References

[1] M.J.D.Powell, *Some global convergence properties of a variable metric algorithm for minimization without exact linesearches* , in Nonlinear Programming, SIAM-AMS Proceedings, Vol.IX., R.W.Cottle and C.E.Lemke, eds., American Mathematical Society, Providence, RI, 1976.

[2] M.J.D.Powell, *Some properties of the variable metric algorithm* , in Numerical Methods for Nonlinear Optimization, F.A.Lootsma, ed., Academic Press, London, 1972.

[3] M.J.D.Powell, *On the convergence of the variable metric algorithm* , Journal of the Institute of Mathematics and its Applications, **7**(1971), 21-36.

[4] J.D.Pearson, *Variable metric methods of minimization* , The Computer Journal, **12**(1969), 171-178.

[5] J.E.Dennis and J.J.More, *Quasi-Newton methods, motivation and theory* , SIAM Rev., **19**(1977), 46-89.

[6] J.E.Dennis and J.J.More, *A characterization of superlinear convergence and its application to quasi-Newton methods* , Math. Comput., **28**(1974), 549-560.

[7] R.H.Byrd and J.Nocedal, *A tool for the analysis of quasi-Newton methods with application to unconstrained minimization* , SIAM J. Numer. Anal., Vol.26, No.3, 1989, 727-739.

带有广义 Wolfe 线搜索的变尺度算法的收敛性

刘 光 辉

(美国西北大学工业工程与管理学系)

韩 继 业

(中国科学院应用数学所)

徐 中 玲

(国家自然科学基金委员会)

摘 要

本文提出一类广义 Wolfe 线搜索模型，并且把它与著名的BFGS方法相结合，对于所得到的算法证明了：对于凸函数算法具有全局收敛性和超线性收敛速度．这推广了参考文献[1]中的结果．

Journal of Mathematical Research & Exposition
Vol.15, No.4, 499–508, Nov. 1995

Global Convergence of the Variable Metric Algorithms with a Generalized Wolfe Linesearch *

Liu Guanghui
(Dept. of IEMS, Northwestern University, Evanston)

Han Jiye
(Institute of Applied Mathematics, Academia Sinica, Bejing 100080)

Xu Zhongling
(The National Natural Science-Foundation of China)

Abstract In this paper, we present a generalized Wolfe linesearch method, and apply it to the well-known BFGS algorithm, for which we obtain the global convergence and superlinear convergence, our results are extention of those in [1].

Keywords BFGS method, generalized Wolfe linesearch, global convergence, superlinear convergence.

Classification AMS(1991) 90C30/CCL O221.2

1. Introduction

We discuss the unconstrained optimization problem:

$$\min f(x)$$

where $f : R^n \to R^1, f \in C^1$, which is solved by means of iterative methods, $x_{k+1} = x_k + \lambda_k d_k (k = 0, 1, \cdots,)$, in which x_0 is any given starting point, $d_0 = -H_0 g_0, H_0$ is any given $n \times n$ symetric positive definite matrix, and $d_k = -H_k g_k, H_k = B_k^{-1}$ is iteratived by the following BFGS formula

$$B_{k+1} = B_k - \frac{B_k S_k S_k^T B_k}{S_k^T B_k S_k} + \frac{y_k y_k^T}{S_k^T y_k}, \tag{1}$$

where $S_k = x_{k+1} - x_k, y_k = g(x_{k+1}) - g(x_k), g$ denotes the gradient ∇f of f. The BFGS algorithm is regarded as one of the most efficient algorithm in nonlinear optimization, and used more often, so in the following sections we shall discuss its properties.

*Received Nov.12, 1993. This work is supported by the National Natural Science Foundation of China.

It is known to us that the Wolfe linesearch method is often used in both the theoretical analysis and application of algorithms. Applying Wolfe linesearch method to the BFGS algorithm, Powell ([1]) got some nice properties of BFGS algorithm. We extend Wolfe linesearch to a generalized Wolfe linesearch, which is described as **GW linesearch method:** Select the steplength λ_k satisfying

$$f(x_{k+1}) \leq f(x_k) + \varepsilon_1 \lambda_k g_k^T d_k, \tag{2}$$

$$g(x_{k+1})^T d_k \geq \max\{\varepsilon_2, 1 - (\lambda_k \|d_k\|)^p\} g_k^T d_k, \tag{3}$$

where $\varepsilon_1 \in (0,1), \varepsilon_2 \in (0, \frac{1}{2}), p \in (-\infty, 1)$. Using this new linesearch method, we shall extend the results of [1].

2. Some Lemmas

We give the assumption

(H) (i) the level set $L_0 = \{x | f(x) \leq f(x_1)\}$ is bounded.

(ii) the objective function f is convex on L_0.

(iii) $f \in C^2$, moreover, there exists a constant $M > 0$, such that

$$\|G(x)\| \leq M,$$

where $G(x) = \nabla^2 f(x)$.

Lemma 1[2] *Assume that (H) holds. Then there exists a positive number M_1 such that*

$$\frac{\|y_k\|^2}{y_k^T S_k} \leq M_1, \quad k = 1, 2, \cdots.$$

Lemma 2 *If the sequence of nonnegative numbers $m_k, k = 1, 2, \cdots$ is such that*

$$\prod_{j=1}^{k} m_j \geq c_1^k, \quad c_1 > 0, k = 1, 2, \cdots$$

Then $\limsup_k m_k > 0$.

Proof We, by contradiction, assume that $\limsup_k m_k = 0$, then, for $0 < \varepsilon < c_1$, there exist $k_0 > 0$, such that $m_k < \varepsilon$, for all $k \geq k_0$. Hence, for all $k > k_0$,

$$c_1^k \leq \prod_{j=1}^{k_0 - 1} m_j \prod_{j=k_0}^{k} \varepsilon, \quad (\frac{c_1}{\varepsilon})^k \leq (\prod_{j=1}^{k_0-1} m_j)\varepsilon^{1-k_0},$$

$$+\infty = \limsup_k (\frac{c_1}{\varepsilon})^k \leq (\prod_{j=1}^{k_0-1} m_j)\varepsilon^{1-k_0} < +\infty,$$

which is a contradiction, thus, $\limsup p_k m_k > 0$.

Lemma 3[4] $\det(B_{k+1}) = \det(B_k)\frac{y_k^T S_k}{S_k^T B_k S_k}$, *where* $\det(B_k)$ *denotes the determinate of* B_k.

Lemma 4 *Assume that (H) (i),(iii) hold. Then*

$$\lim_k \|g_k\| \cos\theta_k = 0,$$

where $\cos\theta_k = \frac{-g_k^T d_k}{\|g_k\|\|d_k\|}$.

Proof From (H) (iii) it follows that

$$[g(x_k + \lambda_k d_k) - g_k]^T d_k = \lambda_k d_k^T \int_0^1 G(x_k + t\lambda_k d_k)dt \le \lambda_k \|d_k\|^2 M.$$

On the other hand, from (3) it follows that

$$[g(x_k + \lambda_k d_k) - g_k]^T d_k \ge \max\{\varepsilon_2 - 1, -\|S_k\|^p\} g_k^T d_k = -\min\{1 - \varepsilon_2, \|S_k\|^p\} g_k^T d_k. \quad (4)$$

Thus,

$$\lambda_k \|d_k\|^2 M \ge -\min\{1 - \varepsilon_2, \|S_k\|^p\} g_k^T d_k,$$

i.e.,

$$\|S_k\| \ge \frac{1}{M} \min\{1 - \varepsilon_2, \|S_k\|\}\gamma_k, \quad (5)$$

where $\gamma_k = \frac{-g_k^T d_k}{\|d_k\|}$.

Comparing $\|S_k\|^p$ with $1 - \varepsilon_2$, we easily derive from (5) that

$$\|S_k\| \ge \min\{\frac{1-\varepsilon}{M}\gamma_k, (\frac{\gamma_k}{M})^{\frac{1}{1-p}}\}.$$

Substitute it into (2), we have

$$f(x_{k+1}) \le f(x_k) - \varepsilon_1 \lambda_k \|d_k\| \gamma_k \le f(x_k) - \varepsilon_2 \min\{\frac{1-\varepsilon}{M}\gamma_k, (\frac{\gamma_k}{M})^{\frac{1}{1-\gamma}}\}\gamma_k.$$

The assumption (H) (i) implies that $\lim_k [f(x_k) - f(x_{k+1})] = 0$. Therefore,

$$\lim_{k\to\infty} \|g_k\| \cos\theta_k = \lim_{k\to\infty} \gamma_k = 0.$$

Lemma 5 *Assume that (H) (i),(iii) hold. Then*

$$\lim_{k\to\infty} \max\{\|S_k\|, \|S_k\|^{1-p}\} \|g_k\| \cos\theta_k = 0.$$

Proof (2) implies that

$$\begin{aligned} \varepsilon_1 \lambda_k \|d_k\| \gamma_k &\le f(x_k) - f(x_{k+1}), \\ \|S_k\| &\le \frac{f(x_k) - f(x_{k+1})}{\varepsilon_1 \|g_k\| \cos\theta_k}, \end{aligned}$$

hence,

$$0 \le \max\{\|S_k\|, \|S_k\|^{1-p}\}\|g_k\| \cos\theta_k \le \max\{\frac{f(x_k) - f(x_{k+1})}{\varepsilon_1}, (M_0)^{1-p}\|g_k\| \cos\theta_k\},$$

where M_0 is a constant satisefying $\|S_k\| \le M_0$.

From Lemma 4 and (H) (i) it is easy to see that this lemma is true.

3. Global Convergence of BFGS Algorithm Using the GW Linesearch Method

Theorem 1 *Assume that (H) holds. Suppose that x_0 is any starting point, B_0 is any symmetric positive definite matrix, and that the sequence $\{x_k\}$ is generated by he BFGS algorithm, in which the stepsize λ_k is determined by the GW linesearch method (2),(3). Then $\lim_k \inf \|g_k\| = 0$.*

Proof We, by contradiction, assume that $\liminf_k \|g_k\| > 0$, i.e., there exists $c_2 > 0$ such that

$$\|g_k\| \ge c_2, \quad k = 0, 1, \cdots \tag{6}$$

From (1) we have

$$\mathrm{Tr}(B_{k+1}) = \mathrm{Tr}(B_k) - \frac{\|B_k S_k\|^2}{S_k^T B_k S_k} + \frac{\|y_k\|^2}{y_k^T S_k}, \tag{7}$$

where $\mathrm{Tr}(B_k)$ denotes the trace of B_k. From (6), (7) and Lemma 1 we have

$$\begin{aligned} 0 < \mathrm{Tr}(B_{k+1}) &\le \mathrm{Tr}(B_k) - \frac{\|g_k\|^2}{g_k^T H_k g_k} + M_1 \\ &\le \cdots \\ &\le \mathrm{Tr}(B) - \sum_{j=1}^{k} \frac{c_2^2}{g_j^T H_j g_j} + kM_1. \end{aligned}$$

Hence

$$\mathrm{Tr}(B_{k+1}) \le \mathrm{Tr}(B_1) + kM_1. \tag{8}$$

$$\sum_{j=1}^{k} \frac{1}{g_j^T H_j g_j} \le \frac{\mathrm{Tr}(B_1) + kM_1}{c_2^2}$$

From the geometric-arithmetic mean value formula we have

$$\prod_{j=1}^{k} g_j^T H_j g_j \ge [\frac{kc_2^2}{\mathrm{Tr}(B_1) + kM_1}]^k \tag{9}$$

(4) and Lemma 3 imply that

$$\det(B_{k+1}) = \det(B_k)\frac{[g(x_{k+1}) - g_k]^T d_k}{\lambda_k g_k^T H_k g_k}$$

$$\geq \det(B_k)\frac{\min\{1-\epsilon_2, ||S_k||^p\}}{\lambda_k}$$

$$\geq \det(B_1)\prod_{j=1}^{k}\frac{\min\{1-\varepsilon_2, ||S_j||^p\}}{\lambda_j},$$

$$\prod_{j=1}^{k}\max\{\frac{\lambda_j}{1-\varepsilon_2}, \frac{\lambda_j}{||S_j||}\} \geq \frac{\det(B_1)}{\det(B_{k+1})}. \tag{10}$$

Again using the geometric-arithmetic mean value formula, we have

$$\det(B_{k+1}) \leq [\frac{\mathrm{Tr}(B_{k+1})}{n}]^n.$$

Therefore, from (8),(10), we have

$$\prod_{j=1}^{k}\max\{\frac{\lambda_j}{1-\varepsilon_2}, \frac{\lambda_j}{||S_j||}\} \geq \frac{\det(B_1)n^n}{[\mathrm{Tr}(B_1)+kM_1]^n} \geq \frac{1}{k^n}\frac{\det(B_1)n^n}{[\mathrm{Tr}(B_1)+M_1]^n}$$

$$\geq (\frac{1}{e^n})^k \min\{\frac{\det(B_1)n^n}{[\mathrm{Tr}(B_1)+M_1]^n}, 1\} \geq c_3^k, \tag{11}$$

where $c_3 \leq \frac{1}{e^n}\min\{\frac{\det(B_1)n^n}{[\mathrm{Tr}(B_1)+M_1]^n}, 1\}$.

(9) times (11),

$$\prod_{j=1}^{k}\max\{\frac{||S_j||||g_j||\cos\theta_j}{1-\varepsilon_2}, ||g_j||\cos\theta_j||S_j||^{1-p}\}$$

$$\geq c_3^k[\frac{kc_2^2}{\mathrm{Tr}(B_1)+kM_1}]^k \geq [\frac{c_3c_2^2}{\mathrm{Tr}(B_1)+M_1}]^k$$

Since

$$\prod_{j=1}^{k}\max\{\frac{||S_j||||g_j||\cos\theta_j}{1-\varepsilon_2}, ||g_j||\cos\theta_j||S_j||^{1-p}\}$$

$$\leq (\frac{1}{1-\varepsilon_2})^k\prod_{j=1}^{k}\max\{||S_j||, ||S_j||^{1-p}\}||g_j||\}\cos\theta_j.$$

Thus

$$\prod_{j=1}^{k}\max\{||S_j||, ||S_j||^{1-p}\}||g_j\}\cos\theta_j \geq [\frac{(1-\varepsilon_2)c_3c_2^2}{Tr(B_1)+M_1}]^k,$$

from Lemma 2 it immediately follows that

$$\lim_k\sup\max\{||S_j||, 1\}||g_j||\cos\theta_j > 0,$$

which contradicts Lemma 5. Therefore,

$$\liminf_k \|g_k\| = 0.$$

Theorem 2 *Assume that all of the assumptions in Theorem 1 hold, and assume in addition that $G(x)$ is positive definite for all $x \in L_0$. Then $x_k \to x^*$, $g(x^*) = 0$.*

Proof By Theorem 1, we know

$$\liminf_k \|g_k\| = 0.$$

i.e., there exists a subset K of $N = \{1, 2, \cdots,\}$ such that $\lim_{k \to K} \|g_k\| = 0$, since the sequence $\{x_k\}$ is bounded, without loss of generality, we assume that

$$\lim_{k \in K} x_k = x^*, \quad \|g(x^*)\| = 0.$$

We shall prove that $\lim_{k \to \infty} x_k = x^*$. We, by contradiction, assume that,

$$\lim_{k \in K_1} x_k = x^{**}, \quad x^{**} \neq x^*.$$

It is easy to see that $x^* \in L_0, x^{**} \in L_0$, and L_0 is a convex set. Since $\{f(x_k)\}$ is monotonicalliy decreasing, we have

$$f(x^*) = \lim_{k \in K} f(x_k) = \lim_{k \in K_1} f(x_k) = f(x^{**}). \tag{12}$$

However, by Taylor formula and (H) we have

$$f(x^{**}) = f(x^*) + (x^{**} - x^*)^T G(\theta x^* + (1-\theta) x^{**})(X^{**} - x^*) > f(x^*),$$

which contradicts (12). Therefore, $\lim_{k \to \infty} x_k = x^*$.

4. Superlinear convergence analysis

In this section, we shall assume that f is a unform convex function, i.e.,

(H') there exists constants $m > 0$ and $M_2 > 0$ such that

$$M_2 y^T y \geq y^T G(x) y \geq m y^T y.$$

In order to get the superlinear convergence of algorithms, as done by most paper, we make the further assumption ($\bar{\mathrm{H}}$): the Hessian matrix G is Lipschitz continuous at x^*, i.e.,

($\bar{\mathrm{H}}$) there exists a positive constant L such that

$$\|G(x) - G(x^*)\| \leq L \|x - x^*\|,$$

for all x in a neithborhood of x^*

It is easy to verify that the assumption (H') implies the positive definiteness of $G(x)$ on L_0, and stonger than (H). Under the stonger assumption (H'), we get the following better results.

Lemma 6[3] *Under the assumption (H'), we have*

$$\frac{m}{2}\|x_k - x^*\|^2 \leq f(x_k) - f(x^*) \leq \frac{1}{m}\|g_k\|^2.$$

Lemma 7 *Under the assumption (H'), there must exist two constants $c_4 > 0$ and $c_5 > 0$ such that*

$$\frac{\|y_k\|^2}{y_k^T S_k} \leq c_4, \quad \frac{\|S_k\|^2}{y_k^T S_k} \leq c_5.$$

Proof Noticing

$$y_k = \int_0^1 G(x_k + \xi\lambda_k d_k) d\xi S_k,$$

we can easily derive this Lemma from (H').

Lemma 8 *Assume that there exist $c_4 > 0, c_5 > 0$ such that*

$$\frac{\|y_k\|^2}{y_k^T S_k} \leq c_4, \quad \frac{\|S_k\|^2}{y_k^T S_k} \leq c_5.$$

Then there exists $c_6 > 0$ such that for all $k \geq 1, \prod_{j=1}^k \cos^2\theta_j \geq c_6^k$.

Proof From (1) we deduce

$$\begin{aligned}
\mathrm{Tr}(B_{k+1}) &= \mathrm{Tr}(B_k) - \frac{\|g_k\|^2}{g_k^T H_k g_k} + c_4 = \mathrm{Tr}(B_k) - \frac{\|g_k\|}{\|H_k g_k\| \cos\theta_k} + c_4 \\
&\leq \cdots \leq \mathrm{Tr}(B_1) - \sum_{j=1}^k \frac{\|g_j\|}{\|H_j g_j\| \cos\theta_j} + c_4 k \\
\mathrm{Tr}(B_{k+1}) &\leq \mathrm{Tr}(B_1) + c_4 k, \qquad (13) \\
\sum_{j=1}^k \frac{\|g_j\|}{\|H_j g_j\| \cos\theta_j} &\leq \mathrm{Tr}(B_1) + c_4 k, \\
\prod_{j=1}^k \frac{\|g_k\|}{\|H_j g_j\| \cos\theta_j} &\leq [\frac{\mathrm{Tr}(B_1) + c_4 k}{k}]^k \leq [\mathrm{Tr}(B_1) + c_4]^k. \qquad (14)
\end{aligned}$$

From Lemma 3 and the assumptions in this Lemma we have

$$\det(B_{k+1}) = \det(B_k) \frac{y_k^T S_k}{S_k^T B_k S_k} \geq \det(B_k) \frac{c_5 \|S_k\|^2}{S_k^T B_k S_k}$$

$$
\begin{aligned}
&\geq \cdots \geq \det(B_1) c_5^k \prod_{j=1}^{k} \frac{\|S_j\|^2}{S_j^T B_j S_j} \\
&= \det(B_1) c_5^k \prod_{j=1}^{k} \frac{\|H_j g_j\|}{\|g_j\| \cos\theta_j},
\end{aligned}
$$

$$
\prod_{j=1}^{k} \frac{\|H_j g_j\|}{\|g_j\| \cos\theta_j} \leq \frac{\det(B_{k+1})}{c_5^k \det(B_1)}.
$$

Using (11) and (13) we get

$$
\begin{aligned}
\prod_{j=1}^{k} \frac{\|H_j g_j\|}{\|g_j\| \cos\theta_j} &\leq \frac{[Tr(B_{k+1})]^n}{n^n c_5^k \det(B_1)} \leq \frac{[\mathrm{Tr}(B_1) + c_4 k]^n}{n^n \det(B_1) c_5^k} \\
&\leq \frac{[\mathrm{Tr}(B_1) + c_4]^n}{n^n \det(B_1)} \left(\frac{e^n}{c_5}\right)^k \leq c_7^k, \qquad (15)
\end{aligned}
$$

where $c_7 \geq \frac{e^n}{c_5} \max\{1, \frac{[\mathrm{Tr}(B_1)+c_4]^n}{n^n \det(B_1)}\}$.

(14) times (15), we obtain

$$
\begin{aligned}
\prod_{j=1}^{k} \frac{1}{\cos^2\theta_j} &\leq [c_7(\mathrm{Tr}(B_1) + c_4)]^k, \\
\prod_{j=1}^{k} \cos^2\theta_j &\geq \left[\frac{1}{c_7(\mathrm{Tr}(B_1) + c_4)}\right]^k,
\end{aligned}
$$

Select $c_6 = 1/c_7(\mathrm{Tr}(B_1) + c_4)$.

Theorem 3 *Assume that (H') holds. Suppose that x_0 is any starting point, B_0 is any symmetric positive definite matrix, and that the sequence $\{x_k\}$ is generated by the BFGS formula (1), in which the stepsize λ_k satisfies the GW linesearch method. Then*

$$
\sum_{k=1}^{\infty} \|x_k - x^*\| < \infty.
$$

Proof Noticing the proof of Lemma 4 and $\gamma_k \to 0$, we have

$$
f(x_{k+1}) \leq f(x_k) - \varepsilon_3 \min\{\gamma_k^2, (\gamma_k)^{(2-p)/(1-p)}\} = f(x_k) - \varepsilon_3 \gamma_k^2. \qquad (16)
$$

From (16) and Lemma 6 we have for all $k \geq k_1$

$$
\begin{aligned}
0 < f(x_{k+1}) - f(x^*) &\leq f(x_k) - f(x^*) - \varepsilon_3 \|g_k\|^2 \cos^2\theta_k \\
&\leq f(x_k) - f(x^*) - \varepsilon_3 m [f(x_k) - f(x^*)] \cos^2\theta_k \\
&= (1 - \varepsilon_3 m \cos^2\theta_k)[f(x_k) - f(x^*)] \\
&\leq \prod_{j=k_1+1}^{k} (1 - \varepsilon_3 m \cos^2\theta_j)[f(x_j) - f(x^*)]. \qquad (17)
\end{aligned}
$$

We notice that $1-\varepsilon_3 m\cos^2\theta_k>0$ due to $f(x_k)>f(x^*)$, by geometric-arithmetic mean value formula and Lemma 7 we have

$$\begin{aligned}\prod_{j=1}^{k}(1-\varepsilon_3 m\cos^2\theta_j) &\le \left(\frac{k-k_1-\varepsilon_3 m\sum_{j=1}^{k}\cos^2\theta_j}{k-k_1}\right)^{k-k_1}\le\left[\frac{k-\eta_1 mk(\prod_{j=1}^{k}\cos^2\theta_j)^{\frac{1}{2}}}{k}\right]^k\\ &= (1-c_6)^{k-k_1},\end{aligned}$$

substitute it into (17) we get,

$$f(x_{k+1})-f(x^*)\le(1-c_6)^{k-k_1}(f(x_{k_1+1})-f(x^*)),$$

where $1-c_6>0$ due to $f(x_k)>f(x^*)$.

Using Lemma 6,

$$\begin{aligned}\frac{m}{2}\|x_{k+1}-x^*\|^2 &\le (1-c_6)^{k-k_1}(f(x_{k_1+1})-f(x^*))\\ \|x_{k+1}-x^*\| &\le \sqrt{\frac{2}{m}(1-c_6)^{k-k_1}(f(x_{k_1+1})-f(x^*))}\\ \sum_{k=1}^{\infty}\|x_{k+1}-x^*\| &< \infty.\end{aligned}$$

Lemma 9 *Let $\{B_k\}$ is generated by the BFGS formula (1), where B_0 is symmetric and positive definite, and where $y_k^T S_k>0$ for all k. Furthermore assume that $(\bar{H})$ holds. Then*

$$\sum_{k=1}^{k}\|x_k-x^*\|<\infty\Longrightarrow\lim_{k\to\infty}\frac{\|(B_k-G(x^*))S_k\|}{\|S_k\|}=0.$$

Proof See Theorem 3.2 in [7].

From Theorem 3 and Lemma 9 we can easily see that the following Theorems is true.

Theorem 4 *Let x_0 be a starting point for which f satisfies the assumption (H'), and assume that $(\bar{H})$ holds. Then for any positive definite B_0, the BFGS algorithm (1), with the GW linesearch method at each step, gives the Dennis-More condition*

$$\lim_{k\to\infty}\frac{\|(B_k-G(x^*))S_k\|}{\|S_k\|}=0.$$

Dennis and More[5] show that when $\frac{\|(B_k-G(x^*))S_k\|}{\|S_k\|}$ and $\|S_k\|$ are sufficiently small then the steplength $\lambda_k=1$ satisfies the Wolfe conditions, therefore, $\lambda_k=1$ must satisfy the GW linesearch rule. From Theorem 4 and from the well-known characterization result ol Dennis and More[6], we conclude that the rate of convergence of the BFGS algorithm with the GW linesearch method is Q-superlinear if the unit steplength is always tried first.

References

[1] M.J.D.Powell, *Some global convergence properties of a variable metric algorithm for minimization without exact linesearches* , in Nonlinear Programming, SIAM-AMS Proceedings, Vol.IX., R.W.Cottle and C.E.Lemke, eds., American Mathematical Society, Providence, RI, 1976.

[2] M.J.D.Powell, *Some properties of the variable metric algorithm* , in Numerical Methods for Nonlinear Optimization, F.A.Lootsma, ed., Academic Press, London, 1972.

[3] M.J.D.Powell, *On the convergence of the variable metric algorithm* , Journal of the Institute of Mathematics and its Applications, **7**(1971), 21-36.

[4] J.D.Pearson, *Variable metric methods of minimization* , The Computer Journal, **12**(1969), 171-178.

[5] J.E.Dennis and J.J.More, *Quasi-Newton methods, motivation and theory* , SIAM Rev., **19**(1977), 46-89.

[6] J.E.Dennis and J.J.More, *A characterization of superlinear convergence and its application to quasi-Newton methods* , Math. Comput., **28**(1974), 549-560.

[7] R.H.Byrd and J.Nocedal, *A tool for the analysis of quasi-Newton methods with application to unconstrained minimization* , SIAM J. Numer. Anal., Vol.26, No.3, 1989, 727-739.

带有广义 Wolfe 线搜索的变尺度算法的收敛性

刘光辉
(美国西北大学工业工程与管理学系)

韩继业
(中国科学院应用数学所)

徐中玲
(国家自然科学基金委员会)

摘要

本文提出一类广义 Wolfe 线搜索模型，并且把它与著名的BFGS方法相结合，对于所得到的算法证明了：对于凸函数算法具有全局收敛性和超线性收敛速度．这推广了参考文献[1]中的结果．