

一类求解非凸无约束优化问题的改进L-BFGS方法

杨 博¹, 邢宇航², 刘粉干¹, 鲁娅妮¹

(1. 95841部队, 甘肃 酒泉 735018;

2. 火箭军工程大学, 陕西 西安 710025)

摘 要: 针对传统有限记忆BFGS (Limited-memory BFGS, L-BFGS) 方法在求解非凸函数极小值问题时不一定全局收敛的问题, 从非凸目标函数的曲率信息和方法的全局收敛性考虑, 提出了一种新的带参迭代方程, 并用该方程修正了L-BFGS算法; 最后, 对改进算法 (L-MBFGS) 进行了收敛性证明及数值实验验证。理论分析表明: 该方法对于一般函数 (可能非凸) 既保证了Hessian矩阵的正定性, 又具有充分下降性和全局收敛性。数值实验结果表明: 相同数据规模下, L-MBFGS方法在Wood测试函数中的最优值可优于标准L-BFGS方法1个数量级, 优于ML-BFGS相似方法2个数量级; 在Dixon测试函数中的计算效率也明显优于L-BFGS方法和ML-BFGS相似方法。

关键词: 非凸无约束优化; 改进的L-BFGS方法; 全局收敛性; Wolfe线搜索准则

中图分类号: O224

文献标志码: A

文章编号: 2097-2741(2024)06-067-06

A Modified L-BFGS Method for Nonconvex Unconstrained Optimization

YANG Bo¹, XING Yuhang², LIU Feng'an¹, LU Yani¹

(1. Troops No. 95841, Jiuquan 735018, Gansu;

2. Rocket Force University of Engineering, Xi'an 710025, Shaanxi)

ABSTRACT: To solve the problem that the traditional limited-memory BFGS (L-BFGS) does not always complete global convergence in nonconvex function minimum solution, a novel parameterized iterative equation was proposed with the consideration of the curvature information of nonconvex objective functions and the global convergence of the method. Furthermore, the proposed equation was used to modify the L-BFGS algorithm. Additionally, the improved algorithm (L-MBFGS) was verified by convergence and numerical experiments. Theoretical analysis indicated that L-MBFGS not only guarantees the positive definiteness of the Hessian matrix for general functions (which may be non-convex), but also has sufficient descent and global convergence. Numerical experimental results showed that under the same data size, the optimal value of the L-MBFGS in the Wood test function can be one order of magnitude better than the standard L-BFGS. Simultaneously, its computing efficiency in the Dixon test function significantly better than the standard L-BFGS and ML-BFGS similar method.

KEYWORDS: nonconvex unconstrained optimization; improved L-BFGS method; global convergence; Wolfe linear search criteria

收稿日期: 2024-04-08

第一作者: 杨博 (1998—), 女, 工程师, 主要从事雷达测试和数据优化处理研究。E-mail: 24667208@qq.com

引用格式: 杨博, 邢宇航, 刘粉干, 等. 一类求解非凸无约束优化问题的改进L-BFGS方法[J]. 火箭军工程大学学报, 2024, 38(6): 67-72, 90.

BFGS 方法是学者 Broyden^[1]、Fletcher^[2]、Goldfarb^[3]和 Shanno^[4]提出的一种拟牛顿方法，是求解无约束优化问题

$$\min f(x) \quad x \in \mathbb{R}^n \quad (1)$$

最常用也是最有效的方法，其中， $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 是非线性函数。但是 BFGS 算法的修正，仅限于目标函数为凸函数时成立，如果目标函数为非凸函数，Wolfe 线搜索准则无法保证每次迭代都满足 $s_k^T y_k > 0$ ，因此，有必要对 B_{k+1} 重新校正。2018 年，王锋^[5]提出了一类新的求解非凸函数的改进 BFGS 方法，并证明了其收敛性，但未进行数值实验；2020 年，刘克刚^[6]沿用保守修正方法，保证了 $y_k^T s_k > 0$ ，使算法具有了正定遗传性，但在求解大规模优化问题时计算效率不高；2011 年，Biglari 等^[7]提出了一种新的拟牛顿方程；2014 年，Biglari 等^[8]将上述方程推广至有限记忆 BFGS (Limited-Memory BFGS, L-BFGS) 方法的框架内，用于求解大规模无约束优化问题，改善了该方法的性能；2021 年，陈昱含^[9]将 Dehghani^[10]的思想和 Biglari 等^[7-8]基于四阶张量提出的割线方程进行改进，得到一个新的修正割线方程，并基于该方程去修正 L-BFGS 方法，得到一类修正的 L-BFGS 方法，但不能跟踪计算结果在不同数据规模上的动态变化过程。为了克服上述不足，本文设计了一类新的带参修正拟牛顿方程，将其推广应用到求解非凸无约束优化问题的 L-BFGS 方法中，在非凸假设条件下进行了该方法的全局收敛性证明，并用典型函数进行实验验证，取得了较优的数值效果。

1 传统 L-BFGS 方法

求解式 (1) 无约束优化问题的 BFGS 方法，通过迭代公式

$$x_{k+1} = x_k + \alpha_k d_k$$

和搜索方向

$$d_k = -B_k^{-1} g_k$$

满足方程

$$B_{k+1} s_k = y_k$$

式中： α_k 为迭代步长； g_k 为目标函数在点 x_k 处的梯度向量。 B_{k+1} 的更新式为

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k} \quad (2)$$

式中： $s_k = x_{k+1} - x_k$ ； $y_k = g_{k+1} - g_k$ ； B_0 为给定的初始矩阵。

利用 SWM (Sherman-Morrison-Woodbury) 矩阵求逆公式，对式 (2) 进行两次求逆运算，可得 BFGS 的另外一种拟牛顿方程

$$H_{k+1} y_k = s_k$$

$$H_{k+1} = \left(I - \frac{s_k y_k^T}{y_k^T s_k} \right) H_k \left(I - \frac{y_k s_k^T}{y_k^T s_k} \right) + \frac{s_k s_k^T}{y_k^T s_k} \quad (3)$$

式中： H_{k+1} 是 B_{k+1} 的逆。

2 改进的 L-BFGS 算法 (L-MBFGS)

本文结合文献 [5, 7, 9] 的改进思想，提出基于如下修正割线方程的改进 L-BFGS 方法 (L-MBFGS)

$$H_{k+1} y_k^+ = s_k \quad (4)$$

$$y_k^+ = \bar{y}_k + \xi \ln(1 + \|g_k\|^2) s_k + \max \left\{ -\frac{s_k^T \bar{y}_k}{\|s_k\|^2}, 0 \right\} s_k \quad (5)$$

$$\bar{y}_k = y_k + l_k \frac{\theta_k}{s_k^T s_k} s_k \quad (6)$$

$$\theta_k = 4 (f_k - f_{k+1}) + 2 (g_k + g_{k+1})^T s_k$$

$$l_k = \begin{cases} \delta^{-\|s_k\|}, & \|s_k\| \leq 1, \delta \in [2.5, 2.9] \\ 0, & \text{其他} \end{cases}$$

式中： δ 为本文改进方法提出的可调节参数，目的是限制参数 l_k 的取值范围，但不固定其取值，使其具有一定的选择性和可动态跟踪过程性，并且保证了 $(s_k)^T y_k^+ > 0$ ，即算法具有正定遗传性，可用于求解大规模非凸无约束优化问题。

记

$$\rho_k = \frac{1}{(s_k)^T y_k^+}$$

$$V_k = I - \rho_k y_k^+ s_k^T$$

式中： I 为单位矩阵。则

$$H_{k+1} = V_k^T H_k V_k + \rho_k s_k s_k^T$$

若令式 (4) 中 H_{k+1} 的初始矩阵为 H_k^0 ，重复上式 m 次，则

$$H_{k+1} = V_k^T H_k V_k + \rho_k s_k s_k^T =$$

$$V_k^T (V_{k-1}^T H_{k-1} V_{k-1} + \rho_{k-1} y_{k-1}^+ s_{k-1}^T) V_k +$$

$$\rho_k y_k^+ s_k^T = \cdots =$$

$$(V_k^T V_{k-1}^T \cdots V_{k-m+1}^T) H_{k-m+1} (V_{k-m+1} \cdots V_k) +$$

$$\rho_{k-m+1} (V_k^T V_{k-1}^T \cdots V_{k-m+2}^T) s_{k-m+1} s_{k-m+1}^T \cdot$$

$$(V_{k-m+2} \cdots V_{k-1}) + \cdots + \rho_k s_k s_k^T \quad (7)$$

L-MBFGS 算法步骤具体描述如下。

Step1 取初始点 $x_0 \in \mathbb{R}^n$ ，初始对称正定矩阵

$H_0 \in \mathbb{R}^{n \times n}$, $\sigma_1 \in (0, 1/2)$, $\sigma_2 \in (\sigma_1, 1)$, 内存大小 $m > 0$, 精度 $\varepsilon > 0$, 令 $k = 0$ 。

Step2 如果 $\|g_k\| \leq \varepsilon$, 算法终止, 问题的解为 x_k 。否则, 令

$$d_k = -H_k g_k$$

Step3 根据 Wolfe 线性搜索准则式 (8) 来确定步长 α_k , 计算 $x_{k+1} = x_k + \alpha_k d_k$ 。

取 α_k 使其同时满足不等式组:

$$\begin{cases} f(x_k + \alpha_k d_k) \leq f(x_k) + \sigma_1 \nabla f(x_k)^T d_k \\ \nabla f(x_k + \alpha_k d_k)^T d_k \geq \sigma_2 \nabla f(x_k)^T d_k \end{cases} \quad (8)$$

式中: σ_1 、 σ_2 为常数, 满足 $0 < \sigma_1 < \frac{1}{2}$, $\sigma_1 < \sigma_2 < 1$ 。

Step4 令 $\tilde{m} = \min\{k, m-1\}$, 利用式 (4) ~ (7) 计算 H_{k+1} 。

Step5 令 $k = k + 1$, 转至 Step2。

3 L-MBFGS 算法的收敛性分析

在 L-MBFGS 算法中, Hessian 矩阵逆的校正公式可以被看作是式 (7) 更新一个有界初始矩阵 H_k^0 共 m 次得到的, 在算法设计和实际计算时, 使用 H_{k+1} 比较方便, 而证明算法的收敛性时使用 B_{k+1} 则更常用, 所以, 将 L-MBFGS 算法进行如下等价描述 (记为算法 1)。

Step1 取初始点 $x_0 \in \mathbb{R}^n$, 初始对称正定矩阵 $B_0 \in \mathbb{R}^{n \times n}$, $\sigma_1 \in (0, \frac{1}{2})$, $\sigma_2 \in (\sigma_1, 1)$, 内存大小 $m > 0$, 精度 $\varepsilon > 0$, 令 $k = 0$ 。

Step2 如果 $\|g_k\| \leq \varepsilon$, 算法终止, 问题的解为 x_k ; 否则, 令

$$d_k = -B_k^{-1} g_k$$

Step3 根据 Wolfe 线性搜索准则来确定步长 α_k , 计算 $x_{k+1} = x_k + \alpha_k d_k$ 。

Step4 令 $\tilde{m} = \min\{k, m-1\}$, 利用 $\{s_i, y_i^+\}_{i=k-\tilde{m}+1}^k$, 计算

$$\begin{aligned} B_{k+1}^l &= B_k^l - \frac{B_k^l s_l s_l^T B_k^l}{s_l^T B_k^l s_l} + \frac{y_l^+ y_l^{+T}}{s_l^T y_l^+}, \\ l &= k - \tilde{m} + 1, k - m + 2, \dots, k \\ B_{k-\tilde{m}+1}^{k-\tilde{m}+1} &= B_0 \end{aligned} \quad (9)$$

Step5 令 $B_k = B_{k+1}^l$, $k = k + 1$, 转至 Step2。

为了证明算法的收敛性, 提出如下假设。

假设 (A): 目标函数 $f(x)$ 二阶连续可微, 且水平集 $\Omega = \{x | f(x) < f(x_0)\}$ 包含于一个有界集,

x_0 为初始点。

假设 (B): 目标函数 $f(x)$ 的梯度 $g(x)$ 在点 x^* 处 Lipschitz 连续, 即存在正数 $L > 0$, 对于任意 x, y , 都有

$$\|g(x) - g(y)\| \leq L\|x - y\|, \forall x, y \in \Omega \quad (10)$$

引理 1 若假设 (B) 成立, y_k^+ 如式 (5), 则存在常数 $M_1 > 0$, 使得

$$\|y_k^+\| \leq M_1 \|s_k\|$$

证明: 由假设 (B)、中值定理和 θ_k 的定义可知, 存在 $t_k \in [0, 1]$, 使得

$$\begin{aligned} \theta_k &= 4(f_k - f_{k+1}) + 2(g_k + g_{k+1})^T s_k = \\ &= 2(g_k + g_{k+1})^T s_k - 4(f_{k+1} - f_k) = \\ &= 2[(g_k + g_{k+1})^T s_k - 2g_k(x_k + t_k s_k)^T s_k] = \\ &= 2[(g_k - g_k(x_k + t_k s_k))^T s_k + (g_{k+1} - \\ &= g_k(x_k + t_k s_k))^T s_k] \leq 2L\|s_k\|^2 \end{aligned}$$

又由 $t_k (t_k \in [0, 1])$ 的定义可知

$$\begin{aligned} \|\bar{y}_k\| &= \left\| y_k + t_k \frac{\theta_k}{\|s_k\|^2} s_k \right\| \leq \|y_k\| + \frac{|\theta_k|}{\|s_k\|} \leq \\ &= L\|s_k\| + 2L\|s_k\| = 3L\|s_k\| \end{aligned}$$

由假设 (B) 可知, 对于任意的 x , 存在 $\varepsilon_0 > 0$, 使得 $\|g(x)\| \leq \varepsilon_0$, 所以

$$\begin{aligned} \|y_k^+\| &\leq \|\bar{y}_k\| + \xi \ln(1 + \|g_k\|^2) \|s_k\| \leq \\ &= (3L + \xi \ln(1 + \varepsilon_0^2)) \|s_k\| \leq \\ &= (3L + \xi \varepsilon_0^2) \|s_k\| \end{aligned}$$

令 $M_1 = 3L + \xi \varepsilon_0^2$, 引理得证。

引理 2 若假设 (A) 成立, $\{x_k\}$ 由算法 L-MBFGS 的等价算法 1 产生, 给定常数 $\varepsilon_1 > 0$, 对所有的 k 都有 $\|g_k\| > \varepsilon_1$ 成立, 则存在常数 $M_2 > 0$, 使得

$$\frac{\|y_k^+\|^2}{s_k^T y_k^+} \leq M_2 \quad (11)$$

证明: 根据 y_k^+ 的定义, 有

$$\begin{aligned} s_k^T y_k^+ &= s_k^T (\bar{y}_k + \xi \ln(1 + \|g_k\|^2) s_k) + \\ &= \max(-\frac{\bar{y}_k^T s_k}{\|s_k\|^2}, 0) s_k = \\ &= s_k^T \bar{y}_k + \xi \ln(1 + \|g_k\|^2) \|s_k\|^2 + \\ &= \max(-\frac{\bar{y}_k^T s_k}{\|s_k\|^2}, 0) \|s_k\|^2 = \end{aligned}$$

$$\begin{cases} s_k^T \bar{y}_k + \xi \ln(1 + \|g_k\|^2) \|s_k\|^2, & \bar{y}_k^T s_k > 0 \\ \xi \ln(1 + \|g_k\|^2) \|s_k\|^2, & \bar{y}_k^T s_k \leq 0 \end{cases}$$

$$\geq \xi \ln(1 + \|g_k\|^2) \|s_k\|^2$$

根据引理 1, 可得

$$\frac{\|y_k^+\|^2}{s_k^T y_k^+} \leq \frac{M_1^2 \|s_k\|^2}{\xi \ln(1 + \|g_k\|^2) \|s_k\|^2} \leq \frac{M_1^2}{\xi \ln(1 + \varepsilon_1^2)} = M_2$$

引理得证。

引理 3 若假设 (A)、(B) 成立, $\{x_k\}$ 由算法 L-MBFGS 的等价算法 1 产生, 则存在常数 $M_3 > 0$, 使得

$$\text{tr}(B_{k+1}) \leq M_3 \quad (12)$$

证明: 对式 (9) 两边作求迹运算, 则

$$\begin{aligned} \text{tr}(B_{k+1}) &= \text{tr}(B_k^l) - \text{tr}\left(\frac{B_k^l s_k s_k^T B_k^l}{s_k^T B_k^l s_k}\right) + \text{tr}\left(\frac{y_l^+ y_l^{+T}}{s_l^T y_l^+}\right) = \\ &= \text{tr}(B_k^l) - \frac{\|B_k^l s_l\|^2}{s_l^T B_k^l s_l} + \frac{\|y_l^+\|^2}{s_l^T y_l^+} = \\ &= \text{tr}(B_k^{k-\tilde{m}+1}) - \sum_{l=k-\tilde{m}+1}^k \frac{\|B_k^l s_l\|^2}{s_l^T B_k^l s_l} + \\ &\quad \sum_{l=k-\tilde{m}+1}^k \frac{\|y_l^+\|^2}{s_l^T y_l^+} \leq \\ &= \text{tr}(B_k^{k-\tilde{m}+1}) - \sum_{l=k-\tilde{m}+1}^k \frac{\|B_k^l s_l\|^2}{s_l^T B_k^l s_l} + \tilde{m} M_2 \leq \\ &= \text{tr}(B_k^{k-\tilde{m}+1}) + \tilde{m} M_2 \leq M_3 \end{aligned}$$

引理得证。

引理 4 若假设 (A)、(B) 成立, $\{x_k\}$ 由算法 L-MBFGS 的等价算法 1 产生, 则存在常数 $M_4 > 0$, 使得

$$\det(B_{k+1}) \geq M_4 \quad (13)$$

证明: 根据公式 $\det(I + u_1 u_1^T + u_3 u_3^T) = (1 + u_3^T u_4) - (u_1^T u_4)(u_2^T u_4)$, 对式 (9) 取行列式, 有

$$\begin{aligned} \det(B_{k+1}) &= \det(B_k^{k-\tilde{m}+1}) \prod_{l=k-\tilde{m}+1}^k \frac{s_l^T y_l^+}{s_l^T B_k^l s_l} = \\ &= \det(B_k^{k-\tilde{m}+1}) \prod_{l=k-\tilde{m}+1}^k \frac{s_l^T y_l^+}{s_l^T s_l} \frac{s_l^T y_l^+}{s_l^T B_k^l s_l} \end{aligned}$$

由引理 1 可知

$$s_k^T y_k^+ \leq \|s_k\| \|y_k^+\| \leq (3L + \xi) \|s_k\|^2$$

由引理 3 可知, B_k^l 的特征值小于 M_3 , 故

$$\det(B_{k+1}) = \det(B_k^{k-\tilde{m}+1}) \left(\frac{3L + \xi}{M_3}\right)^{\tilde{m}} \geq M_4$$

引理得证。

定理 1 若假设 (A)、(B) 成立, α_k 满足 Wolfe 准则式 (8) 时, 有

$$\sum_{k \geq 1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty \quad (14)$$

证明: 由 Wolfe 准则式 (8) 的第 2 个不等式和假设 (A), 可知

$$L \alpha_k \|d_k\|^2 \geq (g_{k+1} - g_k)^T d_k \geq -(1 - \sigma_2) g_k^T d_k$$

所以有

$$\alpha_k \geq \frac{-(1 - \sigma_2) g_k^T d_k}{L \|d_k\|^2}$$

又结合 Wolfe 线搜索准则式 (8) 的第 1 个不等式, 有

$$f_{k+1} - f_k \leq \sigma_1 (\sigma_2 - 1) \frac{(g_k^T d_k)^2}{L \|d_k\|^2}$$

对上述不等式两边同时求和, 又因为目标函数 f 有下界, 则 $\sum_{k \geq 1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty$, 得证。

引理 5 若假设 (A)、(B) 成立, $\{x_k\}$ 由算法 L-MBFGS 的等价算法 1 产生, 则

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0 \quad (15)$$

证明: 由 Wolfe 准则式 (8)、定理 1, 可得

$$\begin{aligned} f_k - f_{k+1} &\geq -\alpha_k \sigma_1 g_k^T d_k = \\ &= \alpha_k \sigma_1 \|d_k\| \|g_k\| \cos \varphi_k \geq \\ &= \sigma_1 \|g_k\| \cos \varphi_k \frac{(1 - \sigma_2)}{L} \|g_k\| \cos \varphi_k = \\ &= \frac{\sigma_1 (1 - \sigma_2)}{L} \|g_k\|^2 \cos^2 \varphi_k \end{aligned} \quad (16)$$

令 $\xi = \frac{\sigma_1 (1 - \sigma_2)}{L}$, 则有

$$f_k - f_{k+1} \geq \xi \|g_k\|^2 \cos^2 \varphi_k \quad (17)$$

对式 (17) 两边求和, 则对于所有的 $k > 0$, 都有

$$\begin{aligned} f(x_0) - f(x_k) &= \sum_{i=0}^{k-1} [f(x_i) - f(x_{i+1})] \geq \\ &= \xi \min_{0 \leq i \leq k} \|g_i\|^2 \sum_{i=0}^{k-1} \cos^2 \varphi_i \end{aligned}$$

即

$$\liminf_{k \rightarrow \infty} \|g_k\| \cos \varphi_k = 0$$

又由式 (12) 和 (13) 可知, 存在一个常数 μ , 使得

$$\cos \varphi_k = \frac{s_k^T B_k s_k}{\|s_k\| \|B_k s_k\|} \geq \mu \quad (18)$$

由式 (18) 可知, 式 (15) 成立。引理得证。

4 数值实验

本实验软件在 Windows 11 中文版、处理器参数为 11th Gen Intel (R) Core (TM) i5-1155G7

2.50 GHz、8.00 GB 内存环境下运行。

4.1 参数设置及终止条件

1) $\delta \in [2.5, 2.9]$, 当 δ 为自然常数 e 时, 与文献[9]的基于混合割线方程修正的 L-BFGS (ML-BFGS) 方法相似, 区别是 y_k^* 修正第二项取值稍有不同, 其余取值为改进算法, 即本文的 L-MBFGS 算法。每个函数任意选取 δ 取值范围内的 4 个等间隔值进行实验, 标准的 L-MBFGS 不涉及参数 δ , $\xi = 1 \times 10^{-5}$, $m = 5$ 。

2) 终止算法收敛控制参数设置为 $\|g_k\| < \varepsilon_1 = 1 \times 10^{-5}$, 即目标函数梯度的二范数优于 ε_1 时, 算法收敛到最小值。

3) 迭代次数超过 400 次, 算法终止运行。

4.2 数值计算

为了验证算法的有效性, 用如下两个典型的求解大规模无约束优化问题的测试函数进行实验, 结果如表 1 所示。

1) 函数 1: Wood

初始点 $(-3, -1, -3, -1, \dots, -3, -1)$

最优值 0;

2) 函数 2: Dixon

初始点 $(-2, -2, \dots, -2)$ 最优值 0。

4.3 结果分析

由表 1 可知, 对 2 个经典测试函数进行计算, 本文所提的 L-MBFGS 方法均得到了正确解, 求得了最小值, 整体平稳且有改善, 说明本文所提出的 L-MBFGS 方法是有效的。通过表 1 可以看出, 在可变参数 δ 的调节下, 算法对不同维度的函数计算性能都比标准 L-BFGS 与 ML-BFGS 相似方法有较好的数值表现。对于 Wood 函数, 计算的最优值结果有提高, 迭代次数和计算用时有不同程度的减少, $\delta = 2.55$ 、数据规模为 500 时, 该函数最优值达到 $5.866\,096 \times 10^{-12}$, 优于标准 L-BFGS 方法 1 个数量级, 优于 ML-BFGS 相似方法 2 个数量级; Dixon 函数在 3 个不同的数据规模上, 虽然求解的最优值与标准 L-BFGS 方法处在相当水平上, 但是迭代次数和计算用时明显减少, 表明新改进的方法是有竞争力的。

表 1 数值计算结果比对
Tab. 1 Comparison of numerical calculation results

方法类型	数据规模	函数 Wood			Dixon		
		迭代次数	计算时间 /ms	目标函数值	迭代次数	计算时间 /ms	目标函数值
标准 L-BFGS	500	90	42	2.395 135E-11	141	56	2.530 615E-09
	2 000	87	71	1.469 701E-12	145	110	3.459 744E-09
	5 000	97	151	5.308 599E-11	140	183	3.650 253E-08
L-MBFGS: $\delta = 2.55$	500	82	24	5.866 096E-12	136	53	4.130 302E-09
	2 000	92	84	3.102 911E-11	134	104	8.849 559E-09
	5 000	91	142	5.057 501E-10	135	170	1.964 729E-07
L-MBFGS: $\delta = 2.65$	500	91	30	1.816 652E-11	141	53	7.695 786E-09
	2 000	90	57	2.990 986E-11	132	96	7.680 419E-09
	5 000	91	157	3.257 972E-11	141	185	7.553 484E-08
L-MBFGS: $\delta = 2.75$	500	89	33	4.023 632E-13	134	40	1.523 488E-08
	2 000	89	62	8.467 131E-14	136	100	5.951 180E-08
	5 000	90	141	2.753 623E-11	136	176	3.778 563E-08
L-MBFGS: $\delta = 2.85$	500	87	30	1.781 101E-13	132	45	2.505 612E-08
	2 000	91	73	2.084 329E-11	136	98	2.649 615E-08
	5 000	93	143	7.231 487E-11	138	175	1.800 047E-08
ML-BFGS ^[9] 相似方法	500	88	30	1.630 188E-10	135	50	8.716 632E-09
	2 000	91	69	2.176 899E-10	140	105	9.638 055E-09
	5 000	93	143	1.929 894E-12	135	171	2.084 260E-07

5 结论

针对传统 L-BFGS 方法在用于求解非凸函数极小值时不一定全局收敛的问题, 本文提出了求解非凸函数带参方程的改进 L-BFGS 方法, 建立了算法的收敛性证明, 进行了数值实验和分析, 选取 2 个典型求解无约束最优化问题的测试函数, 将可调节参数分段, 并进行等间隔取值, 从 500、2 000、5 000 数据规模进行了数值实验。实验分析表明, 本文提出的 L-MBFGS 方法优于标准 L-BFGS 方法和 ML-BFGS 相似方法。

参考文献

- [1] BROYDEN C G. The convergence of a class of double-rank minimization algorithms: The new algorithm[J]. IMA Journal of Applied Mathematics, 1970, 6: 222-231.
- [2] FLETCHER R. A new approach to variable metric algorithms[J]. The Computer Journal, 1970, 13(3): 317-322.
- [3] GOLDFARD D. A family of variable-metric methods derived by variational means[J]. Mathematics of Computation, 1970, 24(109): 23-26.
- [4] SHANNO D F. Conditioning of quasi-Newton methods for function minimization[J]. Mathematics of Computation, 1970, 24(111): 647-656.
- [5] 王锋. 非凸优化的一类改进 BFGS 算法的收敛性研究[J]. 萍乡学院学报, 2018, 35(3): 1-6.
WANG F. The research on convergence of modified BFGS algorithm for non-convex optimization[J]. Journal of Pingxiang University, 2018, 35(3): 1-6.
- [6] 刘克刚. 基于二阶信息的优化算法[D]. 上海: 华东师范大学, 2020: 12, 16-17.
LIU K G. Optimization algorithm based on second-order information[D]. Shanghai: East China Normal University, 2020: 12, 16-17.
- [7] 武思蕊, 李斌, 赵梁成, 等. 磁性石墨烯/聚氨酯柔性复合材料的制备及自修复效能[J]. 化工进展, 2020, 39(4): 1422-1430.
WU S R, LI B, ZHAO L C, et al. Preparation and self-repairing efficiency of magnetic graphene/polyurethane flexible composites[J]. Chemical Industry and Engineering Progress, 2020, 39(4): 1422-1430.
- [8] XIE M H, LI S X, QI X M, et al. Thermal and infrared light self-repairing, high sensitivity, and large strain sensing range shape memory MX-ene/CNTs/EVA composites fiber strain sensor for human motion monitoring[J]. Sensors and Actuators A: Physical, 2022, 347: 113939.
- [9] 肖楷彬. 基于微胶囊化环氧-胺化学的自修复环氧基复合材料研究[D]. 广州: 华南理工大学, 2022.
XIAO K B. Self-healing epoxy composites based on microencapsulated epoxy-amine chemistry[D]. Guangzhou: South China University of Technology, 2022.
- [10] 韩闯, 杨帆, 郭兴魁, 等. 树脂基复合材料微胶囊自修复技术综述[J]. 纤维复合材料, 2021, 38(1): 61-64.
HAN C, YANG F, GUO X K, et al. Review of microencapsulated self-healing technology for resin matrix composite[J]. Fiber Composites, 2021, 38(1): 61-64.