VOLUME 16, NOMOR 1, APRIL 2014 ISSN 1410-9883

CAKRAWALA PENDIDIKAN

FORUM KOMUNIKASI ILMIAH DAN EKSPRESI KREATIF ILMU PENDIDIKAN

ISSN 1410-9883

CAKRAWALA PENDIDIKAN

Forum Komunikasi Ilmiah dan Ekspresi Kreatif Ilmu Pendidikan

Terbit dua kali setahun pada bulan April dan Oktober
Terbit pertama kali April 1999

Alamat Penerbit/Redaksi: STKIP PGRI Blitar, Jalan Kalimantan No. 49 Blitar,Telepon (0342)801493. Langganan 2 nomor setahun Rp 50.000,00 ditambah ongkos kirim Rp 5.000,00. Uang langganan dapat dikirim dengan wesel ke alamat Tata Usaha.

CAKRAWALA PENDIDIKAN diterbitkan oleh Sekolah Tinggi Keguruan dan Ilmu Pendidikan PGRI Blitar. **Ketua:** Dra. Hj. Karyati, M.Si, **Pembantu Ketua:** M. Khafid Irsyadi, ST.,S.Pd

Penyunting menerima sumbangan tulisan yang belum pernah diterbitkan dalam media cetak lain. Syarat-syarat, format, dan aturan tata tulis artikel dapat diperiksa pada *Petunjuk bagi Penulis* di sampul belakang-dalam jurnal ini. Naskah yang masuk ditelaah oleh Penyunting dan Mitra Bestari untuk dinilai kelayakannya. Penyunting melakukan penyuntingan atau perubahan pada tulisan yang dimuat tanpa mengubah maksud isinya.

ISSN 1410-9883

CAKRAWALA PENDIDIKAN

Forum Komunikasi Ilmiah dan Ekspresi Kreatif Ilmu Pendidikan

Volume 16, Nomor 1, April 2014

Daftar Isi

Desain sampul: H. Prawoto

Setting dan Cetak: IDC Malang, Telp./Faks. (0341)576 446, email: idc.malindo@gmail.com

Petunjuk Penulisan Cakrawala Pendidikan

1. Naskah belum pernah diterbitkan dalam media cetak lain, diketik spasi rangkap pada kertas kuarto, panjang 10–20 halaman, dan diserahkan paling lambat 3 bulan sebelum penerbitan, dalam bentuk ketikan di atas kertas sebanyak 2 eksemplar dan pada disket komputer IBM PC atau kompatibel. Berkas naskah pada disket komputer diketik dengan menggunakan pengolah kata *Microsoft Word*.
2. Artikel yang dimuat dalam jurnal ini meliputi tulisan tentang hasil penelitian, gagasan konseptual, kajian dan aplikasi teori, tinjauan kepustakaan, dan tinjauan buku baru.
3. Semua karangan ditulis dalam bentuk *esai*, disertai judul subbab *(heading)* masing-masing bagian, kecuali bagian pendahuluan yang disajikan tanpa judul subbab. Peringkat judul sub-bab dinyatakan dengan jenis huruf yang berbeda, letaknya rata tepi kiri halaman, dan tidak menggunakan nomor angka, sebagai berikut.

PERINGKAT 1 (HURUF BESAR SEMUA TEBAL, RATA TEPI KIRI)
Peringkat 2 (Huruf Besar-kecil Tebal, Rata Tepi Kiri)
Peringkat 3 (Huruf Besar-kecil Tebal, Miring, Rata Tepi Kiri)

4. Artikel konseptual meliputi (a) judul, (b) nama penulis, (c) abstrak (50–75 kata), (d) kata kunci, (e) identitas peulis (tanpa gelar akademik), (f) pendahuluan (tanpa judul subbab) yang berisi latar belakang dan tujuan atau ruang lingkup tulisan, (g) isi/pembahasan (terbagi atas sub-subjudul), (h) penutup, dan (i) daftar rujukan. Artikel hasil penelitian disajikan dengan sistematika: (a) judul, (b) nama (-nama) peneliti, (c) abstrak, (d) kata kunci, (e) identitas peneliti (tanpa gelar akademik) (f) pendahuluan (tanpa judul subbab) berisi pembahasan kepustakaan dan tujuan penelitian, (g) metode, (h) hasil, (i) pembahasan, (j) kesimpulan dan saran, dan (k) daftar rujukan.
5. Daftar rujukan disajikan mengikuti tatacara seperti contoh berikut dan diurutkan secara alfabetis dan kronologis.

Anderson, D.W., Vault, V.D., dan Dickson, C.E. 1993. *Problems and Prospects for the Decades Ahead: Competency Based Teacher Education*. Berkeley: McCutchan Publishing Co.

Huda, N. 1991. *Penulisan Laporan Penelitian untuk Jurnal*. Makalah disajikan dalam Lokakarya Penelitian Tingkat Dasar bagi Dosen PTN dan PTS di Malang Angkatan XIV, Pusat Penelitian IKIP MALANG, Malang, 12 Juli.

Prawoto. 1988. *Pengaruh Penginformasian Tujuan Pembelajaran dalam Modul terhadap Hasil Belajar Siswa SD PAMONG Kelas Jauh*. Tesis tidak diterbitkan. Malang: FPS IKIP MALANG..

Russel, T. 1993. An Alternative Conception: Representing Representation. Dalam P.J. Black & A. Lucas (Eds.). *Children's Informal Ideas in Science* (hlm. 62-84). London: Routledge.

Santosa, R. Gunawan. 2002. *Aplikasi Teorema Polya Pada Enumerasi Graf sederhana*, (online), (http://home.unpar.ac.id/integral.pdf.html, diakses 29 Desember 2006)

Sihombing, U. 2003. *Pendataan Pendidikan Berbasis Masyarakat*. http://www.puskur.or.id. Diakses 21 April 2006

Zainuddin, M.H. 1999. Meningkatkan Mutu Profesi Keguruan Indonesia. *Cakrawala Pendidikan*, 1(1):45–52.

6. Naskah diketik dengan memperhatikan aturan tentang penggunaan tanda baca dan ejaan yang dimuat dalam *Pedoman Umum Ejaan Bahasa Indonesia yang Disempurnakan* (Depdikbud, 1987).

GLOBAL CONVERGENCE OF THE MODIFIED FLETCHER-REEVES CONJUGATE GRADIENT METHOD WITH THE MODIFIED ARMIJO-TYPE LINE SEARCH

Dahliatul Hasanah
Jurusan Matematika Universitas Negeri Malang
e-mail: dahlia.jatmiko@gmail.com

Abstract: A conjugate gradient method is well-known for solving large scale unconstrained optimization problem. However, the direction generated by a conjugate gradient method may not be a descent direction. We propose an algorithm utilizing the modified Fletcher-Reeves conjugate gradient method and the modified Armijo-type line search. We prove that the direction generated is a descent direction and the algorithm is globally convergent if the objective function has Lipschitz continuous gradient.

Keywords: Conjugate gradient method, Descent direction, Fletcher-Reeves conjugate gradient method, modified Armijo-type line search, Global convergent.

Abstrak: Metode Conjugate Gradient merupakan metode yang terkenal untuk menyelesaikan masalah optimasi tanpa kendala dalam skala besar. Akan tetapi arah yang dihasilkan metode ini dimungkinkan bukan merupakan arah yang menurun. Peneliti mengusulkan suatu algoritma yang menggabungkan metode Fletcher-Reeves Conjugate Gradient dan pencarian arah Armijo termodifikasi. Dalam artkel ini akan ditunjukkan bahwa arah yang dihasilkan metode gabungan ini merupakan arah yang menurun dan algoritmanya konvergen global jika fungsi objektifnya mempunyai gradient yang memenuhi kondisi Lipschitz dan kontinyu.

Kata kunci: Metode Conjugate Gradient, Arah yang menurun, Metode Fletcher-Reeves Conjugate Gradient, Pencarian garis Armijo termodifikasi, Konvergen global.

INTRODUCTION

The conjugate gradient method is a well-known method for solving large scale unconstrained optimization problems due to its low memory requirements and strong local and global properties. Many types of conjugate gradient methods had been developed to provide a method with more robust and faster optimization algorithm for nonlinear problems. In general, the nonlinear conjugate gradient method is designed to solve the following unconstrained optimization problem:

min $f(\boldsymbol{x})$, $\boldsymbol{x} \in \boldsymbol{R}^n$. **(1.1)**

where f is nonlinear function whose gradient is denoted by g. Let $\boldsymbol{x}_0$ be the initial guess of the solution of (1.1). The iterative formula of the conjugate gradient method is given by $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k$

where α_k is the step length is obtained by carrying out some line search, and the direction $\boldsymbol{d}_k$ is defined by

$$d_k = \begin{cases} -g_k & \text{if } k = 0 \\ -g_k + \beta_k d_{k-1}, & \text{if } k > 0 \end{cases} \quad (1.2)$$

where β_k is a parameter such that when applied to minimize a strictly convex quadratic function, the directions d_k and d_{k-1} are conjugate with respect to the Hessian of the objective function.

Fletcher-Reeves conjugate gradient method formulates parameter β_k which is denoted as β_{k+1}^{FR} as follows:

$$\beta_{k+1}^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2}. \quad (1.3)$$

For $k \geq 1$, the directional derivative of f at x_k along the direction d_k is given by

$$g_k^T d_k = -\|g_k\|^2 + \beta_k^{FR} g_k^T d_{k-1}.$$

It can be seen that if the step length is carried out by the exact line search, then for any $k \geq 0$, we have

$$g_k^T d_k = -\|g_k\|^2 < 0.$$

Zoutendijk (1970) had proved that the Fletcher-Reeves method with exact line search is globally convergent. A direction generated by the Fletcher-Reeves method is also guaranteed to be a descent direction if the line search is carried out by the strong Wolfe-Powell line search. Global convergence of this method has been proved by Al-Baali (1985).

However, if the line search is Armijo-type line search or Wolfe-type line search, the descent property of d_k given in (1.2) is not guaranteed in general. In the case that d_k is not a descent direction, Al-Baali suggested to use the steepest descent direction $-g_k$ instead of d_k given by (1.2).

Birgin and Martinez (2001) proposed three kinds of spectral conjugate gradient methods by combining conjugate gradient method and spectral gradient method. The direction d_k is given by

$$d_k = -\theta_k g_k + \beta_k s_{k-1}$$

where $s_{k-1} = x_k - x_{k-1}$ and parameter β_k is computed in three ways as follows:

$$\beta_k = \frac{(\theta_k y_{k-1} - s_{k-1})^T g_k}{s_{k-1}{}^T y_{k-1}}, \quad \beta_k = \frac{\theta_k y_{k-1}{}^T g_k}{\alpha_{k-1}\theta_{k-1} g_{k-1}{}^T g_{k-1}}, \quad \beta_k = \frac{\theta_k g_k{}^T g_k}{\alpha_{k-1}\theta_{k-1} g_{k-1}{}^T g_{k-1}},$$

θ_k is the spectral gradient which is evaluated by

$$\theta_k = \frac{s_{k-1}{}^T s_{k-1}}{s_{k-1}{}^T y_{k-1}},$$

where $y_k = g_{k+1} - g_k$. In the numerical results, these methods perform very effectively. However, the direction generated by these methods may not be a descent direction.

Motivated by the success of the spectral conjugate gradient method, Liu et.al (2012) proposed a new method by combining the conjugate gradient method and the spectral gradient method. The direction is generated in the same way as in the conjugate gradient method, and β_k and θ_k are specified in the following way

$$\beta_k = \begin{cases} \beta_k^{CD}, & \text{if } d_{k-1}{}^T g_k \leq 0 \\ 0, & \text{else} \end{cases}$$

$$\theta_k = 1 - \frac{g_k{}^T d_{k-1}}{g_{k-1}{}^T d_{k-1}}.$$

Under some mild conditions, the global convergence of this method has been guaranteed with the strong Wolfe line search.

Yu et.al (2010) proposed a spectral conjugate gradient method for impulse noise removal. The search direction generated by this method is guaranteed to be descent direction. Moreover, under the strong Wolfe line search, this method is globally convergent.

The global convergence of the Fletcher-Reeves and the Polak-Ribiere-Polyak methods with Armijo inexact line search has been discussed in a systematic way. For finding an effective and efficient step length, considerable researches have been made provided an iterate x_k and a descent direction d_k. There are at least four types of inexact line search procedures. The Armijo-type line search is one of several inexact line search procedures which guarantees a sufficient degree of accuracy to ensure the algorithm convergence.

The Armijo-type line search is finding $\alpha_k = \rho^{j_k}$ such that j_k is the smallest non-negative integer *j* satisfying

$$f(x_k + \rho^j d_k) \leq f(x_k) + c_0 \rho^j g_k^T d_k,$$

where $\rho \in (0,1)$ and $c_0 \in (0, \frac{1}{2})$. The iterative scheme in the Armijo-type line search is often referred to as backtracking. This tends to make finding the step length vary in the predictable manner.

THE MODIFIED FLETCHER-REEVES CONJUGATE GRADIENT METHOD

Zhang et.al (2006) proposed a conjugate gradient method by modifying a Fletcher-Reeves conjugate gradient method to be similar with a spectral conjugate gradient method but with different parameters θ_k and β_k^{FR}. This modification to the Fletcher-Reeves conjugate gradient method ensures that the direction generated is always a descent direction.

In the modified Fletcher-Reeves method, the direction is defined by the following way

$$d_k = \begin{cases} -g_k & \text{if } k = 0, \\ -\theta_k g_k + \beta_k^{FR} d_{k-1}, & \text{if } k > 0, \end{cases} \quad (2.1)$$

where β_k^{FR} is described by (1.3) and

$$\theta_k = \frac{d_{k-1}^T y_{k-1}}{\|g_{k-1}\|^2}. \quad (2.2)$$

From (1.3), (2.1), and (2.2), we have

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

By induction, we can get

$$d_k^T g_k = -\|g_k\|^2.$$

It is clear that if the exact line search is used to determine the step length, then $g_k^T d_{k-1} = 0$. Thus,

$$\theta_k = \frac{d_{k-1}^T g_k - d_{k-1}^T g_{k-1}}{\|g_{k-1}\|^2} = -\frac{d_{k-1}^T g_{k-1}}{\|g_{k-1}\|^2} = 1$$

In this case, the modified Fletcher-Reeves method reduces to the standard Fletcher-Reeves method.

THE MODIFIED ARMIJO-TYPE LINE SEARCH

Modification of Armijo line search is based on using the function

$$f_k = f + \frac{1}{2}(x - x_k)^T B_k (x - x_k)$$

where B_k is a simple symmetric and positive definite matrix proposed by Wei et.al (2000). This modified line search can be applied by using μI for $\mu > 0$ instead of B_k for convenience.

Wei, et.al (2008) proposed the modified Armijo-type line search in the following way.

Let $c_0 \in (0, \frac{1}{2})$, $\rho \in (0,1)$, $c \in (0,1)$, $\mu > 0$, and $\varphi_k > 0$ be given. Let we denote $\alpha_k = \rho^{j_k} \varphi_k$ The Armijo-type line search is to find α_k where j_k is the smallest nonnegative integer *j* such that

$$f(x_k + \rho^j \varphi_k d_k) \leq f(x_k) + c_0 \rho^j \varphi_k g_k^T d_k - \frac{\mu}{2}(\rho^j \varphi_k)^2 \|d_k\|^2 \quad (3.1)$$

and

$$g(x_k + \rho^j \varphi_k d_k)^T Q_k(j) \leq -c\|g(x_k + \rho^j \varphi_k d_k)\|^2 \quad (3.2)$$

where $Q_k(j)$ is defined as

$$Q_k(j) = -g(x_k + \rho^j \varphi_k d_k) + \frac{g(x_k + \rho^j \varphi_k d_k)^T \left(g(x_k + \rho^j \varphi_k d_k)^T - g_k\right)}{\|g_k\|^2} d_k. \quad (3.3)$$

Parameter φ_k plays an important rule for improving the initial step length. It has been proved that if the direction generated is a descent direction, then it is guaranteed that there exists a nonnegative integer *j* satisfying the Armijo-type line search.

Wei, et.al (2008) also introduced a reasonable choice for selecting φ_k based on the quadratic model $q_k(t)$ which is a Taylor series of order two of the function $f(x_k + td_k)$ around $t = 0$. If $\epsilon > 0$ is sufficiently small, introduce z_k by $z_k = \frac{g(x_k + \epsilon d_k) - g_k}{\epsilon}$. Let $\eta > 0$ be a very small real number, then we select φ_k in the following way

$$\varphi_k = \begin{cases} \frac{-g_k^T d_k}{d_k^T z_k}, & \text{if } \frac{-g_k^T d_k}{d_k^T z_k} \geq \eta, \\ 1, & \text{else,} \end{cases}$$

It is reported that this choice works quite well by selecting $\epsilon = 1e - 8$ and $\eta = 1e - 10$.

ALGORITHM

The following algorithm is the algorithm of the modified Fletcher-Reeves conjugate gradient method in which its step length is chosen by the Armijo-type line search.

Algorithm MFR-MA

Step 1 Given constants $c_0 \in (0, \frac{1}{2})$, $\rho \in (0,1)$, $\mu > 0$, $\varepsilon > 0$, $\eta > 0$;

Step 2 Take a starting point x_0 and let *k* = 0;

Step 3 Set $d_0 = -g_0$

Step 4 While $\|g_0\| > \varepsilon$ do

- Find α_k such that $f(x_k + \alpha_k d_k) \leq f(x_k) + c_0 \alpha_k g_k^T d_k - \frac{\mu}{2}(\alpha_k)^2 \|d_k\|^2$

○ Compute $z_k = \frac{g(x_k + \varepsilon d_k) - g_k}{\varepsilon}$;

○ Compute $\frac{-g_k^T d_k}{d_k^T z_k}$;

○ If $\frac{-g_k^T d_k}{d_k^T z_k} \geq \eta$

$\varphi := \frac{-g_k^T d_k}{d_k^T z_k}$

else

$\varphi := 1$

end

○ Let $\alpha = \varphi$;

○ Set *j* = 0;

○ Take $c_0 \in (0, \frac{1}{2})$, $\rho \in (0,1)$;

○ While $f(x_k + \alpha d_k) > f(x_k) + c_0 \alpha g_k^T d_k - \frac{\mu}{2}(\alpha)^2 \|d_k\|^2$

$\alpha = \rho * \alpha;$

$j = j + 1;$

end

○ Output α

- Let the next iterate $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k$;
- Evaluate $\boldsymbol{g}_{k+1} = \nabla f(\boldsymbol{x}_{k+1})$;
- Compute the spectral gradient θ_{k+1} by $\theta_{k+1} = \frac{\boldsymbol{d}_k^T(\boldsymbol{g}_{k+1} - \boldsymbol{g}_k)}{\|\boldsymbol{g}_k\|^2}$;
- Compute β_{k+1}^{FR} by $\beta_{k+1}^{FR} = \frac{\|\boldsymbol{g}_{k+1}\|^2}{\|\boldsymbol{g}_k\|^2}$;
- Generate $\boldsymbol{d}_{k+1}$ by $\boldsymbol{d}_{k+1} = -\theta_{k+1}\boldsymbol{g}_{k+1} + \beta_{k+1}^{FR}\boldsymbol{d}_k$;
- Let $k = k + 1$;
- Output k and $\boldsymbol{x}_k$

GLOBAL CONVERGENCE

Global convergence of the Algorithm MFR-MA will be proved in this section under the following assumption.

Assumption MFR-MA

1) The level set $\Omega = \{\boldsymbol{x} \in \boldsymbol{R}^n : f(\boldsymbol{x}) \leq f(\boldsymbol{x}_0)\}$ is bounded where $\boldsymbol{x}_0$ is the starting point.
2) The gradient of the objective function $\boldsymbol{g}$ satisfies the Lipschitz condition, i.e. there exists $L > 0$ such that for all $\boldsymbol{x}, \boldsymbol{y} \in \Omega$, $\|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{y})\| \leq L\|\boldsymbol{x} - \boldsymbol{y}\|$.

By the modified Armijo condition, $\{f(\boldsymbol{x}_k)\}$ is a decreasing sequence. This implies that $\{\boldsymbol{x}_k\}$ generated by Algorithm MFR-MA is contained in Ω. It also implies that there exists a constant f^* such that

$$\lim_{k\to\infty} f(\boldsymbol{x}_k) = f^* \tag{5.1}$$

In addition, by the Assumption MFR-MA, there exists a constant $M > 0$ such that for all k satisfy

$$\|\boldsymbol{g}_k\| \leq M. \tag{5.2}$$

Lemma 1. Suppose that Assumption MFR-MA holds. Then $\lim_{k\to\infty} \alpha_k \|\boldsymbol{d}_k\| = 0$ and $\lim_{k\to\infty} \alpha_k \|\boldsymbol{g}_k\|^2 = 0$ (5.3)

Proof : From (5.1) we obtain

$$\sum_{k=0}^{\infty}\left(f(\boldsymbol{x}_k) - f(\boldsymbol{x}_{k+1})\right) = \lim_{N\to\infty}\sum_{k=0}^{N}\left(f(\boldsymbol{x}_k) - f(\boldsymbol{x}_{k+1})\right)$$
$$= \lim_{N\to\infty}\left(f(\boldsymbol{x}_0) - f(\boldsymbol{x}_{N+1})\right)$$
$$= f(\boldsymbol{x}_0) - f^*.$$

Therefore, $\sum_{k=0}^{\infty}\left(f(\boldsymbol{x}_k) - f(\boldsymbol{x}_{k+1})\right) < \infty$.

We have the Modified Armijo condition, that is

$$f(\boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k) - f(\boldsymbol{x}_k) \leq c_0 \alpha_k \boldsymbol{g}_k^T \boldsymbol{d}_k - \frac{\mu}{2}(\alpha_k)^2 \|\boldsymbol{d}_k\|^2$$

then we get

$$\sum_{k=0}^{\infty}\left(-c_0\alpha_k \boldsymbol{g}_k^T \boldsymbol{d}_k + \frac{\mu}{2}(\alpha_k)^2\|\boldsymbol{d}_k\|^2\right) \leq \sum_{k=0}^{\infty}\left(f(\boldsymbol{x}_k) - f(\boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k)\right) < \infty.$$

This implies $\sum_{k=0}^{\infty} -\alpha_k \boldsymbol{g}_k^T \boldsymbol{d}_k < \infty$ and $\sum_{k=0}^{\infty} (\alpha_k)^2 \|\boldsymbol{d}_k\|^2 < \infty$.

We know that $\boldsymbol{g}_k^T \boldsymbol{d}_k = -\|\boldsymbol{g}_k\|^2$ for all $k \geq 0$, and the two series are series of non-negative real numbers. Hence, the series are decreasing. This leads to
$\lim_{k\to\infty} \alpha_k \|\boldsymbol{d}_k\| = 0$ and
$\lim_{k\to\infty} \alpha_k \|\boldsymbol{g}_k\|^2 = 0$.
This property is very important for proving the global convergence of Algorithm MFR-MA.

Lemma 2. Suppose that the Assumption MFR-MA holds. If there exists a constant $\varepsilon > 0$ such that for all $k \geq 0$, $\|\boldsymbol{g}_k\| \geq \varepsilon$ then there exists a constant $M_1 > 0$ such that for all *k* satisfy $\|\boldsymbol{d}_k\| \leq M_2$. (5.4)

Proof : Using (1.2), (1.3), (eq. of spectral), and triangle inequality of the Euclidean norm, we obtain

$$\|\boldsymbol{d}_k\| \leq \|\boldsymbol{\theta}_k\| \|\boldsymbol{g}_k\| + \|\beta_k^{FR}\| \|\boldsymbol{d}_{k-1}\|$$
$$= \frac{\|\boldsymbol{d}_{k-1}\| \|\boldsymbol{g}_k - \boldsymbol{g}_{k-1}\| \|\boldsymbol{g}_k\|}{\|\boldsymbol{g}_{k-1}\|^2} + \frac{\|\boldsymbol{g}_k\|^2 \|\boldsymbol{d}_{k-1}\|}{\|\boldsymbol{g}_{k-1}\|^2}$$

Since $\boldsymbol{g}$ satisfies the Lipschitz condition and $\|\boldsymbol{x}_k - \boldsymbol{x}_{k-1}\| = \alpha_{k-1}\|\boldsymbol{d}_{k-1}\|$, then

$$\|\boldsymbol{d}_k\| \leq \frac{\|\boldsymbol{d}_{k-1}\| \|\boldsymbol{g}_k\|}{\|\boldsymbol{g}_{k-1}\|^2} (L\alpha_{k-1}\|\boldsymbol{d}_{k-1}\| + \|\boldsymbol{g}_k\|)$$
$$\leq \frac{\|\boldsymbol{d}_{k-1}\| \|\boldsymbol{g}_k\|}{\varepsilon^2} (L\alpha_{k-1}\|\boldsymbol{d}_{k-1}\| + \|\boldsymbol{g}_k\|).$$

Since $\|\boldsymbol{g}_k\| \leq M$ for all $k \geq 0$, then

$$\|\boldsymbol{d}_k\| \leq \frac{M}{\varepsilon^2} L\alpha_{k-1}\|\boldsymbol{d}_{k-1}\|^2 + \frac{M^2}{\varepsilon^2}\|\boldsymbol{d}_{k-1}\|.$$

From (5.3), this implies that there exists a constant $q \in (0,1)$ and an integer *K* such that for all $k \geq K$,

$$\frac{ML}{\varepsilon^2}\alpha_{k-1}\|\boldsymbol{d}_{k-1}\| \leq q.$$

For any $k > K$, we have

$$\|\boldsymbol{d}_k\| \leq \left(\frac{ML}{\varepsilon^2}\alpha_{k-1}\|\boldsymbol{d}_{k-1}\|\right)\|\boldsymbol{d}_{k-1}\| + \frac{M^2}{\varepsilon^2}\|\boldsymbol{d}_{k-1}\|$$
$$= q\|\boldsymbol{d}_{k-1}\| + \frac{M^2}{\varepsilon^2}\|\boldsymbol{d}_{k-1}\|$$
$$= \left(q + \frac{M^2}{\varepsilon^2}\right)\|\boldsymbol{d}_{k-1}\|.$$

Therefore, we obtain

$$\|\boldsymbol{d}_k\| \leq \left(q + \frac{M^2}{\varepsilon^2}\right)^{k-K}\|\boldsymbol{d}_K\|.$$

Letting $M_2 = \max\{\|\boldsymbol{d}_1\|, \|\boldsymbol{d}_2\|, \dots, \|\boldsymbol{d}_K\|, \left(q + \frac{M^2}{\varepsilon^2}\right)^{k-K}\|\boldsymbol{d}_K\|\}$ gives $\|\boldsymbol{d}_k\| \leq M_2$ for all *k*.

Lemma 3. Suppose that the Assumption MFR-MA holds. Let $\{\boldsymbol{x}_k\}$ be the sequence of points generated by the Algorithm MFR-MA. Then there exists a constant $M_3 > 0$ such that for all *k*,

$$\alpha_k \geq M_3 \frac{\|\boldsymbol{g}_k\|^2}{\|\boldsymbol{d}_k\|^2}. \qquad (5.5)$$

Proof : We prove (5.5) by considering the following cases:

Case 1. $\alpha_k = 1$. By the modified Fletcher-Reeves method, we have $\boldsymbol{g}_k^T\boldsymbol{d}_k = -\|\boldsymbol{g}_k\|^2$. Then $\|\boldsymbol{g}_k\|\|\boldsymbol{d}_k\| = \|\boldsymbol{g}_k\|^2$.

Hence, the inequality (5.5) is satisfied with $M_3 = 1$.

Case 2. $\alpha_k < 1$. From the definition of α_k, (3.1) and (3.2) cannot simultaneously satisfied for $\frac{\alpha_k}{\rho} = \varphi_k \rho^{j_k - 1}$. If $\frac{\alpha_k}{\rho}$ does not satisfy (3.1) then we have

$$f\left(\boldsymbol{x}_k + \left(\frac{\alpha_k}{\rho}\right)\boldsymbol{d}_k\right) - f(\boldsymbol{x}_k) > c_0\left(\frac{\alpha_k}{\rho}\right)\boldsymbol{g}_k^T\boldsymbol{d}_k - \frac{\mu}{2}\left(\frac{\alpha_k}{\rho}\right)^2\|\boldsymbol{d}_k\|^2.$$

By the Mean Value theorem, there exists $\gamma_k \in (0,1)$ such that

$$f\left(\boldsymbol{x}_k + \left(\frac{\alpha_k}{\rho}\right)\boldsymbol{d}_k\right) - f(\boldsymbol{x}_k) = \left(\frac{\alpha_k}{\rho}\right)\boldsymbol{d}_k^T\boldsymbol{g}\left(\boldsymbol{x}_k + \gamma_k\left(\frac{\alpha_k}{\rho}\right)\boldsymbol{d}_k\right).$$

Then we get

$$\left(\frac{\alpha_k}{\rho}\right)\boldsymbol{d}_k^T\boldsymbol{g}\left(\boldsymbol{x}_k + \gamma_k\left(\frac{\alpha_k}{\rho}\right)\boldsymbol{d}_k\right) > c_0\left(\frac{\alpha_k}{\rho}\right)\boldsymbol{g}_k^T\boldsymbol{d}_k - \frac{\mu}{2}\left(\frac{\alpha_k}{\rho}\right)^2\|\boldsymbol{d}_k\|^2.$$

Dividing by $\frac{\alpha_k}{\rho}$ on both sides gives

$$\boldsymbol{d}_k^T\boldsymbol{g}\left(\boldsymbol{x}_k + \gamma_k\left(\frac{\alpha_k}{\rho}\right)\boldsymbol{d}_k\right) > c_0\boldsymbol{g}_k^T\boldsymbol{d}_k - \frac{\mu}{2}\left(\frac{\alpha_k}{\rho}\right)\|\boldsymbol{d}_k\|^2.$$

Now we subtract $\boldsymbol{d}_k^T\boldsymbol{g}_k$ from both sides, then

$$\boldsymbol{d}_k^T\left(\boldsymbol{g}\left(\boldsymbol{x}_k + \gamma_k\left(\frac{\alpha_k}{\rho}\right)\boldsymbol{d}_k\right) - \boldsymbol{g}_k\right) > (c_0 - 1)\boldsymbol{g}_k^T\boldsymbol{d}_k - \frac{\mu}{2}\left(\frac{\alpha_k}{\rho}\right)\|\boldsymbol{d}_k\|^2.$$

By the Assumption MFR-MA and Cauchy-Schwarz inequality, we have

$$(c_0 - 1)\boldsymbol{g}_k^T\boldsymbol{d}_k - \frac{\mu}{2}\left(\frac{\alpha_k}{\rho}\right)\|\boldsymbol{d}_k\|^2 < L\gamma_k\left(\frac{\alpha_k}{\rho}\right)\|\boldsymbol{d}_k\|^2.$$

Consequently,

$$(c_0 - 1)\boldsymbol{g}_k^T\boldsymbol{d}_k < \left(L\gamma_k + \frac{\mu}{2}\right)\left(\frac{\alpha_k}{\rho}\right)\|\boldsymbol{d}_k\|^2.$$

Hence we obtain

$$\alpha_k > \frac{(1 - c_0)\rho(-\boldsymbol{g}_k^T\boldsymbol{d}_k)}{\left(L\gamma_k + \frac{\mu}{2}\right)\|\boldsymbol{d}_k\|^2}.$$

Since $-\boldsymbol{g}_k^T\boldsymbol{d}_k = \|\boldsymbol{g}_k\|^2$, then

$$\alpha_k > \frac{(1 - c_0)\rho\|\boldsymbol{g}_k\|^2}{\left(L\gamma_k + \frac{\mu}{2}\right)\|\boldsymbol{d}_k\|^2}.$$

If $\frac{\alpha_k}{\rho}$ does not satisfy (3.2), we have

$$-c\left\|\boldsymbol{g}\left(\boldsymbol{x}_k + \left(\frac{\alpha_k}{\rho}\right)\boldsymbol{d}_k\right)\right\|^2 < \boldsymbol{g}\left(\boldsymbol{x}_k + \left(\frac{\alpha_k}{\rho}\right)\boldsymbol{d}_k\right)^T \boldsymbol{Q}_k(j_k - 1).$$

By definition of Q_k we obtain

$$-c\left\|g\left(x_k+\left(\frac{\alpha_k}{\rho}\right)d_k\right)\right\|^2 < -\left\|g\left(x_k+\left(\frac{\alpha_k}{\rho}\right)d_k\right)\right\|^2 + \frac{\left\|g\left(x_k+\left(\frac{\alpha_k}{\rho}\right)d_k\right)\right\|^2}{\|g_k\|^2} g\left(x_k+\left(\frac{\alpha_k}{\rho}\right)d_k\right)^T d_k.$$

Dividing both sides of the inequality by $\left\|g\left(x_k+\left(\frac{\alpha_k}{\rho}\right)d_k\right)\right\|^2$ yields

$$-c < -1 + \frac{g\left(x_k+\left(\frac{\alpha_k}{\rho}\right)d_k\right)^T d_k}{\|g_k\|^2}.$$

Next, we multiply both sides by $\|g_k\|^2$. So we get

$$g\left(x_k+\left(\frac{\alpha_k}{\rho}\right)d_k\right)^T d_k > (1-c)\|g_k\|^2.$$

Subtracting both sides by $g_k^T d_k$ gets

$$\left(g\left(x_k+\left(\frac{\alpha_k}{\rho}\right)d_k\right)-g_k\right)^T d_k > (1-c)\|g_k\|^2 - g_k^T d_k.$$

By the Assumption MFR-MA and the fact that $-g_k^T d_k = \|g_k\|^2$, then

$$\begin{aligned}(2-c)\|g_k\|^2 &< \left(g\left(x_k+\left(\frac{\alpha_k}{\rho}\right)d_k\right)-g_k\right)^T d_k \\ &\le \left\|g\left(x_k+\left(\frac{\alpha_k}{\rho}\right)d_k\right)-g_k\right\| \|d_k\| \\ &\le \left(\frac{\alpha_k}{\rho}\right)L\|d_k\|^2.\end{aligned}$$

Then

$$\alpha_k > \frac{(2-c)\rho\|g_k\|^2}{L\|d_k\|^2}.$$

Now we introduce a constant $M_3 > 0$ by

$$M_3 = \min\left\{1, \frac{(1-c_0)\rho}{\left(L\gamma_k+\frac{\mu}{2}\right)}, \frac{(2-c)\rho}{L}\right\},$$

Then we get $\alpha_k \ge M_3 \frac{\|g_k\|^2}{\|d_k\|^2}$ as required.

Theorem 1. Suppose that the Assumption MFR-MA holds. Let $\{x_k\}$ be the sequence of points generated by the Algorithm MFR-MA. Then

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

Proof : We now prove this theorem by a contradiction. Assume that the conclusion is not true. Then there exist a constant $\varepsilon > 0$ such that for all *k*,

$$\|g_k\| \ge \varepsilon.$$

By lemma 2, we can obtain a constant $M_2 > 0$ such that for all *k*,

$$\|d_k\| \le M_2.$$

On the other hand, by lemma 3, we have

$$\alpha_k \ge M_3 \frac{\|g_k\|^2}{\|d_k\|^2}.$$

If we combine the results of both lemmas, we have

$$||\boldsymbol{g}_k||^2 \leq \frac{M_2}{M_3}\alpha_k ||\boldsymbol{d}_k||.$$

We know that by letting $k \to \infty$, we have $\lim_{k\to\infty} \alpha_k ||\boldsymbol{d}_k|| = 0$.

Then this contradicts with $||\boldsymbol{g}_k|| \geq \varepsilon$ for all *k*. Therefore, there holds $\lim_{k\to\infty} \inf ||\boldsymbol{g}_k|| = 0$.

This completes the proof.

REFERENCES

Al-Baali, A., *Descent property and global convergence of the Fletcher_reeves method with inexact line search, IMA J. Numer. Anal*, 5 (1985), pp. 121-124.

Birgin, E., Martinez, J., *A spectral conjugate gradient method for unconstrained optimization, Appl. Math. Optim*., 43 (2001), pp. 117-128.

Liu, J., Jiang, Y., *Global convergence of a spectral conjugate gradient method for unconstrained optimization, Abstract and Applied Analysis*, 2012 (2012).

Wei, Z., Qi, L., Ito, S., *New step-size rules for optimization problems*, Department of Mathematics and Information Science, Guangxi University, Nanning, Guangxin, P. R. China, October, 2000.

Wei, Z., Li, G., Qi, L., *Global convergence of the Polak-Ribiere-Polyak conjugate gradient method with an Armijo-type inexact line search for nonconvex unconstrained optimization problems, Mathematics of Computation*, 77 (2008), pp. 2178-2193.

Yu, G., Huang, J., Zhou, Y., *A descent spectral conjugate gradient method for impulse noise removal, Applied Mathematics Letters*, 23 (2010), pp. 555-560.

Zhang, L., Zhou, W., Li, D., *Global convergence of a modified Fletcher-Reeves conjugate gradient method with Armijo-type line search,* Numerische Mathematik, 104 (2006), pp. 561-572.

Zoutendijk, G., *Nonlinear programming, computational methods*, Abadie,J. Integer and Nonlinear Programming, (1970), pp. 37-86.