Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2009, Article ID 875097, 13 pages
doi:10.1155/2009/875097

Research Article

A Truncated Descent HS Conjugate Gradient Method and Its Global Convergence

Wanyou Cheng[1] and Zongguo Zhang[2]

[1] *College of Software, Dongguan University of Technology, Dongguan 523000, China*
[2] *College of Mathematics and Physics, Shandong Institute of Light Industry, Jinan 250353, China*

Correspondence should be addressed to Wanyou Cheng, chengwanyou421@yahoo.com.cn

Received 2 December 2008; Accepted 8 April 2009

Recommended by Ekaterina Pavlovskaia

Recently, Zhang (2006) proposed a three-term modified HS (TTHS) method for unconstrained optimization problems. An attractive property of the TTHS method is that the direction generated by the method is always descent. This property is independent of the line search used. In order to obtain the global convergence of the TTHS method, Zhang proposed a truncated TTHS method. A drawback is that the numerical performance of the truncated TTHS method is not ideal. In this paper, we prove that the TTHS method with standard Armijo line search is globally convergent for uniformly convex problems. Moreover, we propose a new truncated TTHS method. Under suitable conditions, global convergence is obtained for the proposed method. Extensive numerical experiment show that the proposed method is very efficient for the test problems from the CUTE Library.

1. Introduction

Consider the unconstrained optimization problem:

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

where f is continuously differentiable. Conjugate gradient methods are very important methods for solving (1.1), especially if the dimension n is large. The methods are of the form

$$x_{k+1} = x_k + \alpha_k d_k, \quad k = 0, 1, \ldots, \tag{1.2}$$

$$d_k = \begin{cases} -g_k, & \text{if } k = 0, \\ -g_k + \beta_k d_{k-1}, & \text{if } k > 0, \end{cases} \tag{1.3}$$

where g_k denotes the gradient of f at x_k, α_k is the step length obtained by a line search and β_k is a scalar. The strong Wolfe line search is to find a step length α_k such that

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k, \tag{1.4}$$

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

where $\delta \in (0, 1/2)$ and $\sigma \in (\delta, 1)$. In the conjugate gradient methods field, it is also possible to use the Wolfe line search [1, 2], which calculates an α_k satisfying (1.4) and

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k. \tag{1.6}$$

In particular, some conjugate gradient methods admit to use the Armijo line search, namely, the step length α_k can be obtained by letting $\alpha_k = \max\{\beta\rho^j,\ j = 0, 1, 2, \ldots\}$ satisfy

$$f\left(x_k + \beta\rho^j d_k\right) \leq f(x_k) + \delta_1 \beta\rho^j g_k^T d_k, \tag{1.7}$$

where $0 < \beta \leq 1$, $0 < \rho < 1$, and $0 < \delta_1 < 1$. Varieties of this method differ in the way of selecting β_k. In this paper, we are interested in the HS method [3], namely,

$$\beta_k^{\mathrm{HS}} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}. \tag{1.8}$$

Here and throughout the paper, without specification, we always use $\|\cdot\|$ to denote the Euclidian norm of vectors, $y_{k-1} = g_k - g_{k-1}$ and $s_k = \alpha_k d_k$.

We refer to a book [4] and a recent review paper [5] about progress of the global convergence of conjugate gradient methods. We know that the study in the HS method has made great progress. In practical computation, the HS method is generally believed to be one of the most efficient conjugate gradient methods. Theoretically, the HS method has the property that the conjugacy condition

$$d_k^T y_{k-1} = 0, \tag{1.9}$$

always holds, which is independent of line search used. Expecting the fast convergence of the method, Dai and Liao [6] modified the numerator of the HS method to obtain DL method by using the secant condition of quasi-Newton methods. Due to Powell's [7] example, the DL method may not converge with exact line search for general function. Similar to the PRP+ method [8], Dai and Liao [6] proposed the DL+ method from a view of global convergence. In a further development of this update strategy, Yabe and Takano [9] used another modified secant condition in [10, 11] and proposed the YT and YT+ methods. Recently, Hager and Zhang [5] modified the HS method to propose a new conjugate gradient method called CG_DESCENT method. A good property of the CG_DESCENT method lies in that the direction d_k satisfies sufficient descent property $g_k^T d_k \leq -(7/8)\|g_k\|^2$ which is independent of the line search used. Hager and Zhang [5] proved that the CG_DESCENT method with Wolfe line search is globally convergent even for nonconvex problems. Zhang [12] proposed the TTHS method. The sufficient descent property of the TTHS method is also independent of line search used. In order to obtain the global convergence of the TTHS method, Zhang truncated the search direction of the TTHS method. Numerical experiments in [12] show the truncated TTHS method is not very effective. In this paper, we further study the TTHS method. We prove that the TTHS method with standard Armijo line search is globally convergent for uniformly convex problems. To improve the efficiency of the truncated TTHS method, we propose a new truncated strategy to the TTHS method. Under suitable conditions, global convergence is obtained for the proposed method. Numerical experiments show that the proposed method outperforms the known CG_DESCENT method.

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

2. Global Convergence Analysis

Recently, Zhang [12] proposed a three-term modified HS method as follows

$$d_k = \begin{cases} -g_k, & \text{if } k = 0, \\ -g_k + \beta_k^{\mathrm{HS}} d_{k-1} - \theta_k y_{k-1}, & \text{if } k > 0, \end{cases} \tag{2.1}$$

where $\theta_k = g_k^T d_{k-1} / d_{k-1}^T y_{k-1}$. An attractive property of the TTHS method is that the direction always satisfies

$$g_k^T d_k = -\|g_k\|^2, \tag{2.2}$$

which is independent of the line search used. In order to obtain the global convergence of the TTHS method, Zhang truncated the TTHS method as follows

$$d_k = \begin{cases} -g_k, & \text{if } s_k^T y_k < \varepsilon_1 \|g_k\|^r s_k^T s_k, \\ -g_k + \beta_k^{\mathrm{HS}} d_{k-1} - \theta_k y_{k-1}, & \text{if } s_k^T y_k \geq \varepsilon_1 \|g_k\|^r s_k^T s_k, \end{cases} \tag{2.3}$$

where ε_1 and r are positive constants. Zhang proved that the truncated TTHS method converges globally with the Wolfe line search (1.4) and (1.6). However, numerical results show the truncated TTHS method is not very effective. In this paper, we will study the TTHS method again. In the rest of this section, we will establish two preliminary convergent results for the TTHS method.

(i) Uniformly convex functions: converge globally with the standard Armijo line search (1.7).

(ii) General functions: converge globally with the strong Wolfe line search (1.4) and (1.5) by using a new truncated strategy to the TTHS method.

In order to establish the global convergence of our method, we need the following assumption.

Assumption 2.1. (i) The level set $\Omega = \{x \in R^n \mid f(x) \le f(x_0)\}$ is bounded.

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

$$\|g(x) - g(y)\| \le L\|x - y\|, \quad \forall x, y \in N. \tag{2.4}$$

Under Assumption 2.1, It is clear that there exist positive constants B and γ such that

$$\|x - y\| \le B \quad \forall x, y \in \Omega, \tag{2.5}$$

$$\|g(x)\| \le \gamma \quad \forall x \in \Omega. \tag{2.6}$$

Lemma 2.2. *Suppose that Assumption 2.1 holds. Consider $\{x_k\}$ be generated by the TTHS method, where α_k is obtained by the Armijo line search* (1.7)*, one has*

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty. \tag{2.7}$$

Proof. If $\alpha_k = \beta$, then

$$\delta_1 \|g_k\|^2 = -\delta_1 g_k^T d_k \le \frac{1}{\beta}\left[f(x_k) - f(x_{k+1})\right]. \tag{2.8}$$

Combining with

$$\left\|g_k^T d_k\right\|^2 \le \|g_k\|^2 \|d_k\|^2, \tag{2.9}$$

yields

$$\frac{\|g_k\|^4}{\|d_k\|^2} \le \|g_k\|^2 \le \frac{1}{\beta\delta_1}\left(f(x_k) - f(x_{k+1})\right). \tag{2.10}$$

On the other hand, if $\alpha_k \ne \beta$, by the line search rule, then $\rho^{-1}\alpha_k$ does not satisfy (1.7). This implies

$$f\left(x_k + \rho^{-1}\alpha_k d_k\right) > f(x_k) + \delta_1 \rho^{-1}\alpha_k g_k^T d_k. \tag{2.11}$$

By the mean-value theorem, there exists $\mu_k \in (0, 1)$ such that

$$f\left(x_k + \rho^{-1}\alpha_k d_k\right) = f(x_k) + \rho^{-1}\alpha_k g\left(x_k + \mu_k \rho^{-1}\alpha_k d_k\right)^T d_k. \tag{2.12}$$

This together with (2.11) implies

$$\left(g\left(x_k + \mu_k \rho^{-1} \alpha_k d_k\right) - g_k\right)^T d_k \geq -(1 - \delta_1) g_k^T d_k. \tag{2.13}$$

Since g is Lipschitz continuous, the last inequality shows

$$\alpha_k \geq \frac{-(1 - \delta_1)\rho g_k^T d_k}{L\|d_k\|^2} = \frac{(1 - \delta_1)\rho\|g_k\|^2}{L\|d_k\|^2}. \tag{2.14}$$

That is

$$f(x_{k+1}) - f(x_k) \leq \delta_1 \alpha_k g_k^T d_k = -\frac{(1 - \delta_1)\delta_1 \rho}{L} \frac{\|g_k\|^4}{\|d_k\|^2}. \tag{2.15}$$

This implies that there is a constant $M_1 > 0$ such that

$$\frac{\|g_k\|^4}{\|d_k\|^2} \leq M_1 \left(f(x_k) - f(x_{k+1})\right). \tag{2.16}$$

Inequality (2.10) together with (2.16) shows that

$$\frac{\|g_k\|^4}{\|d_k\|^2} \leq M_2 \left(f(x_k) - f(x_{k+1})\right), \tag{2.17}$$

with some constant $M_2 > 0$. Summing these inequalities, we obtain (2.7). □

The following theorem establishes the global convergence of the TTHS method with the standard Armijo line search (1.7) for uniformly convex problems.

Theorem 2.3. *Suppose that Assumption 2.1 holds and f is a uniformly convex function. Consider the TTHS method, where α_k is obtained by the Armijo line search* (1.7)*, one has that*

$$\lim_{k \to \infty} \inf \|g_k\| = 0. \tag{2.18}$$

Proof. We proceed by contradiction. If (2.18) does not hold, there exists a positive constant ε such that for all k

$$\|g_k\| \geq \varepsilon. \tag{2.19}$$

From Lemma 2.2, we get

$$\sum_{k=0}^{\infty} \frac{1}{\|d_k\|^2} < \infty. \tag{2.20}$$

Since f is a uniformly convex function, there exists a constant $\mu > 0$ such that

$$\left(g(x) - g(y)\right)^T (x - y) \geq \mu \|x - y\|^2, \quad \forall x, y \in N. \tag{2.21}$$

This means

$$d_{k-1}^T y_{k-1} \geq \mu \alpha_{k-1} \|d_{k-1}\|^2. \tag{2.22}$$

By (2.1), (2.4), (2.6), and (2.22), one has

$$\begin{aligned}
\|d_k\| &\leq \|g_k\| + \left|\beta_k^{\mathrm{HS}}\right| \|d_{k-1}\| + |\theta_k| \|y_{k-1}\| \\
&\leq \|g_k\| + \left|\beta_k^{\mathrm{HS}}\right| \|d_{k-1}\| + \frac{\|g_k\| \|d_{k-1}\|}{\left|d_{k-1}^T y_{k-1}\right|} \|y_{k-1}\| \\
&\leq \|g_k\| + 2 \frac{\|g_k\| \|d_{k-1}\|}{\left|d_{k-1}^T y_{k-1}\right|} \|y_{k-1}\| \\
&\leq \|g_k\| + 2 \frac{L \|g_k\| \|s_{k-1}\|}{\mu \alpha_{k-1} \|d_{k-1}\|^2} \|d_{k-1}\| \\
&\leq \frac{2L + \mu}{\mu} \gamma.
\end{aligned} \tag{2.23}$$

This implies

$$\sum_{k=0}^{\infty} \frac{1}{\|d_k\|^2} \geq \sum_{k=0}^{\infty} \frac{\mu^2}{(\mu + 2L)^2 \gamma^2} \longrightarrow \infty. \tag{2.24}$$

This yield a contradiction with (2.20). □

We are going to investigate the global convergence of the TTHS method with the strong Wolfe line search (1.4) and (1.5). Similar to the PRP+ method [8], we restrict $\beta_k^{\mathrm{HS}} = \max\{\beta_k^{\mathrm{HS}}, 0\}$. In this case, the search direction (2.1) may not be a descent direction. Noting the search direction (2.1) can be rewritten as

$$d_k = \begin{cases} -g_k, & \text{if } k = 0, \\ -g_k + \beta_k d_{k-1} - \beta_k \dfrac{g_k^T d_{k-1}}{g_k^T y_{k-1}} y_{k-1}, & \text{if } k > 0, \end{cases} \tag{2.25}$$

where $\beta_k = \beta_k^{\mathrm{HS}}$. Since the term $g_k^T y_{k-1}$ may be zero in practice computation, we consider the following search direction

$$d_k = \begin{cases} -g_k, & \text{if } \left|g_k^T y_{k-1}\right| < c\left\|g_k\right\|^2, \\ -g_k + \beta_k^{\mathrm{HS}+} d_{k-1} - \beta_k^{\mathrm{HS}+} \dfrac{g_k^T d_{k-1}}{g_k^T y_{k-1}} y_{k-1}, & \text{if } \left|g_k^T y_{k-1}\right| \geq c\left\|g_k\right\|^2, \end{cases} \tag{2.26}$$

where c is a positive constant and $\beta_k^{\mathrm{HS}+} = \max\{\beta_k^{\mathrm{HS}}, 0\}$. It is clear that the relation (2.2) always holds. For simplicity, we regard the method defined by (1.2) and (2.26) as the method (2.26).

Now, we describe a lemma for the search directions, which shows that they change slowly, asymptotically. The lemma is similar to [8, Lemma 3.4].

Lemma 2.4. *Suppose that Assumption 2.1 holds. Consider* $\{x_k\}$ *be generated the method* (2.26), *where* α_k *is obtained by the strong Wolfe line search* (1.4) *and* (1.5). *If there exists a constant* $\varepsilon > 0$ *such that for all* k

$$\left\|g_k\right\| \geq \varepsilon, \tag{2.27}$$

then $d_k \neq 0$ *and*

$$\sum_{k\geq 0} \|u_{k+1} - u_k\|^2 < \infty, \tag{2.28}$$

where $u_k = d_k / \|d_k\|$.

Proof. Noting that $d_k = 0$, for otherwise (2.2) would imply $g_k = 0$. Therefore, u_k is well defined. Now, let us define $r_k = v_k / \|d_k\|$ and $\delta_k = \beta_k^{\mathrm{HS}+}(\|d_{k-1}\| / \|d_k\|)$, where

$$v_k = -\left(1 + \beta_k^{\mathrm{HS}+} \frac{g_k^T d_{k-1}}{g_k^T y_{k-1}}\right) g_k. \tag{2.29}$$

From (2.26), we have

$$u_k = r_k + \delta_k u_{k-1}. \tag{2.30}$$

Since u_k are unit vectors, we have

$$\|r_k\| = \|u_k - \delta_k u_{k-1}\| = \|\delta_k u_k - u_{k-1}\|. \tag{2.31}$$

Since $\delta_k > 0$, it follows that

$$\begin{aligned} \|u_k - u_{k-1}\| &\leq \|(1 + \delta_k)(u_k - u_{k-1})\| \\ &\leq \|u_k - \delta_k u_{k-1}\| + \|\delta_k u_k - u_{k-1}\| \\ &= 2\|r_k\|. \end{aligned} \tag{2.32}$$

Then we have

$$\|u_k - u_{k-1}\|^2 \le 4r_k^2. \tag{2.33}$$

Now, we evaluate the quantity v_k. If $g_k^T y_{k-1} \ge c\|g_k\|^2$, by (1.5), we have

$$d_{k-1}^T y_{k-1} = d_{k-1}^T (g_k - g_{k-1}) \ge (\sigma - 1) g_{k-1}^T d_{k-1} = (1-\sigma)\|g_{k-1}\|^2. \tag{2.34}$$

By the strong Wolfe condition (1.5) and the relation (2.2), we obtain

$$\left|g_k^T d_{k-1}\right| \le \sigma \left|g_{k-1}^T d_{k-1}\right| = \sigma \|g_{k-1}\|^2. \tag{2.35}$$

Inequalities (2.34) and (2.35) yield

$$\frac{\left|g_k^T d_{k-1}\right|}{\left|d_{k-1}^T y_{k-1}\right|} \le \frac{\sigma}{1-\sigma}. \tag{2.36}$$

This implies

$$\|v_k\| \le \left(1 + \beta_k^{\mathrm{HS+}} \left|\frac{g_k^T d_{k-1}}{g_k^T y_{k-1}}\right|\right)\|g_k\| \le \left(1 + \left|\frac{g_k^T d_{k-1}}{d_{k-1}^T y_{k-1}}\right|\right)\|g_k\| \le \frac{1}{1-\sigma}\|g_k\|. \tag{2.37}$$

If $g_k^T y_{k-1} < c\|g_k\|^2$, then $\|v_k\| = \|g_k\|$. The relation (2.37) also holds. It follows from the definition of r_k, Lemma 2.2, (2.27) and (2.37) that

$$\sum_{k\ge 0} r_k^2 \le \sum_{k\ge 0} \frac{\|g_k\|^4}{(1-\sigma)^2 \varepsilon^2 \|d_k\|^2} < \infty. \tag{2.38}$$

By (2.33), we get the conclusion (2.28). □

The next theorem establishes the global convergence of method (2.26) with the strong Wolfe line search (1.4) and (1.5). The proof of the theorem is similar to [15, Theorem 3.2].

Theorem 2.5. *Suppose that Assumption 2.1 holds. Consider* $\{x_k\}$ *be generated by the method* (2.26), *where* α_k *is obtained by the strong Wolfe line search* (1.4) *and* (1.5), *one has*

$$\lim_{k\to\infty} \inf \|g_k\| = 0. \tag{2.39}$$

Proof. We assume that the conclusion (2.39) is not true, then there exists a constant $\varepsilon > 0$ such that for all k

$$\|g_k\| \ge \varepsilon. \tag{2.40}$$

The proof is divided into the following three steps.

Step 1. A bound for $\beta_k^{\mathrm{HS}+}$. From (2.4), (2.6), and (2.34), we get

$$\left|\beta_k^{\mathrm{HS}+}\right| \leq \left|\frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}\right| \leq \frac{L\|g_k\| \cdot \|s_{k-1}\|}{(1-\sigma)\|g_{k-1}\|^2} \leq \frac{L\gamma\|s_{k-1}\|}{(1-\sigma)\varepsilon^2} \triangleq C_1\|s_{k-1}\|. \tag{2.41}$$

Step 2. A bound on the steps s_k. This is a modified version of [8, Theorem 4.3]. Observe that for any $l \geq k$,

$$x_l - x_k = \sum_{j=k}^{l-1} x_{j+1} - x_j = \sum_{j=k}^{l-1}\|s_j\|u_j = \sum_{j=k}^{l-1}\|s_j\|u_k + \sum_{j=k}^{l-1}\|s_j\|(u_j - u_k). \tag{2.42}$$

Taking norms and by the triangle inequality to the last equality, we get from (2.5) that

$$\sum_{j=k}^{l-1}\|s_j\| \leq \|x_l - x_k\| + \sum_{j=k}^{l-1}\|s_j\|\|u_j - u_k\| \leq B + \sum_{j=k}^{l-1}\|s_j\|\|u_j - u_k\|. \tag{2.43}$$

Let Δ be a positive integer, chosen large enough that

$$\Delta \geq 4BC, \tag{2.44}$$

where $C = (1 + \sigma\gamma^2/\varepsilon^2)C_1$. By Lemma 2.4, we can chose k_0 large enough that

$$\sum_{i \geq k_0}\|u_{i+1} - u_i\|^2 \leq \frac{1}{4\Delta}. \tag{2.45}$$

If $j > k \geq k_0$ and $j - k \leq \Delta$, then by (2.45) and the Cauchy-Schwarz inequality, we have

$$\begin{aligned}\|u_j - u_k\| &\leq \sum_{i=k}^{j-1}\|u_{i+1} - u_i\| \\ &\leq \sqrt{j-k}\left(\sum_{i=k}^{j-1}\|u_{i+1} - u_i\|^2\right)^{1/2} \\ &\leq \sqrt{\Delta}\left(\frac{1}{4\Delta}\right)^{1/2} = \frac{1}{2}.\end{aligned} \tag{2.46}$$

Combining this with (2.43) yields

$$\sum_{j=k}^{l-1}\|s_j\| \leq 2B, \tag{2.47}$$

where $l > k \geq k_0$ and $l - k \leq \Delta$.

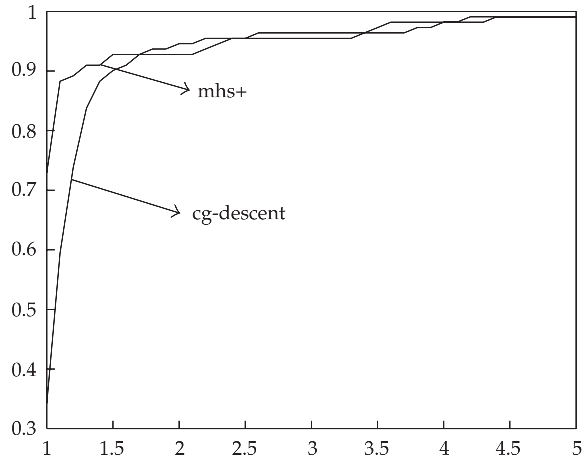

Figure 1: Performance based on the number of iteration.

Step 3. A bound on the direction d_l determined by (2.26). If $g_l^T y_{l-1} \geq c\|g_l\|^2$, from (2.26), (2.27), (2.35), and (2.41), we have

$$
\begin{aligned}
\|d_l\|^2 &\leq \left(\|g_l\| + \beta_l^{\text{HS+}} \|d_{l-1}\| + \beta_l^{\text{HS+}} \frac{|g_l^T d_{l-1}|}{\|g_l\|^2} y_{l-1} \right)^2 \\
&\leq \left(\|g_l\| + \left(1 + \frac{LB\sigma\gamma^2}{\varepsilon^2}\right) \beta_l^{\text{HS+}} \|d_{l-1}\| \right)^2 \\
&\leq 2\gamma^2 + 2\left(1 + \frac{LB\sigma\gamma^2}{\varepsilon^2}\right)^2 C_1^2 \|s_{l-1}\|^2.
\end{aligned} \tag{2.48}
$$

If $g_l^T y_{l-1} < c\|g_l\|^2$, then $d_l = -g_l$, we know that the relation (2.48) also holds. Define $S_i = 2C^2\|s_i\|^2$, we conclude that for $l > k_0$,

$$
\|d_l\|^2 \leq 2\gamma^2 \left(\sum_{i=k_0+1}^{l} \prod_{j=i}^{l-1} S_j \right) + \|d_{k_0}\|^2 \prod_{j=k_0}^{l-1} S_j. \tag{2.49}
$$

Proceeding the similar proof as the case III of [15, Theorem 3.2], we get the conclusion. □

3. Numerical Experiments

In this section, we report some numerical results. We tested 111 problems that are from the CUTE [13] library. We compared the performance of the method (2.26) with the CG_DESECENT method. The CG_DESECNT code can be obtained from Hager's web page at http://www.math.ufl.edu/hager/papers/CG.

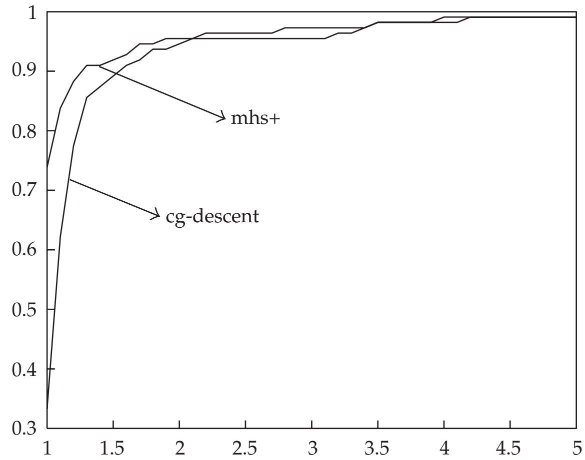

Figure 2: Performance based on the number of function evaluations.

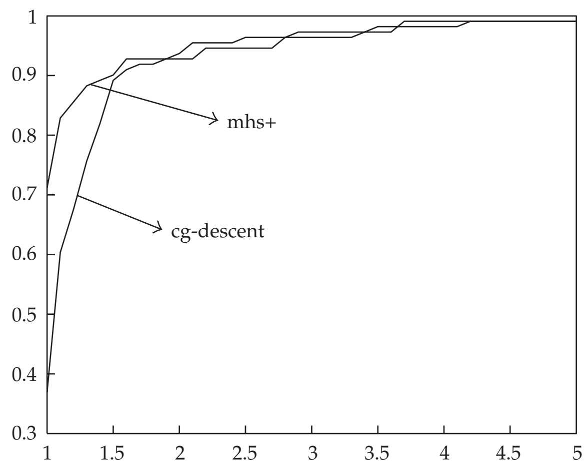

Figure 3: Performance based on the number of gradient evaluations.

In the numerical experiments, we used the latest version—Source code Fortran 77 Version 1.4 (November 14, 2005) with default parameters. We implemented the method (2.26) with the approximate Wolfe line search in [5]. Namely, the method (2.26) used the same line search and parameters as the CG_DESECENT method. The stop criterion is that the inequality $\|g(x)\|_\infty \leq \max\{10^{-8}, 10^{-12}\|\nabla f(x_0)\|_\infty\}$ is satisfied or the iteration number exceeds 4×10^4. All codes were written in Fortran 77 and run on a PC with PIII 866 processor and 192 RAM memory and Linux operation system. Detailed results are posted at the following web site: http://hi.814e.com/wanyoucheng/results.htm.

We adopt the performance profiles by Dolan and Moré [14] to compare the performance between different methods. That is, 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

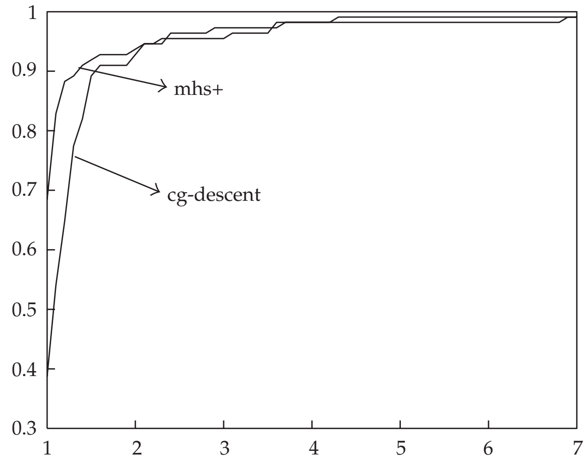

Figure 4: Performance based on CPU time.

top curve is the method that solved the most problems in a time that is within a factor τ of the best time.

The curves in Figures 1, 2, 3, and 4 have the following meaning:

(i) cg-descent: the CG_DSCENT method with the approximate Wolfe line search proposed by Hager and Zhang [15];

(ii) mhs+: the method (2.26) with the same line search as "cg-descent" and $c = 10^{-8}$.

From Figures 1–4, it is clear that the "mhs+" method outperforms the "cg-descent" method.

Acknowledgments

The authors are indebted to the anonymous referee for his helpful suggestions which improved the quality of this paper. The authors are very grateful also to Professor W. W. Hager and Dr. H. Zhang for their CG_DESCENT code and line search code. This paper is supported by the NSF of China via Grant 10771057.

References

[1] P. Wolfe, "Convergence conditions for ascent methods," *SIAM Review*, vol. 11, no. 2, pp. 226–235, 1969.

[2] P. Wolfe, "Convergence conditions for ascent methods—II: some corrections," *SIAM Review*, vol. 13, no. 2, pp. 185–188, 1971.

[3] 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.

[4] Y.-H. Dai and Y. Yuan, *Nonlinear Conjugate Gradient Methods*, Shanghai Science and Technology, Shanghai, China, 2000.

[5] W. W. Hager and H. Zhang, "A survey of nonlinear conjugate gradient methods," *Pacific Journal of Optimization*, vol. 2, no. 1, pp. 35–58, 2006.

[6] Y.-H. Dai and L.-Z. Liao, "New conjugacy conditions and related nonlinear conjugate gradient methods," *Applied Mathematics and Optimization*, vol. 43, no. 1, pp. 87–101, 2001.

[7] M. J. D. Powell, "Convergence properties of algorithms for nonlinear optimization," *SIAM Review*, vol. 28, no. 4, pp. 487–500, 1986.

[8] 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.
[9] H. Yabe and M. Takano, "Global convergence properties of nonlinear conjugate gradient methods with modified secant condition," *Computational Optimization and Applications*, vol. 28, no. 2, pp. 203–225, 2004.
[10] J. Z. Zhang, N. Y. Deng, and L. H. Chen, "New quasi-Newton equation and related methods for unconstrained optimization," *Journal of Optimization Theory and Applications*, vol. 102, no. 1, pp. 147–167, 1999.
[11] J. Z. Zhang and C. Xu, "Properties and numerical performance of quasi-Newton methods with modified quasi-Newton equations," *Journal of Computational and Applied Mathematics*, vol. 137, no. 2, pp. 269–278, 2001.
[12] L. Zhang, *Nonlinear conjugate gradient methods for optimization problems*, Ph.D. thesis, College of Mathematics and Econometrics, Hunan University, Changsha, China, 2006.
[13] 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.
[14] E. D. Dolan and J. J. Moré, "Benchmarking optimization software with performance profiles," *Mathematical Programming*, vol. 91, no. 2, pp. 201–213, 2002.
[15] 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.