第25卷第1期 数 学 研 究 与 评 论 Vol.25, No.1
2005年2月 JOURNAL OF MATHEMATICAL RESEARCH AND EXPOSITION Feb., 2005

文章编号: 1000-341X(2005)01-0114-08 文献标识码: A

一类广义拟牛顿算法的收敛性

焦宝聪， 陈兰平
(首都师范大学数学系， 北京 100037)

摘 要: 本文提出一类广义拟牛顿算法，新类算法降低了关于目标函数的假设条件，将线搜索扩展到一般形式，它概括了若干种常用的非精确线搜索技术. 此外，算法对迭代校正公式中的参数 Φ_k 的选取范围做了较大扩展 (可以取负值).

关键词: 广义拟牛顿算法；无约束最优化；一般线搜索原则；全局收敛性.
MSC(2000): 90C
中图分类: O174.13

1 引 言

对于无约束最优化问题

$$\min f(x), \quad x \in R^n \tag{1.1}$$

拟牛顿算法是目前应用最广泛的解法之一. 关于带非精确线搜索的非拟牛顿算法的研究见文献 [1,2,5,6] 等. 近年来，基于修正拟牛顿方程的非拟牛顿算法的研究吸引了不少国内外学者. 1991 年， Yuan Ya-xiang[9] 提出了一种修正的 BFGS 方法，同年，杨晓光 [11] 在一定条件下扩大了对 BFGS 方法中的参数 Φk 的选取范围. 1995 年， Yuan Ya-xiang, Richard, H.Byrd[10] 又给出了一类非拟牛顿算法算法. 1995 年，柯小伍 [12] 证明了 Broyden 非凸族的收敛性. 1997 年，陈兰平等 [8] 提出了一类非拟牛顿算法，证明了非拟牛顿凸族的收敛性. 1999 年，焦宝聪 [3] 提出了一类广义拟牛顿算法，在目标函数一致凸的条件下，结合 Wolfe 线搜索原则，证明了所给算法的全局收敛性和局部超线性收敛性. 本文改进了文献 [3] 中给出的广义拟牛顿算法，改进后的广义拟牛顿算法降低了关于目标函数的假设条件，将线搜索扩展到一般形式，它概括了若干种常用的非精确线搜索技术，使文献 [13,14] 中的算法成为本文算法的特例. 此外，算法对迭代校正公式中的参数 Φ_k 的选取范围做了较大扩展 (可以取负值). 我们对改进后的新算法，给出了全局收敛性的证明. 数值算例表明，适当选择参数 Φ_k 的值，可以提高计算的效率.

对问题 (1.1)，我们假设 $f(x)$ 二阶连续可微，记 $f_k = f(x_k), g_k = \nabla f(x_k), \gamma_k = g_{k+1} - g_k, G_k = \nabla^2 f(x_k), \delta_k = x_{k+1} - x_k$. 当 $\|\delta_k\|$ 很小时，由

$$\frac{1}{2}\delta_k^T G_k \delta_k \approx f_{k+1} - f_k - g_k^T \delta_k,$$

$$\delta_k^T G_k \delta_k \approx \delta_k^T \gamma_k,$$

有

$$\delta_k^T G_k \delta_k \approx \theta \delta_k^T \gamma_k + 2(1-\theta) R_k, \quad \theta \in [0,1],$$

收稿日期: 2002-06-27
基金项目: 国家自然科学基金 (10371101), 北京市教委科研基金 (KM200310028117)

对二次函数上式等号成立. 为书写方便，记 $Q_k(\theta)=\theta\delta_k^T\gamma_k+2(1-\theta)R_k$, 其中 $R_k=f_{k+1}-f_k-g_k^T\delta_k$.

考虑如下迭代:

$$x_{k+1}=x_k+\lambda_k d_k, \tag{1.2}$$

$$d_k=-B_k^{-1}g_k, \tag{1.3}$$

$$B_{k+1}(\theta,\Phi_k)=B_k-\frac{B_k\delta_k\delta_k^TB_k}{\delta_k^TB_k\delta_k}+\frac{Q_k(\theta)}{(\delta_k^T\delta_k)^2}\delta_k\delta_k^T+\Phi_k(\delta_k^TB_k\delta_k)Z_kZ_k^T, \tag{1.4}$$

其中 B_k 为 G_k 的近似矩阵， λ_k 为搜索步长， $\Phi_k\in[0,1),x_1$ 为初始点,

$$Z_k=\frac{\delta_k}{\delta_k^T\delta_k}-\frac{B_k\delta_k}{\delta_k^TB_k\delta_k}.$$

由 B_k 产生的修正矩阵 B_{k+1} 满足非拟牛顿方程

$$\delta_k^TB_{k+1}\delta_k=Q_k(\theta),\quad \theta\in[0,1]. \tag{1.5}$$

设问题 (1.1) 中的目标函数 $f(x)$ 在集合 $H\subset R^n$ 上一致凸，即存在正数 $M>m>0$ 使得对所有的 $x\in H$ 及 $Z\in R^n$, 有不等式

$$m\|Z\|^2\le Z^TG(x)Z\le M\|Z\|^2,$$

并假设存在 x_1 使水平集 $L=\{x|f(x)\le f(x_1)\}\subseteq H$ 有界.

文献 [3] 在这些假设下，将迭代 (1.2)–(1.5) 与 Wolfe 线搜索结合，给出了一类广义拟牛顿算法，并证明了算法的全局收敛性. 本文改进了文献 [3] 的结果，主要减弱了目标函数的条件，扩大了 Φ_k 的取值范围，并使得该算法对更一般的线搜索原则所产生的点列 $\{x_k\}$ 满足

$$\liminf_{x\to\infty}\|g_k\|=0,$$

其中 $\|\cdot\|$ 为欧氏范数.

2 有关的假设、定义及引理

本文假设

(H) 水平集 $L=\{x|f(x)\le f(x_1)\}$ 有界，目标函数 $f(x)$ 在 L 上二次连续可微，且 $Q_k(\theta)>0,k=1,2,\cdots$

引理 2.1 设

$$D_k=B_k-\frac{B_k\delta_k\delta_k^TB_k}{\delta_k^TB_k\delta_k}+\frac{Q_k(\theta)}{(\delta_k^T\delta_k)^T}\delta_k\delta_k^T, \tag{2.1}$$

若 B_k 正定，则 D_k 也正定，且有

$$\det(D_k)=\det(B_k)\frac{Q_k(\theta)}{\delta_k^TB_k\delta_k}, \tag{2.2}$$

其中 $\det(B)$ 表示 B 的行列式.

引理 2.2 设 A 为 n 阶非奇异矩阵，α 是任意 n 维向量，λ 为任意常数，则有

$$\det(A+\lambda\alpha\alpha^T)=(1+\lambda\alpha^T A^{-1}\alpha)\det(A). \tag{2.3}$$

证明 $\alpha=0$ 时结论显然. 下面假设 $\alpha\neq 0$. 由于

$$\det(A+\lambda\alpha\alpha^T)=\det(A)\det(I+\lambda\alpha\alpha^T A^{-1}\alpha),$$

而

$$(I+\lambda\alpha\alpha^T A^{-1})\alpha=(1+\lambda\alpha^T A^{-1}\alpha)\alpha,$$

所以，向量 α 是矩阵 $I+\lambda\alpha\alpha^T A^{-1}$ 对应于特征值 $1+\lambda\alpha^T A^{-1}\alpha$ 的一个特征向量. 对任一非零向量 $x\in R^n$, 若满足 $x^T A^{-T}\alpha=0$, 则有 $(I+\lambda\alpha\alpha^T A^{-1})x=x$. 这表明 x 是矩阵 $I+\lambda\alpha\alpha^T A^{-1}$ 的一个特征向量，相应的特征值为 1. 由于 A 非奇异，$\alpha\neq 0$, 故 $A^{-T}\alpha$ 的正交空间是 $n-1$ 维子空间，所以矩阵 $I+\lambda\alpha\alpha^T A^{-1}$ 有一个特征值为 $1+\lambda\alpha^T A^{-1}\alpha$, 其余 $n-1$ 个特征值为 1. 所以

$$\det(I+\lambda\alpha\alpha^T A^{-1})=1+\lambda\alpha^T A^{-1}\alpha,$$

从而有

$$\det(A+\lambda\alpha\alpha^T)=(1+\lambda\alpha^T A^{-1}\alpha)\det(A).$$

引理 2.3 校正公式 (1.4) 中的 B_{k+1} 满足

$$\det(B_{k+1})=[1+\Phi_k(\delta_k^T B_k\delta_k)(Z_k D_k^{-1}Z_k)]\det(B_k)\frac{Q_k(\theta)}{\delta_k^T B_k\delta_k}, \tag{2.4}$$

其中 D_k 由 (2.1) 式定义.

证明 由 (1.4) 式与 (2.1) 式可知

$$B_{k+1}=D_k+\Phi_k\delta_k^T B_k\delta_k)Z_k Z_k^T.$$

由引理 2.1, 当 B_k 正定时 D_k 正定，从而 D_k 非奇异. 再由引理 2.2 有

$$\begin{aligned}\det(B_{k+1})&=[1+\Phi_k(\delta_k^T B_k\delta_k)(Z_k^T D_k^{-1}Z_k)]\det(D_k)\\&=[1+\Phi_k(\delta_k^T B_k\delta_k)(Z_k^T D_k^{-1}Z_k)]\det(B_k)\frac{Q_k(\theta)}{\delta_k^T B_k\delta_k}.\end{aligned}$$

引理 2.4 当 $1+\Phi_k(\delta_k^T B_k\delta_k)(Z_k^T D_k^{-1})>0$ 时，由 B_k 正定可保证 B_{k+1} 正定.

证明 由 B_k 正定及引理 2.1 知 D_k 正定，要证 B_{k+1} 正定，只须证明：对任意非零向量 $x\in R^n$, 有 $x^T B_{k+1}x>0$. 事实上，

$$x^T B_{k+1}x=x^T D_k x+\Phi_k(\delta_k^T B_k\delta_k)(x^T Z_k)^2, \tag{2.5}$$

由 D_k 和 B_k 的正定性可知：当 $\Phi_k\geq 0$ 时，(2.5) 右端为正. 但现在只要求 Φ_k 满足 $1+\Phi_k(\delta_k^T B_k\delta_k)(Z_k^T D_k^{-1}Z_k)>0$, 即

$$\Phi_k>-\frac{1}{(\delta_k^T B_k\delta_k)(Z_k^T D_k^{-1}Z_k)}.$$

此时由 (2.5) 式知

$$x^T(B_{k+1})x > x^T D_k x - \frac{(x^T Z_k)^2}{(Z_k^T D_k^{-1} Z_k)}.$$

由 D_k 的对称正定性和柯西不等式有

$$(x^T D_k x)((Z_k^T D_k^{-1} Z_k) \geq (x^T Z_k)^2.$$

从而 $x^T B_{k+1} x > 0$, 即推理 B_{k+1} 正定.

以下我们总假定 $1 + \Phi_k(\delta_k^T B_k \delta_k)(Z_k^T D_k^{-1} Z_k) > 0$. 由 B_1 正定及校正公式 (1.4), 我们得到正定对称序列 $\{B_k\}$, 从而保证在每步迭代中由公式 (1.3) 确定的搜索方向 d_k 为下降方向.

引理 2.5 若假设 (H) 成立，则存在常数 $M_1 > 0$, 使得对所有下标 $k \geq 1$ 有

$$\frac{Q_k(\theta)}{\delta_k^T \delta_k} \leq M_1. \tag{2.6}$$

证明 令 L^* 表示水平集 L 的凸包的闭集，即包含 L 的最小闭凸集. 由 L 的有界性及 $G(x)$ 的连续性，必存在常数 M_2, 使得对任意的 $x \in L^*$ 有 $\|G(x)\| \leq M_2$. 又由我们给出的算法为下降算法，故由算法产生的迭代点列 $\{x_k\} \subset L^*$, 故对任意 $t \in [0,1]$, 有 $x_k + t\delta_k \in L^*$. 所以

$$\begin{aligned} R_k = f_{k+1} - f_k - g_k^T \delta_k &= \int_0^1 \int_0^t \delta_k^T G(x_k + \tau\delta_k)\delta_k \mathrm{d}\tau \mathrm{d}t \\ &\leq \int_0^1 \int_0^t M_2 \|\delta_k\|^2 \mathrm{d}\tau \mathrm{d}t = \frac{M_2}{2}\|\delta_k\|^2. \end{aligned}$$

又因为

$$\delta_k^T \gamma_k = \delta_k^T(g_{k+1} - g_k) = \int_0^1 \delta_k^T G(x_k + t\delta_k)\delta_k \mathrm{d}t \leq M_2\|\delta_k\|^2.$$

从而存在 $M_1 > 0$ 使

$$\frac{Q_k(\theta)}{\delta_k^T \delta_k} = \frac{\theta\delta_k^T\gamma_l + 2(1-\theta)R_k}{\|\delta_k\|^2} \leq M_1.$$

下面引入强迫函数定义 [7]:

定义 映射 $\sigma : [0,+\infty) \to [0,+\infty)$ 称为强迫函数，如果对任意 $t_j > 0$, $\lim_{j\to\infty} \sigma(t_j) = 0$, 则 $\lim_{j\to\infty} t_j = 0$.

在本文中，我们将文 [3] 中的 Wolfe 线搜索原则扩展为如下一般线搜索原则 [11]:

假设 $\sigma_1(\cdot), \sigma_2(\cdot)$ 是两个强迫函数， c_1, c_2 是两个非负常数，且 $c_1 + c_2 > 0, r \in (0,1)$, 步长 λ_k 满足

(1)

$$f(x_k) - f(x_k + \lambda_k d_k) \geq c_1\sigma_1(-\delta_k^T g_k) + c_2\sigma_2(\|\delta_k\|), \tag{2.7}$$

(2) 存在 ξ_k, β_k 使得 $|\xi_k| < \beta_k \leq \lambda_k$

$$f(x_k + \beta_k d_k) \geq f(x_k + \xi_k d_k) + (\beta_k - \xi_k) r g_k^T d_k. \tag{2.8}$$

由于 $f(x)$ 为 L 上的有界函数，且迭代公式 (1.3) 确定的搜索方向 d_k 为下降方向，故按此线搜索原则有

$$\lim_{k\to\infty}\frac{g_k^Td_k}{\|d_k\|}=0.$$

且有以下关系成立:

$$\lim_{k\to\infty} g_k^T\delta_k=0. \tag{2.9}$$

$$g_{k+1}^Td_k\geq rg_k^Td_k,\quad r\in(0,1). \tag{2.10}$$

$$Q_k(\theta)\geq\theta(r-1)g_k^T\delta_k. \tag{2.11}$$

不难看出，该一般搜索原则包括了 Goldstein,Wolfe,Cuury-Altmenn 线搜索原则.

3 算法及其收敛性

迭代 (1.2)–(1.5) 与一般线搜索 (2.7)–(2.8) 构成求解问题 (1.1) 的新一类广义拟牛顿算法. 与文 [3] 提出的算法相比，将关于目标函数 $f(x)$ 的假设由在 L 上一致凸减弱为凸函数; 参数 Φ_k 的取值范围由 $\Phi_k\in[0,1)$ 扩展到 $\Phi_k\in(\frac{-1}{\delta_k^TB_k\delta_kZ_k^TD_k^{-1}Z_k},1)$; 确定步长 λ_k 的原则由 Wolfe 原则扩展为一般线搜索 (2.7)–(2.8). 本节将讨论在所给条件下新算法的全局收敛性.

定理 3.1 假设 (H) 成立，对由迭代 (1.2)–(1.5) 与一般线搜索 (2.7)–(2.8) 构成的算法，若对任意的 $\tau_1,\tau_2\in(0,1)$, 选择参数 Φ_k 使其满足:

$$\overline{\lim_{k\to\infty}}\Phi_k\leq\tau_2,$$

$$\lim_{k\to\infty}[1+\Phi_k(\delta_k^TB_k\delta_k)(Z_k^TD_k^{-1}Z_k)]\geq\tau_1. \tag{3.1}$$

则算法是收敛的，即有

$$\lim_{k\to\infty}\inf\|g_k\|=0. \tag{3.2}$$

证明 反设 (3.2) 不成立，则存在 $\varepsilon>0$, 使得对任意 $k\geq1$ 有 $\|g_k\|\geq\varepsilon$. 记 Ψ_k 满足

$$-\frac{\delta_k^TB_k\delta_k}{\delta_k^T\delta_k}=\Psi_k\frac{\|B_k\delta_k\|^2}{\delta_k^TB_k\delta_k}, \tag{3.3}$$

则由关系式 $B_k\delta_k=-\lambda_kg_k$ 有

$$\Psi_k=-\frac{(\delta_k^Tg_k)^2}{\|\delta^k\|^2\|g_k\|^2}. \tag{3.4}$$

利用 (2.8) 并注意到 $\|g_k\|\geq\varepsilon$ 可知 $\lim_{k\to\infty}\Psi_k=0$. 由定理条件 (3.1), 存在 $K>0$, 当 $k\geq K$ 时有

$$1-\Phi_k-\Psi_k\Phi_k\geq\frac{1-\tau_2}{2}, \tag{3.5}$$

$$\det(B_{k+1})\geq\frac{\tau_1}{2}\det(B_k)\frac{Q_k(\theta)}{\delta_k^TB_k\delta_k}. \tag{3.6}$$

对校正公式 (1.4) 两边求迹，并利用 (3.3), 有

$$\begin{aligned}Tr(B_{k+1}) &= Tr(B_k) - (1-\Phi_k)\frac{\|B_k\delta_k\|^2}{\delta_k^T B_k\delta_k}\frac{Q_k(\theta)}{\delta_k^T\delta_k} - \Phi_k\frac{\delta_k^T B_k\delta_k}{\delta_k^T\delta_k}\\ &= Tr(B_k) - (1-\Phi_k-\Psi_k\Phi_k)\frac{\|B_k\delta_k\|^2}{\delta_k^T B_k\delta_k} + \frac{Q_k(\theta)}{\delta_k^T\delta_k}.\end{aligned}$$

由引理 2.5 及 (3.5) 式

$$Tr(B_{k+1}) < Tr(B_k) - \frac{1-\tau_2}{2}\frac{\|B_k\delta_k\|^2}{\delta_k^T B_k\delta_k} + M_1,$$

所以存在常数 N_1, N_2, 使得

$$Tr(B_{k+1}) \le N_1 k,\quad \sum_{j=1}^{k}\frac{\|B_j\delta_j\|^2}{\delta_j^T B_j\delta_j} \ge N_2 k.$$

再利用几何平均与算术平均不等式，有

$$\det(B_{k+1}) \le (\frac{N_1 k}{n})^n,$$

$$\prod_{j=1}^{k}\frac{\|B_j\delta_j\|^2}{\delta_j^T B_j\delta_j} \le N_2^k.$$

由 (3.6) 又有

$$\frac{\det(B_{k+1})}{\det(B_1)} \ge (\frac{\tau_2}{2})^k\prod_{j=1}^{k}\frac{Q_j(\theta)}{\delta_j^T B_j\delta_j},$$

所以

$$\prod_{j=1}^{k}\frac{Q_j(\theta)\|B_j\delta_j\|^2}{(\delta_j^T B_j\delta_j)^2} \le \frac{N_2^k(\frac{2}{\tau_1})^k(\frac{N_1 k}{n})^n}{\det(B_1)} \le N_3^k,$$

这里 N_3 是个充分大的常数.

由关系式

$$B_k\delta_k = -\lambda_k g_k,$$

$$Q_k(\theta) \ge \theta(r-1)g_k^T\delta_k,\quad r\in(0,1),$$

有

$$-\frac{\|B_k\delta_k\|^2 Q_k(\theta)}{(\delta_k^T B_k\delta_k)^2} \le \frac{\theta(1-r)\|g_k\|^2}{-\delta_k^T g_k}.$$

由 $\|g_k\| \ge \varepsilon(k=1,2,\cdots)$ 知

$$\prod_{j=1}^{k}\frac{\varepsilon^2}{|\delta_j^T g_j|} \le \prod_{j=1}^{k}\frac{\|g_j\|^2}{|\delta_j^T g_j|} < (\frac{N_3}{1-r})^k,$$

这与 (2.9) 式矛盾. 故有

$$\lim_{k\to\infty}\inf\|g_k\|=0.$$

关于 Φ_k 选取的说明：注意到 B_k,D_k 都是正定对称矩阵，又 $\delta_k\neq 0$, 故对任意的 $\tau_1,\tau_2\in(0,1)$, 当 $Z_k=0$ 时，可选取 $\Phi_k\in(-\infty,\tau_2)$; 否则，可选取

$$\Phi_k\in[\frac{\tau_1-1}{\delta_k^TB_k\delta_kZ_k^TD_k^{-1}Z_k},\tau_2]$$

关于假设 (H) 中的 $Q_k(\theta)>0$ 条件的讨论：对二次连续可微一致凸目标函数， $Q_k(\theta)>0$ 自然满足，对一般二次连续可微非凸目标函数，若采用 Goldstein 线搜索，则在取 $\theta=0$ 时条件满足；若采用 Wolfe 线搜索同时取 $\theta=1$ 时条件亦能满足；而对一般目标函数采用一般的非精确线搜索，能否保证条件 $Q_k(\theta)>0$ 成立是值得进一步研究的问题.

4 数值算例

为检验算法的实算效果，对算法进行了一定的数值试验，数值结果表明算法是有效的. 以下算例中均取参数 $\theta=0$. 计算结果见表 4.1.

1[13]. $\min f(x)=2x_1^2+x_2^2+2x_3^2-x_1x_3-x_1x_2+3x_1+2x_2+x_3$.

2[14]. $\min f(x)=5x_1^2+7.5x_2^2+10x_3^2+2x_1x_2+4x_2x_3+x_1x_3+e^{x_1+x_2+x_3}$.

表 4.1

算例	x_0	k_1	k_2	x_k	f_k
1.	$(-1.0,-1.0,-1.0)$	13	6	$(-1.307692,-1.653846,-0.576923)$	-3.903846
2.	$(-1.0,1.5,-0.5)$	13	3	$(-0.075419,-0.039118,-0.031607)$	0.9271699

其中 x_0 是初始迭代点且 k_1 是文献 [13,14] 中算法的迭代次数， k_2 则表示按本文算法确定 Φ_k 时的迭代次数. 关于算法的超线性收敛性还有等进一步的研究.

参考文献:

[1] BYRD R H, NOCEDAL J, YUAN Y. *Global convergence of a class of quasi-Newton methods on convex problems* [J]. SIAM J. Numer. Anal., 1987, **24**: 1171–1189.

[2] POWELL M J D. *Some Global Convergence Properties of a Variable Metric Algorithm for Minimization without Exact Line Searches, in Nonlinear Programming* [M]. SIAM-AMS Proceeding, Vol.IX,R.W.Cottle and C.E.Lemke, eds, American Mathematical Society Providence. RI,1976.

[3] 焦宝聪. 一类超线性收敛的广义拟牛顿算法 [J]. 高等学校计算数学学报， 1999, **21**(6): 178–188.
JIAO Bao-cong. *A class of generalized quasi-Newton algorithm with super linear convergence* [J]. Numer. Math. J. Chinese Univ., 1999, **21**(6): 178–188.

[4] 陈兰平，焦宝聪. 非拟牛顿非凸族的收敛性 [J]. 计算数学， 2000, **22**(3): 369–378.
CHEN Lan-ping, JIAO Bao-cong. *Convergence properties of the preconvex part of Non-quasi-Newton's family* [J]. Math. Numer. Sin., 2000, **22**(3): 369–378.

[5] 刘光辉，韩继业. 带一类非精确搜索的 Byoyden 族的全局收敛性 [J]. 计算数学， 1996, **18**(3): 233–240.
LIU Guang-hui, HAN Ji-ye. *Global convergence of the Broyden's family with a class inexact linessearches* [J]. Math. Numer. Sin., 1996, **18**(3): 233–240.

[6] R H Byrd, NOCEDAL J. *A tool for the methods with application to unconstrained minimization* [J]. SIAM J. Numer. Anal., 1989, **26**: 727–739.

[7] ORTEGA J M, RHEINBOLDT W C. *Iterative Solution of Nonlinear Equations in Several Variables* [M]. Academic Press, New York, 1976.

[8] 陈兰平，焦宝聪．一类非拟 Newton 算法及其收敛性 [J]. 应用数学与计算数学学报， 1997, **11**(2): 9–17.
CHEN Lan-ping, JIAO Bao-cong. *A calss of Non-quasi-Newton methods and its Convergence* [J]. Comm. Appl. Math. Comput., 1997, **11**(2): 9–17.
[9] YUAN Ya-xiang. *A modified BFGS alogorithm for unconstrained optimization* [J]. IMA Journal of Numerical Analysis, 1991, **11**: 325–332.
[10] YUAN Ya-xiang, BYRD R H. *Non-quais-Newton updates for unconstrained optimization* [J]. J. of Compu. Math., 1995, **2**: 95–107.
[11] 杨晓光．一类拟牛顿算法的收敛性 [C]. 清华大学应用数学论文集， 1991, 113–117.
YANG Xiao-guang. *Global convergence of a class quais-Newton methods* [C]. Tsinghua Univ. Appl. Math. Theses, 1991, 113–117.
[12] 柯小伍． Broyden 非凸族的收敛性 [J]. 北京师范大学学报， 1995, **31**(1): 6–9.
KE Xiao-wu. *Convergence of the preconvex part of Broyden's family* [J]. J. Beijing Normal Univ. (Natural Science), 1995, **31**(1): 6–9.
[13] 赵云彬，段虞荣．伪 Newton-δ 族算法对一般目标函数的收敛性 [J]. 数值计算与计算机应用， 1996, **17**(1): 36–47.
ZHAO Yun-bin, DUAN Yu-rong. *Convergence of the pseudo-Newton-δ calss methods for general object functions* [J]. J. Numer. Methods Comput. Appl., 1996, **17**(1): 36–47.
[14] 赵云彬，易正俊．伪 Newton-δ 族的导出和全局收敛性 [J]. 数值计算与计算机应用， 1995, **16**(1): 53–62.
ZHAO Yun-bin, YI Zheng-jun. *Derivation and global convergence for pseudo-Newton-δ class* [J]. J. Numer. Methods Comput. Appl., 1995, **16**(1): 53.

Convergence Properties of a Class of Generalized Quasi-Newton Methods

Jiao Bao-cong, CHEN Lan-ping
(Dept. of Math., Captial Normal University, Beijing 100037, China)

Abstract: In this paper,we present a calss of the generlized quasi-Newton methods for unconstrained optimization, and study the global convergence properties of the methods when applied to a general objective function. We assume that line search satisfies the general form of stepsize selection rules which summarizes many known stepsize selection rules as its special caese, and that the parameter Φ_k in the matrices update formulae by the methods may be a negative value.

Key words: quasi-Newton methods; unconstrained optimization; global convergence.