Filomat 31:6 (2017), 1813–1825
DOI 10.2298/FIL1706813D

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

New Hybrid Conjugate Gradient Method as a Convex Combination of LS and CD methods

Snežana S. Djordjević[a]

[a] *College for Textile, Vilema Pušmana 17, 16000 Leskovac, Serbia*

Abstract. A new hybrid conjugate gradient algorithm is considered. The conjugate gradient parameter β_k is computed as a convex combination of β_k^{CD} and β_k^{LS}. The parameter θ_k is computed in such a way that the conjugacy condition is satisfied.

The strong Wolfe line search conditions are used.

Numerical comparisons show that the present hybrid conjugate gradient algorithm often behaves better than some known algorithms.

1. Introduction

We consider the nonlinear unconstrained optimization problem

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

where $f : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function, bounded from below.

There exist many different methods for solving the problem (1.1).

Here we are interested in conjugate gradient methods, which have low memory requirements and strong local and global convergence properties [8].

To solve the problem (1.1), starting from an initial point $x_0 \in \mathbb{R}^n$, the conjugate gradient method generates a sequence $\{x_k\} \subset \mathbb{R}^n$ such that

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

where $t_k > 0$ is a step size, received from the line search, and the directions d_k are given by [3], [4]

$$d_0 = -g_0, \quad d_{k+1} = -g_{k+1} + \beta_k s_k. \tag{1.3}$$

In the last relation, β_k is the conjugate gradient parameter, $s_k = x_{k+1} - x_k$, $g_k = \nabla f(x_k)$.

Let the norm $\|\cdot\|$ be the Euclidean norm.

Now, we denote

$$y_k = g_{k+1} - g_k. \tag{1.4}$$

2010 *Mathematics Subject Classification*. 90C30.
Keywords. Hybrid conjugate gradient method; Convex combination; Conjugacy condition.
Received: 23 April 2015; Accepted: 29 September 2015
Communicated by Predrag Stanimirović
Email address: snezanadjordjevic1971@gmail.com (Snežana S. Djordjević)

An excellent survey of conjugate gradient methods is given by Hager and Zhang [14]. Different conjugate gradient methods correspond to different values of the scalar parameter β_k.

A hybrid conjugate gradient method is a certain combination of different conjugate gradient methods; it is made to improve the behavior of these methods and to avoid the jamming phenomenon.

In order to choose the parameter β_k for the method in the present paper, we mention the following choices of β_k [2]:

$$\text{Fletcher and Reeves: [12] } \beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2}; \tag{1.5}$$

$$\text{Dai and Yuan: [7] } \beta_k^{DY} = \frac{\|g_{k+1}\|^2}{y_k^T s_k}; \tag{1.6}$$

$$\text{Conjugate Descent, proposed by Fletcher: [11] } \beta_k^{CD} = \frac{\|g_{k+1}\|^2}{-g_k^T s_k}. \tag{1.7}$$

The conjugate gradient methods with the choice of β_k taken in (1.5), (1.6) and (1.7), have strong convergence properties, but they may have the modest practical performance, due to jamming [2], [3].

From the other side, methods of Polak-Ribiére [19] and Polyak [20], Hestenes and Stiefel [15] and also Liu and Storey [17] in general may not be convergent, but usually they have better computer performances [3]. The choices of β_k in these methods are, respectively [2]:

$$\beta_k^{PRP} = \frac{g_{k+1}^T y_k}{\|g_k\|^2}, \tag{1.8}$$

$$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{y_k^T s_k}, \tag{1.9}$$

$$\beta_k^{LS} = \frac{g_{k+1}^T y_k}{-g_k^T s_k}. \tag{1.10}$$

One important difference between FR and CD is that with CD, the sufficient descent holds for a strong Wolfe line search (the constraint $\sigma \leq \frac{1}{2}$ that arose with FR, is not needed for CD) [14].

Moreover, for a line search that satisfies the generalized Wolfe conditions with $\sigma_1 < 1$ and $\sigma_2 = 0$, it can be shown that CD method is globally convergent [14].

On the other hand, not much research has been done on the choice β_k^{LS} for the update parameter, except for the paper [17], but we expect that the techniques developed for the analysis of the PRP method should apply to the LS method [14].

2. Convex combination

The parameter β_k is chosen here such that it presents the convex combination of (1.7) and (1.10).

Now we define the next conjugate gradient parameter

$$\beta_k^{hyb} = (1 - \theta_k) \cdot \beta_k^{LS} + \theta_k \cdot \beta_k^{CD}. \tag{2.1}$$

So, we can write

$$d_0 = -g_0, \;\; d_{k+1} = -g_{k+1} + \beta_k^{hyb} s_k. \tag{2.2}$$

The parameter θ_k is a scalar parameter which we have to determine. We use here the strong Wolfe line search, i.e., we are going to find a step length t_k, such that:

$$f(x_k + t_k d_k) - f(x_k) \leq \delta t_k g_k^T d_k, \tag{2.3}$$

and

$$|g(x_k + t_k d_k)^T d_k| \leq -\sigma g_k^T d_k. \tag{2.4}$$

It is obvious that, if $\theta_k = 0$, then $\beta_k^{hyb} = \beta_k^{LS}$, and if $\theta_k = 1$, then $\beta_k^{hyb} = \beta_k^{CD}$.
On the other side, if $0 < \theta_k < 1$, then β_k^{hyb} is a proper convex combination of the parameters β_k^{CD} and β_k^{LS}.
Having in view the relations (1.7) and (1.10), the relation (2.1) becomes

$$\beta_k^{hyb} = (1-\theta_k)\cdot\frac{g_{k+1}^T y_k}{-g_k^T s_k} + \theta_k \cdot \frac{g_{k+1}^T g_{k+1}}{-g_k^T s_k}, \tag{2.5}$$

so the relation (2.2) becomes

$$d_0 = -g_0, \;\; d_{k+1} = -g_{k+1} + (1-\theta_k)\frac{g_{k+1}^T y_k}{-g_k^T s_k}\cdot s_k + \theta_k \cdot \frac{g_{k+1}^T g_{k+1}}{-g_k^T s_k}\cdot s_k. \tag{2.6}$$

We shall find the value of the parameter θ_k in such a way that the conjugacy condition

$$y_k^T d_{k+1} = 0 \tag{2.7}$$

holds.

Firstly, we multiply both sides of the relation (2.6) by y_k^T from the left:

$$y_k^T d_{k+1} = -y_k^T g_{k+1} + (1-\theta_k)\frac{g_{k+1}^T y_k}{-g_k^T s_k}\cdot y_k^T s_k + \theta_k \frac{\|g_{k+1}\|^2}{-g_k^T s_k}\cdot y_k^T s_k.$$

Using (2.7), we get

$$\theta_k = \frac{-(g_{k+1}^T y_k)(g_{k+1}^T s_k)}{(g_{k+1}^T g_k)(y_k^T s_k)}. \tag{2.8}$$

Finally, having in view the relation (2.8), we define:

$$\theta_k = \begin{cases} \dfrac{-(g_{k+1}^T y_k)(g_{k+1}^T s_k)}{(g_{k+1}^T g_k)(y_k^T s_k)}, & 0 \leq \dfrac{-(g_{k+1}^T y_k)(g_{k+1}^T s_k)}{(g_{k+1}^T g_k)(y_k^T s_k)} \leq 1; \\ 0, & \dfrac{-(g_{k+1}^T y_k)(g_{k+1}^T s_k)}{(g_{k+1}^T g_k)(y_k^T s_k)} < 0; \\ 1, & \dfrac{-(g_{k+1}^T y_k)(g_{k+1}^T s_k)}{(g_{k+1}^T g_k)(y_k^T s_k)} > 1. \end{cases} \tag{2.9}$$

For some later considerations, we remind to the next lemma.

Lemma 2.1. *[18] Let $f \in C(\mathbb{R}^n)$. Let d_k be a descent direction in the point x_k, and suppose that the function f is bounded from below along the direction $\{x_k + td_k | t > 0\}$. Then if $0 < \delta < \sigma < 1$, there exist the intervals inside which the step size satisfies Wolfe conditions and strong Wolfe conditions.*

3. Algorithm and the Sufficient Descent Condition

Algorithm LSCDCC
Input parameters: $\epsilon > 0$, x_0, $k := 0$, $0 < \delta \leq \sigma < 1$, $\beta < 1$, $t_0 = 1$.

Step 1. *If* $\|g_k\| \leq \epsilon$, *STOP.*

Step 2. *Determine the biggest* j_k, *such that for* $t_k = \beta^{j_k}$ *it holds*

$$f(x_k + t_k d_k) - f(x_k) \leq \delta t_k g_k^T d_k,$$
$$|g_{k+1}^T d_k| \leq -\sigma g_k^T d_k.$$

Set $x_{k+1} = x_k + t_k d_k$.

Step 3. *If* $(g_{k+1}^T g_k)(y_k^T s_k) = 0$, *then* $\theta_k = 0$, *else find* θ_k *from (2.9).*

Step 4.

$$\beta_k^{hyb} = (1 - \theta_k) \cdot \frac{g_{k+1}^T y_k}{-g_k^T s_k} + \theta_k \cdot \frac{g_{k+1}^T g_{k+1}}{-g_k^T s_k}. \tag{3.1}$$

Step 5. *If*

$$|g_{k+1}^T g_k| > a\|g_{k+1}\|^2, \tag{3.2}$$

then $d_{k+1} = -g_{k+1}$, *else* $d_{k+1} = -g_{k+1} + \beta_k^{hyb} \cdot s_k$.

Step 6. $k := k + 1$, *go to Step 1.*

For further considerations we need the next assumptions.

Assumption 3.1. *The level set* $S = \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}$ *is bounded.*

Assumption 3.2. *In a neighborhood* $\mathcal{N}$ *of* S, *the function* f *is continuously differentiable and its gradient is Lipschitz continuous, i.e. there exists a constant* $L > 0$, *such that* $\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$, *for all* $x, y \in \mathcal{N}$.

Under Assumption 3.1 and Assumption 3.2 on f, there exists a constant $\Gamma \geq 0$, such that

$$\|\nabla f(x)\| \leq \Gamma, \tag{3.3}$$

for all $x \in S$ [2].

Theorem 3.1. *Let Assumptions 3.1 and 3.2 hold. Let the constant* a *in the algorithm LSCDCC be such that*

$$0 < a < \frac{1}{\sigma} - 1. \tag{3.4}$$

Then algorithm LSCDCC is well defined and d_k *satisfies the sufficient descent condition for all* k.

Proof. From Lemma 2.1 we know that Step 2 of the algorithm LSCDCC is well defined if d_k is a descent direction. We shall show that d_k satisfies the sufficient descent condition, and that will yield that d_k is a descent direction.

For $k = 0$, it holds $d_0 = -g_0$, so $g_0^T d_0 = -\|g_0\|^2$, and we conclude that sufficient descent condition holds for $k = 0$.

Next it holds

$$d_{k+1} = -g_{k+1} + \beta_k^{hyb} s_k,$$

i.e.

$$d_{k+1} = -g_{k+1} + ((1-\theta_k)\beta_k^{LS} + \theta_k \beta_k^{CD}) s_k.$$

We can write

$$d_{k+1} = -(\theta_k g_{k+1} + (1-\theta) g_{k+1}) + ((1-\theta_k)\beta_k^{LS} + \theta_k \beta_k^{CD}) s_k.$$

It follows that

$$d_{k+1} = \theta_k(-g_{k+1} + \beta_k^{CD} s_k) + (1-\theta_k)(-g_{k+1} + \beta_k^{LS} s_k),$$

wherefrom

$$d_{k+1} = \theta_k d_{k+1}^{CD} + (1-\theta_k) d_{k+1}^{LS}. \tag{3.5}$$

Multiplying (3.5) by g_{k+1}^T from the left, we get

$$g_{k+1}^T d_{k+1} = \theta_k g_{k+1}^T d_{k+1}^{CD} + (1-\theta_k) g_{k+1}^T d_{k+1}^{LS}. \tag{3.6}$$

Firstly, let $\theta_k = 0$. Then $d_{k+1} = d_{k+1}^{LS}$. Remind that

$$d_{k+1}^{LS} = -g_{k+1} + \beta_k^{LS} s_k$$

$$\Rightarrow g_{k+1}^T d_{k+1}^{LS} \le -\|g_{k+1}\|^2 + \frac{(g_{k+1}^T y_k)(g_{k+1}^T s_k)}{-g_k^T s_k}. \tag{3.7}$$

We shall prove that $\left|\dfrac{(g_{k+1}^T y_k)(g_{k+1}^T s_k)}{-g_k^T s_k}\right| \le \mu \|g_{k+1}\|^2$, where $0 < \mu < 1$.

Consider the absolute value of the expression

$$T := \frac{(g_{k+1}^T y_k)(g_{k+1}^T s_k)}{-g_k^T s_k}.$$

So,

$$|T| = \left|\frac{(g_{k+1}^T y_k)(g_{k+1}^T s_k)}{-g_k^T s_k}\right| \le \left|\frac{g_{k+1}^T s_k}{-g_k^T s_k}\right| |g_{k+1}^T y_k|.$$

From the second strong Wolfe line search condition, it holds

$$\left|\frac{g_{k+1}^T s_k}{-g_k^T s_k}\right| \le \sigma.$$

Now it holds

$$|T| \le \sigma |g_{k+1}^T y_k|. \tag{3.8}$$

If (3.2) holds, then $d_{k+1} = -g_{k+1}$, so $g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2$, and so it is proved that d_{k+1} satisfies the sufficient descent condition.

If the relation (3.2) doesn't hold, then it holds

$$|g_{k+1}^T g_k| \le a\|g_{k+1}\|^2. \tag{3.9}$$

Because of $y_k = g_{k+1} - g_k$, from (3.8) we get

$$\begin{aligned} |T| &\le \sigma|g_{k+1}^T y_k| \qquad (3.10)\\ &= \sigma|g_{k+1}^T(g_{k+1} - g_k)| \qquad (3.11)\\ &\le \sigma\|g_{k+1}\|^2 + \sigma|g_{k+1}^T g_k|, \qquad (3.12)\end{aligned}$$

wherefrom, applying the relation (3.9), we get

$$|T| \le \sigma\|g_{k+1}\|^2 + \sigma a\|g_{k+1}\|^2,$$

and having in view the relation (3.4), we can write

$$|T| \le \mu\|g_{k+1}\|^2, \text{ where } 0 < \mu = \sigma(1 + a) < 1.$$

Now, using the relation (3.7), we get

$$g_{k+1}^T d_{k+1}^{LS} \le -\|g_{k+1}\|^2 + \mu\|g_{k+1}\|^2,$$

and

$$g_{k+1}^T d_{k+1}^{LS} \le -(1 - \mu)\|g_{k+1}\|^2.$$

Let's denote $K_1 = (1 - \mu)$; then we can write

$$g_{k+1}^T d_{k+1}^{LS} \le -K_1\|g_{k+1}\|^2. \tag{3.13}$$

Now let $\theta_k = 1$. Then $d_{k+1} = d_{k+1}^{CD}$.

Further, we are going to prove that the sufficient decent condition holds for CD method in the presence of the strong Wolfe conditions, and this fact is mentioned in [14].

For $k = 0$ the proof is trivial one, having in view that $d_0^{CD} = -g_0$ and so $g_0^T d_0^{CD} = -\|g_0\|^2$.

Having in view that

$$d_{k+1}^{CD} = -g_{k+1} + \beta_k^{CD} s_k, \tag{3.14}$$

multiplying the relation (3.14) by g_{k+1}^T from the left and using the expression (1.7), we get $g_{k+1}^T d_{k+1}^{CD} = -\|g_{k+1}\|^2 + \dfrac{\|g_{k+1}\|^2}{-g_k^T s_k} \cdot (g_{k+1}^T s_k)$, wherefrom

$$g_{k+1}^T d_{k+1}^{CD} = -\|g_{k+1}\|^2(1 - \frac{g_{k+1}^T s_k}{-g_k^T s_k}) = -\|g_{k+1}\|^2 \cdot \frac{-g_k^T s_k - g_{k+1}^T s_k}{-g_k^T s_k}.$$

Using the strong Wolfe line search, now it holds

$$\frac{-g_k^T s_k - g_{k+1}^T s_k}{-g_k^T s_k} \ge \frac{(\sigma - 1)g_k^T s_k}{-g_k^T s_k} = 1 - \sigma > 0.$$

Now we have

$$g_{k+1}^T d_{k+1}^{CD} \le -(1 - \sigma)\|g_{k+1}\|^2.$$

Let's denote $1 - \sigma = K_2 > 0$.

So,

$$g_{k+1}^T d_{k+1}^{CD} \le -K_2\|g_{k+1}\|^2. \tag{3.15}$$

Now suppose that $0 < \theta_k < 1$, i.e., $0 < a_1 \le \theta_k \le a_2 < 1$.
From the relation (3.6), now we conclude

$$g_{k+1}^T d_{k+1} \le a_1 g_{k+1}^T d_{k+1}^{CD} + (1 - a_2) g_{k+1}^T d_{k+1}^{LS}. \tag{3.16}$$

Denote $K = a_1K_1 + (1 - a_2)K_2$; then we finally get

$$g_{k+1}^T d_{k+1} \le -K\|g_{k+1}\|^2. \tag{3.17}$$

□

4. Convergence Analysis

Let Assumptions 3.1 and 3.2 hold.

In [7] it is proved that for any conjugate gradient method with the strong Wolfe line search, it holds:

Lemma 4.1. *[7] Let Assumptions 3.1 and 3.2 hold. Consider the method (1.2), (1.3), where d_k is a descent direction, and t_k is received from the strong Wolfe line search. If*

$$\sum_{k\ge1} \frac{1}{\|d_k\|^2} = \infty, \tag{4.1}$$

then

$$\liminf_{k\to\infty} \|g_k\| = 0. \tag{4.2}$$

Theorem 4.1. *Consider the iterative method, defined by algorithm LSCDCC. Let all conditions of Theorem 3.1 hold. Then either $g_k = 0$, for some k, or*

$$\liminf_{k\to\infty} \|g_k\| = 0. \tag{4.3}$$

Proof. Suppose that $g_k \ne 0$, for all k. Then we have to prove (4.3).
Suppose, on the contrary, that (4.3) doesn't hold. Then there exists a constant $c > 0$, such that

$$\|g_k\| \ge c. \tag{4.4}$$

Let D be the diameter of the level set $\mathcal{S}$.
From (3.1), we get

$$\|\beta_k^{hyb}\| \le |\beta_k^{LS}| + |\beta_k^{CD}|. \tag{4.5}$$

It holds

$$|\beta_k^{LS}| = \left|\frac{g_{k+1}^T y_k}{-g_k^T s_k}\right| \le \frac{\|g_{k+1}\|\|y_k\|}{|-g_k^T s_k|} \le \frac{\Gamma\|y_k\|}{|-g_k^T s_k|},$$

where we have used (3.3). Applying the Lipschitz assumption, we get

$$|\beta_k^{LS}| \le \frac{\Gamma L \|s_k\|}{|-g_k^T s_k|}.$$

Because of $\|s_k\| \le D$, it yields that

$$|\beta_k^{LS}| \le \frac{\Gamma L D}{|-g_k^T s_k|}.$$

Using Theorem 3.1, we know that for LS method the sufficient descent condition holds, so it is possible to satisfy strong Wolfe conditions.

Now we are going to prove that there exists $t_* > 0$, such that $t_k \ge t_* > 0$, for all k.

Suppose, on the contrary, that there doesn't exist any t_*, such that $t_k \ge t_* > 0$. Then there exists an infinite subsequence $t_k = \beta^{j_k}, \ k \in K_1$ such that

$$\lim_{k \in K_1} t_k = 0. \tag{4.6}$$

Then

$$\lim_{k \in K_1} \beta^{j_k - 1} = 0,$$

i.e.

$$\lim_{k \in K_1} (j_k - 1) = \infty.$$

But, now we get

$$f(x_k + \beta^{j_k} d_k) - f(x_k) \le \delta \beta^{j_k} g_k^T d_k, \tag{4.7}$$

$$f(x_k + \beta^{j_k - 1} d_k) - f(x_k) > \delta \beta^{j_k - 1} g_k^T d_k. \tag{4.8}$$

Remind that $\delta < 1$. From (4.8), we have

$$\frac{f(x_k + \beta^{j_k - 1} d_k) - f(x_k)}{\beta^{j_k - 1}} > \delta g_k^T d_k. \tag{4.9}$$

But, using the relation (4.6), from (4.9), we conclude that

$$g_k^T d_k \ge \delta g_k^T d_k. \tag{4.10}$$

But, LS method satisfies the sufficient descent condition, so $g_k^T d_k \le 0$. Also, $\delta < 1$. So, the relation (4.10) is correct only if $g_k^T d_k = 0$. Then, from the second strong Wolfe condition, we get that $g_{k+1}^T d_k = 0$, and then it is the exact line search. So, we have a contradiction.

Now we can write

$$|-g_k^T s_k| = |-t_k g_k^T d_k| \ge t_* |-g_k^T d_k|.$$

So, from the sufficient descent condition we can get

$$|\beta_k^{LS}| \le \frac{\Gamma L D}{K' \|g_{k+1}\|^2}, \ K' > 0. \tag{4.11}$$

Now we use (4.4) and we get

$$|\beta_k^{LS}| \le \frac{\Gamma L D}{K' c^2}. \tag{4.12}$$

Using the relation (4.12), we get

$$\|d_{k+1}^{LS}\| \le \|g_{k+1}\| + \frac{\Gamma L D}{K' c^2} D \le \Gamma + \frac{\Gamma L D^2}{K' c^2}. \tag{4.13}$$

On the other side,

$$\|d_{k+1}^{CD}\| \le \|g_{k+1}\| + |\beta_k^{CD}|\|s_k\| \le \Gamma + |\beta_k^{CD}|D. \tag{4.14}$$

It also holds that

$$|\beta_k^{CD}| = \frac{\|g_{k+1}\|^2}{|-g_k^T s_k|} \le \frac{\Gamma^2}{|-g_k^T s_k|}. \tag{4.15}$$

The sufficient descent condition holds for CD method too, so, analogically, we can get

$$|\beta_k^{CD}| \le \frac{\Gamma^2}{K''\|g_k\|^2} \le \frac{\Gamma^2}{K''c^2},\ K'' > 0,$$

so

$$\|d_{k+1}^{CD}\| \le \Gamma + \frac{\Gamma^2 D}{K''c^2}. \tag{4.16}$$

Applying (3.5), we find that

$$\|d_{k+1}\| \le \Gamma + \frac{\Gamma L D^2}{K'c^2} + \Gamma + \frac{\Gamma^2 D}{K''c^2},$$

wherefrom

$$\sum_{k\ge 1} \frac{1}{\|d_k\|^2} = \infty, \tag{4.17}$$

so, applying Lemma 4.1, we conclude that

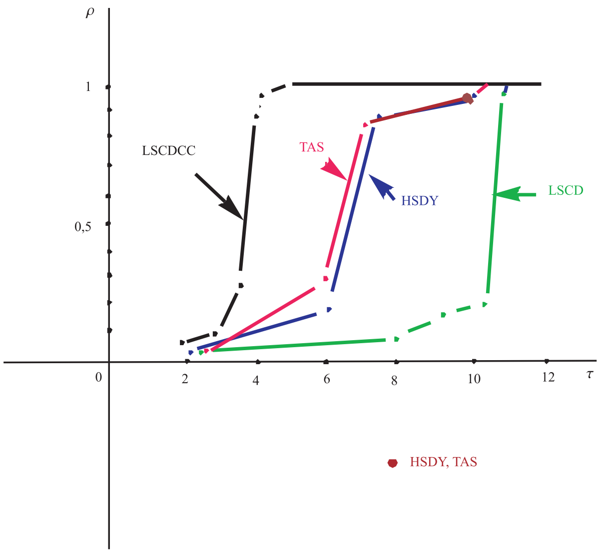
$$\liminf_{k\to\infty} \|g_k\| = 0.$$

This is a contradiction with (4.4), so we have proved (4.3). □

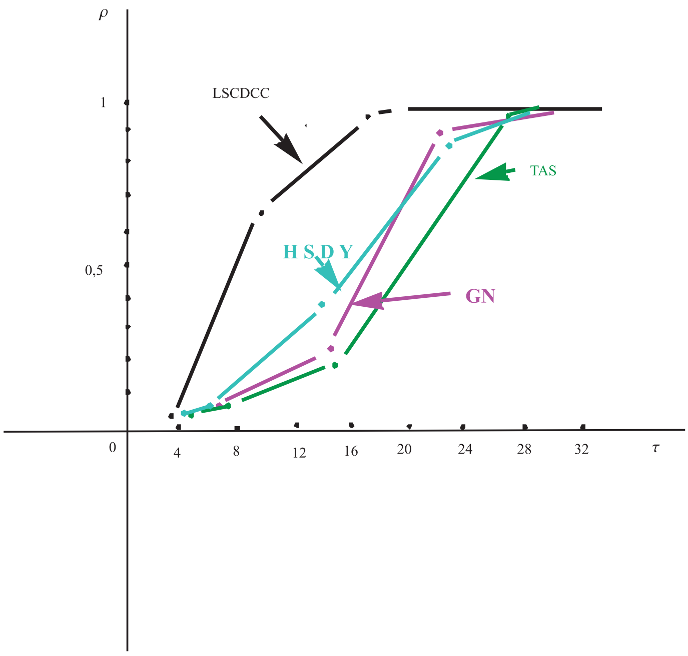
5. Numerical Experiments

In this section we present the computational performance of a Mathematica implementation of LSCDCC algorithm on a set of unconstrained optimization test problems from [5]. Each problem is tested for a number of variables: $n = 50$, $n = 70$, $n = 80$, $n = 90$, $n = 110$, $n = 120$. The criterion used here is CPU time.

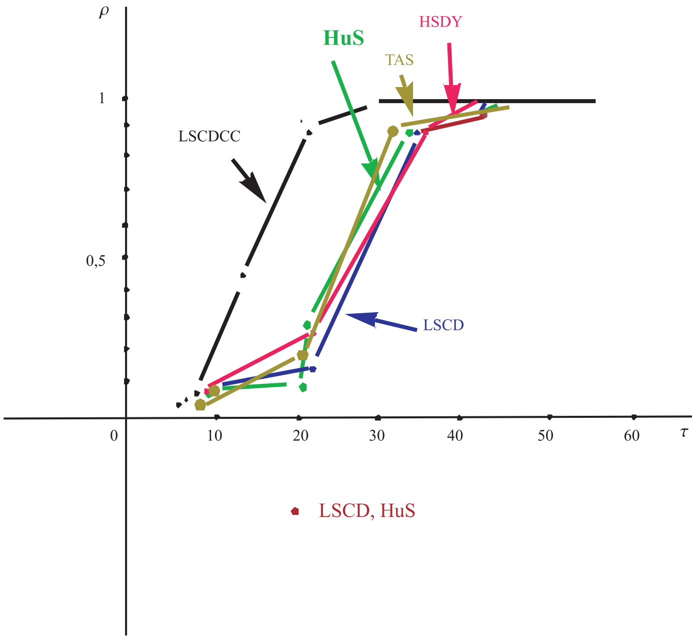
We present comparisons with CCOMB from [2], HYBRID from [3], LSCD from [23], the algorithm from [13], which we call GN here, the algorithm from [16], which we call HuS here, the algorithm from [22], which we call TAS here, using the performance profiles of Dolan and Moré [10]. The stopping criterion of all algorithms is $\epsilon < 10^{-6}$. From the pictures below, we can conclude that LSCDCC algorithm behaves similar to or better than CCOMB, LSCD, GN, HuS, HYBRID and TAS.



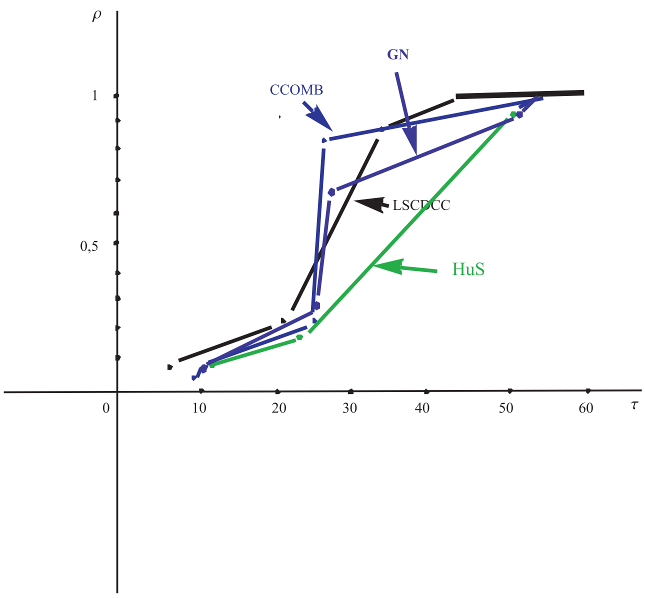
PICTURE 1. (n=50)



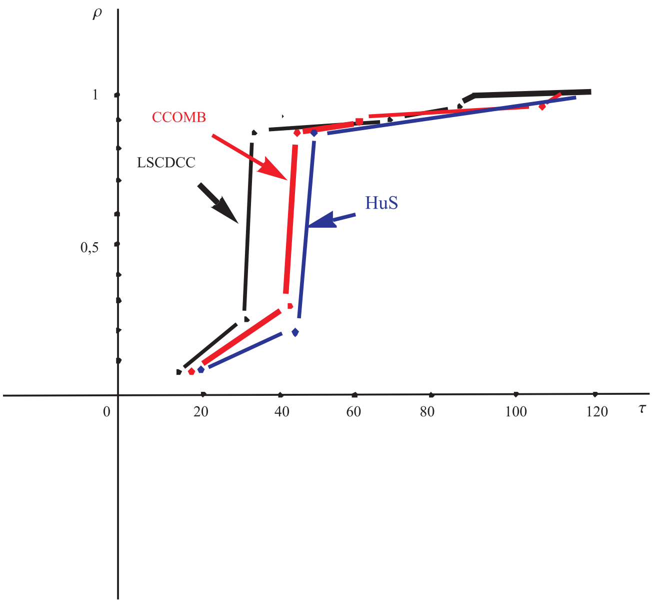
PICTURE 2. (n=70)



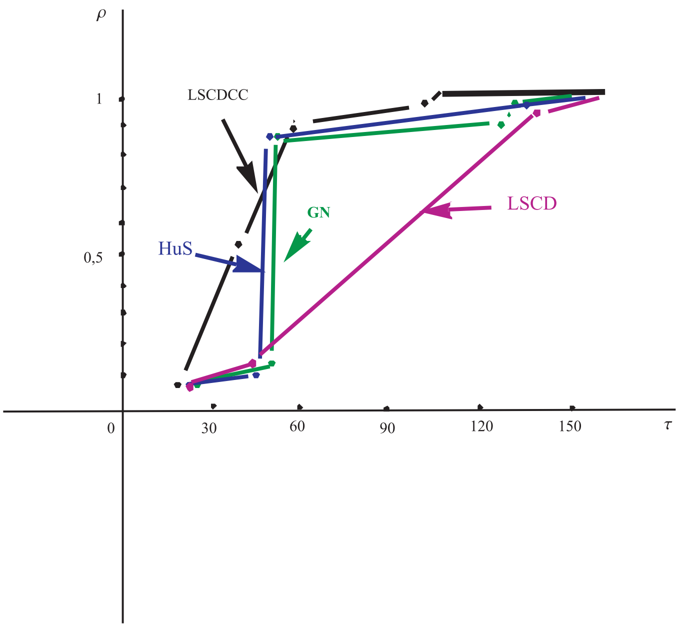
PICTURE 3 (n=80)



PICTURE 4 (n=90)



PICTURE 5 (n=110)



PICTURE 6 (n=120)

Further, these methods are compared for $n = 1000$, $n = 5000$ and $n = 10000$ where the criterion is CPU time again, and the corresponding results (the average CPU times of these methods) are given in the next table.

Table 1.

n	1000	5000	10000
LSCDCC	$4.2 \cdot 10^3$	$109,1 \cdot 10^3$	$438,1 \cdot 10^3$
LSCD	$5.2 \cdot 10^3$	$139.9 \cdot 10^3$	$564.7 \cdot 10^3$
HSDY	$5 \cdot 10^3$	$128.7 \cdot 10^3$	$516.4 \cdot 10^3$
CCOMB	$4.9 \cdot 10^3$	$124.2 \cdot 10^3$	$498.2 \cdot 10^3$
HuS	$5 \cdot 10^3$	$129.2 \cdot 10^3$	$518.3 \cdot 10^3$
TAS	$4.9 \cdot 10^3$	$126,7 \cdot 10^3$	$508.4 \cdot 10^3$
GN	$5 \cdot 10^3$	$127.2 \cdot 10^3$	$510,4 \cdot 10^3$

From Table 1, we see that our method LSCDCC is comparable with the mentioned methods.

References

[1] N. Andrei, *40 Conjugate Gradient Algorithms for unconstrained optimization, A survey on their definition*, ICI Technical Report, 13/08, 2008.

[2] N. Andrei, *New hybrid conjugate gradient algorithms for unconstrained optimization*, Encyclopedia of Optimization, 2560-2571, 2009.

[3] N. Andrei, *A hybrid conjugate gradient algorithm for unconstrained optimization as a convex combination of Hestenes-Stiefel and Dai-Yuan*, Studies in Informatics and Control, 17, 1 (2008), 55-70.

[4] N. Andrei, *Another hybrid conjugate gradient algorithm for unconstrained optimization*, Numerical Algorithms, 47, 2 (2008), 143-156.

[5] N. Andrei, *An unconstrained optimization test functions*, Advanced Modeling and Optimization. An Electronic International Journal, 10 (2008), 147-161.

[6] Y.H. Dai, Y.Yuan, *Convergence properties of the Fletcher-Reeves method*, IMA J. Numer. Anal., 16 (1996), 155-164.

[7] Y. H. Dai, Y. Yuan, *A nonlinear conjugate gradient method with a strong global convergence property*, SIAM J. Optim., 10 (1999), 177-182.

[8] Y.H.Dai, Han J.Y., Liu G.H., Sun D.F., Yin X., Yuan Y., *Convergence properties of nonlinear conjugate gradient methods*, SIAM Journal on Optimization, 10 (1999), 348-358.

[9] S. S. Đorđević, *New hybrid conjugate gradient method as a convex combination of FR and PRP methods*, Filomat, in press.

[10] E. D. Dolan, J. J. Moré, *Benchmarking optimization software with performance profiles*, Math. Programming, 91 (2002), 201-213.

[11] R. Fletcher, *Practical methods of optimization vol. 1: Unconstrained Optimization*, John Wiley and Sons, New York, 1987.

[12] R. Fletcher and C. Reeves, *Function minimization by conjugate gradients*, Comput. J., 7 (1964), 149-154.

[13] J. C. Gilbert, J. Nocedal, *Global convergence properties of conjugate gradient methods for optimization*, SIAM Journal of Optimization, 2 (1992), 21-42.

[14] W.W. Hager and H. Zhang, *A survey of nonlinear conjugate gradient methods*, Pacific journal of Optimization, 2 (2006), 35-58.

[15] M. R. Hestenes, E. L. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Research Nat. Bur. Standards, 49 (1952), 409-436.

[16] Y. F. Hu and C. Storey, *Global convergence result for conjugate gradient methods*, J. Optim. Theory Appl., 71 (1991), 399-405.

[17] Y. Liu and C. Storey, *Efficient generalized conjugate gradient algorithms, Part 1: Theory*, JOTA, 69 (1991), 129-137.

[18] J. Nocedal, S. J. Wright, *Numerical Optimization*, Springer, 1999.

[19] E. Polak, G. Ribiére, *Note sur la convergence de méthodes de directions conjugués*, Revue Française d'Informatique et de Recherche Opérationnelle, 16 (1969), 35-43.

[20] B. T. Polyak, *The conjugate gradient method in extreme problems*, USSR Comp. Math. Math. Phys., 9 (1969), 94-112.

[21] M. J. D. Powell, *Restart procedures of the conjugate gradient method*, Mathematical Programming, 2 (1977), 241-254. Numerical Analysis (Dundee, 1983), Lecture Notes in Mathematics, Vol. 1066, Springer-Verlag, Berlin, 1984, 122-141.

[22] D. Touati-Ahmed, C. Storey, *Efficient hybrid conjugate gradient techniques*, J. Optim. Theory Appl., 64 (1990), 379-397.

[23] X. Yang, Z. Luo, X. Dai, *A Global Convergence of LS-CD Hybrid Conjugate Gradient Method*, Advances in Numerical Analysis, 2013 (2013), Article ID 517452, 5 pages.

[24] P. Wolfe, *Convergence conditions for ascent methods*, SIAM Review, 11 (1969), 226-235.

[25] P. Wolfe, *Convergence conditions for ascent methods. II: Some corrections*, SIAM Review, 11 (1969), 226-235.

[26] G. Zoutendijk, *Nonlinear programming, computational methods*, in Integer and Nonlinear Programming, J. Abadie, ed., North-Holland, Amsterdam, (1970), 37-86.