonov regularization method with the seminorm $|\cdot|$ in the regularization term was introduced and the solution x_α^δ of the minimization problem

$$\min_{x \in D(F) \cap D(L)} \{\|F(x) - y_\delta\|_Y^2 + \alpha \|Lx - Lx^*\|_Z^2\} \tag{2}$$

was used to approximate the x^*-MSS of problem (1). By suitable choice of the regularization parameter α, convergence and convergence rate of x_α^δ were obtained. The practical advantage of allowing for regularization with an operator L is given by the fact that one can realize seminorm regularization terms, which penalize undesired oscillations in the numerical solution without significantly affecting its low modes.

In this paper we will present a finite dimensional method for solving nonlinear ill-posed problems. We describe the method in Section 2 and show that this method is well-defined and prove the existence of the approximate solutions. It is easy to see that our method can be viewed as a modified form of the (generalized) Marti's method if F is a linear operator[10,15]. The analysis for convergence and convergence rates are presented in Section 3 and two examples from the parameter estimation problems of differential equations are given in Section 4 to illustrate the reasonability of our assumptions and the applicability of our method. For the finite-dimensional approximation of Tikhonov regularization of nonlinear ill-posed problems, F is required to be compact[11,13]. For our method, this is not necessary.

2. The Description of the Method

Let x_0 be the sought x^*-MSS of (1). Let $\{P_n\}$ be a sequence of bounded linear operators of finite rank on X such that $P_n x_0 \in D(F) \cap D(L)$ for sufficiently large n, and we can choose positive number sequences $\{b_n\}$ and $\{c_n\}$ such that

$$\|P_n x_0 - x_0\|_X = o(b_n), \quad \lim_{n\to\infty} b_n = 0, \tag{3}$$

$$\|L(P_n x_0 - x_0)\|_Z = O(c_n), \quad \lim_{n\to\infty} c_n = 0. \tag{4}$$

Let F_n be the approximations of F with the properties that $D(F_n) = D(F)$ for all n and $\forall \epsilon > 0$ there is a constant $c(\epsilon)$ depending only on ϵ such that

$$\|F_n(x) - F(x)\|_Y \le C(\epsilon)\rho_n, \quad \forall x \in D(F) \cap U_\epsilon(x_0), \tag{5}$$

where $\rho_n = o(1)$ and $U_\epsilon(x_0) := \{x \in X \mid \|x - x_0\|_X \le \epsilon\}$.

Let y_n be the observation data of y_0 such that

$$\|y_n - y_0\|_Y \le \delta_n, \tag{6}$$

where δ_n is assumed to be known and

$$\lim_{n\to\infty} \delta_n = 0.$$

Now we can define the set

$$S_n := \{x \mid \|F_n(x) - y_n\|_Y \le b_n + \delta_n + c(1)\rho_n, \quad x \in P_n(X) \cap D(F) \cap D(L)\} \tag{7}$$

and construct x_n as follows:

$$x_n \in S_n \quad \text{and} \quad \|Lx_n - Lx^*\|_Z = \inf\{\|Lx - Lx^*\|_Z \mid x \in S_n\}. \tag{8}$$

We will use the sequence $\{x_n\}$ to approximate the x^*-MSS x_0 of (1).

To show our method is well defined, we need the following assumptions.

Assumption 1. 1) $F : D(F) \subset X \to Y$ is Fréchet differentiable at each point $x \in D(F)$ with Fréchet derivative $F'(x) \in \mathcal{L}(X, Y)$, the adjoint of $F'(x)$ is denoted by $F'(x)^*$;

2) Let $0 < \gamma_1 \le 2$, $0 \le \gamma_2 \le 1$ and $\gamma_1 + \gamma_2 \ge 1$. $\forall \epsilon > 0$, there exists $\eta := \eta(\epsilon) > 0$ such that for all $x \in D(F) \cap U_\epsilon(x_0)$ there holds

$$\|F(x) - F(x_0) - F'(x_0)(x - x_0)\|_Y \le \eta\|x - x_0\|_X^{\gamma_1}\|F(x) - F(x_0)\|_Y^{\gamma_2}.$$

Assumption 2. 1) For each fixed n, $(F_n, L) : D(F) \cap D(L) \subset X \to Y \times Z$ is weakly closed, i.e., for any sequence $\{x_k\} \subset D(F) \cap D(L)$, if $x_k \rightharpoonup x$ in X, $(F_n(x_k), Lx_k) \rightharpoonup (y, z)$ in $Y \times Z$, then $x \in D(F) \cap D(L)$ and $F_n(x) = y$, $Lx = z$. (Here "$\rightharpoonup$" denotes the weakly convergence).

2) For each fixed n, let $\{x_k\}$ be a sequence in $D(F) \cap D(L)$, if $\{(F_n(x_k), Lx_k)\}$ is bounded in $Y \times Z$, then $\{x_k\}$ is bounded in X.

Now we are in a position to prove the well-definedness of our method. we first prove that for sufficiently large n there holds

$$\|F(P_n x_0) - F(x_0)\|_Y \le b_n. \tag{9}$$

In fact, equation (3) implies that $\|P_n x_0 - x_0\|_X \le 1$ for sufficiently large n. Therefore, Assumption 1 2) can be applied to obtain (here $\eta := \eta(1)$)

$$(\|F(P_n x_0) - F(x_0)\|_Y^{1-\gamma_2} - \eta\|P_n x_0 - x_0\|_X^{\gamma_1})\|F(P_n x_0) - F(x_0)\|_Y^{\gamma_2}$$

$$\leq \|F'(x_0)(P_n x_0 - x_0)\|_Y. \tag{10}$$

If $\|F(P_n x_0) - F(x_0)\|_Y^{1-\gamma_2} \leq 2\eta \|P_n x_0 - x_0\|_X^{\gamma_1}$, then

$$\|F(P_n x_0) - F(x_0)\|_Y \leq (2\eta)^{\frac{1}{1-\gamma_2}} \|P_n x_0 - x_0\|_X^{\frac{\gamma_1}{1-\gamma_2}}.$$

Since $\dfrac{\gamma_1}{1-\gamma_2} \geq 1$, from (3) we immediately obtain (9) providing n large enough.

If $\|F(P_n x_0) - F(x_0)\|_Y^{1-\gamma_2} \geq 2\eta \|P_n x_0 - x_0\|_X^{\gamma_1}$, then (10) gives

$$\frac{1}{2}\|F(P_n x_0) - F(x_0)\|_Y \leq \|F'(x_0)(P_n x_0 - x_0)\|_Y.$$

Therefore

$$\|F(P_n x_0) - F(x_0)\|_Y \leq 2\|F'(x_0)\|_{\mathcal{L}(X,Y)} \|P_n x_0 - x_0\|_X,$$

and we again obtain (9) providing n large enough.

Thus from (5), (6) and (9) we get for sufficiently large n that

$$\begin{aligned}\|F_n(P_n x_0) - y_n\|_Y \leq& \|F(P_n x_0) - F(x_0)\|_Y + \|y_n - y_0\|_Y + \|F_n(P_n x_0) - F(P_n x_0)\|_Y \\ \leq& b_n + \delta_n + c(1)\rho_n.\end{aligned}$$

This implies that $P_n x_0 \in S_n$. Therefore, without loss of generality, in the following we always assume that $S_n \neq \emptyset$ for all n.

Now for each fixed n, we define

$$d_n := \inf\{\|Lx - Lx^*\|_Z \mid x \in S_n\},$$

and let $\{x_n^{(k)}\}$ be a minimizing sequence. Therefore $x_n^{(k)} \in P_n(X) \cap D(F) \cap D(L)$, and

$$\|F_n(x_n^{(k)}) - y_n\|_Y \leq b_n + \delta_n + c(1)\rho_n, \quad \lim_{k\to\infty} \|Lx_n^{(k)} - Lx^*\|_Z = d_n.$$

This implies that $\{(F_n(x_n^{(k)}), Lx_n^{(k)})\}$ is bounded in $Y \times Z$. Thus from Assumption 2 2) we know that $\{x_n^{(k)}\}$ is bounded in X. Since a bounded set in Hilbert space always has a weakly convergent subsequence, there is a subsequence $\{x_n^{(k_l)}\}$ and elements $x_n \in X$, $(\bar{y}_n, \bar{z}_n) \in Y \times Z$ such that

$$x_n^{(k_l)} \rightharpoonup x_n \text{ in } X, \quad (F_n(x_n^{(k_l)}), Lx_n^{(k_l)}) \rightharpoonup (\bar{y}_n, \bar{z}_n) \text{ in } Y \times Z.$$

Hence Assumption 2 1) implies that $x_n \in D(F) \cap D(L)$, $F_n(x_n) = \bar{y}_n$ and $Lx_n = \bar{z}_n$. By the weak lower semicontinuity of Hilbert space norm we have

$$\|F_n(x_n) - y_n\|_Y \leq \liminf_{l\to\infty} \|F_n(x_n^{(k_l)}) - y_n\|_Y \leq b_n + \delta_n + c(1)\rho_n.$$

Since a closed subspace in Banach space is always weakly closed, from $x_n^{(k_l)} \in P_n(X)$ and $x_n^{(k_l)} \rightharpoonup x_n$ we know $x_n \in P_n(X)$. Hence $x_n \in S_n$. Note that

$$\|Lx_n - Lx^*\|_Z \leq \liminf_{l\to\infty} \|Lx_n^{(k_l)} - Lx^*\|_Z = d_n,$$

we have $\|Lx_n - Lx^*\|_Z = d_n$. Therefore we obtain the existence of x_n defined by (8).

3. Convergence and Convergence Rates

In this section we present the analysis for convergence and convergence rates of our method. The following additional assumptions are needed.

Assumption 3. $(F, L) : D(F) \cap D(L) \subset X \mapsto Y \times Z$ is weakly closed.

Assumption 4. Let $\{x_k\} \subset D(L)$. If $x_k \rightharpoonup x$ in X and $\|Lx_k\|_Z \to \|Lx\|_Z$, then $x_k \to x$ in X.

We first give the convergence result for $\{x_n\}$.

Theorem 1. *Let Assumptions* 1–4 *be fulfilled and* $D(F)$ *be bounded in* X*, let* $\{x_n\}$ *be the sequence defined by* (8)*. Then there is a subsequence of* $\{x_n\}$ *convergent in* $(D(L), \|\cdot\|_L)$ *and the limit is an* x^**-MSS of* (1)*. If in addition, the* x^**-MSS* x_0 *of* (1) *is unique, then*

$$\lim_{n\to\infty} \|x_n - x_0\|_L = 0.$$

Proof. According to the analysis in Section 2 we know there is an N such that $P_n x_0 \in S_n$ for all $n \geq N$. Hence from the definition of x_n it follows that

$$\|Lx_n - Lx^*\|_Z \leq \|L(P_n x_0 - x^*)\|_Z \leq \|L(P_n x_0 - x_0)\|_Z + \|L(x_0 - x^*)\|_Z. \tag{11}$$

Since $x_n \in S_n$, we have

$$\|F_n(x_n) - y_n\|_Y \leq b_n + \delta_n + c(1)\rho_n.$$

Due to the boundedness of $D(F)$ in X, there is a constant M such that $\|x_n - x_0\|_X \leq M$, Therefore (5) implies that

$$\begin{aligned} \|F(x_n) - y_n\|_Y &\leq \|F_n(x_n) - F(x_n)\|_Y + \|F_n(x_n) - y_n\|_Y \\ &\leq b_n + \delta_n + c(1)\rho_n + c(M)\rho_n. \end{aligned} \tag{12}$$

Combining the above we know that $\{(x_n, F(x_n), Lx_n)\}$ is bounded in $X \times Y \times Z$. Thus there is a subsequence $\{x_{n_k}\}$ and elements $\bar{x} \in X$, $\bar{y} \in Y$, $\bar{z} \in Z$ such that

$$x_{n_k} \rightharpoonup \bar{x} \text{ in } X, \quad (F(x_{n_k}), Lx_{n_k}) \rightharpoonup (\bar{y}, \bar{z}) \text{ in } Y \times Z.$$

Now Assumptiom 3 implies that $\bar{x} \in D(F) \cap D(L)$, $F(\bar{x}) = \bar{y}$ and $L\bar{x} = \bar{z}$. But from (12) we know

$$\lim_{n\to\infty} \|F(x_n) - y_0\|_Y = 0.$$

Therefore $\bar{y} = y_0$. Using (4), (11) and the weak lower semicontinuity of Hilbert space norm we have

$$\begin{aligned} \|L\bar{x} - Lx^*\|_Z &\leq \liminf_{k\to\infty} \|Lx_{n_k} - Lx^*\|_Z \\ &\leq \limsup_{k\to\infty} \|Lx_{n_k} - Lx^*\|_Z \leq \|Lx_0 - Lx^*\|_Z. \end{aligned}$$

Since x_0 is the x^*-MSS of (1), we have

$$\|L\bar{x} - Lx^*\|_Z = \|Lx_0 - Lx^*\|_Z, \quad \lim_{k\to\infty} \|Lx_{n_k} - Lx^*\|_Z = \|L\bar{x} - Lx^*\|_Z. \tag{13}$$

This implies that $\bar{x}$ is an x^*-MSS of (1), and from

$$\|Lx_{n_k} - L\bar{x}\|_Z^2 = \|Lx_{n_k} - Lx^*\|_Z^2 - 2(Lx_{n_k} - Lx^*, L\bar{x} - Lx^*)_Z + \|L\bar{x} - Lx^*\|_Z^2$$

we know

$$\lim_{k\to\infty} \|Lx_{n_k} - L\bar{x}\|_Z = 0.$$

On the other hand, since $x_{n_k} \rightharpoonup \bar{x}$ in X, from (13) and Assumption 4 it follows that

$$\lim_{k\to\infty} \|x_{n_k} - \bar{x}\|_X = 0.$$

Hence

$$\lim_{k\to\infty} \|x_{n_k} - \bar{x}\|_L = 0.$$

If the x^*-MSS x_0 of (1) is unique, from above we know that each subsequence of $\{x_n\}$ has a subsequence convergent to x_0 in $(D(L), \|\cdot\|_L)$. Therefore

$$\lim_{n\to\infty} \|x_n - x_0\|_L = 0.$$

The above theorem provides us only the convergence result of $\{x_n\}$. When the x^*-MSS x_0 of (1) satisfies suitable smoothness condition, we can derive the convergence rate for $\{x_n\}$.

Lemma 1. *Let X, Y and Z be Hilbert spaces, let $L : D(L) \subset X \to Z$ be a closed linear operator with $R(L)$ closed in Z. If $N(L) \cap N(F'(x_0)) = \{0\}$ and there is a constant $\beta > 0$ such that $\|F'(x_0)v\|_Y \geq \beta\|v\|_X$ for all $v \in N(L)$, then there exists a constant $\kappa > 0$ such that*

$$\|F'(x_0)x\|_Y^2 + \|Lx\|_Z^2 \geq \kappa^2 \|x\|_L^2, \quad \forall x \in D(L).$$

Proof. Please refer to [9].

Theorem 2. *Let Assumptions 1–3 hold and $D(F)$ be bounded in X, let $q_n := \max\{b_n, c_n\}$, let $L : D(L) \subset X \to Z$ be a closed, densely defined linear operator with $R(L)$ closed in Z and $N(L) \cap N(F'(x_0)) = \{0\}$, and there is a constant $\beta > 0$ such that $\|F'(x_0)v\|_Y \geq \beta\|v\|_X$ for all $v \in N(L)$. Let $x^* \in D(F) \cap D(L)$ be chosen such that $\|P_n x^* - x^*\|_L \leq O(q_n)$. If $x_0 - x^* \in D(L^*L)$ and there is an $w \in Y$ such that $L^*L(x_0 - x^*) = F'(x_0)^*w$, then for the sequence $\{x_n\}$defined by (10), we have*

$$\|x_n - x_0\|_L \leq O(\sqrt{q_n + \delta_n + \rho_n}) \quad as \quad n \to \infty$$

if one of the following conditions holds:

i) *$\gamma_1 = 2$, $\gamma_2 = 0$, $2\eta\|w\|_Y < \kappa^2$ with $\eta := \eta(2\|x_0 - x^*\|_X)$ and $x_n \to x_0$ in X as $n \to \infty$;*

ii) *$0 < \gamma_1 \leq 2$, $0 < \gamma_2 \leq 1$ and $\gamma_1 + 2\gamma_2 \geq 2$.*

Proof. Obviously we always have $\gamma_1 + \gamma_2 \geq 1$. Thus from the argument in Section 2 it follows that there is an N_0 such that $P_n x_0 \in S_n$ for all $n \geq N_0$. Therefore, from (11) and (4) we can obtain

$$\|Lx_n - Lx^*\|_Z \leq \|Lx_0 - Lx^*\|_Z + O(q_n).$$

Hence

$$\begin{aligned}\|Lx_n - Lx_0\|_Z^2 &= \|Lx_n - Lx^*\|_Z^2 + \|Lx_0 - Lx^*\|_Z^2 - 2(Lx_n - Lx^*, Lx_0 - Lx^*)_Z \\ &\leq 2\|Lx_0 - Lx^*\|_Z^2 - 2(Lx_n - Lx^*, Lx_0 - Lx^*)_Z + O(q_n) \\ &= 2(Lx_0 - Lx_n, Lx_0 - Lx^*)_Z + O(q_n) \\ &= 2(x_0 - x_n, L^*L(x_0 - x^*))_X + O(q_n).\end{aligned}$$

Applying the assumptions on $x_0 - x^*$ yields

$$\begin{aligned}\|Lx_n - Lx_0\|_Z^2 &\leq 2(x_0 - x_n, F'(x_0)^* w)_X + O(q_n) \\ &= 2(F'(x_0)(x_0 - x_n), w)_Y + O(q_n) \\ &\leq 2\|w\|_Y \|F'(x_0)(x_n - x_0)\|_Y + O(q_n).\end{aligned} \tag{14}$$

Since $D(F)$ is bounded in X, there is a constant M such that $\|x_n - x_0\|_X \leq M$. Hence by Assumption 1 2) we have (here $\eta := \eta(M)$)

$$\|F'(x_0)(x_n - x_0)\|_Y \leq D_n + \eta E_n^{\gamma_1} D_n^{\gamma_2}. \tag{15}$$

where we use the abbreviations $D_n := \|F(x_n) - F(x_0)\|_Y$ and $E_n := \|x_n - x_0\|_X$. Therefore, by adding $\|F'(x_0)(x_0 - x_n)\|_Y^2$ to the both sides of (14) and using Lemma 1 and (15) we can get

$$\begin{aligned}\kappa^2\|x_n - x_0\|_L^2 \leq& \|Lx_n - Lx_0\|_Z^2 + \|F'(x_0)(x_n - x_0)\|_Y^2 \\ \leq& 2\|w\|_Y(D_n + \eta E_n^{\gamma_1} D_n^{\gamma_2}) + 2D_n^2 + 2\eta^2 E_n^{2\gamma_1} D_n^{2\gamma_2} + O(q_n) \\ =& 2(\|w\|_Y + D_n)D_n + 2\eta(\|w\|_Y + \eta E_n^{\gamma_1} D_n^{\gamma_2}) D_n^{\gamma_2} E_n^{\gamma_1} + O(q_n).\end{aligned} \tag{16}$$

Now we give the estimate of D_n. Noting the boundedness of $\{x_n\}$ in X, from (5) we can choose a constant τ_0 independent of n such that $\|F_n(x_n) - F(x_n)\|_Y \leq \tau_0 \rho_n$. Therefore

$$\begin{aligned}D_n \leq& \|F_n(x_n) - y_n\|_Y + \|y_n - y_0\|_Y + \|F_n(x_n) - F(x_n)\|_Y \\ \leq& b_n + \delta_n + c(1)\rho_n + \delta_n + \tau_0\rho_n = O(b_n + \delta_n + \rho_n).\end{aligned} \tag{17}$$

In the following we are going to give the proof of the assertion.

i) When $\gamma_1 = 2$ and $\gamma_2 = 0$, we have from (16) that

$$\kappa^2\|x_n - x_0\|_L^2 \leq 2(\|w\|_Y + D_n)D_n + (2\eta\|w\| + 2\eta^2 E_n^2)\|x_n - x_0\|_L^2 + O(q_n). \tag{18}$$

Since $x_n \to x_0$ in X, we have $E_n \leq 2\|x_0 - x^*\|_X$ for sufficiently large n. Therefore η appearing in (18) can be chosen as $\eta := \eta(2\|x_0 - x^*\|_X)$. Since $2\eta\|w\|_Y < \kappa^2$, we can choose a constant $q>0$ and an integer N_1 such that $\kappa^2-(2\eta\|w\|_Y+2\eta^2E_n^2)\geq q$ for all $n\geq N_1$. Therefore $q\|x_n-x_0\|_L^2\leq 2(\|w\|_Y+D_n)D_n+O(q_n)$. By applying (17) we obtain the desired result immediately.

ii) We first assume $\gamma_1<2$. Applying the implication (cf. [12])

$$a,b,c\geq 0,\ p>q\geq 0,\ a^p\leq c+ba^q \Longrightarrow a^p\leq O(c+b^{\frac{p}{p-q}})$$

to (16) and noting $\dfrac{2\gamma_2}{2-\gamma_1}\geq 1$ it follows that

$$\|x_n-x_0\|_L^2\leq O(D_n+D_n^{\frac{2\gamma_2}{2-\gamma_1}}+q_n)\leq O(q_n+\delta_n+\rho_n). \tag{19}$$

Now we consider the case $\gamma_1=2$. From (17) it follows that $2\eta(\|w\|+\eta E_n^2D_n^{\gamma_2})D_n^{\gamma_2}\to 0$ as $n\to\infty$. Therefore, by noting that

$$(\kappa^2-2\eta(\|w\|+\eta E_n^2D_n^{\gamma_2})D_n^{\gamma_2})\|x_n-x_0\|_L^2\leq 2(\|w\|_Y+D_n)D_n+O(q_n),$$

we immediately obtain $\|x_n-x_0\|_L^2\leq O(q_n+\delta_n+\rho_n)$.

Summing up, the proof is complete.

Remark 1. 1) In the above two theorems, we have assume that $D(F)$ is bounded in X. This is frequently used in parameter estimation problems (cf. [1,8]).

2) The assumption that $D(F)$ is bounded in X is only needed to guarantee the boundedness of $\{x_n\}$ in X. If $L:=I$ =identity, this assumption is not necessary.

3) When $L:=I$, the assumptions on L in Theorem 2 and Assumptions 2 2) and 4 hold automatically and (4) is superfluous, the number κ appearing in Theorem 2 should be replaced by $\kappa=1$ and $x_n\to x_0$ in X is not needed, and we also have $\|x_n-x_0\|_X\leq O(\sqrt{b_n+\delta_n+\rho_n})$.

4) To obtain the convegence rates, the following assumption has been assumed in many papers (cf. [4,7,11]):

$\forall\epsilon>0$, there is a constant γ such that

$$\|F'(x)-F'(x_0)\|\leq\gamma\|x-x_0\|_X,\quad \forall x\in D(F)\cap U_\epsilon(x_0).$$

From this we can easily derive

$$\|F(x)-F(x_0)-F'(x_0)(x-x_0)\|\leq\frac{\gamma}{2}\|x-x_0\|^2.$$

Therefore this is a special case of Assumption 1 2) with $\gamma_1=2$, $\gamma_2=0$.

5) When $0<\gamma_1<2$, $0<\gamma_2<1$ and $\gamma_1+\gamma_2\geq 1>\frac{1}{2}\gamma_1+\gamma_2$, we can also obtain the convergence rate for $\{x_n\}$, but now the convergence rate is $O((q_n+\delta_n+\rho_n)^{\frac{2\gamma_2}{2-\gamma_1}})$, not $O(\sqrt{q_n+\delta_n+\rho_n})$. As a matter of fact, by noting that $\dfrac{2\gamma_2}{2-\gamma_1}<1$ we obtain this assertion from the proof of Theorem 2 at once.

6) When the condition ii) in Theorem 2 is fulfilled, the smallness condition $2\eta\|w\|_Y<\kappa^2$ can be removed.

4. Examples

In this section we consider two parameter estimation problems of differential equations to illustrate the applicability of our method. These two problems are both nonlinear and ill-posed and have been studied in [4,13] by Tikhonov regularization.

Example 1. We treat the problem of estimating the parameter c in the two point boundary value problem

$$-u_{xx} + cu = f \quad \text{in } (0,1), \tag{20}$$

$$u(0) = 0 = u(1) \tag{21}$$

from the noise measurement $u_\delta \in L^2[0,1]$. we assume $c_0 \geq 0$, $\|c_0\|_{L^2} \leq K$ be the sought solution corresponding to the unperturbed observation u_0, i.e. $u(c_0) = u_0$. Here $u(c_0) \in H_0^1[0,1] \cap H^2[0,1]$ denotes the solution of (20), (21) with $c = c_0$ and $f \in L^2[0,1]$, $f \neq 0$ and K is a given constant.

To implement of our method, we choose $X = Y = Z := L^2[0,1]$ and $L := \dfrac{\mathrm{d}}{\mathrm{d}x}$, and define the nonlinear operator F by

$$F : D(F) := \{c \in L^2[0,1] \mid c \geq 0 \text{ a.e.}, \|c\|_{L^2} \leq K+1\} \subset L^2[0,1] \mapsto L^2[0,1],$$
$$c \mapsto F(c) := u(c).$$

Clearly $D(L) = H^1[0,1]$ and $L : D(L) \subset L^2[0,1] \mapsto L^2[0,1]$ is closed, densely defined and surjective. Since F is weakly closed[3], Assumption 3 follows. Assumption 2 2) follows immediately from the definition of $D(F)$. It is well known that[4] F is Fréchet differentiable with

$$F'(c)h := -A(c)^{-1}(hu(c)), \qquad c \in D(F),\ h \in L^2[0,1],$$

and satisfies Assumption 1 with $\gamma_1 = 2$ and $\gamma_2 = 0$, where $A(c) : H_0^1 \cap H^2[0,1] \mapsto L^2[0,1]$ is defined by $A(c)u := -u_{xx} + cu$. Note that $\dim N(L) = 1$ and $R(L) = Z$, Assumption 4 follows fron [7, Lemma 1].

Now suppose that $h \in N(L) \cap N(F'(c_0))$. This implies that h is a constant and $hA(c)^{-1}(u(c)) = 0$. If $h \neq 0$ then this implies that $u(c_0) = 0$ which can happen only if $f = 0$, which is excluded. Since $\dim N(L) = 1$, there is a constant $\beta > 0$ such that $\|F'(c_0)h\|_{L^2} \geq \beta \|h\|_{L^2}$ for all $h \in N(L)$.

To give the finite dimensional approximation, let $P_n(X)$ be the space of the linear splines on a uniform grid of $n+1$ points in $[0,1]$. If c_0 satisfies: $c_0 \in H^2[0,1]$, $c_0 > 0$, then from [14, Corollary 7.3] we have

$$\lim_{n\to\infty} \|P_n c_0 - c_0\|_{L^\infty} = 0.$$

This implies that $P_n c_0 \in D(F) \cap D(L)$ for sufficiently large n.

From [14] we also have

$$\|P_n c_0 - c_0\|_{L^2} \leq O(n^{-2}\|c_0\|_{H^2}),$$

$$\|L(P_n c_0 - c_0)\|_{L^2} \leq O(n^{-1}\|c_0\|_{H^2}).$$

Thus we can choose the quantities b_n and c_n appearing in our method to be $b_n = n^{-2}\log n$ and $c_n = n^{-1}$, hence $q_n = O(n^{-1})$. To define the approximation F_n of F, we choose Y_n be the space of linear splines on a uniform grid of $n+1$ points in $[0,1]$, vanishing at 0 and 1, and define F_n by

$$\begin{aligned} F_n : & D(F) \subset L^2[0,1] \mapsto L^2[0,1], \\ & c \mapsto F_n(c) := u_n(c), \end{aligned}$$

where $u_n(c)$ is the unique solution of the variational equation

$$((u_n)_x, v_x)_{L^2} + (cu_n, v)_{L^2} = (f, v)_{L^2}, \quad \forall v \in Y_n.$$

Then we have (cf.[2,13])

$$\|F_n(c) - F(c)\|_{L^2} \leq O((1 + \|c\|_{L^2})n^{-2}).$$

Thus we can choose $\rho_n = n^{-2}$. To show the applicability of our method, now we only need to prove the weakly closedness of F_n for each fixed n. Suppose $\{c_k\} \subset D(F)$ be a sequence such that $c_k \rightharpoonup c$ in $L^2[0,1]$ and $u_n(c_k) \rightharpoonup u$ in $L^2[0,1]$. By the weakly closedness of $D(F)$, we have $c \in D(F)$. Note that $\{c_k\}$ is bounded in $L^2[0,1]$, $\{u_n(c_k)\}$ is bounded in $H^1[0,1]$ by the theory of elliptic equations. Since a bounded set in a Hilbert space always has a weakly convergent subsequence, and by the embedding theorem of Sobolev space and Ascoli-Arzela theorem we know there is a subsequence, still denote it by $\{u_n(c_k)\}$, such that $u_n(c_k) \rightharpoonup \tilde{u}$ in $H^1[0,1]$ and $u_n(c_k) \to \bar{u}$ in the maximum norm. Obviously $\tilde{u} = \bar{u} = u$. Now letting $k \to \infty$ in

$$((u_n(c_k))_x, v_x)_{L^2} + (c_k u_n(c_k), v)_{L^2} = (f, v)_{L^2}, \quad \forall v \in Y_n$$

we can obtain

$$(u_x, v_x)_{L^2} + (cu, v)_{L^2} = (f, v)_{L^2}, \quad \forall v \in Y_n.$$

Since $u \in Y_n$ due to the weakly closedness of Y_n, it follows that $u = u_n(c)$.

Example 2. Consider the problem of estimating the diffusion coefficient a in

$$-(au_x)_x = f \quad \text{in } (0,1), \tag{22}$$

$$u(0) = 0 = u(1) \tag{23}$$

with $f \in L^2[0,1]$ from the noise data u_δ of the state variable u_0, $\|u_\delta - u_0\|_{L^2} \leq \delta$. Let a_0 be the sought solution and $u_0 = u(a_0)$. To put this problem into our framework, we choose $X = H^1[0,1]$, $Y = L^2[0,1]$, and define the nonlinear operator F by

$$\begin{aligned} F : D(F) : = \{a \in H^1[0,1] \mid a(x) \geq \nu \subset H^1[0,1] \mapsto L^2[0,1], \\ a \mapsto F(a) := u(a), \end{aligned}$$

where $u(a)$ is the unique solution of (22), (23) and $\nu > 0$ is a given constant. It is well known that[4] F is weakly closed, continuous and Fréchet differentiable with

$$F'(a)h = A(a)^{-1}(hu(a)_x)_x, \quad a \in D(F), \quad h \in H^1[0,1],$$

and $A(a) : H_0^1 \cap H^2[0,1] \mapsto L^2[0,1]$ is defined by $A(a)u := -(au_x)_x$. We can show that[6] there is an $\epsilon_0 > 0$ such that for every $0 < \epsilon \le \epsilon_0$ there is an $\eta := \eta(\epsilon)$ such that for all $a \in D(F) \cap U_\epsilon(a_0)$,

$$\|F(a) - F(a_0) - F'(a_0)(a - a_0)\|_{L^2} \le \eta \|a - a_0\|_{H^1} \|F(a) - F(a_0)\|_{L^2}.$$

If we choose $L := I$, then Assumptions 2 and 4 and the assumptions on L in Theorem 2 hold automatically. To give the finite dimensional approximation, let X_n and Y_n be as in Example 1, and let $u_n(a)$ be the unique solution of the variational equation

$$(a(u_n)_x, v_x)_{L^2} = (f, v)_{L^2}, \quad \forall v \in Y_n,$$

and define the approximation F_n of F by

$$\begin{aligned} F_n : & D(F) \subset H^1[0,1] \mapsto L^2[0,1], \\ & a \mapsto F_n(a) := u_n(a). \end{aligned}$$

Then we can show that F_n is weakly closed for each fixed n as in Example 1 and have the estimate (cf.[2,13])

$$\|F_n(a) - F(a)\|_{L^2} \le O((1 + \|a\|_{H^1})n^{-2}).$$

Hence we can choose $\rho_n = n^{-2}$. If the sought solution a_0 satisfies: $a_0 \in H^2[0,1]$, $a_0 > \nu$, then [14, Corollary 7.3] implies $P_n a_0 \in D(F)$ for sufficiently large n and

$$\|P_n a_0 - a_0\|_{H^1} \le O(n^{-1}\|a_0\|_{H^2}).$$

Hence we can choose $b_n = n^{-1}\log n$. Thus we also verify the applicability of our method to this parameter estimation problem.

References

[1] H.T. Banks, K. Kunisch, Estimation techniques for distributed parameter systems, Boston, Birkhauser, 1989.

[2] P. Ciarlet, Lectures on the finite element method, Inst. of Fund. Research, Bombay, 1975.

[3] F. Colonius, K. Kunisch, Stability for parameter estimation in two point boundary value problems, *J. Reine Angewandte Math.,* **370** (1989), 1–29.

[4] H.W. Engl, K. Kunisch, A. Neubauer, Convergence rates for Tikhonov regularization of nonlinear ill-posed problems, *Inverse Problems,* **5** (1989), 523–540.

[5] H.W. Engl, Regularization methods for the stable solutions of inverse problems, *Surv. Math. Ind.,* **3** (1993), 71–143.

[6] B. Hofmann, O. Scherzer, Factors influencing the ill-posedness of nonlinear problems, *Inverse Problems,* **10** (1994), 1277–1297.

[7] K. Kunisch, W. Ring, Regularization of nonlinear ill-posed problems with closed operators, *Numer. Funct. Anal. Optimiz.,* **14** (1993), 389–404.

[8] K. Kunisch, L.W. White, Regularity properties in parameter estimation of diffusion coefficients in one dimensional elliptic boundary value problems, *Applicable Anal.,* **21** (1986), 71–87.

[9] J. Locker, P.M. Prenter, Regularization with differential operators I. General theory, *J. Math. Anal. Appl.,* **74** (1980), 504–529.

[10] J.T. Marti, On the convergence of an algorithm computing minimum-norm solutions of ill-posed problems, *Math. Comput.,* **34** (1980), 521–527.

[11] A. Neubauer, Tikhonov regularization for nonlinear ill-posed problems: Optimal convergence rates and finite dimensional approximation, *Inverse Problems,* **5** (1989), 541–557.

[12] A. Neubauer, Tikhonov regularization of nonlinear ill-posed problems in Hilbert scales, *Applicable Anal.,* **46** (1992), 59–72.

[13] A. Neubauer, O. Scherzer, Finite-dimensional approximation of Tikhonov regularized solutions of nonlinear ill-posed problems, *Numer. Funct. Anal. Optimiz.,* **11** (1990), 85–99.

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

[15] M.R. Trummer, A method for solving ill-posed linear operator equations, *SIAM J. Numer. Anal.,* **21** (1984), 729–736.