Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 814912, 8 pages
http://dx.doi.org/10.1155/2013/814912

Research Article

A Self-Adjusting Spectral Conjugate Gradient Method for Large-Scale Unconstrained Optimization

Yuanying Qiu,[1] Dandan Cui,[2] Wei Xue,[2] and Gaohang Yu[2]

[1] *School of Foreign Languages, Gannan Normal University, Ganzhou 341000, China*
[2] *School of Mathematics and Computer Sciences, Gannan Normal University, Ganzhou 341000, China*

Correspondence should be addressed to Gaohang Yu; maghyu@163.com

Received 21 February 2013; Accepted 17 March 2013

Academic Editor: Guoyin Li

This paper presents a hybrid spectral conjugate gradient method for large-scale unconstrained optimization, which possesses a self-adjusting property. Under the standard Wolfe conditions, its global convergence result is established. Preliminary numerical results are reported on a set of large-scale problems in CUTEr to show the convergence and efficiency of the proposed method.

1. Introduction

Consider the following unconstrained optimization problem:

$$\min \left\{ f(x) \mid x \in \mathfrak{R}^n \right\}, \tag{1}$$

where $f : \mathfrak{R}^n \to \mathfrak{R}$ is a nonlinear smooth function and its gradient is available. Conjugate gradient methods are very efficient for solving (1), especially when the dimension n is large, and have the following iterative form:

$$x_{k+1} = x_k + \alpha_k d_k, \tag{2}$$

where $\alpha_k > 0$ is a steplength obtained by a line search, and d_k is the search direction defined by

$$d_k = \begin{cases} -g_k, & \text{for } k = 1, \\ -g_k + \beta_k d_{k-1}, & \text{for } k \geq 2, \end{cases} \tag{3}$$

where β_k is a scalar and g_k denotes the gradient of f at point x_k.

There are at least six formulas for β_k, which are given below:

$$\beta_k^{\text{FR}} = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}}, \qquad \beta_k^{\text{CD}} = -\frac{g_k^T g_k}{d_{k-1}^T y_{k-1}},$$
$$\beta_k^{\text{DY}} = \frac{g_k^T g_k}{d_{k-1}^T y_{k-1}}, \qquad \beta_k^{\text{PR}} = \frac{g_k^T y_{k-1}}{g_{k-1}^T g_{k-1}}, \tag{4}$$
$$\beta_k^{\text{HS}} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, \qquad \beta_k^{\text{LS}} = -\frac{g_k^T y_{k-1}}{d_{k-1}^T g_{k-1}},$$

where $y_{k-1} = g_k - g_{k-1}$ and $\|\cdot\|$ denotes the Euclidean norm. In the above six methods, HS, PR, and LS methods are especially efficient in real computations, but one may not globally converge for general functions. FR, CD, and DY methods are globally convergent, but they perform much worse. To combine the good numerical performance of HS method and the nice global convergence property of DY method, Dai and Yuan [1] proposed an efficient hybrid formula for β_k which is defined as the following form:

$$\beta_k^{\text{HSDY}} = \max \left\{ 0, \min \left\{ \beta_k^{\text{DY}}, \beta_k^{\text{HS}} \right\} \right\}. \tag{5}$$

Their studies suggested that the HSDY method (5) has the same advantage of avoiding the propensity of short steps as the HS method [1]. They also proved that the HSDY method with the standard wolfe line search produces a descent search direction at each iteration and converges globally. Descent condition may be crucial for the convergence analysis of conjugate gradient methods with inexact line searches [2, 3]. Further, there are some modified conjugate gradient methods [4–7] which possess the sufficiently descent property without any line search condition. Recently, Yu [8] proposed a spectral version of HSDY method:

$$\beta_k^{\text{S-HSDY}} = \max\left\{0, \min\left\{\beta_k^{\text{SDY}}, \beta_k^{\text{SHS}}\right\}\right\}, \tag{6}$$

where

$$\beta_k^{\text{SDY}} = \frac{\|g_k\|^2}{\delta_k y_{k-1}^T d_{k-1}}, \qquad \beta_k^{\text{SHS}} = \frac{g_k^T y_{k-1}}{\delta_k y_{k-1}^T d_{k-1}}, \tag{7}$$

with $\delta_k = y_{k-1}^T s_{k-1}/\|s_{k-1}\|^2$, $s_{k-1} = x_k - x_{k-1}$. The numerical experiments show that this simple preconditioning technique benefits to its performance.

In this paper, based on a new conjugate condition [9], we propose a new hybrid spectral conjugate gradient method with β_k defined by

$$\beta_k^{\text{DS-HSDY}} = \max\left\{0, \min\left\{\beta_k^{\text{DSDY}}, \beta_k^{\text{DSHS}}\right\}\right\}, \tag{8}$$

where

$$\beta_k^{\text{DSDY}} = \frac{\|g_k\|^2}{\delta_k {y_{k-1}^*}^T d_{k-1}}, \qquad \beta_k^{\text{DSHS}} = \frac{g_k^T y_{k-1}^*}{\delta_k {y_{k-1}^*}^T d_{k-1}},$$
$$y_{k-1}^* = y_{k-1} + \frac{\max\{\vartheta_k, 0\}}{\|s_{k-1}\|^2} s_{k-1}, \tag{9}$$
$$\vartheta_k = 2\{f(x_k) - f(x_{k-1})\} + [g(x_k) + g(x_{k-1})]^T s_{k-1}.$$

A full description of DS-HSDY method is formally given as follows.

Algorithm 1 (DS-HSDY conjugate gradient method).

Data. Choose constants $0 < \rho < \sigma < 1$, $\mu > 1$, and $0 \le \epsilon \ll 1$. Given an initial point $x_1 \in R^n$, set $d_1 = -g_1$. Let $k := 1$.

Step 1. If $\|g_k\| \le \varepsilon$, then stop.

Step 2. Determine α_k satisfying the standard Wolfe condition:

$$g(x_k + \alpha_k d_k)^T d_k > \sigma g_k^T d_k, \tag{10}$$

$$f(x_k + \alpha_k d_k) - f(x_k) \le \rho \alpha_k g_k^T d_k. \tag{11}$$

Then update $x_{k+1} = x_k + \alpha_k d_k$.

Step 3. Compute g_{k+1}, δ_{k+1} and $\beta_{k+1}^{\text{DS-HSDY}}$. Then update d_{k+1} such as

$$d_{k+1} = -\frac{1}{\delta_{k+1}} g_{k+1} + \beta_{k+1}^{\text{DS-HSDY}} d_k. \tag{12}$$

Set $k := k + 1$ and go to Step 1.

The rest of the paper is organized as follows. In the next section, we show that the DS-HSDY method possesses a self-adjusting property. In Section 3, we establish its global convergence result under the standard Wolfe line search conditions. Section 4 gives some numerical results on a set of large-scale unconstrained test problems in CUTEr to illustrate the convergence and efficiency of the proposed method. Finally we have a Conclusion section.

2. Self-Adjusting Property

In this section, we prove that the DS-HSDY method possesses a self-adjusting property. To begin with, we assume that

$$g_k \ne 0, \quad \forall k \ge 1, \tag{13}$$

otherwise, a stationary point has been found, and define the two following important quantities:

$$q_k = \frac{\|d_k\|^2}{(g_k^T d_k)^2}, \qquad \gamma_k = -\frac{\delta_k g_k^T d_k}{\|g_k\|^2}. \tag{14}$$

The quantity q_k shows the size of d_k, where γ_k is a quantity showing the descent degree of d_k. In fact, if $\gamma_k > 0$, d_k is a descent direction. Furthermore, if $\gamma_k \ge C$ for some constant $C > 0$, then we have the sufficient descent condition

$$g_k^T d_k \le -C\|g_k\|^2. \tag{15}$$

On the other hand, it follows from (12) that

$$d_k + \frac{1}{\delta_k} g_k = \beta_k^{\text{DS-HSDY}} d_{k-1}. \tag{16}$$

Hence

$$\|d_k\|^2 = \left(\beta_k^{\text{DS-HSDY}}\right)^2 \|d_{k-1}\|^2 - \frac{2}{\delta_k} g_k^T d_k - \frac{1}{\delta_k^2}\|g_k\|^2. \tag{17}$$

Combining $|\beta_k^{\text{DS-HSDY}}| \le |\beta_k^{\text{DSDY}}| \le |\beta_k^{\text{SDY}}|$ with (17) yields

$$\begin{aligned} \|d_k\|^2 &= \left(\beta_k^{\text{DS-HSDY}}\right)^2 \|d_{k-1}\|^2 - \frac{2}{\delta_k} g_k^T d_k - \frac{1}{\delta_k^2}\|g_k\|^2 \\ &\le \left(\beta_k^{\text{SDY}}\right)^2 \|d_{k-1}\|^2 - \frac{2}{\delta_k} g_k^T d_k - \frac{1}{\delta_k^2}\|g_k\|^2. \end{aligned} \tag{18}$$

Dividing both sides of (18) by $(g_k^T d_k)^2$ and using (7), we obtain

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} \le \frac{\|d_{k-1}\|^2}{(g_{k-1}^T d_{k-1})^2} - \frac{2}{\delta_k}\frac{1}{g_k^T d_k} - \frac{1}{\delta_k^2}\frac{\|g_k\|^2}{(g_k^T d_k)^2}. \tag{19}$$

It follows from (19) and the definitions of q_k and γ_k that

$$q_k \le q_{k-1} + \frac{1}{\|g_k\|^2}\frac{2}{\gamma_k} - \frac{1}{\|g_k\|^2}\frac{1}{\gamma_k^2}. \tag{20}$$

Additionally, we assume that there exist positive constants γ and $\overline{\gamma}$ such that

$$0 < \gamma \le \|g_k\| \le \overline{\gamma}, \quad \forall k \ge 1, \tag{21}$$

then we have the following result.

Theorem 2. *Consider the method* (2), (8) *and* (12), *where* d_k *is a descent direction. If* (21) *holds, there exist positive constants* ξ_1, ξ_2, *and* ξ_3 *such that relations*

$$-g_k^T d_k \ge \frac{\xi_1}{\sqrt{k}}, \tag{22}$$

$$\|d_k\|^2 \ge \frac{\xi_2}{k}, \tag{23}$$

$$\gamma_k \ge \frac{\xi_3}{\sqrt{k}} \tag{24}$$

hold for all $k \ge 1$.

Proof. Summing (20) over the iterates and noting that $d_1 = -g_1$, we get

$$q_k \ge \sum_{i=1}^{k} \frac{1}{\|g_i\|^2} \left(\frac{2}{\gamma_i} - \frac{1}{\gamma_i^2} \right). \tag{25}$$

Since $q_k \ge 0$, it follows from (25) that

$$\frac{1}{\|g_i\|^2} \left(\frac{2}{\gamma_i} - \frac{1}{\gamma_i^2} \right) \le \sum_{i=1}^{k-1} \frac{1}{\|g_i\|^2} \left(\frac{2}{\gamma_i} - \frac{1}{\gamma_i^2} \right). \tag{26}$$

Equations (21), (26), and $2/\gamma_i - 1/\gamma_i^2 \le 1$ yield

$$\frac{1}{\gamma_k^2} - \frac{2}{\gamma_k} - \frac{\overline{\gamma}^2}{\gamma^2}(k-1) \le 0. \tag{27}$$

Furthermore, we have

$$\frac{1}{\gamma_k} \le 1 + \sqrt{1 + \frac{\overline{\gamma}^2}{\gamma^2}(k-1)} \le 1 + \frac{\overline{\gamma}^2}{\gamma^2}\sqrt{k} \le \frac{2\overline{\gamma}}{\gamma}\sqrt{k}. \tag{28}$$

Thus (24) holds with $\xi_3 = \gamma / 2\overline{\gamma}$.

Noting that $-g_k^T d_k = \|g_k\|^2 \gamma_k$ and $\|d_k\| \ge \|g_k\| \gamma_k$, it is easy to derive that (22) and (23) hold with $\xi_1 = \xi_3 \gamma^2$ and $\xi_2 = \xi_3^2 \gamma^2$, respectively. Hence the proof is complete. □

Theorem 3. *Consider the method* (2), (8), *and* (12), *where* d_k *is a descent direction. If* (21) *holds, then for any* $p \in (0, 1)$, *there exist constants* ξ_4, ξ_5, *and* $\xi_6 > 0$ *such that, for any* k, *the relations*

$$-g_i^T d_i \ge \xi_4,$$
$$\|d_i\|^2 \ge \xi_5, \tag{29}$$
$$\gamma_i \ge \frac{\xi_6}{\sqrt{k}}$$

hold for at least $[pk]$ *values of* $i \in [1, k]$.

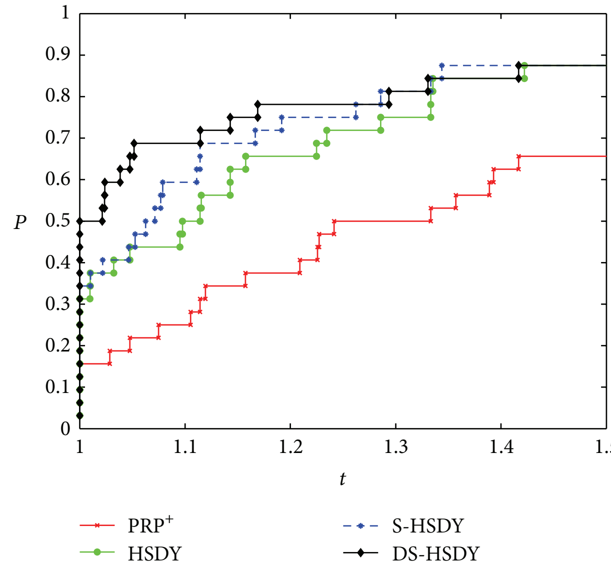

FIGURE 1: Performance profiles for CPU time.

Proof. The proof is similar to the Theorem 2 in [10], so we omit it here. □

Therefore, by Theorems 2 and 3, it was shown that DS-HSDY method possesses a self-adjusting property which is independent of the line search and the function convexity.

3. Global Convergence

Throughout the paper, we assume that the following assumptions hold.

Assumption 1. (1) f is bounded below in the level set $\mathcal{L} = \{x \in R^n : f(x) \le f(x_1)\}$;

(2) in a neighborhood $\mathcal{N}$ of $\mathcal{L}$, f is differentiable and its gradient g is Lipschitz continuous; namely, there exists a constant $L > 0$ such that

$$\|g(x) - g(y)\| \le L\|x - y\|, \quad \forall x, y \in \mathcal{N}. \tag{30}$$

Under Assumption 1 on f, we could get a useful lemma.

Lemma 4. *Suppose that* x_1 *is a starting point for which Assumption 1 holds. Consider any method in the form* (2), *where* d_k *is a descent direction and* α_k *satisfies the weak Wolfe conditions; then one has that*

$$\sum_{k\ge1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty. \tag{31}$$

For DS-HSDY method, one has the following global convergence result.

Theorem 5. *Suppose that* x_1 *is a starting point for which Assumption 1 hold. Consider DS-HSDY method; if* $g_k \ne 0$ *for all* $k \ge 1$, *then one has that*

$$g_k^T d_k < 0 \quad \forall k \ge 1. \tag{32}$$

Table 1: Numerical results for PRP+ method.

Function	n	NI	Nfg	T (0.01 S)	$\|g(x)\|_\infty$
Quadratic QF2	10000	2227	2885	2016	$9.98E-07$
Extended EP1	10000	4	7	3	$6.09E-13$
Extended Tridiagonal 2	10000	39	98	47	$9.29E-07$
ARGLINA	10000	5	15	4	$1.95E-07$
ARWHEAD	10000	7	14	21	$3.10E-07$
BDQRTIC	5000	157	720	526	$1.47E-04$
BDEXP	5000	6	8	7	$1.72E-07$
BRYBND	5000	5	11	1215	$2.60E-07$
COSINE	10000	21	45	39	$9.60E-07$
CRAGGLVY	10000	129	250	444	$5.35E-06$
DIXMAANA	10000	6	12	19	$4.20E-07$
DIXMAANB	10000	8	16	26	$6.72E-07$
DIXMAANC	10000	11	23	38	$3.38E-08$
DIXMAAND	10000	13	29	44	$1.32E-07$
DIXMAANE	5000	558	799	712	$9.96E-07$
DIXMAANF	5000	558	598	525	$8.65E-07$
DIXMAANG	5000	519	784	684	$5.99E-07$
DIXMAANH	5000	379	3488	2469	$8.97E-07$
DIXMAANI	5000	593	854	755	$7.51E-07$
DIXMAANJ	5000	492	751	651	$5.08E-07$
DIXMAANK	5000	653	979	863	$9.61E-07$
DQDRTIC	10000	11	23	19	$9.59E-08$
DQRTIC	10000	33	57	42	$3.44E-07$
EDENSCH	10000	26	90	78	$9.10E-06$
EG2	10000	209	1426	473	$1.10E-03$
ENGVAL1	10000	30	93	21	$1.37E-06$
EXTROSNB	10000	29	63	25	$3.77E-08$
FREUROTH	10000	61	145	81	$2.33E-07$
LIARWHD	10000	25	49	36	$6.54E-09$
NONDIA	10000	9	17	17	$1.03E-09$
NONDQUAR	5000	1786	3258	1752	$7.96E-07$
NONSCOMP	10000	5001	6799	9751	$3.47E-06$

Further, the method converges in the sense that

$$\liminf_{k\to\infty} \|g_k\| = 0. \tag{33}$$

Proof. Since $d_1 = -g_1$, it is obvious that $g_1^T d_1 < 0$. Assume that $g_{k-1}^T d_{k-1} < 0$. By (10) and the definition of the y_k^*, we have $d_{k-1}^T y_{k-1}^* \geq d_{k-1}^T y_{k-1} > 0$, then $\beta_k^{\text{DSDY}} > 0$. In addition, from (8), we have

$$0 \leq \beta_k^{\text{DS-HSDY}} \leq \beta_k^{\text{DSDY}} \leq \beta_k^{\text{SDY}}. \tag{34}$$

Let $\lambda_k = \beta_k^{\text{DS-HSDY}}/\beta_k^{\text{SDY}}$, then we have $0 \leq \lambda_k \leq 1$. By (12) with $k+1$ replaced by k, and multiplying it by g_k, we have

$$g_k^T d_k = \frac{g_{k-1}^T d_{k-1} + (\lambda_k - 1)\, g_k^T d_{k-1}}{\delta_k d_{k-1}^T y_{k-1}} \|g_k\|^2. \tag{35}$$

From this and the formula for β_k^{SDY}, we get

$$\beta_k^{\text{DS-HSDY}} = \lambda_k \beta_k^{\text{SDY}} = \frac{\lambda_k g_k^T d_k}{g_{k-1}^T d_{k-1} + (\lambda_k - 1)\, g_k^T d_{k-1}} = \xi_k \frac{g_k^T d_k}{g_{k-1}^T d_{k-1}}, \tag{36}$$

where

$$\xi_k = \frac{\lambda_k}{1 + (\lambda_k - 1)\, l_{k-1}}, \tag{37}$$

$$l_{k-1} = \frac{g_k^T d_{k-1}}{g_{k-1}^T d_{k-1}}. \tag{38}$$

At the same time, if we define

$$\zeta_k = \frac{1 + (\lambda_k - 1)\, l_{k-1}}{l_{k-1} - 1}, \tag{39}$$

Table 2: Numerical results for HSDY method.

Function	n	NI	Nfg	T (0.01 S)	$\|g(x)\|_\infty$
Quadratic QF2	10000	1593	1902	1876	$7.11E-07$
Extended EP1	10000	4	7	4	$6.09E-13$
Extended Tridiagonal 2	10000	34	55	32	$9.40E-07$
ARGLINA	10000	5	15	4	$1.95E-07$
ARWHEAD	10000	13	58	71	$4.42E-07$
BDQRTIC	5000	171	567	422	$6.31E-04$
BDEXP	5000	6	8	6	$1.72E-07$
BRYBND	5000	5	11	1222	$2.60E-07$
COSINE	10000	21	46	41	$8.02E-07$
CRAGGLVY	10000	109	255	434	$1.45E-06$
DIXMAANA	10000	5	10	16	$5.13E-07$
DIXMAANB	10000	9	18	29	$2.21E-07$
DIXMAANC	10000	10	21	33	$5.42E-07$
DIXMAAND	10000	13	29	45	$1.14E-07$
DIXMAANE	5000	446	541	493	$9.24E-07$
DIXMAANF	5000	389	876	690	$9.60E-07$
DIXMAANG	5000	552	660	602	$9.85E-07$
DIXMAANH	5000	202	5106	3417	$4.05E-04$
DIXMAANI	5000	365	450	409	$9.95E-07$
DIXMAANJ	5000	444	532	484	$4.95E-07$
DIXMAANK	5000	367	452	410	$9.77E-07$
DQDRTIC	10000	8	17	14	$6.35E-07$
DQRTIC	10000	37	68	48	$3.40E-07$
EDENSCH	10000	30	99	85	$7.97E-07$
EG2	10000	305	2811	879	$2.08E-03$
ENGVAL1	10000	30	52	21	$8.62E-07$
EXTROSNB	10000	27	54	21	$6.20E-09$
FREUROTH	10000	143	283	177	$8.01E-07$
LIARWHD	10000	32	62	44	$1.75E-07$
NONDIA	10000	7	14	14	$4.57E-07$
NONDQUAR	5000	2049	3730	2011	$6.68E-07$
NONSCOMP	10000	58	100	98	$5.03E-07$

it follows from (39) that

$$g_k^T d_k = \frac{\zeta_k}{\delta_k}\|g_k\|^2. \tag{40}$$

Then we have by (10), with k replaced by $k-1$, that

$$l_{k-1} \le \sigma. \tag{41}$$

Furthermore, we have

$$1 + (\lambda_k - 1)\, l_{k-1} \ge 1 + \left(-\frac{1-\sigma}{1+\sigma} - 1\right)\sigma = \frac{1-\sigma}{1+\sigma}. \tag{42}$$

The above relation, (40), (41), and the fact that $\sigma < 1$ imply that $g_k^T d_k < 0$. Thus by induction, (32) holds.

We now prove (33) by contradiction and assume that there exists some constant $\gamma > 0$ such that

$$\|g_k\| \ge \gamma \quad \forall k \ge 1. \tag{43}$$

Since $d_k + (1/\delta_k) g_k = \beta_k^{\text{DS-HSDY}} d_{k-1}$, we have that

$$\|d_k\|^2 = \left(\beta_k^{\text{DS-HSDY}}\right)^2 \|d_{k-1}\|^2 - \frac{2}{\delta_k} g_k^T d_k - \frac{1}{\delta_k^2}\|g_k\|^2. \tag{44}$$

Dividing both sides of (44) by $(g_k^T d_k)^2$ and using (36) and (40), we obtain

$$\begin{aligned}\frac{\|d_k\|^2}{(g_k^T d_k)^2} &= \xi_k^2 \frac{\|d_{k-1}\|^2}{(g_{k-1} d_{k-1})^2} - \frac{1}{\|g_k\|^2}\left(\frac{2}{\zeta_k} + \frac{1}{\zeta_k^2}\right) \\ &= \xi_k^2 \frac{\|d_{k-1}\|^2}{(g_{k-1} d_{k-1})^2} + \frac{1}{\|g_k\|^2}\left[1 - \left(1 + \frac{1}{\zeta_k}\right)^2\right].\end{aligned} \tag{45}$$

In addition, since $l_{k-1} < 1$ and $\lambda_k \le 1$, we have that $(1-\lambda_k)(1-l_{k-1}) \ge 0$, or equivalently

$$1 + (\lambda_k - 1)\, l_{k-1} \ge \lambda_k, \tag{46}$$

Table 3: Numerical results for S-HSDY method.

Function	n	NI	Nfg	T (0.01 S)	$\|g(x)\|_\infty$
Quadratic QF2	10000	1582	1941	1836	$6.58E-07$
Extended EP1	10000	4	7	3	$6.09E-13$
Extended Tridiagonal 2	10000	34	55	34	$9.40E-07$
ARGLINA	10000	5	15	3	$1.95E-07$
ARWHEAD	10000	13	58	75	$5.60E-07$
BDQRTIC	5000	111	526	377	$3.39E-04$
BDEXP	5000	6	8	5	$1.72E-07$
BRYBND	5000	5	11	1179	$2.60E-07$
COSINE	10000	21	46	39	$9.72E-07$
CRAGGLVY	10000	103	189	332	$1.94E-06$
DIXMAANA	10000	5	10	15	$5.13E-07$
DIXMAANB	10000	9	18	30	$2.21E-07$
DIXMAANC	10000	10	21	33	$5.42E-07$
DIXMAAND	10000	13	29	43	$1.14E-07$
DIXMAANE	5000	422	514	468	$9.73E-07$
DIXMAANF	5000	310	792	618	$6.77E-07$
DIXMAANG	5000	410	495	449	$9.83E-07$
DIXMAANH	5000	217	6957	4642	$4.13E-04$
DIXMAANI	5000	380	450	417	$9.96E-07$
DIXMAANJ	5000	359	438	402	$9.95E-07$
DIXMAANK	5000	404	485	448	$6.67E-07$
DQDRTIC	10000	8	17	14	$6.35E-07$
DQRTIC	10000	37	68	49	$3.41E-07$
EDENSCH	10000	30	99	84	$1.54E-06$
EG2	10000	242	1731	570	$4.25E-04$
ENGVAL1	10000	29	124	22	$1.78E-06$
EXTROSNB	10000	27	54	22	$3.98E-09$
FREUROTH	10000	214	408	260	$8.08E-07$
LIARWHD	10000	27	54	37	$4.45E-12$
NONDIA	10000	7	14	14	$4.58E-07$
NONDQUAR	5000	1782	3210	1738	$8.99E-07$
NONSCOMP	10000	58	100	100	$5.03E-07$

which with (37) yields

$$\left|\xi_k\right| \leq 1. \tag{47}$$

By (45) and (47), we obtain

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} \leq \frac{\|d_{k-1}\|^2}{(g_{k-1} d_{k-1})^2} + \frac{1}{\|g_k\|^2}. \tag{48}$$

Using (48) recursively and noting that $\|d_1\|^2 = -g_1^T d_1 = \|g_1\|^2$,

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} \leq \sum_{i=1}^{k} \frac{1}{\|g_k\|^2}. \tag{49}$$

Then we get from this and (43) that

$$\frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2} \geq \frac{\lambda^2}{k}, \tag{50}$$

which indicates

$$\sum_{k\geq 1} \frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2} = +\infty. \tag{51}$$

This contradicts the Zoutendijk condition (31). Hence we complete the proof. □

4. Numerical Result

In this section, we compare the performance of DS-HSDY method to PRP$^+$ method [11], HSDY method [1], and S-HSDY method [8]. The test problems are taken from CUTEr (http://hsl.rl.ac.uk/cuter-www/problems.html) with the standard initial points. All codes are written in double precision Fortran and complied with f77 (default compiler settings) on a PC (AMD Athlon XP 2500 + CPU 1.84 GHz). Our line search subroutine computes α_k such that the Wolfe conditions (10) and (11) hold with $\rho = 10^{-4}$ and $\sigma = 0.5$. We

Table 4: Numerical results for DS-HSDY method.

Function	n	NI	Nfg	T (0.01 S)	$\|g(x)\|_\infty$
Quadratic QF2	10000	1623	1978	1783	$9.81E-07$
Extended EP1	10000	4	7	3	$6.09E-13$
Extended Tridiagonal 2	10000	34	55	30	$1.95E-07$
ARGLINA	10000	5	15	3	$5.60E-07$
ARWHEAD	10000	13	58	70	$5.60E-07$
BDQRTIC	5000	165	448	324	$2.93E-03$
BDEXP	5000	6	8	4	$1.72E-07$
BRYBND	5000	5	11	990	$2.60E-07$
COSINE	10000	14	38	28	$5.78E-07$
CRAGGLVY	10000	110	150	266	$8.92E-07$
DIXMAANA	10000	5	10	16	$5.17E-07$
DIXMAANB	10000	9	18	27	$2.21E-07$
DIXMAANC	10000	10	21	31	$5.42E-07$
DIXMAAND	10000	13	29	44	$1.10E-07$
DIXMAANE	5000	410	493	430	$9.57E-07$
DIXMAANF	5000	432	546	469	$3.65E-07$
DIXMAANG	5000	476	582	505	$5.89E-07$
DIXMAANH	5000	442	1204	7792	$4.05E-04$
DIXMAANI	5000	397	467	408	$9.45E-07$
DIXMAANJ	5000	445	594	503	$9.66E-07$
DIXMAANK	5000	403	507	438	$9.05E-07$
DQDRTIC	10000	10	21	17	$1.19E-07$
DQRTIC	10000	35	62	43	$9.73E-07$
EDENSCH	10000	29	87	70	$5.74E-06$
EG2	10000	251	1121	381	$4.02E-03$
ENGVAL1	10000	29	50	19	$4.26E-07$
EXTROSNB	10000	65	122	44	$7.11E-07$
FREUROTH	10000	50	133	67	$1.59E-07$
LIARWHD	10000	47	94	61	$1.16E-08$
NONDIA	10000	7	14	12	$4.60E-07$
NONDQUAR	5000	1831	3262	1665	$9.66E-07$
NONSCOMP	10000	73	126	119	$5.10E-07$

use the condition $\|g(x_k)\|_\infty \leq 10^{-6}$ or $\alpha_k g_k^T d_k < 10^{-20}|f(x_k)|$ as the stopping criterion. The numerical results are presented in Tables 1, 2, 3, and 4 with the form NI/Nfg/T, where we report the dimension of the problem (n), the number of iteration (NI), the number of function evaluations (Nfg), and the CPU time (T) in 0.01 seconds.

Figure 1 shows the performance of these test methods relative to the CPU time, which were evaluated using the profiles of Dolan and Moré [12]. That is, for each method, we plot the fraction P of problems for which the method is within a factor t of the best time. The top curve is the method that solved the most problems in a time that was within a factor t of the best time. Clearly, the left side of the figure gives the percentage of the test problems for which a method is the fastest. As we can see from Figure 1, DS-HSDY method has the best performance which performs better than S-HSDY method, HSDY method, and the well-known PRP$^+$ method.

5. Conclusion

In this paper, we proposed an efficient hybrid spectral conjugate gradient method with self-adjusting property. Under some suitable assumptions, we established the global convergence result for the DS-HSDY method. Numerical results indicated that the proposed method is efficient for large-scale unconstrained optimization problems.

Acknowledgments

This work was partly supported by the National Natural Science Foundation of China (no. 61262026), the JGZX program of Jiangxi Province (20112BCB23027), and the science and technology program of Jiangxi Education Committee (LDJH12088). The authors would also like to thank the editor and an anonymous referees for their comments and suggestions on the first version of the paper, which led to significant improvements of the presentation.

References

[1] Y. H. Dai and Y. Yuan, "An efficient hybrid conjugate gradient method for unconstrained optimization," *Annals of Operations Research*, vol. 103, pp. 33–47, 2001.

[2] Y. H. Dai and Y. Yuan, "A nonlinear conjugate gradient method with a strong global convergence property," *SIAM Journal on Optimization*, vol. 10, no. 1, pp. 177–182, 1999.

[3] Z. X. Wei, G. Y. Li, and L. Q. Qi, "Global convergence of the Polak-Ribière-Polyak conjugate gradient method with an Armijo-type inexact line search for nonconvex unconstrained optimization problems," *Mathematics of Computation*, vol. 77, no. 264, pp. 2173–2193, 2008.

[4] W. W. Hager and H. Zhang, "A new conjugate gradient method with guaranteed descent and an efficient line search," *SIAM Journal on Optimization*, vol. 16, no. 1, pp. 170–192, 2005.

[5] L. Zhang, W. Zhou, and D.-H. Li, "A descent modified Polak-Ribière-Polyak conjugate gradient method and its global convergence," *IMA Journal of Numerical Analysis*, vol. 26, no. 4, pp. 629–640, 2006.

[6] G. Yuan, "Modified nonlinear conjugate gradient methods with sufficient descent property for large-scale optimization problems," *Optimization Letters*, vol. 3, no. 1, pp. 11–21, 2009.

[7] G. Yu, L. Guan, and W. Chen, "Spectral conjugate gradient methods with sufficient descent property for large-scale unconstrained optimization," *Optimization Methods and Software*, vol. 23, no. 2, pp. 275–293, 2008.

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

[9] G. Li, C. Tang, and Z. Wei, "New conjugacy condition and related new conjugate gradient methods for unconstrained optimization," *Journal of Computational and Applied Mathematics*, vol. 202, no. 2, pp. 523–539, 2007.

[10] Y.-H. Dai, "New properties of a nonlinear conjugate gradient method," *Numerische Mathematik*, vol. 89, no. 1, pp. 83–98, 2001.

[11] J. C. Gilbert and J. Nocedal, "Global convergence properties of conjugate gradient methods for optimization," *SIAM Journal on Optimization*, vol. 2, no. 1, pp. 21–42, 1992.

[12] E. D. Dolan and J. J. Moré, "Benchmarking optimization software with performance profiles," *Mathematical Programming*, vol. 91, no. 2, pp. 201–213, 2002.