

文章编号: 1674-8085(2016)01-0026-03

一种基于四阶 Taylor 展开的多目标改进的 拟牛顿算法研究

*王菲菲, 徐 尔, 赵金玲

(北京科技大学数理学院, 北京 100083)

摘 要: 借助于目标函数的四阶 Taylor 展开导出新的拟牛顿方程, 并将其应用到多目标优化问题中, 给出了一种多目标优化改进的拟牛顿算法(称为 M-TBFGS 算法), 同时在一定的假设条件下, 结合 Wolfe 搜索准则, 证明了本文算法的收敛性, 并进行了数值试验, 结果表明, 本文的 M-TBFGS 算法是正确和有效的。

关键词: 多目标优化; 四阶 Taylor 展开; M-TBFGS 方法; Wolfe 线搜索; Pareto 最优解

中图分类号: O221.6

文献标识码: A

DOI:10.3969/j.issn.1674-8085.2016.01.006

NEW QUASI-NEWTON'S METHOD BASED ON FOUR-ORDER TAYLOR SERIES EXPANSION FOR MULTIOBJECTIVE OPTIMIZATION

*WANG Fei-fei, XU Er, ZHAO Jin-ling

(School of Mathematics and Physics, University of Science & Technology Beijing, Beijing 100083, China)

Abstract: The new quasi-Newton's equation is applied to multiobjective optimization, which is derived by using the four-order Taylor series expansion. The improved method for multiobjective optimization without constraints is presented, which is called M-TBFGS algorithm. Furthermore, its global convergence is proved. Under the Wolfe line search, numerical results also show that the proposed method is correct and efficient.

Key words: multiobjective optimization; four-order Taylor series expansion; M-TBFGS method; Wolfe line search; Pareto optimal

0 引言

考虑如下无约束多目标优化问题:

$$\min F(x) = \{f_1(x), f_2(x), \dots, f_p(x)\}, x \in R^n \quad (1)$$

其中 $F(x): R^n \rightarrow R^p$ 二阶连续可微, $j=1, 2, \dots, p$ 。

在多目标规划问题中, 我们使用文献[1]给出的 Pareto 最优解的概念。在此定义基础上, 文献[2]给出了多目标优化的最速下降方向; 文献[3]研究了无约束多目标优化的牛顿法; 文献[4]给出了多目标优

化的拟牛顿法。

文献[5]将目标函数进行四阶 Taylor 展开, 得到了一种新的拟牛顿校正公式, 在每步迭代中, 不仅利用了目标函数的梯度信息, 还利用了目标函数值的信息, 所以其数值结果优于原始拟牛顿法。本文在文献[5]的基础上, 结合多目标优化的拟牛顿法^[4], 给出了一种基于四阶 Taylor 展开的多目标改进的拟牛顿算法, 并证明了算法的全局收敛性, 数值试验也表明本文算法较之文献[4]中的拟牛顿算法具有一定的优越性。

收稿日期: 2015-10-06; 修改日期: 2015-12-20

基金项目: 国家自然科学基金青年基金项目(11101028); 北京高校青年英才计划项目(YETP0385)

作者简介: *王菲菲(1990-), 女, 山东潍坊人, 硕士生, 主要从事运筹学及其应用研究(E-mail: wffustb2013@163.com);

徐 尔(1961-), 女, 北京人, 副教授, 主要从事运筹学及其应用研究(E-mail: xuer_2009@sina.com);

赵金玲(1980-), 女, 北京人, 副教授, 主要从事最优化计算方法研究(E-mail: jlzhao@ustb.edu.cn)。

1 M-TBFGS 算法及其性质

引理 1 为求解多目标优化问题(1), 本文算法设计校正矩阵如下:

$$B_{jk+1} = \begin{cases} B_{jk} - \frac{B_{jk}\delta_k\delta_k^TB_{jk}}{\delta_k^TB_{jk}\delta_k} + \frac{y_{jk}^*(y_{jk}^*)^T}{\delta_k^Ty_{jk}^*}, & \delta_k^Ty_{jk}^* > 0, j=1,2,\dots,p \\ B_{jk} & , \text{其它} \end{cases}$$

$$\text{其中 } y_{jk}^* = y_{jk} + \frac{5\theta_{jk}}{\|\delta_k\|^2} \delta_k,$$

$$\theta_{jk} = 2(f_{jk} - f_{jk+1}) + (g_{jk+1} + g_{jk})^T \delta_k \quad (2)$$

证明 由求解单目标优化问题 $\min f_j(x)$, $x \in R^n$ 的新拟牛顿法^[5], $f_j(x)$ 充分光滑时对其进行四阶 Taylor 展开得到:

$$f_{jk} = f_{jk+1} - g_{jk+1}\delta_k + \frac{1}{2}G_{jk+1}\delta_k^2 - \frac{1}{6}T_{jk+1}\delta_k^3 + \frac{1}{24}V_{jk+1}\delta_k^4 + o(\|\delta_k\|^4)$$

$$g_{jk}\delta_k = g_{jk+1}\delta_k - G_{jk+1}\delta_k^2 + \frac{1}{2}T_{jk+1}\delta_k^3 - \frac{1}{6}V_{jk+1}\delta_k^4 + o(\|\delta_k\|^4)$$

$$G_{jk}\delta_k^2 = G_{jk+1}\delta_k^2 - T_{jk+1}\delta_k^3 + \frac{1}{2}V_{jk+1}\delta_k^4 + o(\|\delta_k\|^4)$$

其中 $T_{jk+1} \in R^{n \times n \times n}$ 和 $V_{jk+1} \in R^{n \times n \times n \times n}$ 是三阶和四阶张量, 记 $f_{jk} = f_j(x_k)$, $g_{jk} = \nabla f_j(x_k)$, $B_{jk} = B_j(x_k)$, $y_{jk} = g_{jk+1} - g_{jk}$, $G_k = \nabla^2 f(x_k)$, $\delta_k = x_{k+1} - x_k = a_k d_k$.

用 B_{jk+1} 近似 G_{jk+1} , 最终得到新拟牛顿方程:

$$\delta_k^TB_{jk+1}\delta_k \approx \delta_k^T \left(y_{jk} + \frac{5\theta_{jk}}{\delta_k^Ts} s \right) = \delta_k^Ty_{jk}^*,$$

其中 s 是使 $\delta_k^Ts \neq 0$ 的向量, $s \in R^n$.

本文在文献[5-6]的启发下, 选取 $s = \delta_k$, 得到:

$$y_{jk}^* = y_{jk} + \frac{5\theta_{jk}}{\|\delta_k\|^2} \delta_k,$$

其中 $\theta_{jk} = 2(f_{jk} - f_{jk+1}) + (g_{jk+1} + g_{jk})^T \delta_k$.

对于单目标改进拟牛顿算法的全局收敛性已经证明, 这里不再赘述. 本文将其引用到多目标优化问题(1), 考虑到 M-TBFGS 算法产生的 B_{jk+1} 必须正定, 故校正公式如下:

$$B_{jk+1} = \begin{cases} B_{jk} - \frac{B_{jk}\delta_k\delta_k^TB_{jk}}{\delta_k^TB_{jk}\delta_k} + \frac{y_{jk}^*(y_{jk}^*)^T}{\delta_k^Ty_{jk}^*}, & \delta_k^Ty_{jk}^* > 0, j=1,2,\dots,p \\ B_{jk} & , \text{其它} \end{cases}$$

引理 2 本文算法下降方向 $d_k = d(x_k)$.

证明 考虑无约束极小极大值问题:

$$\min_{d \in R^n} \max_{j=1,2,\dots,m} \nabla f_j(x)^Td + \frac{1}{2}d^TB_j(x)d$$

根据文献[4], 可求得其最优解 $d(x)$ 和最优值 $\tau(x)$, 这里不再赘述.

引理 3 M-TBFGS 算法采用 Wolfe 搜索准则获取步长 a_k .

$$f_j(x_k + a_k d_k) \leq f_j(x_k) + ca_k \tau(x_k) \quad (3)$$

$$\nabla f_j(x_{k+1})^T d_k \geq \sigma \nabla f_j(x_k)^T d_k, j=1,2,\dots,p \quad (4)$$

求解多目标优化问题(1)的 M-TBFGS 算法设计如下:

步 0 选取初始点 $x_0 \in R^n$, 初始对称正定矩阵 $B_{j1} = I \in R^{n \times n}$, $c \in (0, 1)$, $\sigma \in (c/2, 1)$, 任意小 $\varepsilon > 0$.

置 $k := 0$, 定义 $J = \left\{ \frac{1}{2^n} | n=0, 1, \dots \right\}$.

步 1 若 $\tau(x_0) > -\varepsilon$, 停止迭代. 否则转第 2 步.

步 2 计算 d_k .

步 3 计算搜索步长 a_k , 其中 $a_k \in J = \left\{ \frac{1}{2^n} | n=0, 1, \dots \right\}$

中满足(3)式与(4)式的最大者.

步 4 计算 B_{jk+1} .

步 5 置 $x_{k+1} = x_k + a_k d_k$, $k := k+1$, 转至第 1 步.

2 算法的全局收敛性

假设 H_1 : 水平集 $\Omega = \{x \in R^n | F(x) \leq F(x_0)\}$ 有界.

假设 H_2 : 梯度 $\nabla f_j(x)$ ($j=1,2,\dots,p$) 在包含水平集 Ω 的开凸集上 Lipschitz 连续.

定理 1 若假设成立, 则算法或有限步终止于问题(1)的 Pareto-临界点, 或产生无穷点列 $\{x_k\}$, 其任意聚点都是问题(1)的 Pareto-临界点.

证明 假设算法产生无穷点列 $\{x_k\}$, x^* 为其任意极限点. 由假设 H_1 可知 x^* 为一有限点且存在无穷子列 $\{x_{k_i}\}_{k_i \in N} \rightarrow x^*$, $N \subset \{1, 2, 3, \dots\}$ 为无穷子列.

由假设 H_1 及 $\{F(x_k)\}$ 按分量单调下降且有下界, 可知 $\{F(x_k)\}$ 有极限。

又由步长原则可知:

$$F(x_k) - F(x_k + a_k d_k) \geq -\sigma a_k \tau(x_k) \geq 0。$$

则得到:

$$\lim_{k \rightarrow +\infty} a_k \tau(x_k) = 0 \quad (5)$$

现取无穷子列 $\{x_{k_u}\}$ 收敛于 x^* , 考虑下列两种情况:

$$(i) \liminf_{u \rightarrow \infty} a_{k_u} > 0, \quad (ii) \liminf_{u \rightarrow \infty} a_{k_u} = 0。$$

(i) 存在一个序列 $\{x_{k_i}\}$ 收敛到 x^* 且满足

$\lim_{i \rightarrow \infty} a_{k_i} = \bar{a} > 0$, 根据(5)式, 得 $\lim_{i \rightarrow \infty} \tau(x^{(k_i)}) = 0$ 。因为 τ 是连续函数^[4], 则函数与极限可交换次序, 于是得到 $\tau(x^*) = 0$ 。

(ii) 序列 $\{d_{k_u}\}$ 有界, 取 $\{x_{k_u}\}$ 的一个子序列 $\{x_{k_r}\}$, 使得序列 $\{d_{k_r}\}$ 同样收敛于某个 $\bar{d}$ 。因为 $\liminf_{u \rightarrow \infty} a_{k_u} = 0$, 取某个 $q \in N$, r 足够大有 $a_{k_r} < \frac{1}{2^q}$,

则 Wolfe 条件不满足 $a = \frac{1}{2^q}$, 即存在一个 j , 有

$$f_j(x_{k_r} + (\frac{1}{2^q})d_{k_r}) > f_j(x_{k_r}) + c(\frac{1}{2^q})\tau(x_{k_r}) \text{ (对 } r \text{ 足够大)}。$$

取极限 $r \rightarrow \infty$, 得

$$f_j(x^* + (\frac{1}{2^q})\bar{d}) > f_j(x^*) +$$

$$c(\frac{1}{2^q})(\min_{d \in R^n} \max_{j=1, \dots, m} \nabla f_j(x^*)^T \bar{d} + \frac{1}{2} \bar{d}^T \nabla^2 f_j(x^*) \bar{d}) =$$

$$f_j(x^*) + c(\frac{1}{2^q})\tau(x^*)$$

对至少一个 $j \in \{1, 2, \dots, m\}$ 成立, 于是有 $\tau(x^*) \geq 0$ 。

由文献[4]定理 2 知 $\tau(x^*) \leq 0$, 所以 $\tau(x^*) = 0$ 。

结合(i), (ii), 知 x^* 是 $F(x)$ 的临界点。再由文献[4]定理 1 可知, x^* 为 V 上 $F(x)$ 的 Pareto 最优解。

3 数值试验

为进一步检验算法的实算效果, 选取算例对本文进行数值试验。本文单目标改进的拟牛顿法与原始 BFGS 方法比较; M-TBFGS 新方法 with 文献[4]的多目标拟牛顿法进行比较。

$$\text{例 1}^{[7]} \min_{x \in R^3} (10x_1^4 + 25x_2^4 + 12x_3^4 + 18x_1^2 + 13x_2^2 +$$

$$10x_3^2 + 2x_1x_2 + 2x_2x_3 - 5x_1 - 3x_2 - x_3)。$$

$$\text{例 2} \min_{x \in R^2} (\frac{1}{4}((x_1 - 1)^4 + 2(x_2 - 2)^4), x_1^2 - x_1x_2 + x_2^2 - 2x_1 + x_2)。$$

用 EX 表示算例, IP 表示初始点, M-TB 表示本文提出的新算法, TB 表示本文单目标改进的拟牛顿新算法, M-B 表示文献[4]中的多目标拟牛顿算法。用 c 表示 Pareto-临界点, F 表示对应的函数值, k 表示算法迭代次数, 其中取参数 $\sigma = 0.7, c = 0.1$, 精度设为 $|\tau(x_k)| \leq 10^{-4}$ 。所得到的数据如表 1。

表 1 算例迭代情况

Table 1 The iterative situation of examples

EX	IP	k-(BFGS)	k-(TB)	c
(1)	(-0.4, 1.0, 0.1)	22	15	(0.1309, 0.0985, 0.0399)
EX	IP	M-B(c/F/k)	M-TB(c/F/k)	M-TB(c/F/k)
(2)	(-5, 2)	(-1, 2)/(10, 4)/5	(-1, 2)/(10, 4)/4	(1, 2)/(3, 0)/10
(2)	(-10, 1)	(-1, 1)/(6, 4.5)/11	(-1, 1)/(6, 4.5)/10	(1, 1)/(0, 0.5)/11

从以上数值试验及比较的结果可以看出, 当达到同一精度时, 单目标改进的新拟牛顿法所需要的迭代次数比原始 BFGS 方法要少; 而本文提出的多目标的新拟牛顿算法 M-TB 在迭代得到相同临界点的情况下, 所需的迭代次数要比 M-B 少, 并且 M-TB 继续迭代会得到比 M-B 更好的临界点, 而 M-B 已终止于相同临界点。

参考文献:

- [1] Luc D T. Theory of vector optimization[M]. Berlin: Springer-verlary, 1989.
- [2] Fliege J, Svaiter B F. Steepest descent methods for multicriteria Optimization[J]. Mathematical Methods of Operations Research, 2000, 51(3): 479-494.
- [3] Fliege J, Grana Drummond L M, Svaiter B F. Newton's methods for multiobjective optimization[J]. SIAM Journal on Optimization, 2009, 20(2): 602-626.
- [4] Povalej Ž. Quasi-Newton's method for multiobjective optimization[J]. Journal of Computational and Applied Mathematics, 2014, 255(1): 765-777.
- [5] 王希云, 时平平. 一种基于新拟牛顿方程的拟牛顿法[J]. 运筹学学报, 2007, 11: 190-195.
- [6] Wei Z, Li G, Qi L. New quasi-Newton methods for unconstrained optimization problems[J]. Applied Mathematics and Computation, 2006, 175(2): 1156-1188.
- [7] 陈兰平, 焦宝聪. 非凸无约束优化问题的广义拟牛顿法的全局收敛性[J]. 应用数学, 2005, 18(4): 573-579.