文章编号: 1000-0887(2009) 06-0690-11

求解非线性不适定的混合 Newton-Tikhonov 迭代法*

康传刚, 贺国强

(上海大学 数学系, 上海 200444)

(郭兴明推荐)

摘要: 鉴于 Newton 型方法在实际计算中计算量可能非常大, 因此提出了一种一步 Newton 结合若干步简化 Newton 的混合 Newton-Tikhonov 方法, 并且在一定条件下证明了该方法的收敛性和稳定性. 数值试验表明, 在减少计算量方面该方法相对于经典的 Newton 方法有明显的改善.

关 键 词: 非线性不适定问题; 热传导反问题; 混合 Newton-Tikhonov 方法; 收敛性; 稳定性

中图分类号: O241 文献标识码: A

DOI: 10.3879/j.issn.1000-0887.2009.06.008

引 言

在物理学、力学和工程科学等领域都出现了大量的反问题和不适定问题[1-2], 吸引了越来越多的学者从事这方面的研究. 在过去的 20 多年里, 求解非线性不适定问题的理论和方法取得了很大的发展[3-8]. 由于实际提出的反问题多数以非线性不适定的形式出现, 所以这类问题的解决无论对反问题理论本身还是对实际应用都具有重要意义.

非线性不适定问题在数学上主要归结为这样一类不适定算子方程, 即

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

这里 $F: \mathscr{D}(F) \subset X \to Y$ 是非线性 Fréchet 可微算子, 其中 X, Y 是 Hilbert 空间. 本文主要考虑方程(1) 的不适定性是由解不连续依赖于右端数据引起的. 实际问题中方程右端一般通过测量得到, 常带有误差, 因此上面的方程一般又写为具有扰动右端的形式, 即

$$F(x) = y^{\delta}, \tag{2}$$

其中 $\| y^{\delta} - y \| \leqslant \delta$, $\delta > 0$ 是已知的误差水平.

对于适定的算子方程(2) 的求解而言, Newton 迭代法

$$F'(x_k^{\delta})(x - x_k^{\delta}) = y^{\delta} - F(x_k^{\delta}), \qquad k = 0, 1, 2, \cdots \tag{3}$$

是一种最基本的方法, 它是通过把由上式得到的解 $x_{k+1}^{\delta} = x$ 作为进一步的逼近解实现的. 适定形式的简化 Newton 方法[9] 的算法为

* 收稿日期: 2008-10-14; 修订日期: 2009-05-14
基金项目: 上海市重点学科资助项目(S30104); 上海市教委重点学科建设资助项目(J50101)
作者简介: 康传刚(1978—), 男, 山东人, 博士(E-mail: ckang78@sohu.com);
贺国强(1946—), 男, 上海人, 教授(联系人. E-mail: gqhe@staff.shu.edu.cn).

$$F'(x)(x - x_k^\delta) = y^\delta - F(x_k^\delta), \qquad k = 0, 1, 2, \cdots, \tag{4}$$

其中 x 是某个固定的元.

但是, 如果方程(2)不适定, 那么线性化方程(3)或(4)通常也是不适定的, 因此需要采用正则化策略求解. 本文我们主要采用 Tikhonov 正则化方法. 对方程(3)而言, 即求如下泛函的极小元:

$$J_{\alpha_k}(x) = \|y^\delta - F(x_k^\delta) - F'(x_k^\delta)(x - x_k^\delta)\|^2 + \alpha_k\|x - x_k^\delta\|^2, \qquad k = 0, 1, 2, \cdots.$$

上面的问题等价于求解方程

$$(F'(x_k^\delta)^* F'(x_k^\delta) + \alpha_k I)(x - x_k^\delta) = F'(x_k^\delta)^*(y^\delta - F(x_k^\delta)),$$
$$k = 0, 1, 2\cdots. \tag{5}$$

式(5)即为求解非线性不适定问题的 Newton-Tikhonov 方法(也称 Levenberg-Marquardt 方法[1]).

同理, 应用 Tikhonov 正则化到方程(4)可以得到下面的简化 Newton-Tikhooov 方法:

$$(F'(x)^* F'(x) + \alpha_k I)(x - x_k^\delta) = F'(x)^*(y^\delta - F(x_k^\delta)), \qquad k = 0, 1, 2, \cdots. \tag{6}$$

结合方程(5)和(6), 我们提出求解方程(2)的混合 Newton-Tikhonov 方法:

$$\begin{cases} (F'(x_n^\delta)^* F'(x_n^\delta) + \alpha_{n,k} I)(x_{n,k+1}^\delta - x_{n,k}^\delta) = \\ \qquad F'(x_n^\delta)^*(y^\delta - F(x_{n,k}^\delta)), \qquad k = 0, 1, \cdots, p(n), \\ x_{n,0}^\delta = x_n^\delta, \quad x_{n+1}^\delta = x_{n,p+1}^\delta, \qquad n = 0, 1, \cdots, \end{cases} \tag{7}$$

其中第一步为 Newton-Tikhonov 方法, 接着是 $p(n)$ 步简化 Newton-Tikhonov 方法. 为简明起见, 本文主要考虑 $p(n) \equiv 2$ 的情况, 一般的 $p(n)$ 可类似分析.

下面给出一些基本条件, 这些条件对于理论分析是必要的(参见文献[1,5-7]).

假设 1 假设 x^+ 是方程(1)的真解, 且 F 在以 x^+ 为中心 r 为半径的某闭球域 $\mathscr{B}(x^+, r) \subset \mathscr{D}(F)$ 内可微且满足

1) 存在 $\eta > 0$, 使得 F 的 Taylor 余项满足不等式

$$\|F(x_1) - F(x_2) - F'(x_2)(x_1 - x_2)\| \leqslant$$
$$\eta\|x_1 - x_2\|\|F(x_1) - F(x_2)\|, \qquad \forall x_1, x_2 \in \mathscr{B}(x^+, r); \tag{8}$$

2) $F'(x^+)$ 存在右逆, 因此 $\forall x \in \mathscr{B}(x^+, r)$ 有

$$F'(x) = R_x F'(x^+), \tag{9}$$

进一步假设满足

$$\|R_x - I\| \leqslant C_0\|x - x^+\|, \tag{10}$$

其中 $C_0 > 0$ 是一个常数.

在算法(7)中 $\alpha_{n,k}$ 和迭代指标 (n, k) 起正则化参数的作用. 本文 $\alpha_{n,k}$ 由 Hanke 准则确定[1]:

$$\|y^\delta - F(x_{n,k}^\delta) - F'(x_n^\delta)(x_{n,k+1}^\delta - x_{n,k}^\delta)\| = \rho\|y^\delta - F(x_{n,k}^\delta)\|, \tag{11}$$

其中 $0 < \rho < 1$. 停止指标 $(n(\delta), k(\delta))$ 根据偏差原理确定, 即

$$\|y^\delta - F(x_{n(\delta),k(\delta)}^\delta)\| \leqslant \tau\delta$$

且

$$\forall (n, k) \in \Big\{(n, k) \mid n < n(\delta) \vee (n = n(\delta) \wedge k < k(\delta))\Big\}:$$
$$\|y^\delta - F(x_{n,k}^\delta)\| > \tau\delta, \tag{12}$$

其中 $\tau > 1$ 是一个常数.

注 1 式(12)对于 $\delta = 0$ 也适用, 一般来说此时得到无限次迭代序列 $\{x_{n,k}\}$, 除非 $F(x_{n,k}) = y^\delta$, 其中 $x_{n,k} = x_{n,k}^\delta$. 要强调指出的是, 在 $\delta = 0$ 时上式中的 τ 可取任意的正常数.

为方便计, 下面记

$$A_n = F'(x_n),\ A_{n,k} = F'(x_{n,k}),\ y_{n,k} = y - F(x_{n,k}),$$
$$A_n^\delta = F'(x_n^\delta),\ A_{n,k}^\delta = F'(x_{n,k}^\delta),\ y_{n,k}^\delta = y^\delta - F(x_{n,k}^\delta).$$

1 迭代序列的单调性

定理 1 设 F 满足假设 1, $0<\rho<1<\gamma$, 而且当由方程(7) 和(11) 确定的 $x_{n,k}^\delta (k=0,1,2)$ 不满足外迭代停止准则(12) 时.

(i) 假设

$$\| y^\delta - F(x_n^\delta) - F'(x_n^\delta)(x^+ - x_n^\delta) \| \leqslant \frac{\rho}{\gamma} \| y^\delta - F(x_n^\delta) \|, \tag{13}$$

那么有

$$\| x^+ - x_n^\delta \|^2 - \| x^+ - x_{n,1}^\delta \|^2 > \frac{2(\gamma-1)\rho}{\gamma} \| y^\delta - F(x_n^\delta) \| \, \| (F'(x_n^\delta) F'(x_n^\delta)^* + \alpha_{n,0} I)^{-1} (y^\delta - F(x_n^\delta)) \|, \tag{14}$$

而且还有

$$\| x^+ - x_n^\delta \|^2 - \| x^+ - x_{n,1}^\delta \|^2 > \frac{2(\gamma-1)\rho(1-\rho)}{\gamma \| F'(x_n^\delta) \|^2} \| y^\delta - F(x_n^\delta) \|^2. \tag{15}$$

(ii) 在(i)的条件下, 若还有

$$\eta \| x^+ - x_n^\delta \| \leqslant 1, \tag{16}$$

$$\| y^\delta - F(x_{n,1}^\delta) - F'(x_{n,1}^\delta)(x^+ - x_{n,1}^\delta) \| \leqslant \frac{\rho}{\gamma} \| y^\delta - F(x_{n,1}^\delta) \|, \tag{17}$$

$$D_{n,1} := \| R_{x_{n,1}^\delta} - R_{x_n^\delta} \| \leqslant \frac{\tau(\gamma-1)\rho}{2(\tau+1)\gamma}, \tag{18}$$

其中 R_x 满足条件(9) 和(10), 那么有

$$\| x^+ - x_{n,1}^\delta \|^2 - \| x^+ - x_{n,2}^\delta \|^2 \geqslant \frac{2(\gamma-1)\rho - 4D_{n,1}\gamma((\tau+1)/\tau)}{\gamma} \| y^\delta - F(x_{n,1}^\delta) \| \, \| (F'(x_n^\delta) F'(x_n^\delta)^* + \alpha_{n,1} I)^{-1} (y^\delta - F(x_{n,1}^\delta)) \|, \tag{19}$$

且有

$$\| x^+ - x_{n,1}^\delta \|^2 - \| x^+ - x_{n,2}^\delta \|^2 \geqslant \frac{2(\gamma-1)\rho - 4D_{n,1}\gamma((\tau+1)/\tau)}{\gamma \| F'(x_n^\delta) \|^2} (1-\rho) \| y^\delta - F(x_{n,1}^\delta) \|^2. \tag{20}$$

(iii) 在(ii)的条件下, 若还有

$$\| y^\delta - F(x_{n,2}^\delta) - F'(x_{n,2}^\delta)(x^+ - x_{n,2}^\delta) \| \leqslant \frac{\rho}{\gamma} \| y^\delta - F(x_{n,2}^\delta) \|, \tag{21}$$

$$D_{n,2} := \| R_{x_{n,2}^\delta} - R_{x_n^\delta} \| \leqslant \frac{\tau(\gamma-1)\rho}{2(\tau+1)\gamma}, \tag{22}$$

那么有

$$\| x^+ - x_{n,2}^\delta \|^2 - \| x^+ - x_{n,3}^\delta \|^2 \geqslant \frac{2(\gamma-1)\rho - 4D_{n,2}\gamma((\tau+1)/\tau)}{\gamma} \| y^\delta - F(x_{n,2}^\delta) \| \, \| (F'(x_n^\delta) F'(x_n^\delta)^* + \alpha_{n,2} I)^{-1} (y^\delta - F(x_{n,2}^\delta)) \|, \tag{23}$$

且有

$$\|x^+ - x_{n,2}^\delta\|^2 - \|x^+ - x_{n,3}^\delta\|^2 \geqslant \frac{2(\gamma-1)\rho - 4D_{n,2}\gamma((\tau+1)/\tau)}{\gamma\|F'(x_n^\delta)\|^2}(1-\rho)\|y^\delta - F(x_{n,2}^\delta)\|^2. \tag{24}$$

证明 (i)已由Hanke证明[1],下面主要证明(ii).由迭代格式(7)有

$$x^+ - x_{n,2}^\delta = x^+ - x_{n,1}^\delta - ((A_n^\delta)^* A_n^\delta + \alpha_{n,1}I)^{-1}(A_n^\delta)^* y_{n,1}^\delta,$$

所以

$$\|x^+ - x_{n,2}^\delta\|^2 = \|x^+ - x_{n,1}^\delta\|^2 - 2\langle x^+ - x_{n,1}^\delta, ((A_n^\delta)^* A_n^\delta + \alpha_{n,1}I)^{-1}(A_n^\delta)^* y_{n,1}^\delta\rangle + \|((A_n^\delta)^* A_n^\delta + \alpha_{n,1}I)^{-1}(A_n^\delta)^* y_{n,1}^\delta\|^2.$$

因此

$$\begin{aligned}
&\|x^+ - x_{n,2}^\delta\|^2 - \|x^+ - x_{n,1}^\delta\|^2 = \\
&-2\langle x^+ - x_{n,1}^\delta, ((A_n^\delta)^* A_n^\delta + \alpha_{n,1}I)^{-1}(A_n^\delta)^* y_{n,1}^\delta\rangle + \\
&\langle y_{n,1}^\delta, A_n^\delta(A_n^\delta)^*(A_n^\delta(A_n^\delta)^* + \alpha_{n,1})^{-2} y_{n,1}^\delta\rangle = \\
&-2\langle A_n^\delta(x^+ - x_{n,1}^\delta), (A_n^\delta(A_n^\delta)^* + \alpha_{n,1}I)^{-1} y_{n,1}^\delta\rangle + \\
&\langle y_{n,1}^\delta, A_n^\delta(A_n^\delta)^*(A_n^\delta(A_n^\delta)^* + \alpha_{n,1}I)^{-2} y_{n,1}^\delta\rangle = \\
&2\langle y_{n,1}^\delta - A_{n,1}^\delta(x^+ - x_{n,1}^\delta), (A_n^\delta(A_n^\delta)^* + \alpha_{n,1}I)^{-1} y_{n,1}^\delta\rangle + \\
&2\langle (A_{n,1}^\delta - A_n^\delta)(x^+ - x_{n,1}^\delta), (A_n^\delta(A_n^\delta)^* + \alpha_{n,1}I)^{-1} y_{n,1}^\delta\rangle - \\
&2\langle y_{n,1}^\delta, (A_n^\delta(A_n^\delta)^* + \alpha_{n,1}I)^{-1} y_{n,1}^\delta)\rangle + \\
&\langle y_{n,1}^\delta, A_n^\delta(A_n^\delta)^*(A_n^\delta(A_n^\delta)^* + \alpha_{n,1}I)^{-2} y_{n,1}^\delta\rangle.
\end{aligned}$$

进一步有

$$\begin{aligned}
&\|x^+ - x_{n,2}^\delta\|^2 - \|x^+ - x_{n,1}^\delta\|^2 = \\
&2\langle y_{n,1}^\delta - A_{n,1}^\delta(x^+ - x_{n,1}^\delta), (A_n^\delta(A_n^\delta)^* + \alpha_{n,1}I)^{-1} y_{n,1}^\delta\rangle + \\
&2\langle (A_{n,1}^\delta - A_n^\delta)(x^+ - x_{n,1}^\delta), (A_n^\delta(A_n^\delta)^* + \alpha_{n,1}I)^{-1} y_{n,1}^\delta\rangle - \\
&2\alpha_{n,1}\langle y_{n,1}^\delta, (A_n^\delta(A_n^\delta)^* + \alpha_{n,1}I)^{-2} y_{n,1}^\delta\rangle - \\
&\langle y_{n,1}^\delta, A_n^\delta(A_n^\delta)^*(A_n^\delta(A_n^\delta)^* + \alpha_{n,1}I)^{-2} y_{n,1}^\delta\rangle.
\end{aligned} \tag{25}$$

另外,下面诸式成立,

$$\langle y_{n,1}^\delta, A_n^\delta(A_n^\delta)^*(A_n^\delta(A_n^\delta)^* + \alpha_{n,1}I)^{-2} y_{n,1}^\delta\rangle = \|((A_n^\delta)^* A_n^\delta + \alpha_{n,1}I)^{-1}(A_n^\delta)^* y_{n,1}^\delta\|^2 \geqslant 0, \tag{26}$$

$$y_{n,1}^\delta - A_n^\delta(x_{n,2}^\delta - x_{n,1}^\delta) = \alpha_{n,1}(A_n^\delta(A_n^\delta)^* + \alpha_{n,1})^{-1} y_{n,1}^\delta, \tag{27}$$

$$\alpha_{n,1}\langle y_{n,1}^\delta, (A_n^\delta(A_n^\delta)^* + \alpha_{n,1}I)^{-2} y_{n,1}^\delta\rangle = \alpha_{n,1}\|(A_n^\delta(A_n^\delta)^* + \alpha_{n,1}I)^{-1} y_{n,1}^\delta\|^2. \tag{28}$$

联立方程(25)~(28)得到

$$\begin{aligned}
&\|x^+ - x_{n,2}^\delta\|^2 - \|x^+ - x_{n,1}^\delta\|^2 \leqslant \\
&2\|y_{n,1}^\delta - A_{n,1}^\delta(x^+ - x_{n,1}^\delta)\|\|(A_n^\delta(A_n^\delta)^* + \alpha_{n,1}I)^{-1} y_{n,1}^\delta\| + \\
&2\|(A_{n,1}^\delta - A_n^\delta)(x^+ - x_{n,1}^\delta)\|\|(A_n^\delta(A_n^\delta)^* + \alpha_{n,1}I)^{-1} y_{n,1}^\delta\| - \\
&2\|y_{n,1}^\delta - A_n^\delta(x_{n,2}^\delta - x_{n,1}^\delta)\|\|(A_n^\delta(A_n^\delta)^* + \alpha_{n,1}I)^{-1} y_{n,1}^\delta\|.
\end{aligned} \tag{29}$$

由条件(9)和(10)得到

$$\|(A_{n,1}^\delta - A_n^\delta)(x^+ - x_{n,1}^\delta)\| = \|(R_{x_{n,1}^\delta} - R_{x_n^\delta})F'(x^+)(x^+ - x_{n,1}^\delta)\| \leqslant \|(R_{x_{n,1}^\delta} - R_{x_n^\delta})\|\|F'(x^+)(x^+ - x_{n,1}^\delta)\|. \tag{30}$$

由式(8)得

$$\|F(x_{n,1}^\delta) - y - F'(x^+)(x_{n,1}^\delta - x^+)\| \leqslant \eta\|x^+ - x_{n,1}^\delta\|\|y - F(x_{n,1}^\delta)\|.$$

结合式(12)可以推出

$$\begin{aligned}\|F'(x^+)(x^+-x_{n,1}^\delta)\| &\leqslant (1+\eta\|x^+-x_{n,1}^\delta\|)\|y-F(x_{n,1}^\delta)\| \leqslant \\ &(1+\eta\|x^+-x_{n,1}^\delta\|)(\delta+\|y^\delta-F(x_{n,1}^\delta)\| \leqslant \\ &(1+\eta\|x^+-x_{n,1}^\delta\|)\frac{\tau+1}{\tau}\|y^\delta-F(x_{n,1}^\delta)\|. \end{aligned}\tag{31}$$

由式(29)～(31)和假设条件(17)得到

$$\begin{aligned}&\|x^+-x_{n,1}^\delta\|^2-\|x^+-x_{n,2}^\delta\|^2 \geqslant \\ &2\left(\left(1-\frac{1}{\gamma}\right)\rho-\frac{\tau+1}{\tau}(1+\eta\|x^+-x_{n,1}^\delta\|)\|R_{x_{n,1}^\delta}-R_{x_n^\delta}\|\right)\times \\ &\|y_{n,1}^\delta\|\|(A_n^\delta(A_n^\delta)^*+\alpha_{n,1}I)^{-1}y_{n,1}^\delta\|.\end{aligned}\tag{32}$$

根据条件(16)和(ⅰ)的结论, 即有 $\eta\|x^+-x_{n,1}^\delta\|\leqslant\eta\|x^+-x_n^\delta\|\leqslant 1$, 代入式(32)得

$$\begin{aligned}&\|x^+-x_{n,1}^\delta\|^2-\|x^+-x_{n,2}^\delta\|^2 \geqslant \\ &2\left(\left(1-\frac{1}{\gamma}\right)\rho-2\frac{\tau+1}{\tau}\|R_{x_{n,1}^\delta}-R_{x_n^\delta}\|\right)\|y_{n,1}^\delta\|\|(A_n^\delta(A_n^\delta)^*+\alpha_{n,1}I)^{-1}y_{n,1}^\delta\|,\end{aligned}$$

此即式(19). 再由式(18)还有

$$\|x^+-x_{n,1}^\delta\|^2>\|x^+-x_{n,2}^\delta\|^2.$$

下面证明式(20). 首先根据文献[1]中的方法易得

$$\alpha_{n,1}\leqslant\frac{\rho}{1-\rho}\|A_n^\delta\|^2.$$

令 $v_{n,1}=(A_n^\delta(A_n^\delta)^*+\alpha_{n,1}I)^{-1}y_{n,1}^\delta$ 得到

$$\begin{aligned}\|y_{n,1}^\delta\| &\leqslant \|A_n^\delta(A_n^\delta)^*+\alpha_{n,1}I\|\|v_{n,1}\|= \\ &(\|A_n^\delta\|^2+\alpha_{n,1})\|v_{n,1}\|\leqslant\frac{1}{1-\rho}\|A_n^\delta\|^2\|v_{n,1}\|.\end{aligned}$$

于是由式(19)可进一步得

$$\begin{aligned}&\|x^+-x_{n,1}^\delta\|^2-\|x^+-x_{n,2}^\delta\|^2 \geqslant \\ &\frac{2(\gamma-1)\rho-4D_{n,1}\gamma((\tau+1)/\tau)}{\gamma\|A_n^\delta\|^2}(1-\rho)\|y_{n,1}^\delta\|^2.\end{aligned}$$

此即式(20). 同理可证(ⅲ)成立. □

注 2 定理 1 对 $\delta=0$ 成立, 此时我们还可以将定理 1 条件和结论中的 $(\tau+1)/\tau$ 替换为 1.

2 数据 y 无扰动时的收敛性分析

在这一节讨论方程右端数据精确的情形, 首先给出定理 1 中的某些假设成立的一个充分条件.

引理 2 设 F 满足假设 1, 令 $0<\rho<1$, $y^\delta=y=F(x^+)$, $x_0\in\mathscr{B}(x^+,r)$, 那么存在 $\gamma_0>1$ 使得当

$$\|x^+-x_0\|<\frac{\rho}{4C_0+\eta}\tag{33}$$

时, 定理 1 的假设(13)、(16)～(18)、(21)和(22)对 $n=0$ 和 $\gamma=\gamma_0$ 成立.

证明 由式(8), 我们有

$$\|y-F(x_0)-F'(x_0)(x^+-x_0)\|\leqslant\eta\|x^+-x_0\|\|y-F(x_0)\|.$$

令 $$\gamma_0=\frac{\rho}{\eta\|x^+-x_0\|},$$

因此

$$\| y - F(x_0) - F'(x_0)(x^+ - x_0) \| \leqslant \frac{\rho}{\gamma_0} \| y - F(x_0) \|,$$

由上式和式(33)可得 $\gamma_0 > 1$, 因此定理假设式(13)对 $n = 0$ 和 $\gamma = \gamma_0$ 成立, 于是有定理1(i)的结论对 $n = 0$ 和 $\gamma = \gamma_0$ 成立, 特别地

$$\| x^+ - x_0 \| > \| x^+ - x_{0,1} \| . \tag{34}$$

下面证明定理1(ii)中的条件成立. 首先根据条件(33)可得式(16)对于 $n = 0, \delta = 0$ 时成立. 根据式(8)和(34)有

$$\| y - F(x_{0,1}) - F'(x_{0,1})(x^+ - x_{0,1}) \| \leqslant \eta \| x^+ - x_{0,1} \| \| y - F(x_{0,1}) \| < \eta \| x^+ - x_0 \| \| y - F(x_{0,1}) \|,$$

即可得到条件(17)对于 $n = 0$ 和 $\gamma = \gamma_0$ 成立.

对于式(18), 根据式(10)、(33)和(34)有

$$\| R_{x_{0,1}} - R_{x_0} \| \leqslant C_0 \| x^+ - x_{0,1} \| + C_0 \| x^+ - x_0 \| \leqslant 2C_0 \| x^+ - x_0 \| \leqslant \frac{2C_0 \rho}{4C_0 + \eta}. \tag{35}$$

由 $\gamma_0 = \rho / (\eta \| x^+ - x_0 \|)$ 和条件(33)可以推出

$$\eta \geqslant \frac{4C_0}{\gamma_0 - 1}, \tag{36}$$

于是有

$$\| R_{x_{0,1}} - R_{x_0} \| \leqslant \frac{\gamma_0 - 1}{2\gamma_0} \rho. \tag{37}$$

此即式(18)中 $\gamma = \gamma_0$, $\tau / (\tau + 1) = 1$ 时的情形. 至此证明了定理1(ii)中的条件在式(33)的假设下是成立的. 同理可证明定理1(iii)中条件也成立. □

定理3　令 $0 < \rho < 1$, $y^\delta = y = F(x^+)$, 设 F 满足假设1并且 $F'(x)$ 在 $\mathscr{B}(x^+, r)$ 内局部有界, 如果 $x_0 \in \mathscr{B}(x^+, r)$ 且满足

$$\| x^+ - x_0 \| \leqslant r_0 < \frac{\rho}{4C_0 + \eta}, \tag{38}$$

那么当 $n \to \infty$ 时, 由迭代格式(7)得到的序列 $\left\{ \left\{ x_{n,k} \right\}_{k=0}^{2} \right\}_{n=0}^{\infty}$ 收敛到 $F(x) = y$ 的一个解.

证明　根据注2和引理2知在式(38)的条件下 $\left\{ \| x^+ - x_{n,k} \| \right\}$ 是单调递减的.

下面证明 $\left\{ x_{n,k} \right\}$ 是 Cauchy 序列. 记

$$E = \left\{ \| y - F(x_{i,k}) \| : k = 0, 1, 2; i = n, n+1, \cdots, m \right\}.$$

不妨设 $\| y - F(x_{n_0,k_0}) \| (n \leqslant n_0 \leqslant m, 0 \leqslant k_0 \leqslant 2)$ 取到 E 中的最小值, 即

$$\| y - F(x_{n_0,k_0}) \| = \min_{i,k} E . \tag{39}$$

令 $e_{n,k} = x^+ - x_{n,k}$, 那么有

$$\| x_{m,p} - x_{n,q} \| = \| e_{n,q} - e_{m,p} \| \leqslant \| e_{n,q} - e_{n_0,k_0} \| + \| e_{m,p} - e_{n_0,k_0} \| .$$

下面首先证明当 $n \to \infty$ 时, $\| e_{m,p} - e_{n_0,k_0} \| \to 0$.

$$\| e_{m,p} - e_{n_0,k_0} \|^2 = \| e_{m,p} \|^2 - \| e_{n_0,k_0} \|^2 + 2 \langle e_{m,p} - e_{n_0,k_0}, e_{n_0,k_0} \rangle .$$

由算法迭代格式(7)可得,

$$| \langle e_{m,p} - e_{n_0,k_0}, e_{n_0,k_0} \rangle | \leqslant \sum_{i=n_0}^{m} \sum_{k=0}^{2} | \langle x_{i,k+1} - x_{i,k}, e_{n_0,k_0} \rangle | \leqslant$$

$$\sum_{i=n_0}^{m}\sum_{k=0}^{2}\|(A_iA_i^*+\alpha_{i,k}I)^{-1}y_{i,k}\|\|A_ie_{n_0,k_0}\|. \tag{40}$$

令

$$\|y-F(x_{i,s_0(i)})\|=\min\left\{\|y-F(x_{i,k})\|: k=0,1,2\right\},$$
$$i=n_0,n_0+1,\cdots,m, \tag{41}$$

其中 $0\leqslant s_0(i)\leqslant 2$ 是与 i 有关的某个整数, 那么对任意的 $n_0\leqslant i\leqslant m$ 有

$$A_ie_{n_0,k_0}=A_ie_{i,s_0(i)}+A_i(e_{n_0,k_0}-e_{i,s_0(i)}).$$

类似于文献[1]的方法, 可以得到估计

$$\|A_ie_{n_0,k_0}\|\leqslant\Big(\rho/\gamma_0+4\eta\|x^+-x_0\|+1+$$
$$4C_0(1+\eta\|x^+-x_0\|)\|x^+-x_0\|\Big)\|y-F(x_{i,s_0(i)})\|.$$

记 $M_0=\rho/\gamma_0+4\eta\|x^+-x_0\|+1+4C_0(1+\eta\|x^+-x_0\|)\|x^+-x_0\|$, 那么有

$$\|A_ie_{n_0,k_0}\|\leqslant M_0\|y-F(x_{i,s_0(i)})\|. \tag{42}$$

把式(42)代入式(40)并且联立式(41)得到

$$|\langle e_{m,p}-e_{n_0,k_0},e_{n_0,k_0}\rangle|\leqslant$$
$$M_0\sum_{i=n_0}^{m}\sum_{k=0}^{2}\|(A_iA_i^*+\alpha_{i,k}I)^{-1}y_{i,k}\|\|y-F(x_{i,k})\|. \tag{43}$$

由条件(10)和(38)得到 $D_{n,j}\leqslant 2C_0\|x^+-x_0\|\leqslant 2C_0r_0$. 令 $D=\sup_{n,j}D_{n,j}$, 那么有 $D\leqslant 2C_0r_0$. 记 $M=2((\gamma_0-1)\rho/\gamma_0-2D)$, 根据式(36)和(38)可以得到 $M>0$. 根据式(14)、(19)、(23)和(43), 得到

$$|\langle e_{m,p}-e_{n_0,k_0},e_{n_0,k_0}\rangle|\leqslant\frac{M_0}{M}\sum_{i=n_0}^{m}\sum_{k=0}^{2}(\|x^+-x_{i,k}\|^2-\|x^+-x_{i,k+1}\|^2)=$$
$$\frac{M_0}{M}(\|x^+-x_{n_0,0}\|^2-\|x^+-x_{m+1,0}\|^2),$$

上式右端当 $n\to\infty$ 时收敛到0, 所以 $|\langle e_{m,p}-e_{n_0,k_0},e_{n_0,k_0}\rangle|\to 0$, 因此 $\|e_{m,p}-e_{n_0,k_0}\|\to 0(n\to\infty)$, 同理可证, 当 $n\to\infty$ 时, $\|e_{n_0,k_0}-e_{n,q}\|\to 0$, 因此, $\left\{\left\{x_{n,k}\right\}_{k=0}^{2}\right\}_{n=0}^{\infty}$ 是 Cauchy 列.

根据 $F'(x)$ 在 $\mathscr{B}(x^+,r)$ 内有界, 不妨设 $\|F'(x)\|\leqslant M$. 由式(15)、(20)和(24)可以得到

$$\|x^+-x_{n,j}\|^2-\|x^+-x_{n,j+1}\|^2\geqslant\frac{2(\gamma_0-1)\rho-4D_{n,j}\gamma_0}{\gamma_0M^2}(1-\rho)\|y_{n,j}\|^2,$$
$$j=0,1,2;\ n=0,1,\cdots,\infty.$$

再由 D 和 M 的定义, 于是可以得到

$$\|x^+-x_{n,j}\|^2-\|x^+-x_{n,j+1}\|^2\geqslant\frac{M}{M^2}(1-\rho)\|y_{n,j}\|^2,$$
$$j=0,1,2;\ n=0,1,\cdots,\infty. \tag{44}$$

由式(44)得知 $\sum_{n=0}^{\infty}\sum_{k=0}^{2}\|y_{n,k}\|^2$ 收敛, 因此当 $n\to\infty$ 时 $\|y_{n,k}\|\to 0$, 即 $F(x_{n,k})\to y$. 若假设 $x_{n,k}\to x$, 那么有 $F(x)=y$. □

3 数据 y^δ 有扰动时的收敛性分析

引理 4 假设 F 满足条件(8), x^+, z 是方程(1)在 $\mathscr{B}(x^+,r)$ 中的两个解, 那么对于任意 $w\in\mathscr{B}(x^+,r)$ 有

$$F'(x^+)(x^+-z)=0,\ F'(x^+)(z-w)=F'(x^+)(x^+-w).$$

证明　根据条件(8)有

$$\|F'(x^+)(z-x^+)\|=\|F(z)-F(x^+)-F'(x^+)(z-x^+)\|\leqslant$$
$$\eta\|z-x^+\|\|F(z)-F(x^+)\|=0.$$

因此有 $F'(x^+)(z-x^+)=0$, 于是可得

$$F'(x^+)(z-w)=F'(x^+)(z-x^+)+F'(x^+)(x^+-w)=F'(x^+)(x^+-w).\qquad\square$$

引理5　假设 F 在 $\mathscr{B}(x^+,r)$ 中满足条件(8)和(9), x^+, z 是方程(1)在 $\mathscr{B}(x^+,r)$ 中的两个解, 那么对于任意的 $x\in\mathscr{B}(x^+,r)$, 有

$$F'(x)(x^+-z)=0.$$

证明　由条件(9)和引理4有 $F'(x)(x^+-z)=R_xF'(x^+)(x^+-z)=0$.　$\square$

引理6　令 $0<\rho<1$, $\tau>1/\rho$, 假设 F 满足假设1, $F'(\cdot)$ 在 $\mathscr{B}(x^+,r)$ 内有界, 如果 $\|y^\delta-y\|\leqslant\delta$,

$$\|x_0-x^+\|\leqslant r_0<\frac{\rho\tau-1}{(4C_0+\eta)(\tau+1)},\tag{45}$$

$\alpha_{n,k}$ 和 $(n(\delta),k(\delta))$ 分别由Hanke准则(11)和残差准则(12)确定, 那么正则化解 $x^\delta_{n(\delta),k(\delta)}$ 当 $\delta\to 0$ 时收敛到 $F(x)=y$ 的解.

证明　为了叙述的方便, 我们重新定义迭代序列的下标: $z^\delta_{n(\delta)}$ 表示正则化解, 并记 $z_n=z^0_n$,

$$z^\delta_0=x^\delta_0,\ z^\delta_1=x^\delta_{0,1},\ z^\delta_2=x^\delta_{0,2},\ \cdots,$$
$$z^\delta_{3m}=x^\delta_m,\ z^\delta_{3m+1}=x^\delta_{m,1},\ z^\delta_{3m+2}=x^\delta_{m,2},\ \cdots,$$

于是我们得到序列 $\left\{x^\delta_{n,k}\right\}$ 重排后的序列 $\left\{z^\delta_j\right\}$.

下面分3步证明定理. 第1步, 证明

$$\|z-z^\delta_{j+1}\|<\|z-z^\delta_j\|,\qquad j=0,1,\cdots,n(\delta)-1.\tag{46}$$

其中 z 是 $F(x)=y$ 的一个解且满足条件 $z_0=z^\delta_0\in\mathscr{B}_0:=\mathscr{B}\left(z,\dfrac{2(\rho\tau-1)}{\eta(\tau+1)}\right)\subset\mathscr{B}(x^+,r)$. 根据条件(8)和引理6得

$$\|y^\delta-F(z_0)-F'(z_0)(z-z_0)\|\leqslant$$
$$(1+\eta\|x^+-z_0\|)\delta+\eta\|x^+-z_0\|\|y^\delta-F(z_0)\|.$$

如果 $n(\delta)>0$, 那么满足 $\|y^\delta-F(z_0)\|>\tau\delta$, 且由上式可得

$$\|y^\delta-F(z_0)-F'(z_0)(z-z_0)\|\leqslant\frac{1+(1+\tau)\eta\|x^+-z_0\|}{\tau}\|y^\delta-F(z_0)\|,$$

令 $\gamma=\rho\tau/(1+(1+\tau)\eta\|x^+-z_0\|)$, 由 $z_0=x_0$ 和式(45)易知 $\gamma>1$, 因此定理1(i)中的条件(13)对于 $n=0$ 成立, 即得到 $\|z-z_0\|>\|z-z^\delta_1\|$, 特别有 $\|x^+-z_0\|>\|x^+-z^\delta_1\|$.

同理可证定理1(ii)和(iii)结论成立, 即: $\|z-z^\delta_1\|\geqslant\|z-z^\delta_2\|\geqslant\|z-z^\delta_3\|$, 特别有 $\|x^+-z^\delta_1\|\geqslant\|x^+-z^\delta_2\|\geqslant\|x^+-z^\delta_3\|$. 类似于上面的方法可以证明

$$\|z-z^\delta_{j+1}\|<\|z-z^\delta_j\|,\qquad j=0,1,\cdots,n(\delta)-1.$$

第2步, 证明对于 $\delta>0$, $n(\delta)$ 是一个有界数. 假定 z_0 充分接近 x^+, 那么式(46)对于 $z=x^+$ 和 $\delta>0$ 成立, 联立方程(15)、(20)和(24), 得到

$$n(\delta)\tau^2\delta^2\leqslant\sum_{j=0}^{n(\delta)-1}\|y^\delta-F(z^\delta_j)\|^2\leqslant\|x^+-x^\delta_0\|^2-\|x^+-x^\delta_{n(\delta)}\|^2,$$

即证得 $n(\delta)$ 是有界数.

第3步, 证明 $\delta\to 0$ 时, $z^{\delta}_{n(\delta)}$ 的收敛性, 先分析两种特殊情况.

第1种情形, 假定当 $\delta\to 0$ 时, $n(\delta)\equiv j$. 那么由 $F(x)$, $F'(x)$ 的连续性, 当 $\delta\to 0$ 时 $z^{\delta}_j\to z_j$, 其中 z_j 对应于精确右端的第 j 步迭代. 由于 $\|y^{\delta}-F(z^{\delta}_j)\|\leqslant\tau\delta$, 当 $\delta\to 0$ 时有 $F(z_j)=y$, 因此, $\delta\to 0$ 时 $z^{\delta}_{n(\delta)}$ 收敛到 $F(x)=y$ 的解 z_j.

第2种情形, 假定 $\delta\to 0$ 时 $n(\delta)\to\infty$. 令 z 是 $\{z_j\}$ 的极限, 易证 $z=z$ 满足式(46), 而且有

$$\|z-z^{\delta}_{n(\delta)}\|\leqslant\|z-z^{\delta}_l\|\leqslant\|z-z_l\|+\|z_l-z^{\delta}_l\|,$$

类似于文献[1]可以证明 $\delta\to 0$ 时 $z^{\delta}_{n(\delta)}\to z$.

考虑一般情形. 如果收敛性结果不成立, 则存在 $\epsilon_0>0$ 和收敛于 0 的子列 $\{\delta_i\}$, 使得

$$\|z^{\delta}_{n(\delta_i)}-z\|>\epsilon_0, \tag{47}$$

由于实数列 $\{n(\delta_i)\}$ 一定存在收敛子列(有限或无限), 我们仍记收敛子列为 $\{n(\delta_i)\}$, 如果当 $i\to\infty$ 时 $n(\delta_i)\to\infty$, 则由第2种情形知, 当 $i\to\infty$ 时,

$$\|z^{\delta_i}_{n(\delta_i)}-z\|\to 0.$$

这与式(47)矛盾. 如果 $n(\delta_i)\to n$, 类似地, 可得到与第1种情形相矛盾的结论, 这就证明了在一般情形下, 当 $\delta\to 0$ 时有 $z^{\delta}_{n(\delta)}\to z$. □

4 数 值 算 例

考虑热传导方程[10]

$$\begin{cases}\dfrac{\partial u}{\partial t}-\dfrac{\partial}{\partial x}\left(a(x)\dfrac{\partial u}{\partial x}\right)=f(x,t), & (x,t)\in(0,1)\times(0,1),\\ u(x,0)=u_0(x), & x\in[0,1],\\ u(0,t)=r_0(t),\ u(1,t)=r_1(t), & t\in[0,1].\end{cases} \tag{48}$$

对于上面方程(48), 根据已知数据 $a(x)$, $f(x,t)$, $u_0(t)$, $r_0(t)$, $r_1(t)$ 确定温度分布函数 $u(x,t)$ 的问题称为热传导正问题. 但在有些情况下, 我们可以得到数据 $f(x,t)$, $u_0(t)$, $r_0(t)$, $r_1(t)$ 和关于解 $u(x,t)$ 的部分信息, 参数函数 $a(x)$ 未知, 这种确定热传导方程参数的问题我们称为热传导反问题, 而且它是不适定的.

下面, 我们分别用 Newton-Tikhonov 方法与本文提出的混合 Newton-Tikhonov 方法 ($p(n)$ 分别取 2 和 5) 求解上面的热传导反问题. 在式(48)中, 令 $f(x,t)=2x-2$, $u_0(x)=x^2$, $r_0(t)=0$, $r_1(t)=2t+1$, 相应的兼容性条件为

$$(Bu)(t)=\frac{\partial u(0,t)}{\partial x}=r(t),$$

其中 $r(t)=2t$, 相应的算子 F 定义为

$$F(a)(t)=Bu(a);$$

那么相应的方程为

$$F(a)=r.$$

如果取 $u(x,t)=x^2+2xt$, 那么反问题的解为 $a^*=1$.

实际计算时, 我们取扰动数据

$$r^{\delta}(t)=r(t)+\sqrt{2}\,\delta\sin 10\pi t,\qquad t\in[0,1].$$

显然有 $\| r^{\delta} - r \|_{L^2[0,1]} = \delta$, 如果 $\delta = 0$, 此时偏差原理修改为

$$\| F(a^{\delta}_{n(\delta),k(\delta)}) - r \|_{L^2[0,1]} \leqslant 10^{-8} < \| F(a^{\delta}_{n,k}) - r \|_{L^2[0,1]},$$
$$n < n(\delta) \vee (n = n(\delta) \wedge k < k(\delta)).$$

初始迭代取

$$a^{\delta}_0 = 1 + \lambda x(1 - x), \qquad \lambda = 1, 2, \dots, 8.$$

我们沿 (x, t) 方向(30, 60)等分的方法对方程(48)进行离散. 计算中令 $\tau = 2.5$, $\rho = 0.6$, δ 分别取 10^{-4} 和 0 两种情况. 数值结果见表1, 2, 两表中每行表示迭代初值 a^{δ}_0 对应于 λ 时相应的计算结果, $e_0 = \| a^{\delta}_0 - a^* \|_{L^2[0,1]}$, $e_n = \| a^{\delta}_{n(\delta),k(\delta)} - a^* \|_{L^2[0,1]}$, NT 代表由Hanke 准则确定 α_n 的Newton-Tikhonov 方法, MNT(P)表示 $p(n) \equiv p$ 时的混合Newton-Tikhonov 方法, $n(\delta)$ 表示停止参数 $(n(\delta), k(\delta))$ 的第1个分量, 也就是迭代停止时的Newton 迭代步数. computational time 代表计算时间, 单位是百分秒(即0.01 s), 本文所得数据基于的计算机平台是AMD2500+, 主频1.83 G, 内存512 M.

表1　$\delta = 10^{-4}$ 时, NT, MNT(2) 和 MNT(5) 方法的比较

λ	e_0	$n(\delta)$			e_n			computational time t / cs		
		NT	MNT(2)	MNT(5)	NT	MNT(2)	MNT(5)	NT	MNT(2)	MNT(5)
1	0.183	16	7	4	0.016	0.016	0.014	16	12	6
2	0.365	21	9	5	0.024	0.023	0.024	23	16	6
3	0.548	27	12	7	0.025	0.023	0.022	28	20	12
4	0.730	32	14	9	0.020	0.023	0.020	31	24	14
5	0.913	40	13	10	0.031	0.017	0.037	42	22	17
6	1.095	65	15	11	0.048	0.056	0.060	70	25	19
7	1.287	111	29	10	0.075	0.092	0.080	116	49	17
8	1.460	234	53	14	0.180	0.133	0.119	239	87	25

表2　$\delta = 0$ 时, NT, MNT(2) 和 MNT(5) 方法的比较

λ	e_0	$n(\delta)$			e_n			computational time t / cs		
		NT	MNT(2)	MNT(5)	NT	MNT(2)	MNT(5)	NT	MNT(2)	MNT(5)
1	0.183	91	35	20	0.002 7	0.002 8	0.002 8	93	58	36
2	0.365	130	48	26	0.005 1	0.005 7	0.005 3	134	80	43
3	0.548	140	52	29	0.007 9	0.009 2	0.011 2	144	86	48
4	0.730	113	25	19	0.010 4	0.020 5	0.017 9	117	42	33
5	0.913	83	27	33	0.013 3	0.015 3	0.015 3	91	46	53
6	1.095	210	74	45	0.018 5	0.020 1	0.025 0	212	117	73
7	1.278	261	115	57	0.023 2	0.040 5	0.032 0	264	180	92
8	1.460	452	136	57	0.021 2	0.031 7	0.032 0	449	214	94

关于本文的几个注释

1) Newton-Tikhonov 方法计算一步 Newton-Tikhonov 迭代时需要计算若干步简化 Newton-Tikhonov 迭代, 但是计算Newton-Tikhonvo 迭代增加的计算量相对于 Newton-Tikhonov 来时是比较少的. 由于混合 Newton-Tikhonov 方法能有效减少Newton 迭代步数, 因此总的计算量会有效减少, 数值试验可以验证上面的结论.

2) 产生的几个问题: 如何选取 $p(n)$ 的值?$p(n)$ 的选取对方法的收敛性和效率有什么影响?本文 $p(n)\equiv 5$ 时的结果明显好于 $p(n)\equiv 2$, 是不是对其它问题也有这样的性质? 这些问题都是非常重要同时也是比较难以回答的问题.

3) 本文提出的简化思想也可以应用于其它 Newton 正则化方法.

[参 考 文 献]

[1] Hanke M. A regularization Levenberg-Marquardt scheme, with application to inverse groundwater filtration problems[J]. *Inverse Problems*, 1997, **13**(1): 79-95.

[2] Engl H W, Hanke M, Neubauer A. *Regularization of Inverse Problem* [M]. Dordrecht: Kluwer Academic, 1996.

[3] Bakushinskii A B, Kokurin M Y. *Iterative Methods for Approximate Solution of Inverse Problems* [M]. Dordrecht: Springer, 2004.

[4] Kaltenbacher B, Neubauer A, Scherzer O. *Iterative Regularization Methods for Nonlinear Ill-Posed Problems* [M]. Walter de Gruyter, 2008.

[5] Jin Q N. On the iteratively regularized Gauss-Newton method for solving nonlinear ill-posed problems [J]. *Math Comp*, 2000, **69**(232): 1603-1623.

[6] Jin Q N. A convergence analysis of the iteratively regularized Gauss-Newton method under the lipschitz condition[J]. *Inverse Problems*, 2008, **24**(4): 1-16.

[7] Hanke M, Neubauer A, Scherzer O. A convergerce analysis of the Landweber iteration for nonlinear ill-posed problems[J]. *Numerical Mathematics*, 1995, **72**(11): 21-37.

[8] Deuflhard P, Engl H W, Scherzer O. A convergence analysis of iterative methods for the solution of nonlinear ill-posed problems under affinely invariant conditions[J]. *Inverse Problems*, 1998, **14**(5): 1081-1106.

[9] Deuflhard P. *Newton Method for Nonlinear Problems: Affine Invariance and Adaptive Algorithms* [M]. Berlin Heidelberg: Springer-Verlag, 2004.

[10] 贺国强, 孟泽红. 求解热传导反问题的一种正则化 Newton 型迭代法[J]. 应用数学和力学, 2007, **28**(4): 479-486.

A Mixed Newton-Tikhonov Method for Nonlinear Ill-Posed Problems

KANG Chuan-gang, HE Guo-qiang

(Department of Mathematics, Shanghai University, Shanghai 200444, P. R. China)

Abstract: Newton type methods are one kind of the efficient methods to solve nonlinear ill-posed problems and attract extensive attention of people. However, the computational cost of Newton type methods may be very large because of the complexity of practical problems. A mixed Newton-Tikhonov method, i. e., one step Newton-Tikhonov method with several other steps of simplified Newton-Tikhonov method was proposed. The convergence and stability of this method were proved under some conditions. Numerical experiments show that the new method has obvious improvement over the classical Newton method in the reduction of the computational cost.

Key words: nonlinear ill-posed problem; inverse heat conduction problem; mixed Newton-Tikhonov method; convergence; stability