Journal of Applied Mathematics and Computation (JAMC), 2018, 2(9), 366-378
https://www.hillpublisher.com/journals/JAMC/
ISSN Online:2576-0645
ISSN Print:2576-0653

New Hybrid Conjugate Gradient Method as A Convex Combination of HS and FR Methods

Snezana S. Djordjevic

[1]Faculty of Technology, University of Nis, 16000 Leskovac, Serbia

How to cite this paper: Snezana, S.D. (2018) New Hybrid Conjugate Gradient Method as A Convex Combination of HS and FR Methods. *Journal of Applied Mathematics and Computation*, **2(9)**, 366-378.
DOI: 10.26855/jamc.2018.09.002

***Corresponding author**: Snezana S. Djordjevic, Faculty of Technology, University of Nis, 16000 Leskovac, Serbia
Email: snezanadjordjevic1971@gmail.com

Abstract

In this paper we present a new hybrid conjugate gradient algorithm for unconstrained optimization. This method is a convex combination of Hestenes-Stiefel conjugate gradient method and Fletcher-Reeves conjugate gradient method. The parameter θ_k is chosen in such a way that the search direction satisfies the condition of the Newton direction. The strong Wolfe line search conditions are used. The global convergence of new method is proved.

Numerical comparisons show that the present hybrid conjugate gradient algorithm is the efficient one.

Keywords

Conjugate gradient method; Convex combination; Newton direction.

2010 Mathematics Subject Classification: 90C30.

1. Introduction.

We consider the nonlinear unconstrained optimization problem

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

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

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 [12].

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

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

where $t_k \succ 0$ is a step size, received from the line search, and the directions d_k are given by

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

In the last relation, β_k is the conjugate gradient parameter, $g_k = \nabla f(x_k)$.

Now, we denote

$$y_k = g_{k+1} - g_k. \quad (1.4)$$

An excellent survey of conjugate gradient methods is given by Hager and Zhang [20]. Different conjugate gradient methods correspond to different values of the scalar parameter β_k, for example, Fletcher-Reeves method (FR) [16], Polak-Ribiere -Polyak method (PRP) [27], Hestenes-Stiefel method (HS) [21], Liu-Storey method (LS) [25], Dai-Yuan method (DY) [11], Conjugate-Descent method (CD) [15], etc.

These methods are identical for a strongly convex quadratic function and the exact line search, but they behave differently for general nonlinear functions and the inexact line search.

FR, DY and CD conjugate gradient methods are characterized by the strong convergence and bad practical behavior. By the other hand, LS, HS and PRP conjugate gradient methods are well-known per its good practical behavior and in the same time, they may not converge in general.

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

Here we consider a convex combination of HS and FR conjugate gradient methods.

The corresponding conjugate gradient parameters are

$$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k} \quad [21] \quad (1.5)$$

and

$$\beta_k^{FR} = \frac{\| g_{k+1} \|^2}{\| g_k \|^2} \quad [16]. \quad (1.6)$$

The first global convergence result of the FR method for an inexact line search was given by Al-Baali [1] in 1985. Under the strong Wolfe conditions with $\sigma \prec \frac{1}{2}$, he proved that the FR method generates sufficient descent directions.

As a consequence, global convergence was established using the Zoutendijk condition. For $\sigma = \frac{1}{2}$, d_k is a descent direction, but this analysis did not establish sufficient descent.

Liu et al. [24] extended the global convergence proof of Al-Baali to the case $\sigma = \frac{1}{2}$.

Further, Dai and Yuan [9] showed that in the consecutive FR iterations, at least one iteration satisfies the sufficient descent property, i.e.

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

On the other hand, HS method, as well as LS and PRP methods, possess a built-in restart feature that addresses the jamming problem: when the step $x_{k+1} - x_k$ is small, the factor $y_k = g_{k+1} - g_k$ in the numerator of β_k^{HS} tends to zero. Hence, β_k^{HS} becomes small and the new search direction d_{k+1} is essentially the steepest descent direction $-g_{k+1}$.

The HS method has the property that the conjugacy condition $d_{k+1}^T y_k = 0$ always holds, independently on the line search.

Also, $\beta_k^{HS} = \beta_k^{PRP}$ holds for an exact line search.

So, the convergence properties of the HS method should be similar to the convergence properties of the PRP method. Especially, by Powell's example [29] the HS method with an exact line search may not converge for a general nonlinear function. The fact is that if the search directions satisfy the sufficient descent condition and if a standard Wolfe line search is employed, then the HS method satisfies Property (*). Similarly to the PRP+ method, we can let $\beta_k^{HS+} = \max\{\beta_k^{HS}, 0\}$, then it follows [20] that the HS+ method is globally convergent.

Generally, the performance of HS, LS and PRP methods is better than the performance of methods with $\|g_{k+1}\|^2$ in the numerator of β_k.

2. Convex Combination

The parameter β_k in the presented method, denoted as β_k^{hyb}, is computed as a convex combination of β_k^{HS} and β_k^{FR}, i.e.

$$\beta_k^{hyb} = (1-\theta_k)\beta_k^{HS} + \theta_k \beta_k^{FR}. \tag{2.1}$$

So, we can write

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

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:

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

and

$$\left| g(x_k + t_k d_k)^T d_k \right| \le -\sigma g_k^T d_k \tag{2.4}$$

Obviously, if $\theta_k = 0$, then $\beta_k^{hyb} = \beta_k^{HS}$, and if $\theta_k = 1$, then $\beta_k^{hyb} = \beta_k^{FR}$.

On the other side, if $0 \prec \theta_k \prec 1$, then β_k^{hyb} is a proper convex combination of the parameters β_k^{FR} and β_k^{HS}.

Having in view the relations (1.5) and (1.6), the relation (2.1) becomes

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

$$+\theta_k \cdot \frac{\| g_{k+1} \|^2}{\| g_k \|^2}, \tag{2.6}$$

so the relation (2.2) becomes

$$d_0 = -g_0, \tag{2.7}$$

$$d_{k+1} = -g_{k+1} + (1-\theta_k) \cdot \frac{g_{k+1}^T y_k}{d_k^T y_k} \cdot d_k + \theta_k \cdot \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \cdot d_k \tag{2.8}$$

We shall choose the value of the parameter θ_k in such a way that the search direction d_k is the Newton direction.

We prove that under the strong Wolfe line search, our algorithm is globally convergent. According to the numerical experiments, our algorithm is efficient.

The paper is organized as follows. In the next section, we find the value of the parameter θ_k and we give the algorithm of our method. We also consider the sufficient descent property. In Section 3, we establish the global convergence of our method. The numerical results are given in Section 4.

3. A New Hybrid Conjugate Gradient Algorithm

We remind to the fact that the Newton method has the quadratical convergence property for solving unconstrained optimization problems depending in part on its search direction.

So, assuming that $\nabla^2 f(x)^{-1}$ exists at each iterative point for the objective function f, we are going to choose the parameter θ_k such that the search direction d_{k+1}, defined by (2.8), satisfies the condition of the Newton direction, i.e.,

$$-\nabla^2 f(x_{k+1})^{-1} g_{k+1} = -g_{k+1} + \beta_k^{hyb} d_k \tag{3.1}$$

This is the idea introduced in [6] and [8].

Using β_k^{hyb} from (2.5)-(2.6) and multiplying (3.1) by $s_k^T \cdot \nabla^2 f(x_{k+1})$ from the left, we get

$$-s_k^T g_{k+1} = -s_k^T \cdot \nabla^2 f(x_{k+1}) g_{k+1} + (1-\theta_k) \cdot \beta_k^{HS} s_k^T \nabla^2 f(x_{k+1}) \cdot d_k + \theta_k \beta_k^{FR} s_k^T \nabla^2 f(x_{k+1}) \cdot d_k .$$

Further, we use the secant condition $\nabla^2 f(x_{k+1}) s_k = y_k$ and we get

$$-s_k{}^T g_{k+1} = -y_k{}^T g_{k+1} + (1-\theta_k)\beta_k^{HS} y_k{}^T d_k \tag{3.2}$$

$$+\theta_k \beta_k^{FR} y_k{}^T d_k . \tag{3.3}$$

From (3.2)-(3.3) we get the value for θ_k, which we denote by θ_k^{NT}:

$$\theta_k^{NT} = \frac{-s_k{}^T g_{k+1} + y_k{}^T g_{k+1} - \beta_k{}^{HS} y_k{}^T d_k}{(\beta_k{}^{FR} - \beta_k{}^{HS}) y_k{}^T d_k}, \tag{3.4}$$

i.e.

$$\theta_k^{NT} = \frac{(-s_k{}^T g_{k+1}) \| g_k \|^2}{(-g_{k+1}{}^T y_k) \| g_k \|^2 + (y_k{}^T d_k) \| g_{k+1} \|^2}. \tag{3.5}$$

HHSFR method

Step 1. Select $x_0 \in Dom(f)$, $\varepsilon \succ 0$. Let $k = 0$.

Calculate $f_0 = f(x_0)$, $g_0 = \nabla f(x_0)$.

Set $d_0 = -g_0$, and the initial guess $t_0 = 1$.

Step 2. Test a criterion for stopping iterations, for example,

if

$$\| g_k \| \le \varepsilon, \quad \ldots\ldots\ldots\ldots \tag{3.6}$$

then stop.

Step 3. Compute t_k by the strong Wolfe line search, i.e., t_k satisfies

$$f(x_k + t_k d_k) - f(x_k) \le \delta t_k g_k{}^T d_k, \quad \ldots\ldots\ldots\ldots \tag{3.7}$$

$$| g(x_k + t_k d_k)^T d_k | \le -\sigma g_k{}^T d_k, \quad \ldots\ldots\ldots\ldots \tag{3.8}$$

where $0 \prec \delta \prec \sigma \prec 0.5.$

Step 4. Let $x_{k+1} = x_k + t_k d_k$, $g_{k+1} = g(x_{k+1})$. Compute s_k and y_k.

Step 5. If $(-g_{k+1}{}^T y_k) \| g_k \|^2 + \| y_k{}^T d_k g_{k+1} \|^2 = 0,$

then $\theta_k = 0,$ else compute $\theta_k = \theta_k{}^{NT}$.

Step 6. If $\theta_k \le 0,$ then compute $\beta_k = \beta_k^{HS}$.

If $\theta_k \ge 1,$ then compute $\beta_k = \beta_k^{FR}$.

If $0 < \theta_k < 1,$ then compute β_k by (2.5)-(2.6).

Step 7. Generate $d = -g_{k+1} + \beta_k d_k$. If the restart criterion of Powell

$$| g_{k+1}{}^T g_k | \ge 0.2 \| g_{k+1} \|^2 \quad .... \tag{3.9}$$

is satisfied, then set $d_{k+1} = -g_{k+1}$, otherwise define $d_{k+1} = d$.

Set the initial guess $t_k = 1$.

Step 8. $k = k+1,$ go to Step 2.

The next theorem claims that HHSFR method satisfies the sufficient descent condition.

Theorem 3.1. Let the sequences $\{g_k\}$ and $\{d_k\}$ be generated by HHSFR method. Then the search direction satisfies the sufficient descent condition:

$$g_k{}^T d_k \le -c \| g_k \|^2, \quad \textit{for all} \quad k, \quad .... \tag{3.10}$$

where $\sigma \prec \dfrac{5}{11}$.

Proof. From HHSFR method, we know that if the restart criterion (3.9) holds, then $d_k = -g_k$ and (3.10) holds. So, we assume that (3.9) doesn't hold. Then we have

$$| g_{k+1}{}^T g_k | \prec 0.2 \| g_{k+1} \|^2. \quad .... \tag{3.11}$$

The following proof is by induction.

For $k = 0,$ $g_0{}^T d_0 = -\| g_0 \|^2$, so (3.10) holds.

Next it holds

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

i.e.

$$d_{k+1} = -g_{k+1} + ((1-\theta_k)\beta_k^{HS} + \theta_k \beta_k^{FR}) d_k.$$

We can write

$$d_{k+1} = -(\theta_k g_{k+1} + (1-\theta_k) g_{k+1}) + ((1-\theta_k)\beta_k^{HS} + \theta_k \beta_k^{FR}) d_k.$$

It follows that

$$d_{k+1} = \theta_k(-g_{k+1} + \beta_k^{FR} d_k) + (1-\theta_k)(-g_{k+1} + \beta_k^{HS} d_k),$$

wherefrom

$$d_{k+1} = \theta_k d_{k+1}^{FR} + (1-\theta_k) d_{k+1}^{HS}. \tag{3.12}$$

Multiplying (3.12) by g_{k+1}^T from the left, we get

$$g_{k+1}^T d_{k+1} = \theta_k g_{k+1}^T d_{k+1}^{FR} + (1-\theta_k) g_{k+1}^T d_{k+1}^{HS}. \quad (3.13)$$

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

$$d_{K+1}^{HS} = -g_{k+1} + \beta_k^{HS} d_k .$$

$$\Rightarrow g_{k+1}^T d_{k+1}^{HS} = -\| g_{k+1} \|^2 + \frac{(g_{k+1}^T y_k)(g_{k+1}^T d_k)}{d_k^T y_k}. \quad (3.14)$$

Let denote $T = \left| \frac{(g_{k+1}^T y_k)(g_{k+1}^T d_k)}{d_k^T y_k} \right|$.

From the second strong Wolfe line search condition, it holds

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

i.e.

$$\sigma g_k^T d_k \le g_{k+1}^T d_k \le -\sigma g_k^T d_k.$$

Now

$$y_k^T d_k \ge \sigma g_k^T d_k - g_k^T d_k = -(1-\sigma) g_k^T d_k.$$

So,

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

Now

$$T \le \frac{\sigma}{1-\sigma} \left| g_{k+1}^T y_k \right| \le \frac{\sigma}{1-\sigma} \| g_{k+1} \|^2 + \frac{\sigma}{1-\sigma} \left| g_{k+1}^T g_k \right|.$$

If (3.9) 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 (3.9) doesn't hold, then

$$T \le \frac{\sigma}{1-\sigma} \| g_{k+1} \|^2 + 0.2 \frac{\sigma}{1-\sigma} \| g_{k+1} \|^2 = 1{,}2 \frac{\sigma}{1-\sigma} \| g_{k+1} \|^2 .$$

Now it holds

$$g_{k+1}^T d_{k+1}^{HS} \le -\| g_{k+1} \|^2 (1 - 1.2 \frac{\sigma}{1-\sigma}) = -\frac{1-2{,}2\sigma}{1-\sigma} \| g_{k+1} \|^2 .$$

So, it should hold

$$1 - 2.2\sigma \succ 0,$$

wherefrom

$$\sigma < \frac{5}{11}.$$

So, we have

$$g_{k+1}^T d_{k+1}^{HS} \le -c_1 \| g_{k+1} \|^2, \quad c_1 = \frac{1-2.2\sigma}{1-\sigma}, \quad \sigma \prec \frac{5}{11}. \quad (3.15)$$

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

Let's remind to the fact that the sufficient descent condition holds for FR method in the presence of the strong Wolfe conditions, and this fact is mentioned in [20].

So, there exists a constant $c_2 \succ 0,$ such that

$$g_{k+1}^T d_{k+1}^{FR} \leq -c_2 \|g_{k+1}\|^2 \cdot \tag{3.16}$$

Now suppose that

$$0 \prec \theta_k \prec 1, \text{ i.e., } 0 \prec a_1 \leq \theta_k \leq a_2 \prec 1 \cdot$$

From the relation (3.13), now we conclude

$$g_{k+1}^T d_{k+1} \leq a_1 g_{k+1}^T d_{k+1}^{FR} + (1-a_2) g_{k+1}^T d_{k+1}^{HS} \cdot$$

Denote $c = a_1 c_2 + (1-a_2)c_1;$ then we finally get

$$g_{k+1}^T d_{k+1} \leq -c \|g_{k+1}\|^2 . \tag{3.18}$$

4. Convergence Analysis

For further considerations, we need the following assumptions.

(i) The level set $S = \{x \in R^n : \ f(x) \leq f(x_0)\}$ is bounded, i.e. there exists a constant $B \prec \infty$, such that

$$\|x\| \leq B, \text{ for all } x \in S .$$

(ii) In a neighborhood N of S the function f is continuously differentiable and its gradient $\nabla f(x)$ is Lipschitz continuous, i.e. there exists a constant $0 \prec L \prec \infty$ such that

$$\| f(x) - f(y) \| \leq L \|x - y\| \text{ for all } x, y \in N. \tag{4.1}$$

Under these assumptions, there exists the constant $\Gamma \geq 0$, such that

$$\|g(x) \leq\| \Gamma , \tag{4.2}$$

for all $x \in S$. [4].

Lemma 4.1. [11] Let assumptions (i) and (ii) 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

$$\sum_{k \geq 1} \frac{1}{\|d_k\|^2} = \infty , \tag{4.3}$$

then

$$\lim_{k\to\infty} \inf \|g_k\| = 0. \tag{4.4}$$

Theorem 4.1. Consider the iterative method, defined by HHSFR method. Let the conditions of Theorem 3.1. hold. Assume that the assumptions (i) and (ii) hold. Then either $g_k = 0$, for some k, or

$$\lim_{k\to\infty} \inf \|g_k\| = 0. \tag{4.5}$$

Proof. Suppose that $g_k \neq 0$, for all k. Then we have to prove (4.5).

Suppose, on the contrary, that (4.5) doesn't hold. Then there exists a constant $r \succ 0$, such that

$$\|g_k\| \geq r , \text{ for all } k \tag{4.6}$$

Let D be the diameter of the level set S .

From (2.5)-(2.6) we get

$$\left|\beta_k^{hyb}\right| \le \left|\beta_k^{HS}\right| + \left|\beta_k^{FR}\right| \tag{4.7}$$

$$= \left|\frac{g_{k+1}^T y_k}{d_k^T y_k}\right| + \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \tag{4.8}$$

Further, using the second Wolfe line search condition, we can get

$$y_k^T d_k = g_{k+1}^T d_k - g_k^T d_k \tag{4.9}$$

$$\ge (\sigma - 1) g_k^T d_k = -(1-\sigma) g_k^T d_k (\succ 0). \tag{4.10}$$

So,

$$\frac{1}{y_k^T d_k} \le \frac{1}{-(1-\sigma) g_k^T d_k}. \tag{4.11}$$

But, all conditions of Theorem 3.1. are satisfied, so it holds:

$$g_k^T d_k \le -c\|g_k\|^2 \tag{4.12}$$

$$-g_k^T d_k \ge c\|g_k\|^2 \tag{4.13}$$

$$\frac{1}{-g_k^T d_k} \le \frac{1}{c\|g_k\|^2} \tag{4.14}$$

Now, using (4.6), we get

$$\frac{1}{-g_k^T d_k} \le \frac{1}{cr^2}. \tag{4.15}$$

So,

$$\frac{1}{y_k^T d_k} \le \frac{1}{-(1-\sigma) g_k^T d_k} \le \frac{1}{(1-\sigma) cr^2}. \tag{4.16}$$

On the other hand, using Lipschitz condition (4.1), we have

$$\left|g_{k+1}^T y_k\right| \le \Gamma L \|s_k\| \tag{4.17}$$

$$\le \Gamma L D, \tag{4.18}$$

where D is a diameter of the level set S.

Now, we get

$$\left|\beta_k^{hyb}\right| \le \frac{\Gamma L D}{(1-\sigma) cr^2} + \frac{\Gamma^2}{r^2} = A. \tag{4.19}$$

Next, we are going to prove that there exists $t_* \succ 0$, such that

$$t_k \ge t_* \succ 0, \text{ for all } k. \tag{4.20}$$

Suppose, on the contrary, that there doesn't exist any t_*, such that $t_k \ge t_* \succ 0$.

Then there exists an infinite subsequence $t_k = \beta^{j_k}$, $k \in K_1$ such that

$$\lim_{k \in K_1} t_k = 0 \cdot \tag{4.21}$$

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

$$f(x_k+\beta^{j_k}d_k)-f(x_k)\le \delta\beta^{j_k}g_k^T d_k, \tag{4.22}$$

$$f(x_k+\beta^{j_k-1}d_k)-f(x_k)\succ \delta\beta^{j_k-1}g_k^T d_k. \tag{4.23}$$

Remind that $\delta \prec 1$. From (4.23), we have

$$\frac{f(x_k+\beta^{j_k-1}d_k)-f(x_k)}{\beta^{j_k-1}}\succ \delta g_k^T d_k. \tag{4.24}$$

From (4.24), we conclude that

$$g_k^T d_k \ge \delta g_k^T d_k. \tag{4.25}$$

But, HHSFR method satisfies the sufficient descent, so $g_k^T d_k \le 0$.

Also, $\delta \prec 1$. So, the relation (4.25) 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

$$\|d_{k+1}\| \le \|g_{k+1}\| + \left|\beta_k^{hyb}\right| \|d_k\|. \tag{4.26}$$

It holds $s_k = t_k d_k$, so we can write $d_k = \dfrac{s_k}{t_k}$. So, from (4.26), we have

$$\|d_{k+1}\| \le \Gamma + A\frac{\|s_k\|}{t_k}, \tag{4.27}$$

wherefrom

$$\|d_{k+1}\| \le \Gamma + A\frac{D}{t_*} = P. \tag{4.28}$$

Now we get

$$\sum_{k\ge 1}\frac{1}{\|d_{k+1}\|^2}=\infty, \tag{4.29}$$

so, applying Lemma 4.1, we conclude that

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

This is a contradiction with (4.6), so we have proved (4.5).

5. Numerical Experiments

In this section we present the computational performance of a Mathematica 10 implementation of HHSFR algorithm on a set of unconstrained optimization test problems from [7]. Each problem is tested for a number of variables: n=1000, n=3000 and n=5000. The criterion used here is CPU time.

We present comparisons with PRP [27], the algorithm from [22], which we call HuS here, the algorithm from [32], which we denote by LSCDMAX, the algorithm from [18], [19], here denoted by HzbetaN, using the performance profiles of Dolan and Moré[14]. The stopping criterion of all algorithms is given by (3.6), where $\varepsilon = 10^{-6}$.

Considering Figure 1, we can conclude that HHSFR is better than PRP method in the first part of Figure 1, and in the later part of Figure 1, it behaves similar to PRP method. In the last part of Figure 1, it coincides with PRP method.

From Figure 2 we see that although there are intervals inside which HHSFR method isn't the best method, there exists the part of the graphic with the best behavior of HHSFR method.

Finally, looking into Figure 3, we can see that there exist significant parts of this graphic, inside which our method is better than the others or similar to them.

Generally, from Figure 1, Figure 2 and Figure 3, which are presented below, we can conclude that HHSFR algorithm is comparable to the other algorithms.

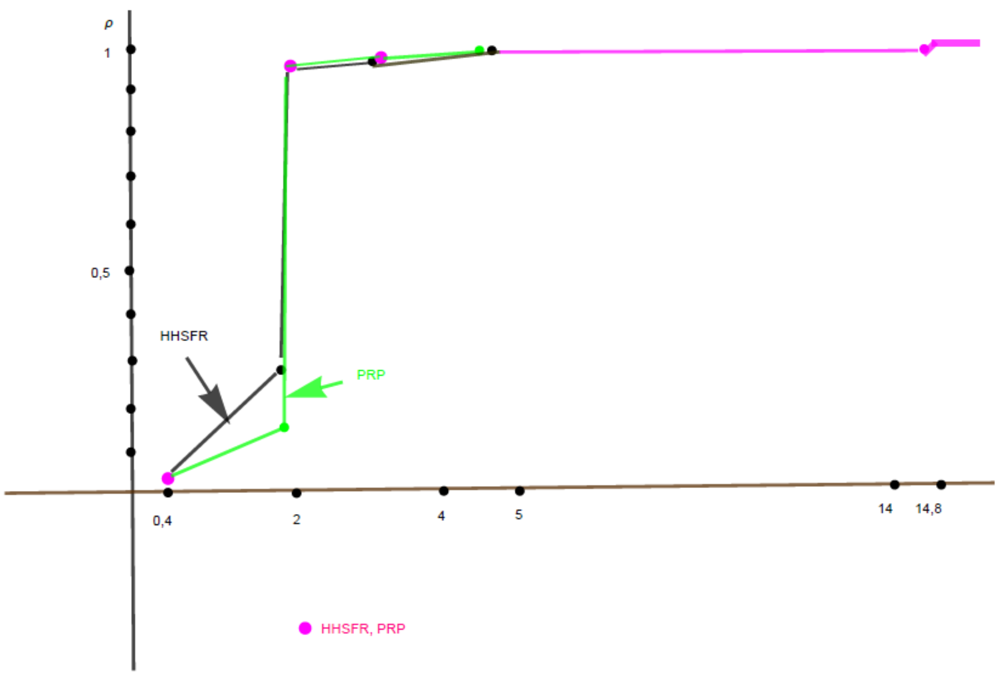

Figure 1. (n = 1000) : Comparison between HHSFR and PRP CG methods

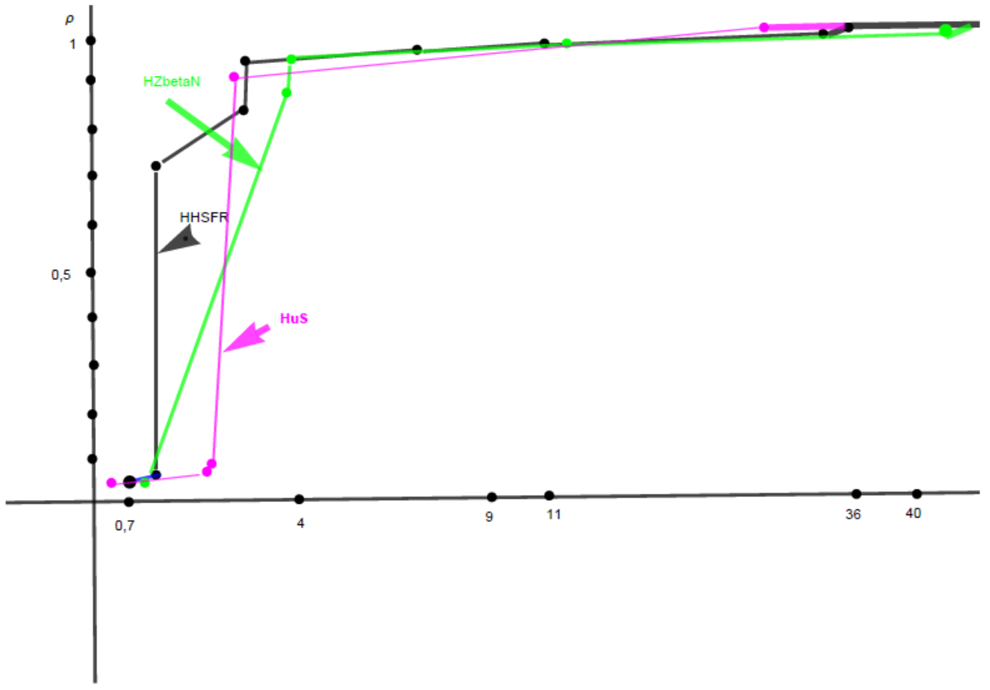

Figure 2. (n = 3000) : Comparison among HHSFR, HZbetaN and HuS CG methods

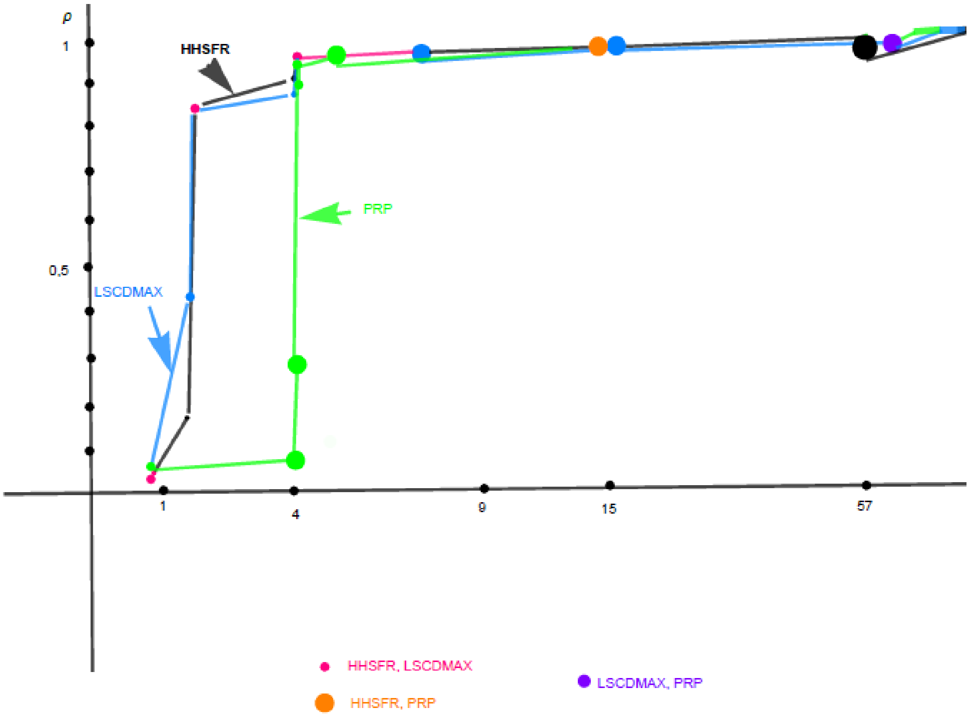

Figure 3. (n = 5000) : Comparison among HHSFR, LSCDMAXand PRP CG methods

References

[1] M. Al-Baali, Descent property and global convergence of the Fletcher-Reeves method with inexact line search, IMA J. Numer. Anal., 5 (1985), 121-124.
[2] N. Andrei, 40 Conjugate Gradient Algorithms for unconstrained optimization, A survey on their definition, ICI Technical Report, 13/08, 2008.
[3] N. Andrei, New hybrid conjugate gradient algorithms for unconstrained optimization, Encyclopedia of Optimization, 2560-2571, 2009.
[4] N. Andrei, A Hybrid Conjugate Gradient Algorithm with Modified Secant Condition for Unconstrained Optimization as a Convex Combination of Hestenes-Stiefel and Dai-Yuan Algorithms, STUDIES IN INFORMATICS AND CONTROL, 17, 4 (2008), 373-392.
[5] 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.
[6] N. Andrei, Another hybrid conjugate gradient algorithm for unconstrained optimization, Numerical Algorithms, 47, 2 (2008), 143-156.
[7] N. Andrei, An unconstrained optimization test functions, Advanced Modeling and Optimization. An Electronic International Journal, 10 (2008), 147-161.
[8] N. Andrei, Accelerated hybrid conjugate gradient algorithm with modified secant condition for unconstrained optimization, Numer. Algorithms 54 (2010) 23-46.
[9] Y.H. Dai, Y.Yuan, Convergence properties of the Fletcher-Reeves method, IMA J. Numer. Anal., 16 (1996), 155-164.
[10] Y. H. Dai,, L. Z. Liao, New conjugacy conditions and related nonlinear conjugate gradient methods, Appl. Math. Optim. 43 (2001), 87-101.
[11] Y. H. Dai, Y. Yuan, A nonlinear conjugate gradient method with a strong global convergence property, SIAM J. Optim., 10 (1999), 177-182.
[12] 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.
[13] S. S. Đorđević, New hybrid conjugate gradient method as a convex combination of FR and PRP methods, Filomat, 30:11 (2016), 3083-3100.
[14] E. D. Dolan, J. J. Moré, Benchmarking optimization software with performance profiles, Math. Programming, 91 (2002), 201-213.
[15] R. Fletcher, Practical methods of optimization vol. 1: Unconstrained Optimization, John Wiley and Sons, New York, 1987.
[16] R. Fletcher and C. Reeves, Function minimization by conjugate gradients, Comput. J., 7 (1964), 149-154.
[17] J. C. Gilbert, J. Nocedal, Global convergence properties of conjugate gradient methods for optimization, SIAM Journal of Optimization, 2 (1992), 21-42.
[18] W. W. Hager, H. Zhang, A new conjugate gradient method with guaranteed descent and an efficient line search, SIAM J. Optim., 16, 1 (2003), 170-192.
[19] W. W. Hager, H. Zhang, CG-DESCENT, a conjugate gradient method with guaranteed descent, ACM Transactions on Mathematical Software, 32, 1 (2006), 113-137.
[20] W.W. Hager and H. Zhang, A survey of nonlinear conjugate gradient methods, Pacific journal of Optimization, 2 (2006), 35-58.
[21] M. R. Hestenes, E. L. Stiefel, Methods of conjugate gradients for solving linear systems, J. Research Nat. Bur. Standards, 49 (1952), 409-436.
[22] Y. F. Hu and C. Storey, Global convergence result for conjugate gradient methods, J. Optim. Theory Appl., 71 (1991), 399-405.
[23] J.K. Liu, S.J. Li, New hybrid conjugate gradient method for unconstrained optimization, Applied Mathematics and Computation, 245 (2014), 36-43.
[24] G. H. Liu, J. Y. Han and H. X. Yin, Global convergence of the Fletcher-Reeves algorithm with an inexact line search, Appl. Math. J. Chinese Univ. Ser. B, 10 (1995), 75-82.
[25] Y. Liu and C. Storey, Efficient generalized conjugate gradient algorithms, Part 1: Theory, JOTA, 69 (1991), 129-137.
[26] J. Nocedal, S. J. Wright, Numerical Optimization, Springer, 1999.
[27] E. Polak, G. Ribière, Note sur la convergence de méthodes de directions conjugués, Revue Française d'Informatique et de Recherche Opérationnelle, 16 (1969), 35-43.
[28] B. T. Polyak, The conjugate gradient method in extreme problems, USSR Comp. Math. Math. Phys., 9 (1969), 94-112.

[29] M. J. D. Powell, Nonconvex minimization calculations and the conjugate gradient method, in D.F. Griffiths, ed., Numerical Analysis Lecture Notes in Mathematics 1066 (Springer-Verlag, Berlin, 1984), 122-141.
[30] M. J. D. Powell, Restart procedures of the conjugate gradient method, Mathematical Programming, 2 (1977), 241-254.
[31] D. Touati-Ahmed, C. Storey, Efficient hybrid conjugate gradient techniques, J. Optim. Theory Appl., 64 (1990), 379-397.
[32] 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.
[33] Y. Yuan, J. Stoer, A subspace study on conjugate gradient algorithm, Z. Angew. Math. Mech. 75 (1995), 69-77.
[34] P. Wolfe, Convergence conditions for ascent methods, SIAM Review, 11 (1969), 226-235.
[35] P. Wolfe, Convergence conditions for ascent methods. II: Some corrections, SIAM Review, 11 (1969), 226-235.
[36] G. Zoutendijk, Nonlinear programming, computational methods, in Integer and Nonlinear Programming, J. Abadie, ed., North-Holland, Amsterdam, (1970), 37-86.