IMA Journal of Numerical Analysis (1996) **16**, 155–164

Convergence properties of the Fletcher–Reeves method

Y. H. DAI AND Y. YUAN

State Key Laboratory of Scientific and Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, PO Box 2719, Beijing 100080, People's Republic of China

[Received 14 April 1994 and in revised form 24 March 1995]

This paper investigates the global convergence properties of the Fletcher–Reeves (FR) method for unconstrained optimization. In a simple way, we prove that a kind of inexact line search condition can ensure the convergence of the FR method. Several examples are constructed to show that, if the search conditions are relaxed, the FR method may produce an ascent search direction, which implies that our result cannot be improved.

1. Introduction

Conjugate gradient methods are very useful for minimizing a smooth function of n variables,

$$\min_{x \in \mathbb{R}^n} f(x), \tag{1.1}$$

especially if n is large. They are of the following form

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

$$d_k = \begin{cases} -g_k, & \text{for } k = 1, \\ -g_k + \beta_k d_{k-1}, & \text{for } k \geqslant 2, \end{cases} \tag{1.3}$$

where $g_k = \nabla f(x_k)$, β_k is a scalar and α_k is a stepsize obtained by means of a one-dimensional search. One well-known formula for β_k is called the Fletcher–Reeves (FR) formula (Fletcher & Reeves (1964)) and is given by

$$\beta_k^{\mathrm{FR}} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}. \tag{1.4}$$

Many results about the FR method have been reported, including Al-Baali (1985), Gilbert & Nocedal (1992), Liu *et al* (1993), Nemirovsky & Yudin (1983), Powell (1984), Zoutendijk (1970), etc. We focus our study on its global convergence properties.

Zoutendijk (1970) (see also Powell (1984)) showed that the FR method converges in the sense that

$$\liminf_{k \to \infty} \|g_k\| = 0 \tag{1.5}$$

if the line search is exact. Al-Baali (1985) extended this result to inexact line searches satisfying the strong Wolfe conditions:

$$f(x_k) - f(x_k + \alpha_k d_k) \geqslant -\delta \cdot \alpha_k g_k^T d_k, \tag{1.6}$$

$$|g(x_k + \alpha_k d_k)^T d_k| \leqslant -\sigma g_k^T d_k, \tag{1.7}$$

where $0 < \delta < \sigma < \frac{1}{2}$. Recently, Liu *et al* (1993) proved that the FR method also gives (1.5) if $\sigma = \frac{1}{2}$. Can or cannot the condition $\sigma \leqslant \frac{1}{2}$ be relaxed? This question motivated the present study.

In this paper, we consider the FR method using the more general search conditions of inequality (1.6) and

$$\sigma_1 g_k^T d_k \leqslant g(x_k + \alpha_k d_k)^T d_k \leqslant -\sigma_2 g_k^T d_k, \tag{1.8}$$

where $0 < \delta < \sigma_1 < 1$ and $\sigma_2 > 0$. If we let $\sigma_1 = \sigma_2 = \sigma$, (1.8) reduces to (1.7). The Wolfe line search conditions are (1.6) and the following inequality

$$g(x_k + \alpha_k d_k)^T d_k \geqslant \sigma g_k^T d_k. \tag{1.9}$$

It is easy to see that (1.9) can be viewed as a special case of (1.8) with $\sigma_1 = \sigma$ and $\sigma_2 = +\infty$. We prove that for the general line search conditions (1.6) and (1.8) the FR method gives (1.5) if

$$\sigma_1 + \sigma_2 \leqslant 1. \tag{1.10}$$

It is also shown by example that the condition (1.10) cannot be relaxed. Our examples in Section 3 show that the FR method with line searches (1.6) and (1.8) may produce an ascent direction if (1.10) is violated.

2. Sufficient conditions

We make the following assumptions on the objective function.

Assumptions 2.1 (1) f is bounded below in $\mathbb{R}^n$ and is continuously differentiable in a neighbourhood E of the level set $L = \{x \mid f(x) \leqslant f(x_1)\}$; (2) The gradient $\nabla f(x)$ is Lipschitz continuous in E.

Under these assumptions, for any $x_k \in \mathbb{R}^n$ and $d_k^T g_k < 0$ there exist stepsizes $\alpha_k > 0$ such that line search conditions (1.6)–(1.7) hold (see Fletcher (1987)). If $d_k^T g_k = 0$, we can let $\alpha_k = 0$, which satisfies (1.6)–(1.7). We state a general convergence result as follows. This result was essentially proved by Zoutendijk (1970) and Wolfe (1969, 1971).

Lemma 2.2 Let x_1 be a starting point for which Assumptions 2.1 are satisfied. Consider any iteration of the form (1.2), if $d_k^T g_k \leqslant 0$ and $\alpha_k \geqslant 0$ satisfies the Wolfe conditions (1.6) and (1.9). Then

$$\sum_{k \geqslant 1, \|d_k\| \neq 0} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty. \tag{2.1}$$

For the FR method, if $g_k^T d_k = 0$ we take $\alpha_k = 0$ which gives $x_{k+1} = x_k$. Thus it follows that $d_{k+1} = -g_{k+1} + d_k$ which implies

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 + g_{k+1}^T d_k = -\|g_{k+1}\|^2. \tag{2.2}$$

Therefore d_{k+1} is a descent direction unless g_k equals zero.

Now we can establish a global convergence result.

THEOREM 2.3 Let x_1 be a starting point for which Assumptions 2.1 are satisfied. Consider the FR method (1.2)–(1.4). If each search direction d_k satisfies

$$g_k^T d_k \leqslant 0, \tag{2.3}$$

if α_k satisfies line search conditions (1.6) and (1.8) with $\sigma_2 < \infty$, and if $\|g_k\| \neq 0$ for all k then we have that

$$\liminf \|g_k\| = 0. \tag{2.4}$$

Proof Because $\|g_k\| \neq 0$ for all k, it follows from (1.3) and (1.4) that

$$-\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} = 1 - \frac{g_{k+1}^T d_k}{\|g_k\|^2}. \tag{2.5}$$

Define

$$\rho_k = -\frac{g_k^T d_k}{\|g_k\|^2}. \tag{2.6}$$

Condition (1.8), (2.3) and relation (2.5) imply that

$$1 - \sigma_2 \rho_k \leqslant \rho_{k+1} \leqslant 1 + \sigma_1 \rho_k. \tag{2.7}$$

The first part of expression (2.7) is the condition $1 \leqslant \sigma_2 \rho_k + \rho_{k+1}$, so the Cauchy–Schwarz inequality provides the bound:

$$\rho_k^2 + \rho_{k+1}^2 \geqslant (1 + \sigma_2^2)^{-1} = c_1 > 0, \tag{2.8}$$

for every k. Using the second inequality of (2.7) repeatedly and noting that $\rho_1 = 1$ and $\rho_k \geqslant 0$, it is easy to deduce that

$$\rho_k < \frac{1}{1 - \sigma_1} \tag{2.9}$$

for all k.

From (1.3), (1.4), (1.8) and (2.9), we have

$$\begin{aligned} \|d_{k+1}\|^2 &= \|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^4}{\|g_k\|^4} \|d_k\|^2 - 2\frac{\|g_{k+1}\|^2}{\|g_k\|^2} g_{k+1}^T d_k \\ &\leqslant \|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^4}{\|g_k\|^4} \|d_k\|^2 + 2\sigma_1 \rho_k \|g_{k+1}\|^2 \\ &\leqslant \frac{\|g_{k+1}\|^4}{\|g_k\|^4} \|d_k\|^2 + \frac{1 + \sigma_1}{1 - \sigma_1} \|g_{k+1}\|^2. \end{aligned} \tag{2.10}$$

Let $t_k = \|d_k\|^2 / \|g_k\|^4$. Thus we obtain

$$t_{k+1} \leqslant t_k + \frac{1 + \sigma_1}{1 - \sigma_1} \frac{1}{\|g_{k+1}\|^2}. \tag{2.11}$$

We now proceed by contradiction and assume that

$$\liminf \|g_k\| \neq 0. \tag{2.12}$$

From (2.12), there exists a constant $\delta_1 > 0$ such that

$$\|g_k\|^2 \geqslant \delta_1 \tag{2.13}$$

for all k. It follows from (2.11) and (2.13) that

$$t_{k+1} \leqslant t_1 + c_2 k \tag{2.14}$$

where

$$c_2 = \frac{1}{\delta_1} \frac{1+\sigma_1}{1-\sigma_1}.$$

(2.14) implies that

$$\frac{1}{t_{2k}} \geqslant \frac{1}{c_2(2k-1)+t_1} \tag{2.15}$$

$$\frac{1}{t_{2k-1}} \geqslant \frac{1}{c_2(2k-2)+t_1} \geqslant \frac{1}{c_2(2k-1)+t_1}. \tag{2.16}$$

Thus from Lemma 2.2 and relations (2.8) and (2.16),

$$\begin{aligned} \infty > \sum_{k\geqslant 1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} &= \sum_{k\geqslant 1} \frac{\rho_k^2}{t_k} \\ &= \sum_{k\geqslant 1} \left(\frac{\rho_{2k-1}^2}{t_{2k-1}} + \frac{\rho_{2k}^2}{t_{2k}} \right) \geqslant \sum_{k\geqslant 1} \frac{\rho_{2k-1}^2 + \rho_{2k}^2}{c_2(2k-1)+t_1} \\ &\geqslant \sum_{k\geqslant 1} \frac{c_1}{c_2(2k-1)+t_1} = \infty, \end{aligned} \tag{2.17}$$

which gives a contradiction. The contradiction shows that $\liminf \|g_k\| = 0$. □

From the above theorem, we can see that the sufficient descent condition

$$g_k^T d_k \leqslant -c \|g_k\|^2, \tag{2.18}$$

where $c > 0$, is not required on every line search. Instead, given that each direction is not ascendent, (2.8) implies that (2.18) holds for at least one line search of any two consecutive line searches and hence ensures (2.4).

Our theorem shows that, provided the search directions are not ascent, the FR method can continue and a subsequence of $\|g_k\|$ converges to zero. In the following we show that $\sigma_1 + \sigma_2 \leqslant 1$ is a sufficient condition for ensuring (2.3) for all k.

LEMMA 2.4 Let x_1 be a starting point for which Assumptions 2.1 are satisfied. If

$$\sigma_1 + \sigma_2 \leqslant 1, \tag{2.19}$$

the FR method (1.2)–(1.4) with the line search conditions (1.6) and (1.8) gives (2.4).

Proof. First we show that (2.3) is true for all k. (2.3) is true for $k=1$ because

$$g_1^T d_1 = -\|g_1\|_2^2 \leqslant 0.$$

Now we assume that (2.3) holds for all $k=1,\ldots,\bar{k}$, which implies that (2.5)–(2.11) hold for all $k=1,\ldots,\bar{k}$. Thus we have

$$\rho_{\bar{k}+1} \geqslant 1-\sigma_2\rho_{\bar{k}} > 1-\sigma_2\frac{1}{1-\sigma_1} = \frac{1-\sigma_1-\sigma_2}{1-\sigma_1} \geqslant 0 \tag{2.20}$$

This shows that (2.3) holds for $\bar{k}+1$ as well. By induction (2.3) is true for all k. Therefore the lemma follows from Theorem 2.3. □

Noting that the search conditions (1.6)–(1.7) correspond to (1.6) and (1.8) with $\sigma_1=\sigma_2=\sigma$, we clearly have the following corollary.

COROLLARY 2.5 Let x_1 be a starting point for which Assumptions 2.1 are satisfied. Consider the FR method (1.2)–(1.4), where α_k satisfies the search conditions (1.6)–(1.7). Then, if

$$\sigma \leqslant \tfrac{1}{2}, \tag{2.21}$$

the FR method gives (2.4).

The above corollary indicates that if $\sigma \leqslant \frac{1}{2}$, the FR method with search conditions (1.6)–(1.7) is globally convergent in the sense of (1.5). In comparison with the approach used in Liu *et al* (1993), ours is much simpler.

3. Necessary conditions

We know from Lemma 2.4 that (2.19) is sufficient for the FR method to give (1.5) if the search conditions are (1.6) and (1.8). The following theorem shows that (2.19) is also necessary.

THEOREM 3.1 Assume that the search conditions are (1.6) and (1.8). Then, for any numbers δ, σ_1 and σ_2 such that $0<\delta<\sigma_1<1$, $0<\sigma_2<\infty$ and

$$\sigma_1+\sigma_2>1, \tag{3.1}$$

there exists an objective function satisfying Assumptions 2.1 such that the FR method (1.2)–(1.4) will produce an ascent direction.

We will show the validity of the above theorem by presenting three counter examples which are given in Lemmas 3.2, 3.3 and 3.4. Our counter examples are constructed by the following technique.

Condition (3.1) and $\sigma_1 \in (0,1)$ allow us to choose a positive integer N such that

$$1+\sigma_2\sigma_1^N < \sigma_1+\sigma_2. \tag{3.2}$$

In our counter examples, we let

$$g_{k+1}^T d_k = \sigma_1 g_k^T d_k, \qquad \text{for } k = 1,\dots,N; \tag{3.3}$$

$$g_{N+1}^T d_N = -\sigma_2 g_N^T d_N. \tag{3.4}$$

Let ρ_k be defined by (2.6); it follows from (3.3)–(3.4) and (2.5) that

$$\rho_{k+1} = 1 + \sigma_1 \rho_k, \qquad \text{for } k = 1,\dots,N-1; \tag{3.5}$$

$$\rho_{N+1} = 1 - \sigma_2 \rho_N. \tag{3.6}$$

Recalling that $\rho_1 = 1$, we have from (3.5)–(3.6) that

$$\rho_k = \frac{1-\sigma_1^k}{1-\sigma_1}, \qquad \text{for } k = 1, 2, \dots, N; \tag{3.7}$$

$$\rho_{N+1} = 1 - \sigma_2 \frac{1-\sigma_1^N}{1-\sigma_1} = \frac{1-\sigma_1-\sigma_2+\sigma_2\sigma_1^N}{1-\sigma_1}. \tag{3.8}$$

It follows from (3.8) and (3.2) that $\rho_{N+1} < 0$, which means that d_{N+1} is an ascent direction.

All our counter examples are one-dimensional functions. Because the number of iterations is finite, we are able to construct several examples. They are seen in the following lemmas.

LEMMA 3.2 Assume that the search conditions are (1.6) and (1.8), where $0 < \delta < \sigma_1 < 1$ and $0 < \sigma_2 < \infty$. If $\sigma_1 + \sigma_2 > 1$ and $\delta < \sigma_1/2$ then for the quadratic convex function

$$f(x) = \tfrac{1}{2}x^2, \qquad x \in \mathbb{R}^1, \tag{3.9}$$

the FR method may produce an ascent direction.

Proof. Because reducing the value of σ_2 makes the line search conditions more strict, there is no loss of generality if we replace σ_2 by the value $\min\{\sigma_2, 1-\delta-\sigma_1/2\}$. Thus we can assume that $\sigma_1 + \sigma_2 > 1$ and

$$1 - \sigma_2 \geqslant \delta + \sigma_1/2 > 2\delta. \tag{3.10}$$

Because $\sigma_1 + \sigma_2 > 1$, we can select N such that (3.2) holds. Given any $x_1 < 0$, $d_1 = -g_1 = -x_1$, we select the stepsize α_k such that

$$x_k = \sigma_1^{k-1} x_1, \qquad \text{for } k = 1,\dots,N; \tag{3.11}$$

$$x_{N+1} = -\sigma_2 \sigma_1^{N-1} x_1. \tag{3.12}$$

Now we show that the above defined x_k $(k = 1,\dots,N+1)$ satisfies the search conditions (1.6) and (1.8). We claim that

$$d_k = -\frac{1-\sigma_1^k}{1-\sigma_1} g_k, \qquad \text{for } k = 1,\dots,N. \tag{3.13}$$

(3.13) is true for $k = 1$ because $d_1 = -g_1$. Now assume that (3.13) is true for some

$k < N$; we prove it is also true for $k+1$. It follows from (3.11), (1.3), (1.4) and $g_k = x_k$ that

$$d_{k+1} = -g_{k+1} + \sigma_1^2 d_k = -g_{k+1} - \sigma_1^2 \frac{1-\sigma_1^k}{1-\sigma_1} g_k$$
$$= -g_{k+1} - \sigma_1 \frac{1-\sigma_1^k}{1-\sigma_1} g_{k+1}$$
$$= -\frac{1-\sigma_1^{k+1}}{1-\sigma_1} g_{k+1}. \tag{3.14}$$

By induction, we have shown that (3.13) is true. From (3.11), the search direction d_k is descent for all $k \leqslant N$. Therefore, due to (3.11)–(3.12), (3.10) and $g_k = x_k$, we have

$$f(x_k) - f(x_{k+1}) = -\tfrac{1}{2}(x_k + x_{k+1})(x_{k+1} - x_k) = -\frac{1}{2}\frac{x_k + x_{k+1}}{g_k} \alpha_k d_k g_k$$
$$= \begin{cases} -\tfrac{1}{2}(1+\sigma_1)\alpha_k g_k^T d_k, & \text{for } k = 1, \dots, N-1; \\ -\tfrac{1}{2}(1-\sigma_2)\alpha_k g_k^T d_k, & \text{for } k = N \end{cases}$$
$$\geqslant -\tfrac{1}{2}(1-\sigma_2)\alpha_k g_k^T d_k > -\delta \alpha_k g_k^T d_k, \tag{3.15}$$

for $k = 1, \dots, N$, and

$$g_{k+1} d_k = \sigma_1 g_k d_k, \qquad \text{for } k = 1, \dots, N-1; \tag{3.16}$$

$$g_{N+1} d_N = -\sigma_2 g_N d_N. \tag{3.17}$$

(3.15)–(3.17) indicate that (1.6) and (1.8) hold for $k = 1, \dots, N$.

It follows from (3.12), (1.3), (1.4) and $g_k = x_k$ that

$$d_{N+1} = -g_{N+1} + \sigma_2^2 d_N = -g_{N+1} - \sigma_2^2 \frac{1-\sigma_1^N}{1-\sigma_1} g_N$$
$$= -g_{N+1} + \sigma_2 \frac{1-\sigma_1^N}{1-\sigma_1} g_{N+1}$$
$$= -\frac{1-\sigma_1-\sigma_2+\sigma_2\sigma_1^N}{1-\sigma_1} g_{N+1}. \tag{3.18}$$

The above relation and (3.2) imply that d_{N+1} is an ascent direction. □

By this lemma, we see that a very simple function satisfies (3.5)–(3.6) if the stepsize α_k is appropriately selected. However, it is usual to try the unit initial stepsize while making line searches. The following example will allow $\alpha_k = 1$ for all k.

LEMMA 3.3 For any numbers δ, σ_1 and σ_2 such that $0 < \delta < \sigma_1 < 1$, $0 < \sigma_2 < \infty$ and $\sigma_1 + \sigma_2 > 1$, there exists an objective function satisfying Assumptions 2.1

such that the FR method with unit stepsizes satisfies line search conditions (1.6) and (1.8) and generates an ascent direction.

Proof. Select N such that (3.2) holds as before. Given x_1, f_1 and $g_1 < 0$, we define a $(2N+1)$-order polynomial $\varphi(x)$, $x \in \mathbb{R}$, such that

$$\varphi(x_k) = f_k, \qquad \nabla\varphi(x_k) = g_k, \qquad \text{for } k = 1, \ldots, N+1, \tag{3.19}$$

where

$$x_{k+1} = x_k + d_k, \qquad \text{for } k = 1, \ldots, N, \tag{3.20}$$

$$g_{k+1} = \begin{cases} \sigma_1 g_k, & \text{for } k = 1, \ldots, N-1; \\ -\sigma_2 g_k, & \text{for } k = N, \end{cases} \tag{3.21}$$

$$f_{k+1} = f_k + \delta g_k d_k, \qquad \text{for } k = 1, \ldots, N, \tag{3.22}$$

$$d_k = \begin{cases} -\sigma_1^{k-1} \dfrac{1-\sigma_1^k}{1-\sigma_1} g_1, & \text{for } k = 1, \ldots, N; \\ \sigma_2 \sigma_1^{N-1} \dfrac{1-\sigma_1-\sigma_2+\sigma_2\sigma_1^N}{1-\sigma_1} g_1, & \text{for } k = N+1. \end{cases} \tag{3.23}$$

Noting that $d_k > 0$ for $k = 1, \ldots, N$, we have

$$x_1 < x_2 < \cdots < x_{N+1}, \tag{3.24}$$

which means that $\varphi(x)$ is well defined. We see that the following function $\bar{\varphi}(x)$ satisfies Assumptions 2.1:

$$\bar{\varphi}(x) = \begin{cases} f_1 + g_1(x - x_1), & \text{for } x < x_1; \\ \varphi(x), & \text{for } x \in [x_1, x_{N+1}]; \\ f_{N+1} + g_{N+1}(x - x_{N+1}), & \text{for } x > x_{N+1} \end{cases} \tag{3.25}$$

Now we use the FR method with the given line search and the starting point x_1 to minimize $\bar{\varphi}(x)$. The choices (3.20)–(3.22) imply that, for all k, $\alpha_k = 1$ satisfies the search conditions (1.6) and (1.8). Further, as in Lemma 3.2, d_k is just given as (3.23). Therefore the FR method will calculate the point x_{N+1}. However, from (3.2), (3.21), (3.23) and $g_1 < 0$, we have

$$g_{N+1} > 0 \qquad \text{and} \qquad d_{N+1} > 0, \tag{3.26}$$

which implies d_{N+1} is an ascent direction. So $\bar{\varphi}(x)$ is the required function. □

It is easy to generalize the stepsize of the above example. Specifically, it is sufficient to replace (3.20) by

$$x_{k+1} = x_k + \bar{\alpha}_k d_k, \tag{3.27}$$

if the number $\bar{\alpha}_k$ is used as the initial stepsize on the kth line search.

Note that g_k in (3.21) is monotonically increasing with k. We make use of this and obtain a convex counter-example for unit initial stepsize as follows.

LEMMA 3.4 For any numbers δ, σ_1 and σ_2 such that $0 < \delta < \sigma_1 < 1$, $0 < \sigma_2 < \infty$ and $\sigma_1 + \sigma_2 > 1$, there exists a convex objective function satisfying Assumptions

2.1 such that the FR method with unit stepsizes satisfies line search conditions (1.6) and (1.8) and generates an ascent direction.

Proof. We assume, without loss of generality, that $\sigma_2 < 1$. Because $\sigma_1 + \sigma_2 > 1$, we select N such that (3.2) holds. Since $\sigma_1 < 1$, there exists an integer $M > 2$ such that

$$M(1-\sigma_1) > (1+\sigma_2). \tag{3.28}$$

The above inequality implies that the numbers μ and ν that are defined by

$$\mu + \nu = 1 + \sigma_2 \tag{3.29}$$

$$\frac{\mu}{2} + \frac{\nu}{M} = 1 - \sigma_1 \tag{3.30}$$

are positive. Pick any x_1, f_1 and $g_1 < 0$. We now define a gradient function:

$$g(x) = \begin{cases} g_1, & \text{for } x \in (-\infty, x_1); \\ g_k + \dfrac{x - x_k}{x_{k+1} - x_k}(g_{k+1} - g_k), & \text{for } x \in [x_k, x_{k+1}),\ k = 1, \ldots, N-1; \\ g_N(1 - t\mu - t^{M-1}\nu), & \text{for } x \in [x_N, x_{N+1}); \\ g_{N+1}, & \text{for } x \in [x_{N+1}, \infty), \end{cases} \tag{3.31}$$

where $t = (x - x_N)/(x_{N+1} - x_N)$, and where x_k, g_k and d_k are given by (3.20), (3.21) and (3.23) respectively. The function

$$f(x) = \int_0^x g(t)\,dt \tag{3.32}$$

is well defined. Because $g(x)$ is increasing, $f(x)$ is convex. In addition, $f(x)$ clearly satisfies Assumptions 2.1. To complete the proof, we need to establish (1.6). In fact, from (3.31), (3.21) and (3.30) we have

$$\begin{aligned} f(x_k) - f(x_{k+1}) &= -\int_{x_k}^{x_{k+1}} g(t)\,dt \\ &= \begin{cases} -\frac{1}{2}(g_k + g_{k+1})d_k, & \text{for } k = 1, \ldots, N-1; \\ -g_N d_N\left(1 - \frac{1}{2}\mu - \dfrac{1}{M}\nu\right) & \text{for } k = N \end{cases} \\ &= \begin{cases} -\frac{1}{2}(1+\sigma_1)g_k^T d_k, & \text{for } k = 1, \ldots, N-1; \\ -\sigma_1 g_N^T d_N, & \text{for } k = N. \end{cases} \end{aligned} \tag{3.33}$$

The above inequality and $0 < \delta < \sigma_1 < 1$ imply that (1.6) holds for all k. Hence this lemma is true. □

Thus we have proved Theorem 3.1 by three examples.

Theorem 3.1 indicates that, if the search conditions are (1.6) and (1.8), the condition (2.19) is also necessary for the FR method to give (1.5). Consequently, if we use the strong Wolfe conditions (1.6) and (1.7) as the search conditions, (2.21) is a requirement. In addition, the Wolfe conditions (1.6) and (1.9) cannot guarantee convergence result (2.4), since they correspond to (1.6) and (1.8) with

$$\sigma_1 = \sigma \qquad \text{and} \qquad \sigma_2 = \infty, \tag{3.34}$$

which implies that (2.19) is not satisfied.

Acknowledgements

We would like to thank the associate editor and two referees for their many valuable comments and constructive suggestions that improved the paper greatly. This work was supported by the state key project 'Scientific and Engineering Computing'.

REFERENCES

AL-BAALI, M. 1985 Descent property and global convergence of the Fletcher–Reeves method with inexact line search. *IMA J. Numer. Anal.* **5,** 121–124.

FLETCHER, R. 1987 *Practical Methods of Optimization,* 2nd edn. New York: Wiley.

FLETCHER, R., & REEVES, C. 1964 Function minimization by conjugate gradients. *Comput. J.* **7,** 149–154.

GILBERT, J. C., & NOCEDAL, J. 1992 Global convergence properties of conjugate gradient methods for optimization. *SIAM J. Optimiz.* **2,** 21–42.

LIN, G. H., HAN, J. Y., & YIN, H. X. 1993 Global convergence of the Fletcher–Reeves algorithm with an inexact line search. *Report,* Institute of Applied Mathematics, Academia Sinica.

NEMIROVSKY, A. S., & YUDIN, D. B. 1983 *Problem Complexity and Method Efficiency.* New York: Wiley.

POWELL, M. J. D. 1984 Nonconvex minimization calculations and the conjugate gradient method. *Lecture Notes in Mathematics,* vol. 1066. Berlin: Springer, pp 122–141.

WOLFE, P. 1969 Convergence conditions for ascent methods. *SIAM Rev.* **11,** 226–235.

WOLFE, P. 1971 Convergence conditions for ascent methods. II: Some corrections. *SIAM Rev.* **13,** 185–188.

ZOUTENDIJK, G. 1970 Nonlinear programming, computational methods. *Integer and Nonlinear Programming* (J. Abadie, ed). Amsterdam: North-Holland, pp 37–86.