文章编号:1000－2367(2016)01－0021－04　　DOI:10.16366/j.cnki.1000－2367.2016.01.004

无约束多目标优化的一种新的拟牛顿法

王菲菲，徐　尔，赵金玲

(北京科技大学 数理学院,北京 100083)

摘　要:基于在新拟牛顿方程形式下无约束单目标优化问题改进的拟牛顿法,提出了无约束多目标优化问题的一种新的拟牛顿法,同时在一定的假设条件下,结合 Wolfe 线性搜索准则,证明了算法具有全局收敛性和超线性收敛性,并进行了数值试验,结果表明,所提的新算法是正确和有效的,并能够迭代得到可使多个目标更优的临界点.

关键词:多目标优化;新拟牛顿法;Wolfe 线性搜索;Pareto 最优解

中图分类号:O221.6　　**文献标志码**:A

对于下列无约束多目标优化问题:

$$(\mathrm{MOP})\ \min F(x)=\{f_1(x),f_2(x),\cdots,f_m(x)\},x\in\mathbf{R}^n,$$

其中 $F(x)$: $\mathbf{R}^n\to\mathbf{R}^m$ 二阶连续可微,设 $F(x)$ 在点 x 处的Jacobi矩阵为 $JF(x)\in\mathbf{R}^{m\times n}$,而 $JF(x)^{\mathrm{T}}=(\nabla f_1,\nabla f_2,\cdots,\nabla f_m)$,$f_j$ 在点 x 处的梯度为 $\nabla f_j(x)\in\mathbf{R}^n$,记 Hessian 矩阵为 $\nabla^2 f_j(x)$,$j=1,2,\cdots,m$.

在多目标优化问题中,不再使用单目标问题中最优解的概念,而是使用文献[1] 给出的 Pareto 最优解的概念. 基于此定义,文献[2] 给出了无约束多目标优化的最速下降法,文献[3] 给出了无约束多目标优化的牛顿法,文献[4] 又研究了无约束多目标优化的拟牛顿法.

考虑下面的无约束极小极大值问题:

$$(\mathrm{P})\ \min_{d\in R^n}\max_{j=1,2,\cdots m}\nabla f_j(x)^{\mathrm{T}}d+\frac{1}{2}d^{\mathrm{T}}B_j(x)d,$$

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

文献[5] 给出了无约束单目标优化改进的拟牛顿法,在每步迭代中,除利用了目标函数的梯度信息还用到了目标函数值的信息,所以其数值结果优于拟牛顿法. 本文在文献[5] 的基础上,结合多目标优化的拟牛顿法[4],给出了无约束多目标优化的一种新的拟牛顿法,并证明了算法的全局收敛性和超线性收敛性,数值试验表明这种新算法较之文献[4] 中的拟牛顿算法具有一定的优越性.

1　算法及其性质

取常数 $c\in(0,1)$,$\sigma\in\left(\frac{c}{2},1\right)$,初始点 $x_0\in\mathbf{R}^n$,初始对称正定矩阵 $B_{j1}=I\in\mathbf{R}^{n\times n}$,$j=1,2,\cdots,m$ 和参数 $\varepsilon>0$. 置 $k:=0$,定义 $J=\left\{\frac{1}{2^n}\middle|n=0,1,\cdots\right\}$.

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

(2) 计算 d_k;$d_k=d(x_k)$.

(3) 计算搜索步长 a_k 其中 $a_k\in J=\left\{\frac{1}{2^n}\middle|n=0,1,\cdots\right\}$中满足下式的最大者.

收稿日期:2015-07-05;修回日期:2015-11-17.

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资助(YETP0385).

第 1 作者简介(通信作者):王菲菲(1990－),女,山东潍坊人,北京科技大学硕士研究生,研究方向为运筹理论及应用,E-mail:1284689742@qq.com.

$$f_j(x_k+a_kd_k)\le f_j(x_k)+ca_k\tau(x_k);\nabla f_j(x_{k+1})^{\mathrm{T}}d_k\ge\sigma\nabla f_j(x_k)^{\mathrm{T}}d_k;j=1,2,\cdots,m.$$

(4) 按校正公式计算 B_{jk+1};$B_{jk+1}=\begin{cases}B_{jk}-\dfrac{B_{jk}p_kp_k{}^{\mathrm{T}}B_{jk}}{p_k{}^{\mathrm{T}}B_{jk}p_k}+\dfrac{y_{jk}^*(y_{jk}^*)^{\mathrm{T}}}{p_k{}^{\mathrm{T}}y_{jk}^*},\ p_k^{\mathrm{T}}y_{jk}^*>0,\\ B_{jk},\text{其他}.\end{cases}$.

记 $B_{jk}=B_j(x_k)$,定义 $y_{jk}=\nabla f_j(x_{k+1})-\nabla f_j(x_k)$,$p_k=x_{k+1}-x_k=a_kd_k$.

其中 $y_{jk}^*=y_{jk}+A_{jk}p_k$,$A_{jk}=\dfrac{2(f_{jk}-f_{jk+1})+(g_{jk+1}+g_{jk})^{\mathrm{T}}p_k}{\|p_k\|^2}I$.

(5) 置 $x_{k+1}=x_k+a_kd_k$,$k:=k+1$,转至第(1)步.

若 $M-N$ 是正定阵,$M,N\in\mathbf{R}^{n\times n}$,本文记为 $M>N$. $B(x,r)$ 表示以点 $x\in\mathbf{R}^n$ 为中心,r 为半径的闭球域.

引理 1 若 $z^{\mathrm{T}}\nabla^2f_j(x)z>0,\forall x,z\in R^nB_{jk}$ 对称正定本文算法产生的 B_{jk+1} 也为对称正定阵.

证明 由于

$$\begin{aligned}p_k^{\mathrm{T}}y_{jk}^*&=p_k^{\mathrm{T}}y_{jk}+p_k^{\mathrm{T}}A_{jk}p_k=2(f_{jk}-f_{jk+1})+2g_{jk+1}^{\mathrm{T}}p_k=2\Big(-g_{jk+1}^{\mathrm{T}}p_k+\\&\frac{1}{2}p_k^{\mathrm{T}}\nabla^2f_j(x_k+\theta(x_{k+1}-x_k))p_k\Big)+2g_{jk+1}^{\mathrm{T}}p_k=\\&p_k^{\mathrm{T}}\nabla^2f_j[x_k+\theta(x_{k+1}-x_k)]p_k>0\ (\theta\in(0,1)).\end{aligned}$$

再由文献[5]中的定理 3.1 可知:当 $p_k^{\mathrm{T}}y_{jk}^*>0$ 时,有 $B_{jk+1}>0$ 成立.

由引理 1 知,若 $z^{\mathrm{T}}\nabla^2f_j(x)z>0$ 不成立,B_{jk+1} 可能失去其正定性,所以采取本文算法步(4) 中的迭代公式.

引理 2 序列$\{(x_k)\}$由本文算法及$\{(B_{jk})\}$迭代得出.对任意的 $x_k,y\in V$ 当 $k\ge k_0$,$\|y-x_k\|<\delta$ 时,有以下不等式成立:

$$\|\nabla f_j(y)-(\nabla f_j(x_k)+B_{jk}(y-x_k))\|<\varepsilon\|y-x_k\|,\tag{1}$$

$$\left|f_j(y)-\left(f_j(x_k)+\nabla f_j(x_k)^{\mathrm{T}}(y-x_k)+\frac{1}{2}(y-x_k)^{\mathrm{T}}B_{jk}(y-x_k)\right)\right|<\frac{\varepsilon}{2}\|y-x_k\|^2.\tag{2}$$

证明 首先,对任意的 $\|x,y\|\in V\subset\mathbf{R}^n$ 当 $y-x<\delta$ 时,有 $\|(\nabla^2f_j(y)-\nabla^2f_j(x))\|<\dfrac{\varepsilon}{2}$,$j=1,2,\cdots,m$ 成立.再根据文献[5]中定理 6.9,有 $\lim\limits_{k\to\infty}\dfrac{\|(B_k-\nabla^2f(x^*)p_k)\|}{\|p_k\|}=0$.拓展到多目标存在 $k_0\in N$,对任意小的 $\varepsilon\in\mathbf{R}_{++}$,当 $k\ge k_0$ 时,下式成立

$$\frac{\|(\nabla^2f_j(x_k)-B_{jk})(y-x_k)\|}{\|(y-x_k)\|}<\frac{\varepsilon}{2},\ j=1,2,\cdots,m,\tag{3}$$

以下证明同文献[4]定理 5,即可证(1) 式(2) 式成立,这里不再赘述.

引理 3 若 $x\in\mathbf{R}^n$,$a,b\in\mathbf{R}_+$,$a\le b$,如果 $aI\le B_j(x)\le bI$,则有

(1) $\dfrac{a}{2}\|d(x)\|^2\le|\tau(x)|\le\dfrac{b}{2}\|d(x)\|^2$;

(2) $|\tau(x)|\le\left(\dfrac{1}{(2a)}\right)\left\|\sum\limits_{j=1}^m\lambda_j\nabla f_j(x)\right\|^2$,对所有的 $\lambda_j\ge0,j=1,2,\cdots,m$,且 $\sum\limits_{j=1}^m\lambda_j=1$.

证明 由引理 1 知本文算法产生的 $B_j(x)>0$. 以下证明见文献[3]中的定理 4.2 与定理 4.3.将其中的 Hessian 矩阵用本文的 B_{jk} 代替.

引理 4 设 $V\subset\mathbf{R}^n$ 是一个凸集,其中 $k_0\in N,a,b,r,\delta,\varepsilon\in\mathbf{R}_+$.序列$\{x_k\}$由本文算法及$\{B_{jk}\}$迭代得出且满足以下条件:

(i) 对所有 $x\in V,j=1,2,\cdots,m,aI\le\nabla^2f_j(x)\le bI$ 且 $aI\le B_j(x)\le bI$;

(ii) 对所有 $x,y\in V$,$\|y-x\|<\delta,j=1,\cdots,m$,有 $\|\nabla^2f_j(y)-\nabla^2f_j(x)\|<\varepsilon/2$;

(iii) 对所有的 $k\ge k_0,y\in V,j=1,2,\cdots,m$,$\|(B_{jk}-\nabla^2f_j(x_k))(y-x_k)\|<(\varepsilon/2)\|y-x_k\|$;

(iv) $B(x_{k_0},r)\subset V$;

(v) $\varepsilon/a\le(1-c)$;

(vi) $\| d_{k_0} \| < \min\{\delta, r(1-\varepsilon/a)\}$.

则当 $k \geq k_0$ 时，有以下结论成立：(a) $\| x_k - x_{k_0} \| \leq (1-(\varepsilon/a)^{k-k_0})/(1-\varepsilon/a) \| d_{k_0} \|$；(b) $\| d_k \| \leq (\varepsilon/a)^{k-k_0} \| d_{k_0} \|$；(c) $a_k = 1$；(d) $\| d_{k+1} \| \leq (\varepsilon/a) \| d_k \|$.

证明　由引理 1 知本文算法产生的 $B_j(x) > 0$ 所以条件(i) 可满足；本文引理 2 中的(3) 式成立则条件(iii) 满足；其他条件同文献[4]，以下证明见文献[4]，即有结论(a)、(b)、(c)、(d) 成立

2　算法的收敛性

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

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

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

证明　$\{F_k\}$ 是一下降序列，显然算法产生的无穷点列 $\{x_k\}$ 包含在集合 Ω 内，假设 H_1 成立，根据假设 H_2，对所有 k，存在 $M > 0$，满足 $\| g_{jk} \| \leq M$. 取 x^* 为其一有限点且存在无穷子列 $\{x_{k_t}\}_{k_t \in N} \to x^*$，$N \subset \{1,2,3,\cdots\}$. 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：$F(x_k) - F(x_k + a_k d_k) \geq -\sigma a_k \tau(x_k) \geq 0$. 于是得到，

$$\lim_{k \to \infty} a_k \tau(x_k) = 0. \tag{4}$$

现取无穷子列收敛于 x^* 考虑下列两种情况：1) $\liminf_{u \to \infty} a_{k_u} > 0$；2) $\liminf_{u \to \infty} a_{k_u} = 0$.

1) 存在一个序列 $\{x_{k_l}\}$ 收敛到 x^* 且满足 $\lim_{l \to \infty} a_{k_l} = \bar{a} > 0$，根据(4) 式，得 $\lim_{l \to \infty} \tau(x^{(k_l)}) = 0$ 因为 τ 是连续函数[4] 则函数与极限可交换次序于是得到 $\tau(x^*) = 0$

2) 由引理 4 知序列 $\{d_{k_u}\}$ 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 $\liminf_{u \to \infty} a_{k_u} = 0$，取某 $q \in N$，对 r 足够大，$a_{k_r} < \frac{1}{2^q}$，这意味着 Wolfe 条件不满足 $a = \frac{1}{2^q}$，即存在一个 j，有 $f_j(x_{k_r} + (\frac{1}{2^q}) d_{k_r}) > f_j(x_{k_r}) + c(\frac{1}{2}) \tau(x_{k_r})$（对 r 足够大）取极限 $r \to \infty$，有

$$f_j\left(x^* + \left(\frac{1}{2^q}\right)\bar{d}\right) > f_j(x^*) + c\left(\frac{1}{2^q}\right)\left(\min_{d \in \mathbf{R}^n} \max_{j=1,\cdots,m} \nabla f_j(x^*)^{\mathrm{T}} \bar{d} + \frac{1}{2}\bar{d}^{\mathrm{T}} \nabla^2 f_j(x^*) \bar{d}\right) = f_j(x^*) + c\left(\frac{1}{2^q}\right)\tau(x^*),$$

对至少一个 $j \in \{1,2,\cdots,m\}$ 成立，所以 $\tau(x^*) \geq 0$. 由文献[4] 中定理 2 有 $\tau(x^*) \leq 0$，所以 $\tau(x^*) = 0$.

结合 1)、2)，知 x^* 是 $F(x)$ 的临界点再由文献[4] 中定理 1 可知 x^* 为 V 上 $F(x)$ 的 Pareto 最优解.

定理 2　本文算法产生的序列 $\{x_k\}$ 超线性收敛于 Pareto 最优解 x^*.

证明　首先，定义 $r_k = \| d_{k_0} \| \dfrac{(\varepsilon/a)^{k-k_0}}{1-\varepsilon/a}$，$\delta_k = \| d_{k_0} \| (\varepsilon/a)^{k-k_0}$，根据三角不等式，假设(iv)，(vi) 及结论(a) 得

$$B(x_k, r_k) \subset B(x_{k_0}, r) \subset V, \tag{5}$$

选择任意的 $\phi \in \mathbf{R}_+$，定义

$$\tilde{\varepsilon} = \min\left\{a \frac{\phi}{1+2\phi}, \varepsilon\right\}, \tag{6}$$

对于 $j = 1,2,\cdots,m$，当 $k \geq k_0$，$\| y - x \| < \delta_k$ 时，对所有的 $x, y \in B(x_k, r_k)$，满足不等式

$$\| \nabla^2 f_j(y) - \nabla^2 f_j(x) \| < \frac{\tilde{\varepsilon}}{2}, \tag{7}$$

同理当 $l \geq k$ 时，对所有的 $y \in B(x_k, r_k)$，满足不等式

$$\| (B_{jl} - \nabla^2 f_j(x_l))(y - x_l) \| < \frac{\tilde{\varepsilon}}{2} \| y - x_l \| . \tag{8}$$

将引理 4 中的条件(i) ~ (vi),分别用 $\tilde{\varepsilon}, r_k, \delta_k, x_k$ 替代 $\varepsilon, r, \delta, x_0$,(ii)、(iii)、(iv) 分别见(7) 式、(8) 式、(5) 式,(i),(v) 不变,(vi) 换为 $r_k, \delta_k, \tilde{\varepsilon}$ 的定义. 由引理 4 结论(a) 知:$\| x_l - x_k \| \leq \| d_k \| \dfrac{1-(\tilde{\varepsilon}/a)^{l-k}}{1-\tilde{\varepsilon}/a}$(用 l 代 k,k 代 k_0) 当 $l \to \infty$ 时,得 $\| x^* - x_k \| \leq \| d_k \| \dfrac{1}{1-\tilde{\varepsilon}/a}$ 再根据引理 4 结论(d) 得:

$$\| x^* - x_{k+1} \| \leq \| d_{k+1} \| \frac{1}{1-\tilde{\varepsilon}/a} \leq \| d_k \| \frac{\tilde{\varepsilon}/a}{1-\tilde{\varepsilon}/a}, \tag{9}$$

根据三角不等式和(9) 式得:

$$\| x^* - x_k \| \geq \| x_{k+1} - x_k \| - \| x^* - x_{k+1} \| \geq \| d_k \| - \| d_k \| \frac{\tilde{\varepsilon}/a}{1-\tilde{\varepsilon}/a} = \| d_k \| \frac{1-2\tilde{\varepsilon}/a}{1-\tilde{\varepsilon}/a}. \tag{10}$$

由于 $\tilde{\varepsilon} = \min\left\{a \dfrac{\phi}{1+2\phi}, \varepsilon\right\}$,可知 $1-2\tilde{\varepsilon}/a > 0, 1-\tilde{\varepsilon}/a > 0$ 再由(9) 式、(10) 式得:

$$\| x^* - x_{k+1} \| \leq \| x^* - x_k \| \frac{\tilde{\varepsilon}/a}{1-2\tilde{\varepsilon}/a}, \tag{11}$$

根据 $\tilde{\varepsilon}$ 的定义:$0 < \dfrac{\tilde{\varepsilon}/a}{1-2\tilde{\varepsilon}/a} < \phi$ 当 $k \to +\infty$ 时,$\tilde{\varepsilon}/a \to 0, 1-2\tilde{\varepsilon}/a \to 1$,那么 $\dfrac{\tilde{\varepsilon}/a}{1-2\tilde{\varepsilon}/a} \to 0$,即

$$\lim_{k\to\infty} \frac{\| x^* - x_{k+1} \|}{\| x^* - x_k \|} = 0 , \tag{12}$$

所以本文序列 $\{x_k\}$ 超线性收敛于 x^*.

3 数值试验

为了进一步检验算法的有效性,选取算例对本文进行数值试验,并与文献[4] 中的多目标拟牛顿算法进行比较.

例:$\min\limits_{x\in\mathbf{R}^2}\left(x_1^2 - x_1x_2 + x_2^2 - 2x_1 + x_2, \dfrac{1}{4}((x_1-1)^4 + 2(x_2-2)^4)\right)$.

用 IP 表示初始点,NM 表示本文提出的新的拟牛顿算法,QN 表示文献[4]中的拟牛顿算法,Critical 表示 Pareto-临界点,Feval 表示对应的函数值,Iter 表示算法迭代次数,T 表示新算法在得到与算法 QN 相同迭代点时的迭代次数. 在本文算例中,取参数 σ=0.7,c=0.1 精度设为 $|\tau_k| \leq 10^{-6}$. 所得到的数据如表 1.

表 1 NM 方法与 QN 方法的数值结果

	IP	Critical	Feval	T	Iter
	(−2,1)	(−1,1)	(6,4.5)	…	1
(QN)	(−5,2)	(−1,2)	(10,4)	…	5
	(−10,1)	(−1,1)	(6,4.5)	…	11
	(−2,1)	(1,1)	(0,0.5)	1	5
(NM)	(−5,2)	(1,2)	(3,0)	4	10
	(−10,1)	(1,1)	(0,0.5)	10	13

从表 1 可以得出当达到同一精度时,本文提出的新的拟牛顿算法 NM 在迭代得到相同临界点的情况下,所需的迭代次数要比 QN 少或相同,但是 NM 继续迭代会得到比 QN 更优(Feval 值更小)的临界点,而 QN 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 0ptimization[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.

(下转第 35 页)

[12] Klement E P, Mesiar R, Pap E. Triangular Norms[M]. Netherlands: Kluwer Academic Plublishers, 2000.

[13] Li G, Liu H W. Continuity of left-continuous triangular norms with special associated negations[J]. Fuzzy Sets and Systems, 2013, 226: 78-88.

[14] Shen Z Q, Zhang D X. A note on the continuity of triangular norms[J]. Fuzzy Sets and Systems, 2014, 252: 35-38.

[15] Gaines B R. Foundations of fuzzy reasoning[J]. Internat J Man-machine Stud, 1976, 6: 623-668.

[16] Trillas E, Valverde L. On some functionally expressable implications for Fuzzy set theory[C]. Proc 3rd Internat Seminar on fuzzy Set Theory, Linz, 1981.

A Study on Equivalence Between R and (S,N)-implications

HAN Yuanliang[1], LI Juan[2]

(1. Department of Basic Curriculum, North China Institute of Science and Technology, Langfang 065201, China;
2. College of Mathematics and Information Science, Henan Normal University, Xinxiang 453007, China)

Abstract: When (T,S,N) is a De Morgan triple, we discussed equivalent relations between R and (S,N)-implications under the conditions that T is a continuous t-norm and a left-continuous t-norm, and obtained several equivalent theorems. Moreover, a sufficient condition has been provided based on the properties of fuzzy implications. On one hand, the results are better for properties of fuzzy implications. On the other hand, it provides the necessary theoretical basis for the application of fuzzy implications.

Keywords: fuzzy implication; triangular norm; strong negation; automorphism

---

(上接第 24 页)

[4] Povalej Ž. Quasi-Newton's method for multiobjective optimization[J]. Journal of Computational and Applied Mathematics, 2014, 255(1): 765-777.

[5] Wei Zengxin, Li Guoyin, Qi Liqun. New Quasi-Newton methods for unconstrained optimization problems[J]. Applied Mathematics and Computation, 2006, 175(2): 1156-1188.

[6] 怀丽波. 利用函数值信息的修正多步拟牛顿法[J]. 南京大学学报数学半年刊, 2007, 24(1): 142-150.

[7] 黄 海, 林穗华. 几种修正拟牛顿法的比较[J]. 广西民族师范学院学报, 2011, 28(3): 8-11.

[8] 袁亚湘, 孙文瑜. 最优化理论与方法[M]. 北京: 科学出版社, 2003.

[9] 郑发美, 刘辉辉. 一类新拟牛顿算法的全局收敛性与数值试验[J]. 河南师范大学学报(自然科学版), 2010, 38(2): 35-38.

[10] 景书杰, 徐莹莹. Wolfe线搜索下一类新的结构拟牛顿算法[J]. 河南理工大学学报(自然科学版), 2015, 34(4): 589-592.

A New Quasi-Newton's Method for Unconstrained Multiobjective Optimization

WANG Feifei, XU Er, ZHAO Jinling

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

Abstract: Based on the improved quasi-Newton's methods for unconstrained optimization problems, we present a new quasi-Newton's method for multiobjective optimization without constraints, and prove its global and superlinear convergence under Wolfe line search. Numerical results also show that the proposed method is efficient, and better critical point can be obtained by iteration.

Keywords: multiobjective optimization; new quasi-Newton's method; Wolfe line search; Pareto optimal