文章编号：1004-4353(2020)04-0302-07

一种改进的求解非线性方程组的 Levenberg-Marquardt 方法

伍珍香， 陈亮*， 周童
(淮北师范大学 数学科学学院，安徽 淮北 235000)

摘要：通过修改 Levenberg-Marquardt 参数，得到了一种改进的求解非线性方程组的 Levenberg-Marquardt 算法. 利用信赖域技术，在不必假设雅克比矩阵非奇异的局部误差界条件下，证明了该算法至少具有超线性收敛性. 数值实验表明，该算法能有效求解非线性方程组问题.

关键词：非线性方程组；局部误差界；二次收敛；Levenberg-Marquardt

中图分类号：O221.1 **文献标识码**：A

A new modified Levenberg-Marquardt method for systems of nonlinear equations

WU Zhenxiang， CHEN Liang*， ZHOU Tong
(*School of Mathematical Sciences*, *Huaibei Normal University*, *Huaibei* 235000, *China*)

Abstract: By modifying the Levenberg-Marquardt parameters, we obtain a new convergent Levenberg-Marquardt algorithm for solving the systems of nonlinear equations. By using trust region technique, under the condition of local error bounds, the convergence of the new Levenberg-Marquardt method is shown to be at least super linearly without the non-singularity assumption of the Jacobi matrix. Numerical experiments show that the new algorithm can solve nonlinear equations effectively.

Keywords: systems of nonlinear equations; local error bound condition; local convergence; Levenberg-Marquardt

0 引言

本文考虑如下非线性方程组

$$F(x)=0 \tag{1}$$

的求解，其中 $F(x): R^n \to R^n$ 是连续可微函数. 在本文中总假设方程(1)的解集为非空，记为 X^*，且 $\|\cdot\|$ 表示 2-范数. 牛顿法是解方程组(1)的一个较为有效的方法，该方法在每次迭代中使用的搜索方向为

$$d_k^N=-\boldsymbol{J}_k^{-1}F_k. \tag{2}$$

其中 $F_k=F(x_k)$，$\boldsymbol{J}_k(\boldsymbol{J}_k=F'(x_k))$ 是 $F(x)$ 在 x_k 处的雅克比矩阵. 当 $\boldsymbol{J}(x)$ 矩阵为 Lipshitz 连续且为非奇异时，牛顿法产生的迭代点列二阶收敛于方程组(1)的解；但当 J_k 是奇异矩阵或者接近于奇异时，牛

收稿日期：2020-09-28 　*** 通信作者**：陈亮(1977—)，男，博士，教授，研究方向为数值最优化和数值代数.
基金项目：安徽省高校优秀青年骨干人才国外访问研修项目(GXGWFX2019022)；安徽省高校自然科学研究项目(KJ2020ZD008，KJ2019A0604)；安徽省省级质量工程项目(2019MOOC158)

顿步(2)无意义，即此时牛顿法不再适用. 为此，Levenberg 和 Marquardt 提出了解决上述问题的一个有效方法——Levenberg-Marquardt 方法(简称 LM 方法)[1-2]. LM 方法的试探步为

$$d_k = -(\boldsymbol{J}_k^{\mathrm{T}}\boldsymbol{J}_k + \lambda_k \boldsymbol{I})^{-1}\boldsymbol{J}_k^{\mathrm{T}}F_k, \tag{3}$$

其中 $\lambda_k(\lambda_k > 0)$ 是迭代参数. 由式(3)显然可知：当 $\lambda_k = 0$ 且 $\boldsymbol{J}_k$ 为非奇异矩阵时，式(3)中的试探步 d_k 等价于牛顿步(2)；当 $\boldsymbol{J}(x)$ 在解为非奇异且 Lipshitz 连续时，LM 法同样具有二阶收敛性. 研究表明，迭代参数 λ_k 的选取对 LM 方法的收敛性具有重要作用. 例如：取 $\lambda_k = \|F(x_k)\|^2$ 时，LM 法在局部误差界的条件约束下具有二阶收敛性[3]. 当采用信赖域技术且取 $\lambda_k = \mu_k\|F(x_k)\|$ 时，LM 法在局部误差界的条件约束下具有局部二阶收敛性；但该选取在 x_k 远离解集时，因 $\|F_k\|$ 较大，因此此时 LM 法的计算效率较低[4]. 取 $\lambda_k = \dfrac{\mu_k\|F_k\|}{1+\|F_k\|}$ 时，LM 法在不必假设 $\boldsymbol{J}_k$ 为非奇异的局部误差界的条件约束下具有二阶收敛性，并且可有效减小因 x_k 远离解集时 $\|F_k\|$ 较大所产生的影响[5]. 为了节省雅克比矩阵的计算量，文献[6-8]的作者还提出了多步迭代的方法，该方法有效地提高了计算效率. 为了讨论 LM 参数的取值范围，受文献[5]的启发，本文令 $\lambda_k = \dfrac{\mu_k\|F_k\|^\delta}{1+\|F_k\|^\delta}$, $\delta \in (0,2]$. 显然，由该式可知当 x_k 远离解集时，$\|F_k\|$ 较大，$\dfrac{\|F_k\|^\delta}{1+\|F_k\|^\delta}$ 趋近于 1，λ_k 接近于 μ_k；当 x_k 接近于解集时，$\|F_k\|$ 趋向于 0，λ_k 接近于 $\mu_k\|F_k\|^\delta$.

定义价值函数 $\phi(x) = \|F(x)\|^2$，则该函数在第 k 次迭代的预测下降量($Pred_k$)和实际下降量($Ared_k$)可定义为如下形式：

$$Ared_k = \|F_k\|^2 - \|F(x_k + d_k)\|^2,\ Pred_k = \|F_k\|^2 - \|F_k + \boldsymbol{J}_k d_k\|^2. \tag{4}$$

其中 d_k 由式(3)得到，其比率为 $r_k = \dfrac{Ared_k}{Pred_k}$. r_k 决定是否接受步长 d_k，以及用于调整新参数 μ_k 的大小. K.Amini 等[5]研究发现，采用非单调策略的算法比采用单调策略的算法效率更高，因此可将实际下降量表示为

$$\overline{A}red_k = F_{l(k)}^2 - \|F(x_k + d_k)^2\|. \tag{5}$$

其中

$$F_{l(k)} = \max_{0 \leqslant j \leqslant n(k)} \{\|F_{k-j}\|\}, \tag{6}$$

$k = 0,1,2,\cdots,n$；$n(k) = \min\{N_0, k\}$，N_0 是一个正整数. 由式(6)可知，实际下降量的改变使得每次迭代得到的新点都与前 $n(k)$ 次迭代中的最差的点进行比较，从而影响比率 $\bar{r}_k$：

$$\bar{r}_k = \frac{\overline{A}red_k}{Pred_k}. \tag{7}$$

1　算法

新修正的 LM 法(算法 1)的计算步骤如下：

步骤 1　起始点 $x_0 \in R^n$，参数 $N_0 > 0$, $\mu_0 > m > 0$, $\varepsilon > 0$, $0 < p_0 < p_1 < p_2 < 1$, $k := 0$.

步骤 2　如果 $\|\boldsymbol{J}_k^{\mathrm{T}}F_k\| \leqslant \varepsilon$，则终止；反之，有

$$\lambda_k = \frac{\mu_k\|F_k\|^\delta}{1+\|F_k\|^\delta},\ \delta \in (0,2]. \tag{8}$$

步骤 3　解方程$(\boldsymbol{J}_k^{\mathrm{T}}F_k + \lambda_k \boldsymbol{I})d = -\boldsymbol{J}_k^{\mathrm{T}}F_k$，求出 d_k.

步骤 4　由式(4)—(7)求出 $Pred_k$、$\overline{A}red_k$、$F_{l(k)}$、$\bar{r}_k$.

步骤 5 令 $x_{k+1}=\begin{cases}x_k+d_k,\ \bar{r}_k\geqslant p_0;\\ x_k,\ \bar{r}_k<p.\end{cases}$

步骤 6 令 $\mu_{k+1}=\begin{cases}4\mu_k,\ \text{如果}\ \bar{r}_k<p_1;\\ \mu_k,\ \text{如果}\ \bar{r}_k\in[p_1,p_2];\\ \max\left\{\dfrac{\mu_k}{4},m\right\},\ \text{其他}.\end{cases}$

步骤 7 令 $k=k+1$,然后转到第 1 步.

算法 1 中要求 $\mu_k\geqslant m$,其中 m 是一个大于零的常数,即

$$\mu_k\geqslant m,\forall k\in\mathbf{N}. \tag{9}$$

这一要求可由步骤 1 给定的初始条件和步骤 6 满足.

2 局部收敛性

定义 1 N 是 R^n 上的一个集合,使得 $N\cap X^*\neq\varnothing$. 如果存在一个常数 $c\,(c>0)$ 使得

$$\|F(x)\|\geqslant c\,\mathrm{dist}(x,X^*),\forall x\in\mathbf{N},$$

$$\mathrm{dist}(x,X)=\inf_{y\in X}\|y-x\|, \tag{10}$$

则称 $\|F(x)\|$ 在 N 上为方程(1) 提供了一个局部误差界.

定义 $\bar{x}_k\in X^*$, $\bar{x}_k$ 表示 X^* 中满足等式 $\|\bar{x}_k-x_k\|=\mathrm{dist}(x_k,X^*)$ 的解.

假设 1[9] (a) $F(x)$ 连续可微并且 $\|F(x)\|$ 在 $N(x^*,b)$ 上为方程(1) 可提供局部误差边界,其中 $0<b<1$, $N(x^*,b)=\{x\in R^n\mid \|x-x^*\|\leqslant b\}$.

(b) $F(x)$ 和 $\boldsymbol{J}(x)$ 在 $N(x^*,b)$ 上 Lipshitz 连续可微,即存在两个常数 $L_1(L_1>0)$ 和 $L_2(L_2>0)$ 使:

$$\|J(y)-J(x)\|\leqslant L_1\|y-x\|,\forall x,y\in N(x^*,b); \tag{11}$$

$$\|F(y)-F(x)\|\leqslant L_2\|y-x\|,\forall x,y\in N(x^*,b). \tag{12}$$

由假设(a) 和(b) 可知

$$\|F(y)-F(x)-J(x)(y-x)\|\leqslant L_1\|y-x\|^2,\forall x,y\in N(x^*,b). \tag{13}$$

引理 1 在假设 1 成立的条件下,对所有充分大的 k 都有 $\|d_k\|\leqslant \mathrm{O}(\|x_k-\bar{x}_k\|)$ 成立.

证明 设 $\varphi_k(d)=\|F_k+J_kd\|^2+\lambda_k\|d\|^2$. 由式(3) 可知 d_k 是 φ_k 的极小值,则再由式(8)、(9)、(12)、(13) 以及 $F(\bar{x}_k)=0$ 可得

$$\|d_k\|^2\leqslant\frac{1}{\lambda_k}\varphi_k(\bar{x}_k-x_k)=\frac{1}{\lambda_k}(\|F_k+J_k(\bar{x}_k-x_k)\|^2+\lambda_k\|\bar{x}_k-x_k\|^2)=$$
$$\frac{1+\|F_k\|^\delta}{\mu_k\|F_k\|^\delta}(\|F_k+J_k(\bar{x}_k-x_k)\|^2)+\|\bar{x}_k-x_k\|^2\leqslant$$
$$\frac{1+L_2^\delta\|\bar{x}_k-x_k\|^\delta}{cm\|\bar{x}_k-x_k\|^\delta}(L_1^2\|\bar{x}_k-x_k\|^4)+\|\bar{x}_k-x_k\|^2=\mathrm{O}(\|\bar{x}_k-x_k\|^2).$$

引理 2 在假设 1 成立的情况下,对所有充分大的 k,有:

(a) 存在一个正的常数 $M>m$,使得

$$\mu_k\leqslant M. \tag{14}$$

$$\text{(b)}\ m\gamma<\lambda_k\leqslant ML_2^\delta\|\bar{x}_k-x_k\|^\delta,\ \gamma=\min\left\{\frac{1}{2},\frac{c}{2}\|\bar{x}_k-x_k\|^\delta\right\}. \tag{15}$$

证明 因引理 2 中的(a) 在文献[5] 中已有具体证明,故在此省略. 下证引理 2 中的(b). 对于任意

足够大的 k，由式(12)、(14) 和 $F(\bar{x}_k)=0$ 有

$$\lambda_k=\frac{\mu_k\|F_k\|^\delta}{1+\|F_k\|^\delta}\leqslant\mu_k\|F_k\|^\delta\leqslant ML_2^\delta\|\bar{x}_k-x_k\|^\delta. \tag{16}$$

下面证明不等式的左半部分. 若 $\|F_k\|\leqslant 1$，则 $\|F_k\|^\delta\leqslant 1$，因此 $1+\|F_k\|^\delta\leqslant 2$. 再根据式(8)、(9) 和局部误差界的条件有 $\lambda_k=\frac{\mu_k\|F_k\|^\delta}{1+\|F_k\|^\delta}\geqslant\frac{\mu_k}{2}\|F_k\|^\delta\geqslant\frac{mc}{2}\|\bar{x}_k-x_k\|^\delta$. 若 $\|F_k\|>1$，则 $1+\|F_k\|^\delta<2\|F_k\|^\delta$，因此 $\lambda_k=\frac{\mu_k\|F_k\|^\delta}{1+\|F_k\|^\delta}>\frac{\mu_k}{2}\geqslant\frac{m}{2}$. 令 $\gamma=\min\left\{\frac{1}{2},\frac{c}{2}\|\bar{x}_k-x_k\|^\delta\right\}$，于是有 $m\gamma<\lambda_k$. 再结合式(16) 可得 $m\gamma<\lambda_k\leqslant ML_2^\delta\|\bar{x}_k-x_k\|^\delta$，证毕.

下面利用奇异值分解(SVD) 证明算法 1 的局部收敛性. 设雅克比矩阵 $\boldsymbol{J}(\bar{x})$ 的奇异值分解为：

$$\boldsymbol{J}(\bar{x})=[\bar{\boldsymbol{U}}_1,\bar{\boldsymbol{U}}_2]\begin{pmatrix}\bar{\sigma}_1 &&&&& \\ &\ddots&&&& \\ &&\bar{\sigma}_r&&& \\ &&&0&& \\ &&&&\ddots& \\ &&&&&0\end{pmatrix}\begin{bmatrix}\bar{\boldsymbol{V}}_1^{\mathrm{T}}\\ \bar{\boldsymbol{V}}_2^{\mathrm{T}}\end{bmatrix}=\bar{\boldsymbol{U}}_1\bar{\boldsymbol{\Sigma}}_1\bar{\boldsymbol{V}}_1^{\mathrm{T}}.$$

其中 $\bar{\sigma}_1\geqslant\bar{\sigma}_2\geqslant\cdots\geqslant\bar{\sigma}_r>0$，$\mathrm{rank}(\bar{\boldsymbol{\Sigma}}_1)=r$，$\bar{\boldsymbol{U}}=[\bar{\boldsymbol{U}}_1,\bar{\boldsymbol{U}}_2]$ 和 $\bar{\boldsymbol{V}}=[\bar{\boldsymbol{V}}_1,\bar{\boldsymbol{V}}_2]$ 是两个正交矩阵. 相应地，设 $\boldsymbol{J}(x)$ 的奇异值分解为

$$\boldsymbol{J}(x)=\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}=$$

$$(\boldsymbol{U}_1,\boldsymbol{U}_2,\boldsymbol{U}_3)\begin{pmatrix}\sigma_1 &&&&&&&& \\ &\ddots&&&&&&& \\ &&\sigma_r&&&&&& \\ &&&\sigma_{r+1}&&&&& \\ &&&&\ddots&&&& \\ &&&&&\sigma_{r+q}&&& \\ &&&&&&\ddots&& \\ &&&&&&&0& \\ &&&&&&&&0\end{pmatrix}\begin{pmatrix}\boldsymbol{V}_1^{\mathrm{T}}\\ \boldsymbol{V}_2^{\mathrm{T}}\\ \boldsymbol{V}_3^{\mathrm{T}}\end{pmatrix}=\boldsymbol{U}_1\boldsymbol{\Sigma}_1\boldsymbol{V}_1^{\mathrm{T}}+\boldsymbol{U}_2\boldsymbol{\Sigma}_2\boldsymbol{V}_2^{\mathrm{T}}. \tag{17}$$

其中，$\boldsymbol{U}=[\boldsymbol{U}_1,\boldsymbol{U}_2,\boldsymbol{U}_3]$ 和 $\boldsymbol{V}=[\boldsymbol{V}_1,\boldsymbol{V}_2,\boldsymbol{V}_3]$ 是两个正交矩阵，$\boldsymbol{\Sigma}_1=\mathrm{diag}(\sigma_1,\sigma_2,\cdots,\sigma_r)$，$\sigma_1\geqslant\sigma_2\geqslant\cdots\geqslant\sigma_r>0$，$\boldsymbol{\Sigma}_2=\mathrm{diag}(\sigma_{r+1},\sigma_{r+2},\cdots,\sigma_{r+q})$，$\sigma_{r+1}\geqslant\sigma_{r+2}\geqslant\cdots\geqslant\sigma_{r+q}>0$. 将 $\boldsymbol{J}_k$ 的 SVD 分解形式写成式(17) 的形式，然后根据 $\boldsymbol{J}_k$ 的 Lipshitz 连续性以及矩阵摄动理论[10]，得

$$\|\mathrm{diag}(\boldsymbol{\Sigma}_1-\bar{\boldsymbol{\Sigma}}_1,\boldsymbol{\Sigma}_2,0)\|\leqslant\|\boldsymbol{J}_k-\bar{\boldsymbol{J}}_k\|\leqslant L_1\|\bar{x}_k-x_k\|.$$

因此有

$$\|\boldsymbol{\Sigma}_1-\bar{\boldsymbol{\Sigma}}_1\|\leqslant L_1\|\bar{x}_k-x_k\|,\ \|\boldsymbol{\Sigma}_2\|\leqslant L_1\|\bar{x}_k-x_k\|. \tag{18}$$

由于 $\{x_k\}$ 可收敛到解集 X^*，因此可假设对于足够大的 k 有 $L_1\|\bar{x}_k-x_k\|\leqslant\frac{\bar{\sigma}_r}{2}$，进而根据式(17) 有

$$\|\boldsymbol{\Sigma}_1^{-1}\|\leqslant\frac{1}{\bar{\sigma}_r-L_1\|\bar{x}_k-x_k\|}\leqslant\frac{2}{\bar{\sigma}_r}. \tag{19}$$

引理 3[5]　在假设 1 成立的条件下，对于充分大的 k 有：

(a) $\|\boldsymbol{U}_1\boldsymbol{U}_1^{\mathrm{T}}F_k\|\leqslant O(\|\bar{x}_k-x_k\|)$；

(b) $\|\boldsymbol{U}_2\boldsymbol{U}_2^{\mathrm{T}}F_k\| \leqslant \mathrm{O}(\|\bar{x}_k - x_k\|^2)$.

证明 (a) 和(b) 的证明可参考文献[5] 中的引理 3.3 的证明,在此省略.

定理 1 在假设 1 成立的条件下,当 $\delta \in (0,1)$ 时,算法 1 产生的序列 $\{x_k\}$ 超线性收敛于方程组(1) 的解;当 $\delta \in [1,2]$ 时,序列 $\{x_k\}$ 二阶收敛于方程组(1) 的解.

证明 根据式(10)、(11)、(15) 和引理 1 有

$$d_k = -\boldsymbol{V}_1(\boldsymbol{\Sigma}_1^2 + \lambda_k\boldsymbol{I})^{-1}\boldsymbol{\Sigma}_1\boldsymbol{U}_1^{\mathrm{T}}F_k - \boldsymbol{V}_2(\boldsymbol{\Sigma}_2^2 + \lambda_k\boldsymbol{I})^{-1}\boldsymbol{\Sigma}_2\boldsymbol{U}_2^{\mathrm{T}}F_k.$$

由式(15)、(19) 以及引理 3 有

$$F_k + \boldsymbol{J}_kd_k = F_k - \boldsymbol{U}_1\boldsymbol{\Sigma}_1(\boldsymbol{\Sigma}_1^2 + \lambda_k\boldsymbol{I})^{-1}\boldsymbol{\Sigma}_1\boldsymbol{U}_1^{\mathrm{T}}F_k - \boldsymbol{U}_2\boldsymbol{\Sigma}_2(\boldsymbol{\Sigma}_2^2 + \lambda_k\boldsymbol{I})^{-1}\boldsymbol{\Sigma}_2\boldsymbol{U}_2^{\mathrm{T}}F_k =$$
$$\lambda_k\boldsymbol{U}_1(\boldsymbol{\Sigma}_1^2 + \lambda_k\boldsymbol{I})^{-1}\boldsymbol{U}_1^{\mathrm{T}}F_k + \lambda_k\boldsymbol{U}_2(\boldsymbol{\Sigma}_2^2 + \lambda_k\boldsymbol{I})^{-1}\boldsymbol{U}_2^{\mathrm{T}}F_k.$$

因此

$$\|F_k + \boldsymbol{J}_kd_k\| \leqslant \lambda_k\|\boldsymbol{\Sigma}_1^{-2}\|\|\boldsymbol{U}_1\boldsymbol{U}_1^{\mathrm{T}}F\|_k + \|\boldsymbol{U}_2\boldsymbol{U}_2^{\mathrm{T}}F_k\| \leqslant$$
$$\frac{4L_1^{\delta}M}{\bar{\sigma}_r^2}\mathrm{O}(\|\bar{x}_k - x_k\|^{1+\delta}) + \mathrm{O}(\|\bar{x}_k - x_k\|^2) = \mathrm{O}(\|\bar{x}_k - x_k\|^{1+\delta}) + \mathrm{O}(\|\bar{x}_k - x_k\|^2). \tag{20}$$

根据公式(10)、(11)、(12)、(19) 和引理 1 有

$$c\|\bar{x}_{k+1} - x_{k+1}\| \leqslant \|F(x_{k+1})\| = \|F(x_k + d_k)\| \leqslant \|F(x_k) + \boldsymbol{J}_kd_k\| + L_1\|d_k\|^2 =$$
$$\mathrm{O}(\|\bar{x}_k - x_k\|^{1+\delta}) + \mathrm{O}(\|\bar{x}_k - x_k\|^2). \tag{21}$$

对于所有充分大的 k,有 $\|\bar{x}_k - x_k\| = \mathrm{dist}(x_k, X^*) \leqslant \|\bar{x}_{k+1} - x_k\| \leqslant \|\bar{x}_{k+1} - x_{k+1}\| + \|d_k\|$. 故由引理 1 知,对所有充分大的 k,有 $\|\bar{x}_k - x_k\| \leqslant 2\|d_k\| \leqslant \mathrm{O}(\|\bar{x}_k - x_k\|)$,所以 $\|d_k\| = \mathrm{O}(\|\bar{x}_k - x_k\|)$. 再由式(21) 可知,若 $\delta \in (0,1)$,则 $\|d_{k+1}\| = \mathrm{O}(\|d_k\|^{1+\delta})$;若 $\delta \in [1,2]$,则 $\|d_{k+1}\| = \mathrm{O}(\|d_k\|^2)$. 即当 $\delta \in (0,1)$ 时,迭代点列 $\{x_k\}$ 超线性收敛于方程组(1) 的解;当 $\delta \in [1,2]$ 时,点列 $\{x_k\}$ 二阶收敛于方程组(1) 的解.

3 数值实验

为了验证算法 1 的有效性,本文利用试验对算法 1 和文献[5] 中的算法 2.1 进行对比. 测验函数 $F(x)$ 是标准非奇异函数,来源于文献[11]. 采用文献[12] 中的方式对测验函数进行改进,结果为:

$$\hat{F}(x) = F(x) - \boldsymbol{J}(x^*)\boldsymbol{A}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}(x - x^*).$$

其中,$\boldsymbol{A} \in \boldsymbol{R}^{n\times k}(1 \leqslant k \leqslant n)$ 是一个列满秩的矩阵,$F(x^*) = 0$. 显然 $\hat{F}(x)$ 在 x^* 处的雅克比矩阵为 $\hat{\boldsymbol{J}}(x^*) = \boldsymbol{J}(x^*)(\boldsymbol{I} - \boldsymbol{A}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}})$,且 $\hat{F}(x^*) = 0$. 注意到 $\hat{F}(x)$ 的根未必是 $F(x)$ 的根,因此本文选择令 $\boldsymbol{A} = [1,1,\cdots,1]^{\mathrm{T}} \in \boldsymbol{R}^{n\times 1}$,$\hat{\boldsymbol{J}}(x^*)$ 的秩为 $n-1$.

算法 1 与文献[5] 中算法 2.1 的参数设置如下:$p_0 = 10^{-4}$,$p_1 = 0.25$,$p_2 = 0.75$,$N_0 = 5$,$\mu_1 = 1$,$m = 10^{-8}$. 算法的停止条件是 $\|\boldsymbol{J}_k^{\mathrm{T}}F_k\| \leqslant 10^{-5}$,或者迭代次数超过 1 000 次. 两种算法的数值结果见表 1. 在表 1 中:第 3 列表示迭代初始点为 $-10x_0$、$-x_0$、x_0、$10x_0$、$100x_0$,其中 x_0 是文献[11] 中的标准初始点;NF 表示函数的计算次数;NJ 表示雅可比矩阵的计算次数;NT ($NT = NF + NJ \times n$) 表示计算的总数;在 NS? 中,Y 表示算法收敛到解 x^*,N 表示算法收敛到其他解.

从表 1 中可以看出,在大部分测试问题中,算法 1 中的函数的计算次数、雅可比矩阵的计算次数以及总的计算次数大部分小于文献[5]中算法 2.1 的计算次数、雅可比矩阵的计算次数以及总的计算次数,由此说明算法 1 的计算量小于文献[5]中算法 2.1 的计算量,即算法 1 对非线性方程组的求解效率高于文献[5]中的算法 2.1.

表 1　数值试验数据和结果

问题	维数	x_0	文献[5]中的算法 2.1				算法 1			
			NF	*NJ*	*NT*	NS?	*NF*	*NJ*	*NT*	NS?
Rosenbrock	2	−10	7	17	51	N	15	15	45	N
		−1	11	11	33	N	10	10	30	N
		1	15	15	45	Y	15	15	45	Y
		10	17	17	51	Y	17	17	51	Y
		100	21	21	63	Y	21	21	63	N
Powell singular	4	−10	13	13	65	Y	12	12	60	Y
		−1	10	10	50	Y	9	9	45	Y
		1	10	10	50	N	10	10	50	N
		10	13	13	65	N	13	13	65	N
		100	16	16	80	N	16	16	80	N
Wood	4	−10	19	19	95	N	17	17	85	N
		−1	15	15	75	N	13	13	65	N
		1	16	16	80	Y	16	16	80	Y
		10	19	19	95	Y	19	19	95	Y
		100	22	22	110	Y	22	22	110	Y
Helical valley	3	−10	5	5	20	N	3	3	12	N
		−1	1	1	4	N	1	1	4	N
		1	8	8	32	N	8	8	32	N
		10	8	8	32	N	8	8	32	N
		100	8	8	32	N	8	8	32	N
Brown almost-linear	10	−10	22	22	242	N	21	21	231	Y
		−1	8	8	88	N	7	7	77	Y
		1	8	8	88	Y	8	8	88	Y
		10	23	23	253	Y	23	23	253	Y
		100	45	45	495	Y	45	45	495	Y
Discrete boundary value	10	−10	6	6	66	Y	7	7	77	Y
		−1	9	9	99	Y	4	4	44	Y
		1	4	4	44	N	3	3	33	N
		10	6	6	66	N	6	6	66	N
		100	9	9	99	N	9	9	99	N
Discrete integral equation	30	−10	11	11	341	N	9	9	279	Y
		−1	8	8	248	N	6	6	186	Y
		1	6	6	186	Y	6	6	186	Y
		10	7	7	217	Y	7	7	217	Y
		100	10	10	310	N	10	10	310	N
Trigonometric	30	−10	30	17	540	Y	19	10	319	Y
		−1	26	14	446	Y	22	12	382	Y
		1	34	18	574	N	27	17	537	N
		10	65	44	1 385	N	57	45	1 407	N
		100	59	49	1 529	N	40	35	1 090	N
Variably dimensioned function	10	−10	17	17	187	N	15	15	165	N
		−1	15	15	165	N	14	14	154	N
		1	14	14	154	Y	14	14	154	Y
		10	16	16	176	Y	16	16	176	Y
		100	19	19	209	Y	19	19	209	Y
Broyden tridiagonal	30	−10	30	18	570	N	21	13	411	N
		−1	20	12	380	N	24	15	474	Y
		1	9	9	279	Y	9	9	279	Y
		10	14	14	434	Y	14	14	434	Y
		100	17	17	527	N	17	17	527	Y
Broyden banded	30	−10	10	10	310	Y	10	10	310	Y
		−1	23	15	473	N	21	12	381	N
		1	12	12	372	Y	12	12	372	Y
		10	18	18	558	Y	18	18	558	Y
		100	24	24	744	Y	24	24	744	Y

（下转第 332 页）

参考文献:

[1] 黄鹏.纳米材料的可控合成、功能化及其在肿瘤诊疗中应用研究[D].上海:上海交通大学,2012.

[2] JNANRANJAN P, BHABANI S S, SUMIT M, et al. Engineered polymeric iron oxide nanoparticles as potential drug carrier for targeted delivery of docetaxel to breast cancer cells[J]. Journal of Magnetism and Magnetic Materials, 2019,485:165-173.

[3] MAEDA H, TSUKIGAWA K, FANG J. A retrospective 30 years after discovery of the enhanced permeability and retention effect of solid tumors: next-generation chemotherapeutics and photodynamic therapy: problems, solutions, and prospects [J]. Microcirculation, 2016,23(3):173-182.

[4] ZHAO T T, LIU Y P, GAO Z, et al. Self-assembly and cytotoxicity study of PEG-modified ursolic acid liposomes[J]. Materials Science and Engineering: C, 2015,53:196-203.

[5] VAN EEIJK G A, VROEGE G, PHILIPSE A. Convenient preparation methods for magnetic colloids[J]. Journal of Magnetism & Magnetic Materials, 1999,201(1-3):31-33.

[6] 杨萍.人参皂苷 Rg_3 长循环脂质体的制备及评价[D].延吉:延边大学,2017.

[7] 国家药典委员会.中华人民共和国药典:一部[S].北京:中国医药科技出版社,2015:8.

[8] 李海量.聚合物修饰的叶酸靶向磁性纳米传递系统的制备及其对鼻咽癌的抑制和成像作用[D].广州:南方医科大学,2017.

[9] MOK H J, ZHANG M Q. Superparamagnetic iron oxide nanoparticle-based delivery systems for biotherapeutics[J]. Expert Opinion on Drug Delivery, 2013,10(1):73-87.

[10] ZHOU H, FAN Z Y, DENG J J, et al. Hyaluronidase embedded in nanocarrier PEG shell for enhanced tumor penetration and highly efficient antitumor efficacy [J]. Nano Letters, 2016, 16 (5):3268-3277.

[11] ZHENG W W, ZHOU K R, CHEN Z W, et al. Characterization of focal hepatic lesions with SPIO-enhanced MRI[J]. World Journal of Gastroenterology, 2002,8(1):82-86.

[12] REYNOLDS P R, LARKMAN D J, HASKARD D O, et al. Detection of vascular expression of E-selectin in vivo with MR imaging[J]. Radiology, 2006,241(2):469-476.

(上接第 307 页)

参考文献:

[1] LEVENBERG K. A method for the solution of certain nonlinear problems in least squares[J]. Quart Appl Math, 1944,2:164-168.

[2] MARQUARDT D W. An algorithm for least-squares estimation of nonlinear parameters[J]. J Soc Ind Appl Math, 1963,11:431-441.

[3] YAMASHITA N, FUKUSHIMA M. On the rate of convergence of the Levenberg-Marquardt method[J]. Computing, 2001,15:239-249 .

[4] FAN J Y, YUAN Y X. On the quadratic convergence of the Levenberg-Marquardt method[J]. Computing, 2005, 74:23-39.

[5] AMINI K, ROSTAMI F, GIUSEPPE G. An efficient Levenberg-Marquardt method with a new LM parameter for systems of nonlinear equations[J]. J Comput Appl Math, 2018,67(5):637-650.

[6] CHEN L, MA Y F. Shamanskii-Like Levenberg-Marquardt method with a new line search for systems of nonlinear equations[J]. J Syst Sci Complex, 2020,33:1694-1707.

[7] LIANG M, ZHENG B, ZHENG Y T, et al. A two-step accelerated Levenberg-Marquardt method for solving multilinear systems in tensor-train format[J]. Journal of Computational and Applied Mathematics, 2021,382:113069.

[8] FAN J Y, HUANG J C, PAN J Y. An adaptive multi-step Levenberg-Marquardt method[J]. Journal of Scientific Computing, 2019,78:531-548.

[9] 杨柳,陈艳萍.求解非线性方程组的一种新的全局收敛的 Levenberg-Marquardt 算法[J].计算数学,2008,30(4):388-396.

[10] STEWART G W, SUN J G. Matrix Perturbation Theory[M]. San Diego (CA): Academic Press, 1990.

[11] MORÉ J J, GARBOW B S, HILLSTROM K E. Testing unconstrained optimization software[J]. ACM Trans Math Softw, 1981,7:17-41.

[12] SCHNABEL R B, FRANK P D. Tensor methods for nonlinear equations[J]. SIAM J Numer Anal, 1984,21:815-843.