Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2012, Article ID 758287, 12 pages
doi:10.1155/2012/758287

Research Article

Global Convergence of a Spectral Conjugate Gradient Method for Unconstrained Optimization

Jinkui Liu and Youyi Jiang

College of Mathematics and Statistics, Chongqing Three Gorges University, Wanzhou 404000, China

Correspondence should be addressed to Jinkui Liu, liujinkui2006@126.com

Received 22 March 2012; Revised 19 June 2012; Accepted 24 June 2012

Academic Editor: Tianshou Zhou

A new nonlinear spectral conjugate descent method for solving unconstrained optimization problems is proposed on the basis of the CD method and the spectral conjugate gradient method. For any line search, the new method satisfies the sufficient descent condition $g_k^T d_k < -\|g_k\|^2$. Moreover, we prove that the new method is globally convergent under the strong Wolfe line search. The numerical results show that the new method is more effective for the given test problems from the CUTE test problem library (Bongartz et al., 1995) in contrast to the famous CD method, FR method, and PRP method.

1. Introduction

Unconstrained optimization problems have extensive applications, for example, in petroleum exploration, aerospace, transportation, and other domains. However, the amount of necessary calculation also grows exponentially with the increasing scale of the problem. Therefore, it is required to develop new methods to solve the large-scale unconstrained optimization problems. The primary objective of this paper is to study the global convergence properties and practical computational performance of a new nonlinear spectral conjugate gradient method for unconstrained optimization problems without restarts, and with suitable conditions.

Consider the following unconstrained optimization problem

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

where $f : R^n \to R$ is a continuously differentiable function and its gradient is available.

Due to need less computer memory especially, conjugate gradient method is very appealing for solving (1.1) when the number of variable n is large. This method can be described by the following

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

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

where x_k is the current iteration, $\alpha_k > 0$ is the step-size which is determined by some line search, d_k is the search direction, g_k is the gradient of $f(x)$ at the point x_k, and β_k is a scalar which determines the different conjugate gradient methods [1, 2]. There are many well-known formulas for β_k, such as the Fletcher-Reeves (FR) [3], Polak-Ribiere-Polyak (PRP) [4], Hestenes-Stiefel (HS) [5], and conjugate-descent (CD) [6]. The conjugate gradient method is a powerful line search method for solving optimization problems, and it remains very popular for engineers and mathematicians who are interested in solving large-scale problems. This method can avoid, like steepest descent method, the computation and storage of some matrices associated with the Hessian of objective functions.

The original CD method proposed by Fletcher [6], in which β_k is defined by the following

$$\beta_k^{\mathrm{CD}} = -\frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}}, \tag{1.4}$$

where $\|\cdot\|$ denotes the Euclidean norm of vectors. An important property of the CD method is that the method will produce a descent direction under the strong Wolfe line search. In the strong Wolfe line search, the step-size α_k is required to satisfy the following:

$$\begin{gathered} f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k, \\ \left| g(x_k + \alpha_k d_k)^T d_k \right| \leq -\sigma g_k^T d_k, \end{gathered} \tag{1.5}$$

where $0 < \delta < \sigma < 1$. Some good results about the CD method have also been reported in recent years [7–10].

Another popular method to solving problem (1.1) is the spectral gradient method, which was developed originally by Barzilai and Borwein [11] in 1988. Raydan [12] further introduced the spectral gradient method for potentially large-scale unconstrained optimization problems. The main feature of this method is that only gradient directions are used at each line search whereas a nonmonotone strategy guarantees global convergence. What is more, this method outperforms sophisticated conjugate gradient method in many problems. Birgin and Martínez [13] proposed three kinds of spectral conjugate gradient methods. The direction d_k is given by the following way

$$d_k = -\theta_k g_k + \beta_k s_{k-1}, \tag{1.6}$$

where the parameter β_k is computed by the following

$$\beta_k^1 = \frac{(\theta_k y_{k-1} - s_{k-1})^T g_k}{s_{k-1}^T y_{k-1}}, \qquad \beta_k^2 = \frac{\theta_k y_{k-1}^T g_k}{\alpha_{k-1}\theta_{k-1} g_{k-1}^T g_{k-1}},$$
$$\beta_k^3 = \frac{\theta_k g_k^T g_k}{\alpha_{k-1}\theta_{k-1} g_{k-1}^T g_{k-1}}, \tag{1.7}$$

respectively, and θ_k is taken to be the spectral gradient and computed by the following

$$\theta_k = \frac{s_{k-1}^T s_{k-1}}{s_{k-1}^T y_{k-1}}, \tag{1.8}$$

where $y_{k-1} = g_k - g_{k-1}, s_{k-1} = x_k - x_{k-1}$. The numerical results show that these methods are very effective. Unfortunately, they cannot guarantee to generate descent directions. Based on the FR conjugate gradient method, Zhang et al. [14] take modification to the FR method such that the direction generated is always a descent direction. The d_k is defined by the following

$$d_k = \begin{cases} -g_k, & \text{for } k = 1; \\ -\theta_k g_k + \beta_k^{\mathrm{FR}} d_{k-1}, & \text{for } k \geq 2, \end{cases} \tag{1.9}$$

where β_k^{FR} is specified in [3], and $\theta_k = d_{k-1}^T y_{k-1} / \|g_{k-1}\|^2$. They prove that this method can guarantee to generate descent directions and is globally convergent.

In this paper, motivated by success of the spectral gradient method, we propose a new spectral conjugate gradient method by combining the CD method and the spectral gradient method. The direction is given by the following way:

$$d_k = \begin{cases} -g_k, & \text{for } k = 1; \\ -\theta_k g_k + \beta_k d_{k-1}, & \text{for } k \geq 2, \end{cases} \tag{1.10}$$

where β_k is specified by the following

$$\beta_k = \begin{cases} \beta_k^{\mathrm{CD}}, & \text{if } g_k^T d_{k-1} \leq 0, \\ 0, & \text{else,} \end{cases} \tag{1.11}$$

$$\theta_k = 1 - \frac{g_k^T d_{k-1}}{g_{k-1}^T d_{k-1}}. \tag{1.12}$$

Under some mild conditions, we give the global convergence of the new spectral conjugate gradient method with the strong Wolfe line search.

This paper is organized as follows. In Section 2, we propose our algorithm, and global convergence analysis is provided under suitable conditions. Preliminary numerical results are presented in Section 3.

2. Global Convergence Analysis

In order to establish the global convergence of our method, we need the following assumption on objective function, which have often been used in the literatures to analyze the global convergence of nonlinear conjugate gradient method and the spectral conjugate gradient method with inexact line searches.

Assumption 2.1. (i) The level set $\Omega=\{x \mid f(x) \leq f(x_1)\}$ is bounded, where x_1 is the starting point.

(ii) In some neighborhood N of Ω, the objective function is continuously differentiable, and its gradient is Lipschitz continuous, that is, there exists a constant $L > 0$ such that

$$\|g(x) - g(y)\| \leq L\|x - y\|, \quad \text{for } \forall x, y \in N. \tag{2.1}$$

Now we present the new spectral conjugate gradient method as follows.

Algorithm 2.2 (SCD method). *Step 1.* Data: $x_1 \in R^n$, $\varepsilon \geq 0$. Set $d_1 = -g_1$, if $\|g_1\| \leq \varepsilon$, then stop.

Step 2. Compute α_k by some line search.

Step 3. Let $x_{k+1} = x_k + \alpha_k d_k$, $g_{k+1} = g(x_{k+1})$, if $\|g_{k+1}\| \leq \varepsilon$, then stop.

Step 4. Compute β_{k+1} by (1.11), and generate d_{k+1} by (1.10).

Step 5. Set $k = k + 1$, go to Step 2.

The following theorem shows that Algorithm 2.2 possesses the sufficient descent condition for any line search.

Theorem 2.3. *Let the sequences $\{g_k\}$ and $\{d_k\}$ be generated by Algorithm 2.2, and let the step-size α_k be determined by any line search, then*

$$g_k^T d_k < -\|g_k\|^2. \tag{2.2}$$

Proof. We can prove the conclusion by induction. From $\|g_1\|^2 = -g_1^T d_1$, the conclusion (2.2) holds for $k = 1$. Now we assume that the conclusion is true for $k - 1$ and $g_k \neq 0$, that is, $g_{k-1}^T d_{k-1} < 0$. In the following, we need to prove that the conclusion holds for k.

If $g_k^T d_{k-1} \leq 0$, then $\beta_k = \beta_k^{\mathrm{CD}}$. From (1.4), (1.10), and (1.12), we have

$$g_k^T d_k = -\left(1 - \frac{g_k^T d_{k-1}}{g_{k-1}^T d_{k-1}}\right) \cdot \|g_k\|^2 - \frac{\|g_k\|^2}{g_{k-1}^T d_{k-1}} \cdot g_k^T d_{k-1} = -\|g_k\|^2. \tag{2.3}$$

If $g_k^T d_{k-1} > 0$, then $\beta_k = 0$. From (1.10), (1.12), and our assumption: $g_{k-1}^T d_{k-1} < 0$, we have

$$g_k^T d_k = -\left(1 - \frac{g_k^T d_{k-1}}{g_{k-1}^T d_{k-1}}\right) \cdot \|g_k\|^2 = -\|g_k\|^2 + \frac{g_k^T d_{k-1}}{g_{k-1}^T d_{k-1}} \cdot \|g_k\|^2 < -\|g_k\|^2. \tag{2.4}$$

From (2.3) and (2.4), we know that the conclusion (2.2) holds for k. □

Remark 2.4. From (1.4) and (2.3), if $g_k^T d_{k-1} \leq 0$, then we have

$$\beta_k^{\mathrm{CD}} = \frac{g_k^T d_k}{g_{k-1}^T d_{k-1}}. \tag{2.5}$$

Remark 2.5. From (1.11), (2.2), and (2.5), we have $\beta_k \leq g_k^T d_k / g_{k-1}^T d_{k-1}$ for $\forall k \geq 1$.

The conclusion of the following lemma, often called the Zoutendijk condition, is used to prove the global convergence of nonlinear conjugate gradient methods. It was originally given by Zoutendijk [15].

Lemma 2.6. *Suppose that Assumption 2.1 holds. Consider any method* (1.1)-(1.2)*, where d_k satisfies $g_k^T d_k < 0$ for $k \in N^+$ and α_k satisfies the Wolfe line search. Then*

$$\sum_{k \geq 1} \frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2} < +\infty. \tag{2.6}$$

Introduction. The strong Wolfe line search is a special case of the Wolfe line search, so the Lemma 2.6 also holds under the strong Wolfe line search. What is more, we can also use the same method to prove the Zoutendijk condition holding for the spectral conjugate gradient method.

The following theorem establishes the global convergence of the new spectral conjugate gradient method with the strong Wolfe line search for the general functions.

Theorem 2.7. *Suppose that (Assumption 2.1) holds. Let the sequences $\{g_k\}$ and $\{d_k\}$ be generated by Algorithm 2.2, and let the step-size α_k be determined by the strong Wolfe line search* (1.5)*. Then*

$$\lim_{k \to +\infty} \inf \|g_k\| = 0. \tag{2.7}$$

Proof. According to the given conditions, Lemma 2.6 all hold. In the following, we will obtain the conclusion (2.7) by contradiction. Suppose by contradiction that there exists a positive constant $r > 0$ such that

$$\|g_k\| \geq r, \tag{2.8}$$

holds for $\forall k \geq 1$. On the one hand, rewriting (1.11) as follows

$$d_k + \theta_k g_k = \beta_k d_{k-1}, \tag{2.9}$$

and squaring both side of it, we get

$$\|d_k\|^2 = \beta_k^2 \|d_{k-1}\|^2 - 2\theta_k g_k^T d_k - \theta_k^2 \|g_k\|^2. \tag{2.10}$$

Table 1: The numerical results of SCD method, CD method, FR method, and PRP method.

Problem	Dim	SCD method	CD method	FR method	PRP method
Extended Freudenstein and Roth	1000	815/0.58	947/1.20	499/0.56	13/0.02
	3000	1450/3.26	1677/6.35	406/1.50	1542/6.10
	5000	1447/6.22	1700/10.92	537/3.18	1566/10.26
Extended trigonometric	5000	32/0.03	817/1.13	342/0.25	73/0.07
	3000	31/0.08	119/0.22	284/0.56	87/0.19
	5000	36/0.16	310/0.97	346/1.10	38/0.15
Extended Beale	1000	24/0.00	44/0.00	24/0.00	13/0.02
	3000	33/0.02	17/0.02	24/0.00	13/0.00
	5000	30/0.03	43/0.05	20/0.04	19/0.02
Extended penalty	1000	65/0.05	29/0.01	16/0.00	96/0.11
	3000	10/0.00	16/0.00	15/0.00	14/0.00
	5000	11/0.02	22/0.05	35/0.13	12/0.01
Perturbed quadratic	1000	334/0.05	1847/0.26	1088/0.17	407/0.08
	3000	679/0.32	1736/0.74	3191/1.27	705/0.31
	5000	922/0.68	1494/1.02	2830/1.84	1081/0.80
Raydan 1	1000	331/0.12	680/0.25	799/0.25	450/0.18
	3000	601/0.77	1542/1.57	1044/0.94	760/0.94
	5000	1229/2.35	1857/3.04	4192/6.34	1058/2.16
Raydan 2	1000	4/0.00	4/0.00	4/0.00	4/0.00
	3000	4/0.01	4/0.02	4/0.02	4/0.00
	5000	4/0.02	4/0.01	4/0.01	4/0.02
Hager	1000	434/4.23	699/2.53	1450/16.01	183/1.48
	3000	1404/47.23	1591/33.20	2193/74.72	1019/34.11
	5000	3163/171.28	946/39.49	4332/234.59	2583/142.81
Generalized tridiagonal 1	1000	26/0.00	72/0.07	87/0.05	43/0.04
	3000	55/0.14	59/0.22	167/0.62	62/0.20
	5000	89/0.56	61/0.31	38/0.09	34/0.11
Extended tridiagonal 1	1000	34/0.00	15/0.00	86/0.02	14/0.00
	3000	10/0.02	67/0.03	22/0.01	12/0.00
	5000	9/0.02	79/0.06	16/0.01	16/0.00
Extended three expo terms	1000	14/0.01	97/1.71	23/0.08	8/0.00
	3000	15/0.05	86/4.27	49/1.72	8/0.03
	5000	64/4.88	123/11.48	247/23.04	8/0.06
Generalized tridiagonal 2	1000	67/0.01	370/0.07	234/0.05	58/0.00
	3000	47/0.04	770/0.40	281/0.18	54/0.04
	5000	62/0.04	261/0.22	139/0.11	60/0.06
Diagonal 4	1000	7/0.02	4/0.00	6/0.00	4/0.00
	3000	4/0.00	4/0.00	6/0.02	4/0.00
	5000	9/0.00	6/0.02	7/0.00	4/0.00
Diagonal 5	1000	4/0.01	4/0.00	4/0.01	4/0.02
	3000	4/0.02	4/0.04	4/0.01	4/0.01
	5000	4/0.03	4/0.05	4/0.04	4/0.03
Extended Himmelblau	1000	10/0.00	17/0.00	17/0.00	23/0.02
	3000	11/0.00	131/0.06	19/0.01	24/0.02
	5000	12/0.02	103/0.07	20/0.01	24/0.02

Table 1: Continued.

Problem	Dim	SCD method	CD method	FR method	PRP method
Extended PSC1	1000	12/0.02	40/0.28	17/0.04	11/0.00
	3000	11/0.03	51/1.18	91/2.36	18/0.22
	5000	9/0.03	67/2.75	39/1.28	11/0.03
Extended block-diagonal BD1	1000	67/0.02	71/0.02	32/0.02	26/0.01
	3000	69/0.06	24/0.03	29/0.03	31/0.07
	5000	71/0.13	23/0.05	35/0.06	28/0.05
Extended Maratos	1000	981/0.14	227/0.03	617/0.08	59/0.00
	3000	781/0.35	313/0.12	672/0.27	55/0.02
	5000	715/0.55	262/0.17	690/0.45	49/0.05
Extended Cliff	1000	41/0.01	48/0.03	154/0.58	27/0.01
	3000	51/0.09	303/2.22	206/4.23	39/0.06
	5000	17/0.04	61/1.08	113/2.89	20/0.06
Quadratic diagonal perturbed	1000	234/0.03	456/0.08	499/0.08	400/0.06
	3000	978/0.45	2999/1.57	702/0.36	817/0.46
	5000	807/0.64	1157/1.03	1092/0.91	1023/0.89
Extended Wood	1000	661/0.13	68/0.02	96/0.02	107/0.03
	3000	491/0.25	113/0.06	60/0.03	150/0.08
	5000	514/0.42	93/0.08	67/0.05	206/0.17
Quadratic QF1	1000	344/0.05	1120/0.14	949/0.11	363/0.05
	3000	607/0.25	1677/0.64	3576/1.29	731/0.28
	5000	1038/0.72	1606/1.04	2467/1.47	1076/0.72
Extended quadratic enalty QP2	1000	521/0.36	98/0.06	2476/1.17	28/0.01
	3000	813/1.65	191/0.30	251/0.37	40/0.10
	5000	561/1.89	156/0.39	242/0.61	35/0.14
Quadratic QF2	1000	1188/0.19	1469/0.20	1679/0.21	433/0.06
	3000	2949/1.34	1867/0.75	2952/1.13	929/0.40
	5000	4161/3.25	2709/1.78	3840/2.39	1236/0.08
Extended EP1	1000	2/V0	2/0.00	2/0.00	2/0.02
	3000	3/0.00	3/0.00	3/0.00	3/0.00
	5000	3/0.00	3/0.00	3/0.00	3/0.00
Extended tridiagonal 2	1000	41/0.00	137/0.10	78/0.01	36/0.00
	3000	179/0.74	865/1.86	236/0.99	109/0.44
	5000	125/0.76	381/2.75	335/2.47	202/1.58
ARWHEAD	1000	5/0.00	58/0.03	44/0.01	5/0.02
	3000	7/0.00	81/0.11	31/0.03	15/0.04
	5000	13/0.03	50/0.18	54/0.16	31/0.14
NONDIA	1000	16/0.00	47/0.02	50/0.00	10/0.01
	3000	16/0.00	10/0.02	11/0.00	12/0.00
	5000	11/0.02	11/0.01	14/0.00	13/0.01
DQDRTIC	1000	33/0.00	30/0.00	7/0.00	16/0.00
	3000	41/0.02	10/0.00	7/0.00	11/0.00
	5000	35/0.02	34/0.02	7/0.00	9/0.02
DIXMAANA	1000	7/0.02	11/0.03	12/0.03	9/0.03
	3000	8/0.02	11/0.01	13/0.02	9/0.01
	5000	9/0.02	11/0.02	13/0.01	9/0.01

Table 1: Continued.

Problem	Dim	SCD method	CD method	FR method	PRP method
DIXMAANB	1000	11/0.00	12/0.00	12/0.02	12/0.02
	3000	12/0.00	12/0.02	12/0.01	12/0.02
	5000	12/0.03	12/0.01	12/0.03	13/0.03
DIXMAANC	1000	14/0.00	14/0.00	16/0.00	15/0.00
	3000	15/0.02	17/0.02	17/0.01	16/0.01
	5000	15/0.01	16/0.01	17/0.03	16/0.04
DIXMAANE	1000	273/0.09	1021/0.31	579/0.16	246/0.07
	3000	521/0.80	756/0.70	712/0.55	484/0.47
	5000	680/1.44	1066/1.50	1013/1.28	666/1.09
Partial perturbed quadratic PPQ1	1000	371/3.19	707/5.31	664/3.86	450/4.08
	3000	370/30.39	543/45.59	802/49.48	517/44.19
	5000	287/65.78	523/113.13	787/133.55	276/63.87
Broyden tridiagonal	1000	41/0.00	346/0.08	2167/0.34	45/0.02
	3000	88/0.05	429/0.20	497/0.22	86/0.05
	5000	86/0.08	734/0.58	344/0.26	81/0.06
Almost perturbed quadratic	1000	410/0.06	809/0.11	990/0.13	384/0.06
	3000	708/0.31	1436/0.57	2255/0.85	768/0.31
	5000	848/0.61	1770/1.19	2400/1.49	967/0.66
Tridiagonal perturbed quadratic	1000	351/0.06	1028/0.15	1265/0.19	358/0.06
	3000	663/0.33	1330/0.59	2074/0.87	549/0.25
	5000	987/0.95	1667/1.22	3034/2.11	1260/0.97
EDENSCH	1000	162/0.28	261/0.25	128/0.21	93/0.16
	3000	45/0.11	178/0.89	145/0.72	47/0.14
	5000	65/0.42	318/1.32	164/1.38	54/0.25
VARDIM	1000	16/0.02	16/0.01	16/0.02	16/0.00
	3000	19/0.01	19/0.02	19/0.01	19/0.01
	5000	14/0.01	14/0.02	14/0.01	14/0.01
Diagonal 6	1000	4/0.00	4/0.00	4/0.02	4/0.01
	3000	4/0.02	4/0.02	4/0.01	4/0.02
	5000	4/0.02	4/0.00	4/0.02	4/0.01
DIXMAANF	1000	279/0.16	587/0.23	338/0.12	239/0.11
	3000	673/1.44	869/0.83	599/0.46	348/0.36
	5000	742/2.45	692/1.01	2588/6.44	584/0.95
DIXMAANG	1000	293/0.12	673/0.29	368/0.11	257/0.08
	3000	989/1.48	1449/1.35	739/0.62	448/0.47
	5000	507/0.97	1671/3.17	997/1.39	742/1.29
DIXMAANH	1000	272/0.11	632/1.41	1010/0.28	247/0.10
	3000	733/1.43	653/0.64	513/0.41	1771/29.08
	5000	898/3.06	604/1.44	1296/1.74	588/2.11
DIXMAANI	1000	231/0.08	876/0.25	426/0.11	268/0.09
	3000	439/0.50	1261/1.06	628/0.50	542/0.52
	5000	468/0.81	1071/1.59	943/1.22	654/1.05
DIXMAANJ	1000	266/0.13	488/0.14	362/0.11	281/0.09
	3000	768/1.56	720/3.23	516/0.41	385/0.39
	5000	471/0.83	1250/2.15	1589/2.08	547/1.64

Table 1: Continued.

Problem	Dim	SCD method	CD method	FR method	PRP method
DIXMAANK	1000	260/0.11	1036/0.36	500/0.14	278/0.11
	3000	420/0.47	1103/1.13	774/0.62	428/0.42
	5000	620/1.59	1386/2.66	784/1.06	620/1.04
ENGVAL1	1000	32/0.00	242/0.14	268/0.38	53/0.03
	3000	80/0.28	366/1.06	950/4.67	229/1.05
	5000	340/2.56	267/1.64	325/2.37	181/1.30
ENSCHNB	1000	10/0.00	9/0.00	9/0.00	9/0.00
	3000	8/0.01	9/0.01	9/0.02	9/0.01
	5000	9/0.02	9/0.02	9/0.02	8/0.00
ENSCHNF	1000	27/0.00	45/0.00	1206/1.51	23/0.00
	3000	28/0.02	85/0.06	487/1.33	26/0.04
	5000	27/0.04	76/0.07	158/0.23	24/0.03

From the above equation and Remark 2.5, we have

$$\|d_k\|^2 \le \left(\frac{g_k^T d_k}{g_{k-1}^T d_{k-1}}\right)^2 \cdot \|d_{k-1}\|^2 - 2\theta_k g_k^T d_k - \theta_k^2 \|g_k\|^2. \tag{2.11}$$

Dividing the above inequality by $(g_k^T d_k)^2$, we have

$$\begin{aligned}\frac{\|d_k\|^2}{(g_k^T d_k)^2} &\le \frac{\|d_{k-1}\|^2}{(g_{k-1}^T d_{k-1})^2} - \frac{2\theta_k}{g_k^T d_k} - \theta_k^2 \cdot \frac{\|g_k\|^2}{(g_k^T d_k)^2}\\ &= \frac{\|d_{k-1}\|^2}{(g_{k-1}^T d_{k-1})^2} - \left(\theta_k \cdot \frac{\|g_k\|}{g_k^T d_k} + \frac{1}{\|g_k\|}\right)^2 + \frac{1}{\|g_k\|^2}\\ &\le \frac{\|d_{k-1}\|^2}{(g_{k-1}^T d_{k-1})^2} + \frac{1}{\|g_k\|^2}.\end{aligned} \tag{2.12}$$

Using (2.12) recursively and noting that $\|d_1\|^2 = -g_1^T d_1 = \|g_1\|^2$, we get

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} \le \sum_{i=1}^{k} \frac{1}{\|g_i\|^2}. \tag{2.13}$$

Then we get from (2.13) and (2.8) that

$$\frac{(g_k^T d_k)^2}{\|d_k\|^2} \ge \frac{r^2}{k}, \tag{2.14}$$

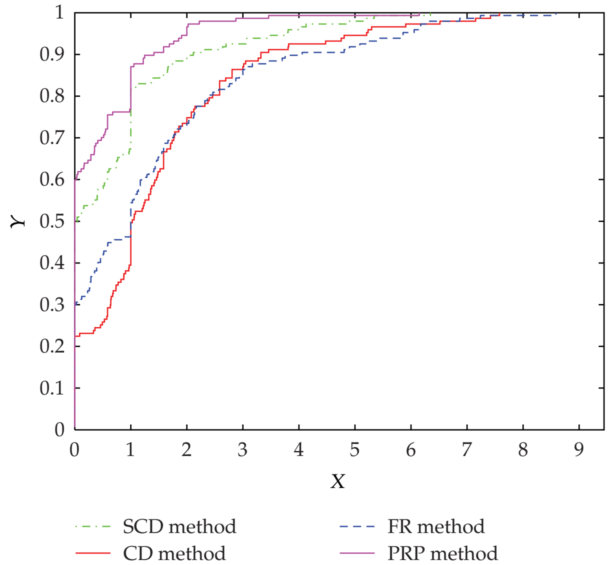


Figure 1: Performance profiles of the given methods with respect to CPU time.

which indicates

$$\sum_{k\geq 1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} = +\infty. \tag{2.15}$$

This contradicts the Zoutendijk condition (2.6). Therefore the conclusion (2.7) holds. □

3. Numerical Experiments

In this section, we report some numerical results. Under the strong Wolfe line search, we compare the performances of the CPU time and the iteration number of the SCD method with that of CD, FR, and PRP methods on the given test problems which come from the CUTE test problem library [16]. The parameters in the strong Wolfe line search are met the following requirements: $\delta = 0.01$ and $\sigma = 0.1$. We stop the iteration if the iteration number exceeds 5000 or the inequity $\|g_k\| \leq 10^{-6}$ is satisfied. All codes were written in FORTRAN 90 and run on a PC with 2.0 GHz CPU processor and 512 MB memory and Windows XP operation system.

In Table 1, the column "Problem" represents the problem's name. "Dim" denotes the problem's dimension. The detailed numerical results are listed in the form NI/CPU, where NI and CPU denote the iteration number and the CPU time in seconds, respectively. From Table 1, some CPU times are zero. This is because that the CPU times are retained two digits in our numerical experiments. In order to compare these methods in the performance of the CPU time, we use a constant $\chi = \min\{\text{CPU}_i(\text{method}) \mid i \in S\}$ instead of $\text{CPU}_{i_0}(\text{method})$ when CPU time of the i_0 problem is zero, where S denotes the set of the test problems whose CPU time is not zero.

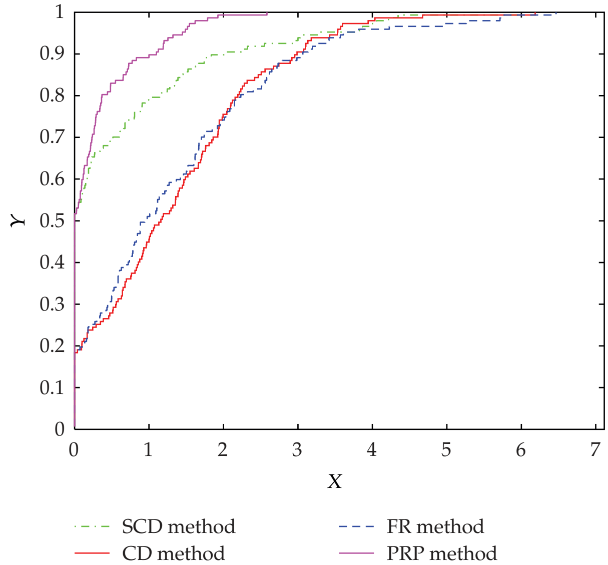


Figure 2: Performance profiles of the given methods with respect to the number of iterations.

In this paper, we adopt the performance profiles by Dolan and Moré [17] to compare the SCD method with CD, FR, and PRP methods. Figure 1 shows the performance profiles with respect to CPU time means that for each method, we plot the fraction P of problems for which the method is within a factor τ of the best time. The left side of the figure gives the percentage of the test problems for which a method is the fastest; the right side gives the percentage of the test problems that are successfully solved by each of the methods. The top curve is the method that solved the most problems in a time that was within a factor τ of the best time. Using the same method, we also test on the iteration number, see Figure 2.

From Figures 1–2, the SCD method performs a little worse than the famous PRP method in the performances of the CPU time and the iteration number. However, the SCD method has absolute potential compared with the famous CD and FR methods in the performances of the CPU time and the iteration number. So the efficiency of the SCD method is encouraging.

Acknowledgments

The authors wish to express their heartfelt thanks to the anonymous referees and the editor for their detailed and helpful suggestions for revising the paper. This work was supported by The Nature Science Foundation of Chongqing Education Committee (KJ121112).

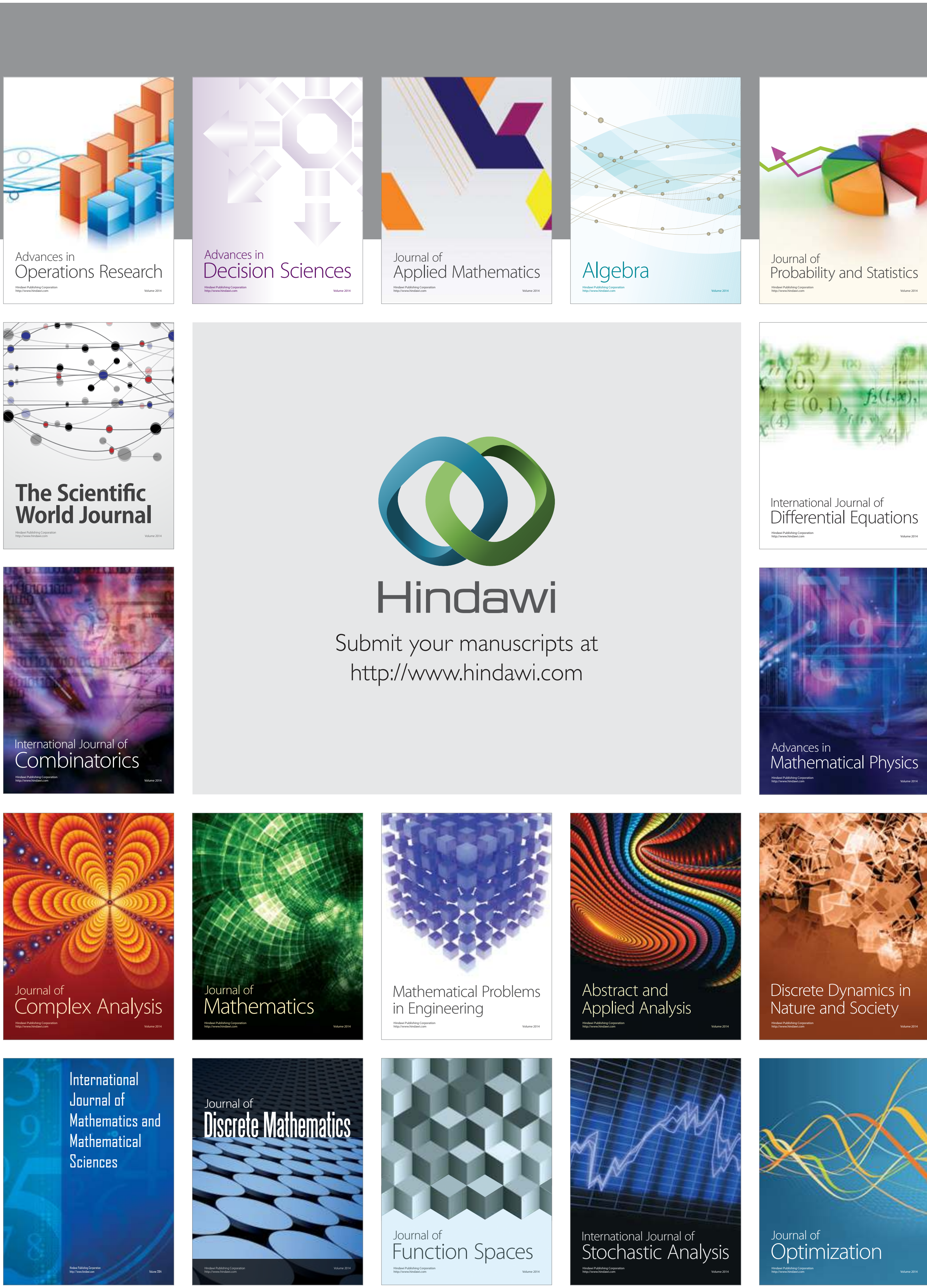
References

[1] G. H. Yu, *Nonlinear self-scaling conjugate gradient methods for large-scale optimization problems [Ph.D. thesis]*, Sun Yat-Sen University, Guangzhou, China, 2007.

[2] G. Yuan and Z. Wei, “New line search methods for unconstrained optimization,” *Journal of the Korean Statistical Society*, vol. 38, no. 1, pp. 29–39, 2009.

[3] R. Fletcher and C. M. Reeves, “Function minimization by conjugate gradients,” *The Computer Journal*, vol. 7, pp. 149–154, 1964.

[4] E. Polak and G. Ribière, "Note sur la convergence de méthodes de directions conjuguées," vol. 3, no. 16, pp. 35–43, 1969.
[5] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, vol. 49, pp. 409–436, 1952.
[6] R. Fletcher, *Practical Methods of Optimization: Unconstrained Optimization*, vol. 1, John Wiley & Sons, New York, NY, USA, 1987.
[7] B. Qu, G.-F. Hu, and X.-C. Zhang, "A global convergence result for the conjugate descent method," *Journal of Dalian University of Technology*, vol. 42, no. 1, pp. 13–16, 2002.
[8] X.-W. Du, L.-Q. Ye, and C.-X. Xu, "Global convergence of a class of unconstrained optimal methods that include the conjugate descent method," *Journal of Applied Mathematics*, vol. 18, no. 2, pp. 120–122, 2001.
[9] C.-Y. Pan and L.-P. Chen, "A class of efficient new descent methods," *Acta Mathematicae Applicatae Sinica*, vol. 30, no. 1, pp. 88–98, 2007.
[10] Y. Dai and Y. Yuan, "Convergence properties of the conjugate descent method," *Advances in Mathematics*, vol. 25, no. 6, pp. 552–562, 1996.
[11] J. Barzilai and J. M. Borwein, "Two-point step size gradient methods," *IMA Journal of Numerical Analysis*, vol. 8, no. 1, pp. 141–148, 1988.
[12] M. Raydan, "The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem," *SIAM Journal on Optimization*, vol. 7, no. 1, pp. 26–33, 1997.
[13] E. G. Birgin and J. M. Martínez, "A spectral conjugate gradient method for unconstrained optimization," *Applied Mathematics and Optimization*, vol. 43, no. 2, pp. 117–128, 2001.
[14] L. Zhang, W. Zhou, and D. Li, "Global convergence of a modified Fletcher-Reeves conjugate gradient method with Armijo-type line search," *Numerische Mathematik*, vol. 104, no. 4, pp. 561–572, 2006.
[15] G. Zoutendijk, "Nonlinear programming, computational methods," in *Integer and Nonlinear Programming*, pp. 37–86, North-Holland, Amsterdam, The Netherlands, 1970.
[16] I. Bongartz, A. R. Conn, N. Gould, and P. L. Toint, "CUTE: constrained and unconstrained testing environment," *ACM Transactions on Mathematical Software*, vol. 21, no. 1, pp. 123–160, 1995.
[17] E. D. Dolan and J. J. Moré, "Benchmarking optimization software with performance profiles," *Mathematical Programming*, vol. 91, no. 2, pp. 201–213, 2002.


Hindawi
Submit your manuscripts at
http://www.hindawi.com
Advances in
Operations Research
Advances in
Decision Sciences
Journal of
Applied Mathematics
Algebra
Journal of
Probability and Statistics
The Scientific
World Journal
International Journal of
Differential Equations
International Journal of
Combinatorics
Advances in
Mathematical Physics
Journal of
Complex Analysis
Journal of
Mathematics
Mathematical Problems
in Engineering
Abstract and
Applied Analysis
Discrete Dynamics in
Nature and Society
International
Journal of
Mathematics and
Mathematical
Sciences
Journal of
Discrete Mathematics
Journal of
Function Spaces
International Journal of
Stochastic Analysis
Journal of
Optimization