

一类求解非线性奇异方程组的牛顿改进算法

吕 巍^{1†}, 魏良亭¹, 冯恩民²

(1. 上海大学 数学系, 上海 200444; 2. 大连理工大学 数学科学学院, 辽宁 大连 116024)

摘 要: 受一个求解非线性奇异方程组迭代格式的启示, 将两种牛顿改进算法推广成一般形式, 并将其发展为一类求解具有奇异雅可比矩阵的非线性方程组的牛顿改进算法. 首先, 描述这类新算法的迭代格式, 并导出其收敛阶, 该新格式每步迭代仅需计算一次函数值和一次导数值; 然后, 对测试函数进行检验, 并与牛顿算法及其他奇异牛顿算法进行比较, 从而验证该算法的快速收敛性; 最后, 通过两个实际问题验证所提出算法的有效性.

关键词: 牛顿算法; 奇异雅可比矩阵; 非线性方程组; 收敛阶

中图分类号: TP301.6 **文献标志码:** A

A modification of Newton's method solving non-linear equations with singular Jacobian

LV Wei^{1†}, WEI Liang-ting¹, FENG En-min²

(1. Department of Mathematics, Shanghai University, Shanghai 200444, China; 2. School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China)

Abstract: Motivated by a singular technique for root-finding, two Newton's methods are generalized to a class of iteration formats, on the basis of which, a modification of the Newton's method solving non-linear equations with singular Jacobian is presented. Firstly, the new rule is described and its convergence order is analyzed. The modified singular method requires one function and one first derivative evaluations per step. Then, numerical examples demonstrate the faster convergence achieved with this modification than Newton's method and some singular schemes. Finally, two practical problems are given to illustrate the effectiveness of the proposed method.

Keywords: Newton's method; singular Jacobian; non-linear equations; order of convergence

0 引 言

牛顿算法因其形式简单和收敛迅速而被广泛应用于求解各类非线性非奇异方程组的问题中.

学者们提出了许多改进的求解非线性非奇异方程组的牛顿算法. 例如, 文献[1]和文献[2]分别提出了求解具有非奇异雅可比矩阵的非线性方程组的牛顿预测-校正格式 PC 和 QMn, 其优点是每步迭代仅需计算两次函数值, 其缺点同牛顿算法类似, 即这些算法要求方程组的雅可比矩阵在方程组根的某个小邻域内可逆, 一旦这些雅可比矩阵在迭代中奇异, 算法就会发散或者失效, 所以这个苛刻的条件限制了这些算法的应用. 基于这个原因, 许多学者进行了针对该问题的改进^[3-7]. 例如文献[3]提出了一种新的二阶迭代格式, 即在奇异的雅可比矩阵上加上一个含参数的一次对角矩阵, 假设函数足够光滑, 使得对角阵及

雅可比阵的和在方程组根的某个邻域内可逆; 文献[4]减弱了文献[3]中函数足够光滑的条件; 文献[5]发展了文献[3]的算法, 提出了一种形式上更为复杂的二阶奇异算法; 文献[6]的二阶算法采用了在初始点用单位阵代替雅可比阵的方法, 避免了雅可比阵在初始点的奇异性; 文献[7]在奇异的雅可比矩阵上加上一个含参数的二次对角阵, 得到了一个三阶收敛的新格式.

本文将算法 PC^[1] 和算法 QMn^[2] 推广成一类算法, 并针对这类算法的弊端, 将文献[3]的技巧应用到这类算法上, 即在奇异的雅可比矩阵上加上一个含参数的一次对角矩阵, 以改变其雅可比矩阵在迭代点不可逆的特点, 并得知这类新奇异的收敛阶为 2. 最后通过数值算例验证新奇异的收敛阶的有效性, 并同牛顿算法和 5 种奇异牛顿改进算法进行了比较.

收稿日期: 2016-09-28; 修回日期: 2016-12-08.

基金项目: 国家自然科学基金青年基金项目(11101262); 上海市重点学科项目(S30104); 上海高校一流学科项目(B类).

作者简介: 吕巍(1978—), 女, 讲师, 博士, 从事分布参数系统辨识及其应用等研究; 魏良亭(1989—), 男, 硕士生, 从事优化算法的研究.

†通讯作者. E-mail: lvwei@t.shu.edu.cn

1 改进的牛顿算法

设 $F(X) : D_l \subset \mathbf{R}^l \rightarrow \mathbf{R}^l$ 是一个非线性向量值函数,且在凸集 D_l 上任意阶 Fréchet 可微, $X = \alpha$ 是方程组 $F(X) = 0$ 的一个单根, $F(X) = (f_1(X), f_2(X), \dots, f_l(X))^T, X = (X_1, X_2, \dots, X_l)^T, F'$ 是 F 的雅可比矩阵, $(F')^{-1}$ 是 F' 的逆矩阵, $l \in \mathbf{Z}^+ = \{1, 2, \dots\}$, 则得到如下牛顿迭代格式:

$$X^{*(n)} = X^{(n)} - [F'(\gamma X^{(n-1)} + (1 - \gamma)X^{*(n-1)})]^{-1}F(X^{(n)}), \quad (1)$$

$$X^{(n+1)} = X^{(n)} - [F'(\gamma X^{(n)} + (1 - \gamma)X^{*(n)})]^{-1}F(X^{(n)}). \quad (2)$$

其中: $\gamma \in \mathbf{R}$, 且 $0 \leq \gamma \leq 1$. 当 $\gamma = 0$ 时, 迭代格式(1)和(2)即为算法 PC; 当 $\gamma = 1/2$ 时, 迭代格式(1)和(2)即为算法 QMn; 当 $\gamma = 1$ 时, 迭代格式(1)和(2)即为标准牛顿算法. 该算法要求方程组的雅可比矩阵在迭代点非奇异, 否则算法就会发散或者失效. 文献[3]提出了一种能有效解决这种困难的迭代格式, 即

$$X^{(n+1)} = X^{(n)} - [\text{diag}(\mu_i^{(n)} f_i(X^{(n)})) + F'(X^{(n)})]^{-1}F(X^{(n)}), \quad (3)$$

其中 $\mu_i^{(n)} \in \mathbf{R} \setminus \{0\}$. $F'(X^{(n)})$ 奇异时 $\text{diag}(\mu_i^{(n)} f_i(X^{(n)}))$ 的加入可以有效避免这种情况的发生. 将式(3)应用到(1)和(2), 则得到一类新的含参数的求解具有奇异雅可比阵的方程组的牛顿迭代格式, 即

$$X^{*(n)} = X^{(n)} - [\text{diag}(\lambda_i^{(n)} f_i(X^{(n)})) + F'(\gamma X^{(n-1)} + (1 - \gamma)X^{*(n-1)})]^{-1}F(X^{(n)}), \quad (4)$$

$$X^{(n+1)} = X^{(n)} - [\text{diag}(\mu_i^{(n)} f_i(X^{(n)})) + F'(\gamma X^{(n)} + (1 - \gamma)X^{*(n)})]^{-1}F(X^{(n)}). \quad (5)$$

其中: $\gamma \in \mathbf{R}$, 且 $0 \leq \gamma \leq 1$; $X^{(n)} = (X_1^{(n)}, X_2^{(n)}, \dots, X_l^{(n)})^T; \lambda_i^{(n)}, \mu_i^{(n)} \in \mathbf{R}$, 且 $0 < |\lambda_i^{(n)}|, |\mu_i^{(n)}| < +\infty, i = 1, 2, \dots, l, n = 0, 1, 2, \dots, l \in \mathbf{Z}^+$. 当雅可比矩阵 $F'(\gamma X^{(n-1)} + (1 - \gamma)X^{*(n-1)})$ 或 $F'(\gamma X^{(n)} + (1 - \gamma)X^{*(n)})$ 奇异时选择 $\lambda_i^{(n)}$ 和 $\mu_i^{(n)}$ 使 $\text{diag}(\lambda_i^{(n)} f_i(X^{(n)})) + F'(\gamma X^{(n-1)} + (1 - \gamma)X^{*(n-1)})$ 或 $\text{diag}(\mu_i^{(n)} f_i(X^{(n)})) + F'(\gamma X^{(n)} + (1 - \gamma)X^{*(n)})$ 非奇异.

新的迭代步骤(4)和(5)需要两个初值 $X^{(0)}$ 和 $X^{*(0)}$, 取 $X^{*(0)} = X^{(0)}$. 因此, 当 $n = 0$ 时, 格式为

$$X^{*(0)} = X^{(0)},$$

$$X^{(1)} = X^{(0)} - [\text{diag}(\mu_i^{(0)} f_i(X^{(0)})) + F'(\gamma X^{(0)} + (1 - \gamma)X^{*(0)})]^{-1}F(X^{(0)}).$$

当 $n \geq 1$ 时, 格式为

$$X^{*(n)} = X^{(n)} - [\text{diag}(\lambda_i^{(n)} f_i(X^{(n)})) + F'(\gamma X^{(n-1)} + (1 - \gamma)X^{*(n-1)})]^{-1}F(X^{(n)}),$$

$$X^{(n+1)} = X^{(n)} - [\text{diag}(\mu_i^{(n)} f_i(X^{(n)})) + F'(\gamma X^{(n)} + (1 - \gamma)X^{*(n)})]^{-1}F(X^{(n)}).$$

由式(4)和(5)定义的新算法的优点在于: 1) 新算法可以求解雅可比矩阵奇异的非线性方程组; 2) 每步迭代仅需要计算两个函数值. 由于预测步(4)中 $F'(\gamma X^{(n-1)} + (1 - \gamma)X^{*(n-1)})$ 的值在上一次迭代中已经得到, 新算法每一次迭代仅需要计算两个函数值 $F(X^{(n)})$ 和 $F'(\gamma X^{(n)} + (1 - \gamma)X^{*(n)})$.

2 收敛性分析

本节对式(4)和(5)定义的新算法进行收敛性分析, 定理1给出了该算法的收敛性并得到它的收敛阶.

定理1 设 $F(X) : D_l \subset \mathbf{R}^l \rightarrow \mathbf{R}^l$ 是定义在凸集 D_l 上的任意阶 Fréchet 可微的非线性向量值函数, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_l)^T \in D_l$ 是方程组 $F(X) = 0$ 的一个单根, $F'(\alpha) \neq 0$, 并且 $[F'(\alpha)]^{-1}$ 有界, 其中 $F(X) = (f_1(X), f_2(X), \dots, f_l(X))^T, X = (X_1, X_2, \dots, X_l)^T, l \in \mathbf{Z}^+$. 对于任意 $i = 1, 2, \dots, l, n = 0, 1, 2, \dots$, 参数 $\lambda_i^{(n)} \in \mathbf{R} \setminus \{0\}$ 和 $\mu_i^{(n)} \in \mathbf{R} \setminus \{0\}$, 使得 $\text{diag}(\lambda_i^{(n)} f_i(X^{(n)})) + F'(\gamma X^{(n-1)} + (1 - \gamma)X^{*(n-1)})$ 或者 $\text{diag}(\mu_i^{(n)} f_i(X^{(n)})) + F'(\gamma X^{(n)} + (1 - \gamma)X^{*(n)})$ 非奇异, 并且 $\lambda_i^{(n)} f'_i(\alpha)$ 和 $\mu_i^{(n)} f'_i(\alpha)$ 有界, 其中 $\gamma \in \mathbf{R}$, 且 $0 \leq \gamma \leq 1$.

1) 令 $I_\varepsilon = \{X \in \mathbf{R}^l : \|X - \alpha\|_\infty \leq \varepsilon\} \subseteq D_l, F(X) \in C^2(I_\varepsilon)$, 定义如下等式:

$$L_1 = \max_{X \in I_\varepsilon} \|F'(X)\|_\infty,$$

$$L_2 = \max_{X \in I_\varepsilon} \left\| \frac{F''(X)}{2} \right\|_\infty,$$

$$L_3 = \max_{X \in I_\varepsilon} \|[\text{diag}(\lambda_i^{(n)} f_i(X^{(n)})) + F'(\gamma X^{(n-1)} + (1 - \gamma)X^{*(n-1)})]^{-1}\|_\infty,$$

$$L_4 = \max_{X \in I_\varepsilon} \|[\text{diag}(\mu_i^{(n)} f_i(X^{(n)})) + F'(\gamma X^{(n)} + (1 - \gamma)X^{*(n)})]^{-1}\|_\infty,$$

$$L_5 = \max_{X \in I_\varepsilon} \|\text{diag}(\lambda_i^{(n)} f_i(X))\|_\infty,$$

$$L_6 = \max_{X \in I_\varepsilon} \|\text{diag}(\mu_i^{(n)} f_i(X))\|_\infty.$$

若 $\max\{L_3[L_5 + (2L_1 + L_2)\varepsilon], L_4[L_6 + (2L_1 + L_2)\varepsilon]\} < 1$, 则对于任意 $X^{(0)}, X^{*(0)}, X^{(1)} \in I_\varepsilon$, 迭代格式(4)和(5)收敛到 $\alpha \in I_\varepsilon$.

2) 在收敛的前提下, 迭代格式(4)和(5)的阶为2.

证明 1) 算法局部收敛性的证明方法与文献[2]相同,此略.

2) 首先证明等式

$$\text{diag}(\lambda_i^{(n)} f_i(X^{(n)}))\varepsilon^{(n)} = O(\|\varepsilon^{(n)}\|^2), \quad (6)$$

$$\text{diag}(\mu_i^{(n)} f_i(X^{(n)}))\varepsilon^{(n)} = O(\|\varepsilon^{(n)}\|^2). \quad (7)$$

对 $f_i(X)$ ($i = 1, 2, \dots, l$) 在点 $X = \alpha$ 进行泰勒展开, 得到

$$f_i(X) = f_i(\alpha) + f'_i(\alpha)(X - \alpha) + O(\|X - \alpha\|^2).$$

取 $X = X^{(n)}$, 由于 $f_i(\alpha) = 0$ ($i = 1, 2, \dots, l$), 有

$$f_i(X^{(n)}) = f'_i(\alpha)\varepsilon^{(n)} + O(\|\varepsilon^{(n)}\|^2).$$

上式两边同乘 $\lambda_i^{(n)}$ ($i = 1, 2, \dots, l$), 可得

$$\lambda_i^{(n)} f_i(X^{(n)}) = \lambda_i^{(n)} f'_i(\alpha)\varepsilon^{(n)} + O(\|\varepsilon^{(n)}\|^2),$$

因此

$$\begin{aligned} \text{diag}(\lambda_i^{(n)} f_i(X^{(n)}))\varepsilon^{(n)} &= \\ \text{diag}(\lambda_i^{(n)} f'_i(\alpha))(\varepsilon^{(n)})^2 + O(\|\varepsilon^{(n)}\|^3) &= O(\|\varepsilon^{(n)}\|^2), \end{aligned}$$

同理可得式(6)和(7).

下面推出迭代格式(4)和(5)的收敛阶. 定义误差项 $\varepsilon^{*(n)} = X^{*(n)} - \alpha$ 和 $\varepsilon^{(n)} = X^{(n)} - \alpha$. 在式(4)和(5)两端减去 α , 可得

$$[\text{diag}(\lambda_i^{(n)} f_i(X^{(n)})) + F'(\gamma X^{(n-1)} + (1 - \gamma)X^{*(n-1)})](\varepsilon^{*(n)} - \varepsilon^{(n)}) = -F(X^{(n)}), \quad (8)$$

$$[\text{diag}(\mu_i^{(n)} f_i(X^{(n)})) + F'(\gamma X^{(n)} + (1 - \gamma)X^{*(n)})] \times (\varepsilon^{(n+1)} - \varepsilon^{(n)}) = -F(X^{(n)}). \quad (9)$$

式(9)两端在 $X = \alpha$ 点泰勒展开, 由于 $F(\alpha) = 0$, 有

$$\begin{aligned} & \left[\text{diag}(\mu_i^{(n)} f_i(X^{(n)})) + \left(F'(\alpha) + F''(\alpha)(\gamma\varepsilon^{(n)} + (1 - \gamma)\varepsilon^{*(n)}) + \frac{1}{2}F'''(\alpha)(\gamma\varepsilon^{(n)} + (1 - \gamma)\varepsilon^{*(n)})^2 + O(\|e^{(n)}\|^3) \right) \right] (\varepsilon^{(n+1)} - \varepsilon^{(n)}) = \\ & -F'(\alpha)\varepsilon^{(n)} - \frac{1}{2}F''(\alpha)(\varepsilon^{(n)})^2 - \frac{1}{6}F'''(\alpha)(\varepsilon^{(n)})^3 + O(\|e^{(n)}\|^4). \end{aligned} \quad (10)$$

其中: 对于所有的 k 和 n , $(e^{(n)})^k$ 表示阶为 k 的 $\varepsilon^{(n)}$ 与 $\varepsilon^{*(n)}$ 的所有组合. 由于 $[F'(\alpha)]^{-1}$ 有界, 在式(10)两端同乘 $[F'(\alpha)]^{-1}$, 可得

$$\begin{aligned} & \left[[F'(\alpha)]^{-1} \text{diag}(\lambda_i^{(n)} f_i(X^{(n)})) + [F'(\alpha)]^{-1} \left(F'(\alpha) + F''(\alpha)(\gamma\varepsilon^{(n)} + (1 - \gamma)\varepsilon^{*(n)}) + \frac{1}{2}F'''(\alpha)(\gamma\varepsilon^{(n)} + (1 - \gamma)\varepsilon^{*(n)})^2 + O(\|e^{(n)}\|^3) \right) \right] (\varepsilon^{(n+1)} - \varepsilon^{(n)}) = \\ & -\varepsilon^{(n)} - \frac{1}{2}[F'(\alpha)]^{-1}F''(\alpha)(\varepsilon^{(n)})^2 - \end{aligned}$$

$$\frac{1}{6}[F'(\alpha)]^{-1}F'''(\alpha)(\varepsilon^{(n)})^3 + O(\|e^{(n)}\|^4).$$

记作 $F_2 = [F'(\alpha)]^{-1}F''(\alpha)$ 和 $F_3 = [F'(\alpha)]^{-1}F'''(\alpha)$, 则 $\varepsilon^{(n+1)}$ 为

$$\begin{aligned} \varepsilon^{(n+1)} &= \\ & \left[I + \gamma F_2 \varepsilon^{(n)} + (1 - \gamma) F_2 \varepsilon^{*(n)} + \frac{\gamma^2}{2} F_3 (\varepsilon^{(n)})^2 + \gamma(1 - \gamma) F_3 \varepsilon^{(n)} \varepsilon^{*(n)} + \frac{(1 - \gamma^2)}{2} F_3 (\varepsilon^{*(n)})^2 + [F'(\alpha)]^{-1} \text{diag}(\lambda_i^{(n)} f_i(X^{(n)})) + O(\|e^{(n)}\|^3) \right]^{-1} \times \\ & \left[\left(\gamma - \frac{1}{2} \right) F_2 (\varepsilon^{(n)})^2 + (1 - \gamma) F_2 \varepsilon^{*(n)} \varepsilon^{(n)} + \frac{1}{2} F_3 (\gamma \varepsilon^{(n)} + (1 - \gamma) \varepsilon^{*(n)})^2 \varepsilon^{(n)} - \frac{1}{6} F_3 (\varepsilon^{(n)})^3 + [F'(\alpha)]^{-1} \text{diag}(\lambda_i^{(n)} f_i(X^{(n)})) \varepsilon^{(n)} + O(\|e^{(n)}\|^4) \right]. \end{aligned} \quad (11)$$

由于 $\|A\| < 1$, 可得 $(I - A)^{-1} = \sum_{j=0}^{\infty} A^j$. 由假设可知, 当 $n \rightarrow \infty$ 时, $\varepsilon^{(n)} \rightarrow 0$, $\varepsilon^{*(n)} \rightarrow 0$. 再由式(6)可知, 存在 $N > 0$, 使得对于所有的 $n \geq N$, 有

$$\begin{aligned} & \left\| -\gamma F_2 \varepsilon^{(n)} - (1 - \gamma) F_2 \varepsilon^{*(n)} - \frac{\gamma^2}{2} F_3 (\varepsilon^{(n)})^2 - \gamma(1 - \gamma) F_3 \varepsilon^{(n)} \varepsilon^{*(n)} - \frac{(1 - \gamma^2)}{2} F_3 (\varepsilon^{*(n)})^2 - [F'(\alpha)]^{-1} \text{diag}(\lambda_i^{(n)} f_i(X^{(n)})) + O(\|e^{(n)}\|^3) \right\| < 1. \end{aligned}$$

根据Neumann级数, 式(11)可化为

$$\begin{aligned} \varepsilon^{(n+1)} &= \\ & \sum_{j=0}^{\infty} \left[\gamma F_2 \varepsilon^{(n)} + (1 - \gamma) F_2 \varepsilon^{*(n)} + \frac{\gamma^2}{2} F_3 (\varepsilon^{(n)})^2 + \gamma(1 - \gamma) F_3 \varepsilon^{(n)} \varepsilon^{*(n)} + \frac{(1 - \gamma^2)}{2} F_3 (\varepsilon^{*(n)})^2 + [F'(\alpha)]^{-1} \text{diag}(\lambda_i^{(n)} f_i(X^{(n)})) + O(\|e^{(n)}\|^3) \right]^j \times \\ & \left[\left(\gamma - \frac{1}{2} \right) F_2 (\varepsilon^{(n)})^2 + (1 - \gamma) F_2 \varepsilon^{*(n)} \varepsilon^{(n)} + \frac{1}{2} F_3 (\gamma \varepsilon^{(n)} + (1 - \gamma) \varepsilon^{*(n)})^2 \varepsilon^{(n)} - \frac{1}{6} F_3 (\varepsilon^{(n)})^3 + [F'(\alpha)]^{-1} \text{diag}(\lambda_i^{(n)} f_i(X^{(n)})) \varepsilon^{(n)} + O(\|e^{(n)}\|^4) \right] = \\ & \left(\gamma - \frac{1}{2} \right) F_2 (\varepsilon^{(n)})^2 + (1 - \gamma) F_2 \varepsilon^{*(n)} \varepsilon^{(n)} + [F'(\alpha)]^{-1} \text{diag}(\lambda_i^{(n)} f_i(X^{(n)})) \varepsilon^{(n)} + \\ & \frac{1}{2} F_3 (\gamma \varepsilon^{(n)} + (1 - \gamma) \varepsilon^{*(n)})^2 \varepsilon^{(n)} - \frac{1}{6} F_3 (\varepsilon^{(n)})^3 + \\ & \gamma \left(\gamma - \frac{1}{2} \right) F_2 \varepsilon^{(n)} F_2 (\varepsilon^{(n)})^2 + \\ & \gamma(1 - \gamma) F_2 \varepsilon^{(n)} F_2 \varepsilon^{*(n)} \varepsilon^{(n)} + \\ & \gamma [F'(\alpha)]^{-1} F_2 \varepsilon^{(n)} \text{diag}(\lambda_i^{(n)} f_i(X^{(n)})) \varepsilon^{(n)} + \\ & (1 - \gamma) \left(\gamma - \frac{1}{2} \right) F_2 \varepsilon^{*(n)} F_2 (\varepsilon^{(n)})^2 + \end{aligned}$$

$$(1-\gamma)^2 F_2 \varepsilon^{*(n)} F_2 \varepsilon^{*(n)} \varepsilon^{(n)} + (1-\gamma)[F'(\alpha)]^{-1} F_2 \varepsilon^{*(n)} \text{diag}(\lambda_i^{(n)} f_i(X^{(n)})) \times \varepsilon^{(n)} + O(\|e^{(n)}\|^4). \tag{12}$$

对式(8)采用相同的处理方法,有

$$\varepsilon^{*(n)} = \left(\gamma - \frac{1}{2}\right) F_2 (\varepsilon^{(n)})^2 + (1-\gamma) F_2 \varepsilon^{(n-1)} \varepsilon^{(n)} + [F'(\alpha)]^{-1} \text{diag}(\mu_i^{(n)} f_i(X^{(n)})) \varepsilon^{(n)} + O(\|e^{(n)}\|^3). \tag{13}$$

将式(13)代入(12),可得

$$\varepsilon^{(n+1)} = \left(\gamma - \frac{1}{2}\right) F_2 (\varepsilon^{(n)})^2 + (1-\gamma) \left(\gamma - \frac{1}{2}\right) F_2 F_2 (\varepsilon^{(n)})^3 + (1-\gamma)^2 F_2 F_2 \varepsilon^{*(n-1)} (\varepsilon^{(n)})^2 + (1-\gamma) F_2 [F'(\alpha)]^{-1} \text{diag}(\mu_i^{(n)} f_i(X^{(n)})) (\varepsilon^{(n)})^2 + [F'(\alpha)]^{-1} \text{diag}(\lambda_i^{(n)} f_i(X^{(n)})) \varepsilon^{(n)} - \frac{1}{6} F_3 (\varepsilon^{(n)})^3 + O(\|e^{(n)}\|^4). \tag{14}$$

由式(7)和 $[F'(\alpha)]^{-1}$ 的有界性可以得到项 $(1-\gamma)F_2[F'(\alpha)]^{-1}\text{diag}(\mu_i^{(n)} f_i(X^{(n)}))(\varepsilon^{(n)})^2$ 的阶为3. 由于收敛阶由方程幂的最低项决定,考虑如下差分方程:

$$\varepsilon^{(n+1)} = \left(\gamma - \frac{1}{2}\right) F_2 (\varepsilon^{(n)})^2 + [F'(\alpha)]^{-1} \text{diag}(\lambda_i^{(n)} f_i(X^{(n)})) \varepsilon^{(n)}.$$

根据式(6),对于 $0 \leq \gamma \leq 1$,有

$$\varepsilon^{(n+1)} = O(\|e^{(n)}\|^2).$$

故式(4)和(5)定义的算法的收敛阶为2. □

注1 对于新迭代格式(4)和(5),选择参数 $\lambda_i^{(n)} \in \mathbf{R} \setminus \{0\}$ 和 $\mu_i^{(n)} \in \mathbf{R} \setminus \{0\} (i = 1, 2, \dots, l, n = 0, 1, 2, \dots, l \in \mathbf{Z}^+)$ 的通常作法是使得它们满足

$$\text{sign}(\lambda_i^{(n)} f_i(X^{(n)})) = \text{sign}\left(\frac{\partial f_i(\gamma X^{(n-1)} + (1-\gamma)X^{*(n-1)})}{\partial X_i}\right), \tag{15}$$

$$\text{sign}(\mu_i^{(n)} f_i(X^{(n)})) = \text{sign}\left(\frac{\partial f_i(\gamma X^{(n)} + (1-\gamma)X^{*(n)})}{\partial X_i}\right), \tag{16}$$

其中 $\text{sign}(\cdot)$ 是符号函数. 这种方案可以使得雅可比矩阵 $F'(\gamma X^{(n-1)} + (1-\gamma)X^{*(n-1)})$ 和 $F'(\gamma X^{(n)} + (1-\gamma)X^{*(n)})$ 的对角项范数更大.

注2 定理1的条件下,在式(4)和(5)中取 $\gamma = 0$,可得

$$X^{*(n)} = X^{(n)} - [\text{diag}(\lambda_i^{(n)} f_i(X^{(n)})) + F'(X^{*(n-1)})]^{-1} F(X^{(n)}), \tag{17}$$

$$X^{(n+1)} = X^{(n)} - [\text{diag}(\mu_i^{(n)} f_i(X^{(n)})) + F'(X^{*(n)})]^{-1} F(X^{(n)}). \tag{18}$$

其中式(17)和(18)是对算法PC^[1]的改进,记为PC-M.

在式(4)和(5)中,取 $\gamma = 1/2$,可得

$$X^{*(n)} = X^{(n)} - \left[\text{diag}(\lambda_i^{(n)} f_i(X^{(n)})) + F'\left(\frac{X^{(n-1)} + X^{*(n-1)}}{2}\right)\right]^{-1} F(X^{(n)}), \tag{19}$$

$$X^{(n+1)} = X^{(n)} - \left[\text{diag}(\mu_i^{(n)} f_i(X^{(n)})) + F'\left(\frac{X^{(n)} + X^{*(n)}}{2}\right)\right]^{-1} F(X^{(n)}), \tag{20}$$

其中式(19)和(20)是对算法QMn^[2]的改进,记为算法QMn-M.

在式(4)和(5)中,取 $\gamma = 1$,可得

$$X^{(n+1)} = X^{(n)} - [\text{diag}(\mu_i^{(n)} f_i(X^{(n)})) + F'(X^{(n)})]^{-1} F(X^{(n)}), \tag{21}$$

则双步迭代格式(4)和(5)变为单步迭代格式(21). 此式是对标准牛顿法的改进,即为迭代格式(3).

下节用数值算例验证算法PC-M和算法QMn-M的有效性.

3 数值实验

本节将通过数值算例来验证算法PC-M和算法QMn-M的有效性,并同牛顿算法和5种奇异算法进行比较:二阶算法^[3-6],三阶算法^[7]. 其中文献[3]中式(6)记为KLW,文献[4]中式(6)记为WX,文献[5]中式(14)记为GN,文献[6]中算法4记为EUNM,文献[7]中式(2)记为KJ. 首先描述所选取的6个数值算例,然后对数值结果进行说明和讨论.

3.1 数值算例及参数选取

表1列出了测试函数:3个二维函数,1个三维函数,1个四维函数及1个五维函数. 这些测试函数涵盖了多项式函数、三角函数、反三角函数、对数函数、指数函数5种类型的组合,具有一定的代表性. 表1给出这些非线性测试函数的根、雅可比矩阵及初值. 除第1个算例外,其余算例所选取的初值都使得每个测试函数的雅可比矩阵在相应的初值点是奇异的.

表2列出了各比较算法的参数值. 对于PC-M算法和QMn-M算法,参数 $\lambda_i^{(n)}$ 和 $\mu_i^{(n)}$ 的选取在很大程度上影响着收敛的速度,最简单的方式是取 $\lambda_i^{(n)}$ 和 $\mu_i^{(n)}$ 为常数 ($i = 1, 2, \dots, l$). 因此,对于同一个算例的每一个 i ,不同的迭代步 n , $\lambda_i^{(n)}$ 和 $\mu_i^{(n)}$ 取为相同的常数. 其中除了算例 $F_1(X)$ 以外,其余测试函数均是按照方案(15)和(16)选取参数. 对于同一个测试函数,PC-M算法和QMn-M算法选取的参数值是相同

表1 测试函数及相关信息

测试函数	根	雅可比矩阵	初值
$F_1(X) = \begin{cases} e^{-X_1} + \arctan(X_2), \\ \ln(X_1) + X_2 \end{cases}$	$(1.316\ 220\ 206\ 5 \dots, -0.274\ 764\ 149\ 0 \dots)^T$	$\begin{bmatrix} -e^{-X_1} & \frac{1}{1+X_2^2} \\ \frac{1}{X_1} & 1 \end{bmatrix}$	$(1, 4)^T$
$F_2(X) = \begin{cases} X_1 - \cos(X_2), \\ \sin(X_1) + 0.5X_2 \end{cases}$	$(0.530\ 388\ 689\ 5 \dots, -1.011\ 737\ 334\ 2 \dots)^T$	$\begin{bmatrix} 1 & \sin(X_2) \\ \cos(X_1) & 0.5 \end{bmatrix}$	$(\frac{\pi}{4}, \frac{\pi}{4})^T$
$F_3(X) = \begin{cases} X_1 + X_2 - 3, \\ X_1^2 + X_2^2 - 9 \end{cases}$	$(0, 3)^T$	$\begin{bmatrix} 1 & 1 \\ 2X_1 & 2X_2 \end{bmatrix}$	$(0, 0)^T$
$F_4(X) = \begin{cases} X_1^3 + X_2^3 - 2, \\ X_2^3 + X_3^3 - 28, \\ X_3^3 + X_1^3 - 28 \end{cases}$	$(1, 1, 3)^T$	$\begin{bmatrix} 3X_1^2 & 3X_2^2 & 0 \\ 0 & 3X_2^2 & 3X_3^2 \\ 3X_1^2 & 0 & 3X_3^2 \end{bmatrix}$	$(0, 0, 0)^T$
$F_5(X) = \begin{cases} X_2X_3 + X_4(X_2 + X_3) + 1, \\ X_1X_3 + X_4(X_1 + X_3) + 1, \\ X_1X_2 + X_4(X_1 + X_2) + 1, \\ X_1X_2 + X_1X_3 + X_2X_3 - 1 \end{cases}$	$(-0.577\ 350\ 269\ 1 \dots, -0.577\ 350\ 269\ 1 \dots, -0.577\ 350\ 269\ 1 \dots, 1.154\ 700\ 538\ 3 \dots)^T$	$\begin{bmatrix} 0 & X_3 + X_4 & X_2 + X_4 & X_2 + X_3 \\ X_3 + X_4 & 0 & X_1 + X_4 & X_1 + X_3 \\ X_2 + X_4 & X_1 + X_4 & 0 & X_1 + X_2 \\ X_2 + X_3 & X_1 + X_3 & X_1 + X_2 & 0 \end{bmatrix}$	$(0, 0, 0, 0)^T$
$F_6(X) = \begin{cases} X_1^2 + X_2 + X_3 + X_4 + X_5 - 5, \\ X_1 + X_2^2 + X_3 + X_4 + X_5 - 5, \\ X_1 + X_2 + X_3^2 + X_4 + X_5 - 5, \\ X_1 + X_2 + X_3 + X_4^2 + X_5 - 5, \\ X_1 + X_2 + X_3 + X_4 + X_5^2 - 5 \end{cases}$	$(1, 1, 1, 1, 1)^T$	$\begin{bmatrix} 2X_1 & 1 & 1 & 1 & 1 \\ 1 & 2X_2 & 1 & 1 & 1 \\ 1 & 1 & 2X_3 & 1 & 1 \\ 1 & 1 & 1 & 2X_4 & 1 \\ 1 & 1 & 1 & 1 & 2X_5 \end{bmatrix}$	$(0.5, 0.5, 0.5, 0.5, 0.5)^T$

表2 比较算法的参数值

测试函数	KLW ^[3]	WX ^[4]	GN ^[5]
$F_1(X)$	$\lambda_{1,2}^{(n)} = 0.1$	$\lambda_{1,2}^{(n)} = 1$	$\nu_{1,2}^{(n)} = 0.1, m = (1, 1)^T$
$F_2(X)$	$\lambda_{1,2}^{(n)} = 1$	$\lambda_{1,2}^{(n)} = 1$	$\nu_{1,2}^{(n)} = 1, m = (1, 1)^T$
$F_3(X)$	$\lambda_{1,2}^{(n)} = -1$	$\lambda_{1,2}^{(n)} = -1$	$\nu_{1,2}^{(n)} = -1, m = (1, 1)^T$
$F_4(X)$	$\lambda_{1,2,3}^{(n)} = -0.1$	$\lambda_{1,2,3}^{(n)} = -1$	$\nu_{1,2,3}^{(n)} = -0.1, m = (1, 1, 1)^T$
$F_5(X)$	$\lambda_{1,2,3,4}^{(n)} = 2.1$	$\lambda_{1,2,3,4}^{(n)} = 1$	$\nu_{1,2,3,4}^{(n)} = 2.1, m = (1, 1, 1, 1)^T$
$F_6(X)$	$\lambda_{1,2,3,4,5}^{(n)} = 0.6$	$\lambda_{1,2,3,4,5}^{(n)} = 1$	$\nu_{1,2,3,4,5}^{(n)} = 0.6, m = (1, 1, 1, 1, 1)^T$

测试函数	KJ ^[7]	PC-M 算法和QMn-M算法
$F_1(X)$	$\lambda_{1,2}^{(n)} = 0.01, \mu_{1,2}^{(n)} = 0.01$	$\lambda_{1,2}^{(n)} = 0.01, \mu_{1,2}^{(n)} = 0.01$
$F_2(X)$	$\lambda_{1,2}^{(n)} = 0.5, \mu_{1,2}^{(n)} = \frac{1}{3}$	$\lambda_{1,2}^{(n)} = 0.5, \mu_{1,2}^{(n)} = 0.9$
$F_3(X)$	$\lambda_{1,2}^{(n)} = -0.05, \mu_{1,2}^{(n)} = 0.05$	$\lambda_{1,2}^{(n)} = -1, \mu_1^{(n)} = -1, \mu_2^{(n)} = -0.3$
$F_4(X)$	$\lambda_{1,2,3}^{(n)} = -1.5, \mu_{1,2,3}^{(n)} = 0.29$	$\lambda_{1,2}^{(n)} = -1.1, \lambda_{1,2,3}^{(n)} = -0.333\ 333, \mu_{1,2,3}^{(n)} = -1$
$F_5(X)$	$\lambda_{1,2,3,4}^{(n)} = 0.98, \mu_{1,2,3,4}^{(n)} = 0.9$	$\lambda_{1,2,3}^{(n)} = 100, \lambda_4^{(n)} = -100$ $\mu_{1,2,3}^{(n)} = 1.732, \mu_4^{(n)} = -0.866$
$F_6(X)$	$\lambda_{1,2,3,4,5}^{(n)} = -0.01, \mu_{1,2,3,4,5}^{(n)} = 0.1$	$\lambda_{1,2,3,4,5}^{(n)} = -0.1, \mu_{1,2,3,4,5}^{(n)} = -0.181\ 8$

的. 总之, 所有算例选取参数的原则是, 使得第1次迭代后所得到的 $X^{(1)}$ 足够接近于根.

3.2 数值结果

所有的程序代码运行的软件环境均为 Matlab R2012a, Windows XP 32位操作系统. 停止准则为

$$\|X^{(n+1)} - X^{(n)}\|_2 + \|F(X^{(n)})\|_2 \leq 10^{-10}.$$

针对每种算法和每个测试函数, 计算迭代次数, 每次迭代结束需要计算函数值的数目(即函数评价指数)

及CPU时间.

对于第1个测试函数 $F_1(X)$, 牛顿算法在第3步迭代时找到了一个虚根, 所以迭代失败; 而对于其他算例, 由于雅可比矩阵 $F'(X)$ 在初值 $X^{(0)}$ 奇异, 牛顿算法失效.

表3给出同其他5种奇异算法的比较结果. 其中: 第1组数字表示迭代次数, 第1组括号内的黑体数字表示函数评价指数, 第2组括号内的数字表示CPU时间, “Failure”表示算法发散或者找到虚根.

表 3 同 5 种奇异算法^[3-7] 的数值比较结果

测试函数	$X^{(0)}$	KLW ^[3]	WX ^[4]	GN ^[5]
$F_1(X)$	$(1, 4)^T$	Failure	Failure	Failure
$F_2(X)$	$\left(\frac{\pi}{4}, \frac{\pi}{4}\right)^T$	8(32)(0.303 351 4)	8(32)(0.299 285 8)	8(32)(0.353 284)
$F_3(X)$	$(0, 0)^T$	9(36)(0.396 820)	9(36)(0.412 864)	9(36)(0.414 975)
$F_4(X)$	$(0, 0, 0)^T$	7(42)(0.769 713 4)	7(42)(0.906 615 4)	7(42)(0.800 969 6)
$F_5(X)$	$(0, 0, 0, 0)^T$	10(100)(0.802 244)	Failure	10(100)(0.925 554)
$F_6(X)$	$(0.5, 0.5, 0.5, 0.5, 0.5)^T$	6(60)(0.615 230)	6(60)(0.626 615)	6(60)(0.735 086)

测试函数	EUNM ^[6]	KJ ^[7]	PC-M 算法	QMn-M 算法
$F_1(X)$	Failure	Failure	8(40)(0.382 267)	8(40)(0.361 900)
$F_2(X)$	7(28)(0.267 695)	5(30)(0.239 074 9)	6(24)(0.270 824)	7(28)(0.335 588)
$F_3(X)$	9(36)(0.371 848)	9(54)(0.531 658)	5(20)(0.259 898)	5(20)(0.245 178)
$F_4(X)$	8(48)(0.990 150)	28(252)(3.115 912 7)	7(42)(0.633 242)	6(36)(0.521 747)
$F_5(X)$	Failure	6(96)(0.685 872)	4(40)(0.437 824)	4(40)(0.371 217)
$F_6(X)$	Failure	Failure	4(40)(0.529 563)	3(30)(0.362 499)

从表 3 可以看出, 对于所比较的奇异算法, 在 30 (6 × 5) 个数值算例中有 21 个算例收敛. 其中 EUNM 有 3 个算例发散, WX 和 KJ 各有两个算例发散, KLW 和 GN 各有 1 个算例发散. 在收敛的算例中, 对于迭代次数, 算法 QMn-M 有 19 个算例优于其他方法, 算法 PC-M 有 17 个算例优于其他方法; 对于函数评价指标, 算法 QMn-M 有 20 个算例优于其他方法, 算法 PC-M 有 18 个算例优于其他方法; 对于 CPU 时间, 算法 QMn-M 有 17 个算例优于其他方法, 算法 PC-M 有 19 个算例优于其他方法.

综上所述, 从以上的数值结果和分析可以看出, 算法 QMn-M 和 PC-M 是有效可行的.

4 实际应用

本节将用算法 QMn-M 和算法 PC-M 计算两个工程问题^[8], 并同上面的 5 种奇异算法进行比较. 第 1 个实例是关于工字形截面梁几何性质的尺寸实现问题, 第 2 个实例是关于薄壁矩形梁截面几何性质的尺寸实现问题. 停止准则

$$\|X^{(n+1)} - X^{(n)}\|_2 + \|F(X^{(n)})\|_2 \leq 10^{-15}.$$

表 4 给出了这两个实例的表达式、雅可比矩阵、初值和根, 其中符号 t, b, h 的实际含义详见文献 [8]. 这两个实例的雅可比矩阵 $F'(X)$ 在初值点均奇异, 因此牛顿法是失效的.

表 4 两个实际问题的相关信息

实例	表达式	雅可比矩阵
实例 1	$\begin{cases} 2tb + t(h - 2t) = 12, \\ \frac{bh^3}{12} - \frac{(b-t)(h-2t)^3}{12} = 12, \\ \frac{tb^3}{6} + \frac{(h-2t)t^3}{12} = 12 \end{cases}$	$\begin{bmatrix} 2b + h - 4t & 2t & t \\ \frac{(h-2t)^2(h-8t+6b)}{12} & \frac{h^3 - (h-2t)^3}{12} & \frac{bh^2 - (b-t)(h-2t)^2}{4} \\ \frac{2b^3 + 3ht^2 - 8t^3}{12} & \frac{tb^2}{2} & \frac{t^3}{12} \end{bmatrix}$
实例 2	$\begin{cases} bh - (b-2t)(h-2t) = 666, \\ \frac{bh^3}{12} - \frac{(b-2t)(h-2t)^3}{12} = 9143, \\ \frac{hb^3}{12} - \frac{(h-2t)(b-2t)^3}{12} = 64783 \end{cases}$	$\begin{bmatrix} 2(h+b-4t) & 2t & 2t \\ \frac{(h-2t)^2(h-8t+3b)}{6} & \frac{h^3 - (h-2t)^3}{12} & \frac{bh^2 - (b-2t)(h-2t)^2}{4} \\ \frac{(b-2t)^2(b-8t+3h)}{6} & \frac{hb^2 - (h-2t)(b-2t)^2}{4} & \frac{b^3 - (b-2t)^3}{12} \end{bmatrix}$
实例	初值	根
实例 1	$(3.46, 3.46, 3.46)^T$	$(3.464\ 101\ 615, 3.464\ 101\ 615, 3.464\ 101\ 615)^T$
实例 2	$(12.90, 9.12, 42.48)^T$	$(12.903\ 487\ 9, 9.126\ 313\ 4, 42.487\ 638\ 2)^T$

算法QMn-M和算法PC-M比较了前面的5种奇异算法,所取的参数值见表5.算法PC-M和QMn-M仍选取了相同的参数值.

表6给出了迭代次数,函数评价指数及CPU时间

的计算结果.

从表6可以看出,KJ对于这两个实际问题求解失败.在收敛算例中,算法QMn-M和PC-M的3个指标的结果分别优于其他算法.

表5 针对两个实例的比较算法的参数值

实例	KLW ^[3]	WX ^[4]	GN ^[5]
实例1	$\lambda_1^{(n)} = -0.240\ 467, \lambda_{2,3}^{(n)} = -0.529\ 142$	$\lambda_{1,2,3}^{(n)} = -1$	$\nu_1^{(n)} = -0.240\ 467, \nu_{2,3}^{(n)} = -0.529\ 142, m = (1, 1, 1)^T$
实例2	$\lambda_{1,2}^{(n)} = -1, \lambda_3^{(n)} = -0.1$	$\lambda_{1,2,3}^{(n)} = -1$	$\nu_{1,2}^{(n)} = -1, \nu_3^{(n)} = -0.1, m = (1, 1, 1)^T$
实例	KJ ^[7]	PC-M算法和QMn-M算法	
实例1	$\lambda_{1,2,3}^{(n)} = 0.01, \mu_{1,2,3}^{(n)} = 0.01$	$\lambda_{1,2,3}^{(n)} = 0.000\ 1, \mu_1^{(n)} = -0.240\ 467, \mu_{2,3}^{(n)} = -0.529\ 142$	
实例2	$\lambda_{1,2,3}^{(n)} = 0.01, \mu_{1,2,3}^{(n)} = 0.1$	$\lambda_{1,2,3}^{(n)} = -1, \mu_{1,2}^{(n)} = -1, \mu_3^{(n)} = -0.1$	

表6 两个实际问题的数值结果比较

实例	指标	KLW ^[3]	WX ^[4]	GN ^[5]	EUNM ^[6]	KJ ^[7]	PC-M算法	QMn-M算法
实例1	迭代次数	23	42	23	53	Failure	6	6
	函数评价指数	276	504	276	636	Failure	72	72
	CPU时间	2.086 161 1	3.768 606	2.349 275	5.111 507	Failure	0.738 877	0.627 894
实例2	迭代次数	41	42	41	66	Failure	30	32
	函数评价指数	451	462	451	726	Failure	330	352
	CPU时间	5.077 575	5.210 622	5.776 266	8.079 101	Failure	4.171 664	4.389 710

5 结论

本文针对具有奇异雅可比矩阵的非线性方程组的求解问题,提出了一种阶为2的牛顿预测-校正格式.采用测试函数和实际问题验证了算法的有效性,并同牛顿算法和5种奇异算法进行了比较.实验结果表明,所提出的奇异算法是有效可行的.此外,对于这类牛顿奇异算法参数的选取,本文仅考虑了令第1次迭代后所得到的 $X^{(1)}$ 足够接近于根.然而,这种选取方案不一定适用于所有的测试函数,对于适用范围更广的参数选取方法,将在以后的工作中作进一步的探讨.

参考文献(References)

[1] 吕巍,隋瑞瑞,冯恩民.改进的牛顿预测-校正格式[J].控制理论与应用,2015,32(12): 1620-1626. (Lv W, Sui R R, Feng E M. Predictor-corrector improvement of Newton method[J]. J of Control Theory and Applications, 2015, 32(12): 1620-1626.)

[2] Howk C L. A class of efficient quadrature-based predictor-corrector methods for solving nonlinear systems[J]. Applied Mathematics and Computation, 2016, 276: 394-406.

[3] Kou J, Li Y, Wang X. Efficient continuation Newton-like

method for solving systems of non-linear equations[J]. Applied Mathematics and Computation, 2006, 174(2): 846-853.

[4] Wu X. Note on the improvement of Newton's method for system of nonlinear equations[J]. Applied Mathematics and Computation, 2007, 189(2): 1476-1479.

[5] Jose L H, Eulalia M, Juan R M. Modified Newtons method for systems of nonlinear equations with singular Jacobian[J]. J of Computational and Applied Mathematics, 2009, 224(1): 77-83.

[6] Aisha H A, Fatima W L, Waziri M Y. Newton's-like method for solving systems of nonlinear equations with singular jacobian[J]. Int J of Computer Applications, 2014, 98(13): 1-3.

[7] Kou J. A third-order modification of Newton method for systems of non-linear equations[J]. Applied Mathematics and Computation, 2007, 191(1): 117-121.

[8] 季振义,吴文渊,冯勇.一类非线性方程组奇异解的计算方法及其应用[J].计算机应用,2013,33(1): 230-233. (Ji Z Y, Wu W Y, Feng Y. Calculation method for singular solutions of a class of nonlinear equations and its application[J]. J of Computer Applications, 2013, 33(1): 230-233.)

(责任编辑: 闫妍)