An optimal parameter for the Generalized Descent Symmetrical Hestenes-Stiefel algorithm

Noureddine Benrabia

n.benrabia@univ-soukahras.dz

Abstract

We have proposed a Generalized Descent Symmetrical Hestenes-Stiefel algorithm [12] , GDSHS for short, which can generate sufficient descent directions for the objective function. Using the Wolfe line search conditions, the global convergence property of the method is also obtained based on the spectral analysis of the conjugate gradient iteration matrix and the Zoutendijk condition for steepest descent methods. I propose in this paper a theoretical choice to improve the performance of the GDSHS algorithm, by the use of an optimal parameter. Based on this, some descent algorithms are developed. 86 numerical experiments are presented to verify their performance and the numerical results show that the new conjugate gradient method GDSHS with the parameter $c = 1$, denoted GDSHS1, is competitive with GDSHS algorithms that have a parameter c chosen in the interval $]0, +\infty[$.

keywords: Conjugate gradient method, Generalized conjugacy condition, Theoretical choice, Global convergence.

MSC Classification: 49M05; 49M30; 90C06; 90C30

1 Introduction

The nonlinear conjugate gradient (NCG) method is one of the most popular methods for solving smooth unconstrained optimization problems due to its simplicity and low requirement. However, the usage of NCG methods are mainly restricted in solving the large-scale problem:

$$\min_{x \in \mathbb{R}^n} f(x). \tag{1}$$

The classical conjugate gradient methods with line searches are as follows:

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

where α_k is the step length of a line search and the directions d_k are given by

$$\begin{cases} d_0 &= -g_0, \\ d_{k+1} &= -g_{k+1} + \beta_k d_k, \quad \forall k > 0, \end{cases} \tag{3}$$

where $g_k = g(x_k) = \bigtriangledown f(x_k)$ and β_k is a scalar. In order to guarantee the global convergence property of the NCG methods, the descent property or the sufficient descent property is necessary and important, namely,[1]

$$d_{k+1}^T g_{k+1} < 0 \qquad \text{(the descent property)} \tag{4}$$

$$d_{k+1}^T g_{k+1} \leq -c_0 \mid\mid g_{k+1} \mid\mid^2, \quad c_0 > 0 \quad \text{(the sufficient descent property).} \tag{5}$$

However, unlike the quasi-Newton methods, in general, the NCG methods may not meet the descent or the sufficient descent property for inexact line searches. efforts have been devoted to investigating the descent property. In 1978, Shanno [2] proposed the memoryless quasi-Newton methods and the selfscaling conjugate gradient algorithms based on Perry's idea [3], the quasi-Newton equation and the self-scaling variable metric technique. Jinwei Wang et. al [4] proposed the Forensics feature analysis in quaternion wavelet domain for distinguishing photographic images and computer graphics. Xuezhi Wen et. al [5] presented A rapid learning algorithm for vehicle classification . In 2016, Yan Kong et. al [6] suggested A belief propagation-based method for Task Allocation in Open and Dynamic Cloud

Environments. Yanhua Zhang et. al [7] presented the Efficient Algorithm for K-Barrier Coverage Based on Integer Linear Programming. Gonglin Yuan, Zehong Meng, and Yong Li [8] suggested A modified Hestenes and Stiefel conjugate gradient algorithm for large-scale nonsmooth minimizations and nonlinear equations. Gonglin Yuan, Maojun Zhang [9] developed A three-terms Polak-*Ribière*-Polyak conjugate gradient algorithm for large-scale nonlinear equations . In 2017, Gonglin Yuan, Zengxin Wei, Xiwen Lu [10] proposed Global convergence of BFGS and PRP methods under a modified weak Wolfe-Powell line search . By applying the symmetrical technique [11] to conjugate gradient methods, a symmetrized conjugate gradient method, satisfies the property (5) for any line search, is introduced in [12] , and this idea can also be applied to other conjugate gradient algorithms.

2 Application of the symmetrical technique to conjugate gradient methods

According to Perry's notation [3], for the HS conjugate gradient method with the CG update parameter β_k ,[13]

$$\beta_k = \frac{y_k^T g_{k+1}}{d_k^T y_k},$$

the line search direction, d_{k+1}, can be rewritten as follows:

$$d_{k+1} = -D_{k+1} g_{k+1}, \tag{6}$$

with,

$$D_{k+1} = (I - \frac{d_k y_k^T}{d_k^T y_k}) = (I - \frac{s_k y_k^T}{s_k^T y_k}), \tag{7}$$

According to Perry's notation [3], for the HS conjugate gradient method with the CG update parameter β_k ,[13]

$$\beta_k = \frac{y_k^T g_{k+1}}{d_k^T y_k},$$

the line search direction, d_{k+1}, can be rewritten as follows:

$$d_{k+1} = -D_{k+1} g_{k+1}, \tag{8}$$

with,

$$D_{k+1} = (I - \frac{d_k y_k^T}{d_k^T y_k}) = (I - \frac{s_k y_k^T}{s_k^T y_k}). \tag{9}$$

Thus, the matrix C_1 can be symmetrized by the matrix C. Then, from the above symmetrizing procedure, we conclud that the conjugate gradient iteration matrix D_{k+1}can be symmetrized by $\bar{D}_{k+1}$as follows:

$$\bar{D}_{k+1} = I - \frac{y_k d_k^T + d_k y_k^T}{d_k^T y_k} + \frac{y_k^T y_k}{(d_k^T y_k)^2} d_k d_k^T = \left(I - \frac{d_k y_k^T}{d_k^T y_k} \right) \left(I - \frac{y_k d_k^T}{d_k^T y_k} \right). \tag{10}$$

Thus, the conjugate gradient directions (3) are rewritten as:

$$\begin{cases} d_0 & = -g_0, \\ d_{k+1} & = -\bar{D}_{k+1} g_{k+1}, \quad \forall k > 0. \end{cases} \tag{11}$$

So d_{k+1} is called *the symmetrical conjugate gradient direction* and $\bar{D}_{k+1}$ is called *the symmetrical conjugate gradient iteration matrix or the symmetrical Hestenes-Stiefel matrix (SHS matrix)* . If the matrix $\bar{D}_{k+1}$ is updated with the rank-1 matrix as follows:

$$\hat{D}_{k+1} = \bar{D}_{k+1} + \frac{s_k \varsigma_k^T}{y_k^T \varsigma_k}, \qquad \forall \varsigma_k \in \mathbb{R}^n$$

then $\hat{D}_{k+1}$ satisfies the quasi-Newton equation, $\hat{D}_{k+1} y_k = s_k$, and under the exact line searches, $d_{k+1} = -\hat{D}_{k+1} g_{k+1}$ satisfies the condition :

$$y_k^T d_{k+1} = 0, \quad \forall k \geq 0 \quad \text{(conjugacy)}. \tag{12}$$

If $\varsigma_k = s_k$, then

$$d_{k+1}^{mBFGS} = -\left(\bar{D}_{k+1} + \frac{s_k s_k^T}{y_k^T s_k} \right) g_{k+1},$$

which just is the formula of the search direction of the memoryless BFGS.

In [12], (11) is substituted by

$$y_k^T d_{k+1} = -\sigma s_k^T g_{k+1}, \tag{13}$$

which is called Dai and Liao conjugacy condition [14] or *the generalized conjugacy condition*, where σ is a parameter. We have supposed that $\bar{D}_{k+1}$ in (10) be updated by a rank one matrix, namely

$$\bar{P}_{k+1} = \bar{D}_{k+1} + uv^T,$$

where u and v two vectors in $\mathbb{R}^n$ such that (12) holds.
Thus, it follows from (9), (12) and $d_{k+1} = -\bar{P}_{k+1} g_{k+1}$ that:

$$\begin{aligned} y_k^T d_{k+1} &= -y_k^T \left(\bar{D}_{k+1} g_{k+1} + uv^T g_{k+1}\right) \\ &= -y_k^T \left(I - \frac{d_k y_k^T}{d_k^T y_k}\right)\left(I - \frac{y_k d_k^T}{d_k^T y_k}\right) g_{k+1} - y_k^T uv^T g_{k+1} \\ &= -\left(y_k^T - y_k^T\right)\left[\left(I - \frac{y_k d_k^T}{d_k^T y_k}\right) g_{k+1}\right] - y_k^T uv^T g_{k+1} \\ &= -y_k^T uv^T g_{k+1} \\ &= -\left(y_k^T u\right) v^T g_{k+1}, \end{aligned}$$

so,

$$-\left(y_k^T u\right) v^T g_{k+1} = -\sigma s_k^T g_{k+1} \Rightarrow \left(\sigma s_k - v y_k^T u\right)^T g_{k+1} = 0.$$

We have selected v such that $v = \dfrac{\sigma s_k}{y_k^T u}$. Hence

$$\bar{P}_{k+1} = \bar{D}_{k+1} + \sigma \frac{u s_k^T}{y_k^T u}, \tag{14}$$

where the vector u is any vector in $\mathbb{R}^n$ such $y_k^T u \neq 0$. The matrix $\bar{P}_{k+1}$ is also called *the symmetrical Hestenes-Stiefel matrix*. So, we have introduced a new line search direction as follows:

$$\begin{cases} d_0 &= -g_0, \\ d_{k+1} &= -\bar{P}_{k+1} g_{k+1}, \quad \forall k > 0, \end{cases} \tag{15}$$

where $\bar{P}_{k+1}$ is defined by (13) with $u = u_k$, i.e.,

$$d_{k+1} = -\bar{D}_{k+1} g_{k+1} - \sigma \frac{u_k s_k^T}{y_k^T u_k} g_{k+1}. \tag{16}$$

Thus, with different σ and $u_k \left(y_k^T u_k \neq 0\right)$, a family of methods can be obtained by (2) and (14) with d_{k+1} defined by (15), which is called *the family of Generalized Symmetrical Hestenes-Stiefel gradient method*, **GSHS**, for short. $\bar{P}_{k+1}$ is also called *the iteration matrix of generalized symmetrical Hestenes-Stiefel gradient method.*
In [12], we have taken $u_k = y_k$ and $\sigma = c\dfrac{y_k^T y_k}{s_k^T y_k}$, $c > 0$. Thus

$$\bar{P}_{k+1} = \bar{D}_{k+1} + c\frac{y_k s_k^T}{s_k^T y_k},$$

and,

$$d_{k+1} = -\bar{D}_{k+1} g_{k+1} - c\frac{d_k^T g_{k+1}}{d_k^T y_k} y_k. \tag{17}$$

We have denoted the iterative scheme (2) and (14) with d_{k+1} calculated by (16) by the **GDSHS**.

3 The sufficient descent property and descent algorithm

In this section, we have considered the sufficient descent property of the **GDSHS** method, that is,

$$d_{k+1}^T g_{k+1} \leq -c_0 \mid\mid g_{k+1} \mid\mid^2, \quad c_0 > 0.$$

By theorem (2.1) in [11] we have $\bar{D}_{k+1}$ is a positive semi-definite matrix and the eigenvalues of this matrix are 0, 1(n - 2 multiplicity) and λ_{max}^{k+1} , respectively, where λ_{max}^{k+1} is the maximum eigenvalue:

$$\lambda_{max}^{k+1} = \frac{\| y_k \|^2 \| d_k \|^2}{(d_k^T y_k)^2}.$$

By (16) it is obtained that:

$$d_{k+1}^T g_{k+1} = -g_{k+1}^T \bar{D}_{k+1} g_{k+1} - c \frac{y_k^T g_{k+1}}{d_k^T y_k} d_k^T g_{k+1},$$

since,

$$g_{k+1}^T \bar{D}_{k+1} g_{k+1} \geq 0,$$

so,

$$d_{k+1}^T g_{k+1} \leq -c \frac{y_k^T g_{k+1}}{d_k^T y_k} d_k^T g_{k+1} \frac{d_k^T y_k}{d_k^T y_k} \quad \Rightarrow$$

$$d_{k+1}^T g_{k+1} \leq -c \frac{((g_{k+1}^T y_k) d_k)^T ((d_k^T y_k) g_{k+1})}{(d_k^T y_k)^2}.$$

From the following inequality,

$$u^T v \leq \frac{1}{2}(a \| u \|^2 + \frac{1}{a} \| v \|^2), \quad \forall a > 0,$$

it can be derived that,

$$\begin{aligned} d_{k+1}^T g_{k+1} &\leq -\frac{c}{2(d_k^T y_k)^2}(a \| (g_{k+1}^T y_k) d_k \|^2 + \frac{1}{a} \| (d_k^T y_k) g_{k+1} \|^2) \\ &\leq -\frac{c}{2(d_k^T y_k)^2}(a (g_{k+1}^T y_k)^2 \| d_k \|^2 + \frac{1}{a} (d_k^T y_k)^2 \| g_{k+1} \|^2) \\ &\leq -\frac{ca \| g_{k+1} \|^2 \| y_k \|^2 \| d_k \|^2}{2(d_k^T y_k)^2} - \frac{c}{2a} \| g_{k+1} \|^2 . \end{aligned}$$

So,

$$d_{k+1}^T g_{k+1} \leq -(ca \frac{\| y_k \|^2 \| d_k \|^2}{2(d_k^T y_k)^2} + \frac{c}{2a}) \| g_{k+1} \|^2 .$$

Thus (5) is true for

$$c_0 = ca \frac{\| y_k \|^2 \| d_k \|^2}{2(d_k^T y_k)^2} + \frac{c}{2a}.$$

From above discussion, we have ,

$$d_{k+1} = -\bar{D}_{k+1} g_{k+1} - c \frac{d_k^T g_{k+1}}{d_k^T y_k} y_k.$$

Thus, the conjugate gradient directions (14) are rewritten as

$$\left\{ \begin{array}{ll} d_0 & = -g_0, \\ d_{k+1} & = -\bar{P}_{k+1} g_{k+1} = -v_{k+1} + \beta_k d_k - c \zeta_k y_k, \end{array} \right. \tag{18}$$

where,

$$v_{k+1} = t_k g_{k+1} + (1 - t_k) g_k = g_{k+1} - (1 - t_k) y_k = g_k + t_k y_k,$$

$$\beta_k = \frac{v_{k+1}^T y_k}{d_k^T y_k} = t_k \frac{g_{k+1}^T y_k}{d_k^T y_k} + (1 - t_k) \frac{g_k^T y_k}{d_k^T y_k},$$

$$\zeta_k = \frac{d_k^T g_{k+1}}{d_k^T y_k},$$

and,

$$t_k = \frac{-d_k^T g_k}{d_k^T y_k}.$$

So, we have obtained *the Generalized Descent Symmetrical Hestenes-Stiefel algorithm*, denoted by **GDSHS**, as follows:
Algorithm 3.1

step 1. Give an initial point x_0 and $\varepsilon \geq 0$. Set $k = 0$.
step 2. Calculate $g_0 = g(x_0)$. If $\| g_k \| \leq \varepsilon$, then stop; otherwise let $d_0 = -g_0$ and continue with **step 3.**.
step 3. Calculate steplength α_k with Wolfe line searches:

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta_1 \alpha_k d_k^T g_k \tag{19}$$

and,

$$d_k^T g(x_k + \alpha_k d_k) \geq \delta_2 d_k^T g_k, \tag{20}$$

where δ_1 and δ_2 are positive constants such that

$$0 < \delta_1 < \delta_2 < 1.$$

step 4. Set $x_{k+1} = x_k + \alpha_k d_k$.
step 5. Calculate $g_{k+1} = g(x_{k+1})$.
step 6. If $\| g_{k+1} \| \leq \varepsilon$, then stop.
step 7. Calculate the direction d_{k+1} via (17). Set $k = k + 1$, then go to **step 3.**.

4 The convergence of the GDSHS algorithm

In this section, to analyze the convergence of **GDSHS** algorithm, first, we have introduced the following assumptions about the objective function $f(x)$.

H1. f is bounded below in $\mathbb{R}^n$ and f is continuously differentiable in a neighborhood $\aleph$ of the level set $\Gamma \stackrel{\text{def}}{=} \{x : f(x) \leq f(x_0)\}$, where x_0 is the starting point of the iteration.

H2. The gradient of f is Lipschitz continuous in $\aleph$, that is, there exists a constant $L > 0$ such that

$$\| \nabla f(\bar{x}) - \nabla f(x) \| \leq L \| \bar{x} - x \|, \qquad \forall \bar{x}, x \in \aleph.$$

Next, we have introduced the spectral condition theorem of the global convergence for an objective function satisfying **H1** and **H2**, which generates Theorem 4.1 in [11].

Theorem 4.1. *Assume that the line search direction of a nonlinear conjugate gradient method satisfies*

$$\begin{cases} d_0 &= -g_0, \\ d_k &= -\bar{P}_k g_k, \quad \forall k > 0. \end{cases} \tag{21}$$

*Let the objective function $f(x)$ satisfy **H1** and **H2**. For a nonlinear conjugate gradient method ((2) and (20)), which satisfies the sufficient descent condition (5), if its line search satisfies the Wolfe conditions (18) and (19), and [15]*

$$\sum_{k=0}^{\infty} \Lambda_k = +\infty, \quad (\textit{the spectral condition}) \tag{22}$$

where Λ_k is the maximum eigenvalue of $\bar{P}_k^T \bar{P}_k$, then:

$$\lim_{x \to \infty} inf \| g_k \| = 0. \tag{23}$$

Moreover, if $\Lambda_k \leq \tilde{\Lambda}$, where $\tilde{\Lambda}$ is a positive constant, then

$$\lim_{x \to \infty} \| g_k \| = 0, \tag{24}$$

see [12].

5 The theoretical choice for parameter c to optimize the GDSHS algorithm

Suppose that $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is twice continuously differentiable , by Taylor's Theorem , which tells us that for any search direction d_k and step-length parameter α, we have :

$$f(x_k + \alpha d_k) = f(x_k) + \alpha d_k^T \nabla f(x_k) + \frac{1}{2}\alpha d_k^T \nabla^2 f(x_k + td_k)d_k, \quad \text{for some } t \in (0, \alpha).$$

The rate of change in f along the direction d_k at x_k is simply the coefficient of α namely, $d_k^T \nabla f(x_k)$.

The unit direction d_k ($\|d_k = 1\|$) of most rapid decrease is the solution to the problem

$$\min_{d \in \mathbb{R}^n} d_k^T \nabla f(x_k) \quad \text{with } \|d\| = 1,$$

and by (16) we have:

$$d_{k+1} = -\bar{D}_{k+1} g_{k+1} - c \frac{d_k^T g_{k+1}}{d_k^T y_k} y_k, \quad c > 0.$$

Let :

$$\Psi(c) = f(-\bar{D}_{k+1} g_{k+1} - c \frac{d_k^T g_{k+1}}{d_k^T y_k} y_k) = f(p_1 + cp_2),$$

such as: $p_1 = -\bar{D}_{k+1} g_{k+1}$ and $p_2 = -\frac{d_k^T g_{k+1}}{d_k^T y_k} y_k$. So, it is naturel to,

$$\min_{c>0} \Psi(c),$$

according to Cauchy's idea, we will try to minimize the derivative of $\Psi(c)$ in 0 .

We have ,

$$\Psi^{'}(0) = p_2^T \nabla f(p_1),$$

and we try to solve the problem :

$$\min_{p_2 \in \mathbb{R}^n} \Psi^{'}(0), \quad \text{with } \|p_2\| = 1.$$

The solution is of course:

$$p_2 = -\frac{\nabla f(p_1)}{\|\nabla f(p_1)\|}.$$

So,

$$d_{k+1} = p_1 - c \frac{\nabla f(p_1)}{\|\nabla f(p_1)\|},$$

and, if $\|d_{k+1}\| = 1$, we will have:

$$1 = \|p_1 - c \frac{\nabla f(p_1)}{\|\nabla f(p_1)\|}\| \Rightarrow c \leq 1 + \|p_1\|.$$

So,

$$c \leq 1 + \|\bar{D}_{k+1}\| \|g_{k+1}\|,$$

and, by Theorem (2.1) in [11] we conclude :

$$0 < c \leq 1 + \lambda_{max}^{k+1} \|g_{k+1}\|.$$

6 Numerical experiment

In this section, we compare the GDSHS algorithms with different values of the parameter c in the interval $]0, +\infty[$, and we also compare the performance of the new conjugate gradient method GDSHS with the parameter $c = 1$, denoted GDSHS1, to the standard FR method and PRP^+ version of the conjugate gradient method developed by Gilbert and Nocedal [16], where the β_k associated with the Polak-Ribiere-Polyak conjugate gradient method [17] is kept nonnegative.

When comparing between the algorithms we have used the backtracking line search . The test problems are the 86 unconstrained problems found in this work, and each test function is made as an experiment with the number of variables being 2, 10, 100, 1000, 2000,..., 10000, respectively.

The starting points used are those given in An Unconstrained Optimization Test Functions Collection [18].

The termination criterion of all algorithms is that $\|g\| < 10^{-7}$. The tests are performed on a PC using a *Pentium Dual-core CPU T4400@2.20GHz, 2.0GB RAM, Mobil Intel 4 Series Express Chipset Family*, using MATLAB codes.

We have adopted the performance profiles of [19] to compare the performance among the tested methods.

6.1 The choice of parameter is far from value c=1

We compare the performance of the new conjugate gradient methods GDSHS with the parameter $c = 0.1$, $c = 1$ and $c = 2$, denoted GDSHS0.1 , GDSHS1 and GDSHS2 respectively.

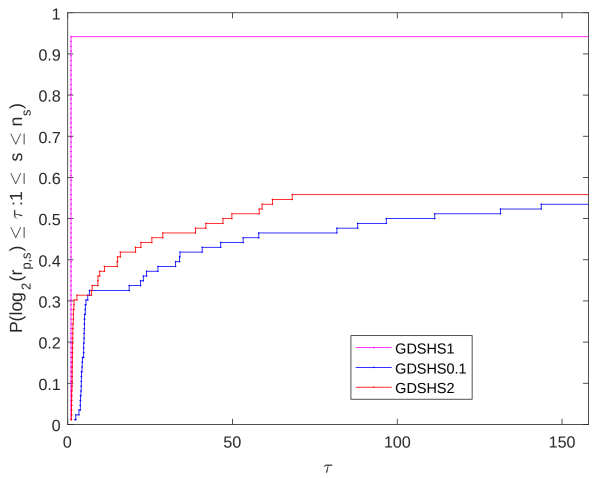

Figure 1: Performance profile by CPU time

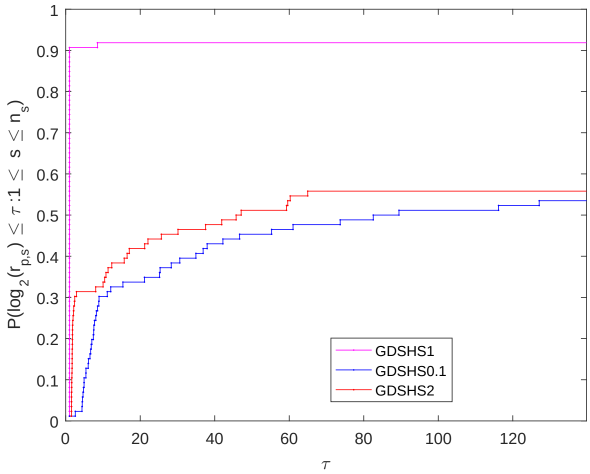

Figure 2: Performance profile by number of iterations

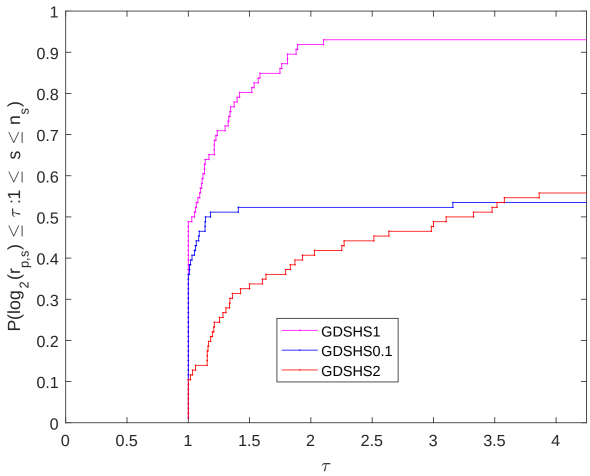

Figure 3: Performance profile by 2-norm of the gradient of the objective function

In Figures 1, 2 and 3, we see that GDSHS0.1 and GDSHS2 algorithms can not be competitive with the GDSHS1 algorithm, especially in performance regarding the the viewpoint of CPU time.

6.2 The choice of the parameter is close to the value c=1

We compare the performance of the new conjugate gradient methods GDSHS with the parameter $c = 0.9$, $c = 1$ and $c = 1.1$, denoted GDSHS0.9 , GDSHS1 and GDSHS1.1 respectively.

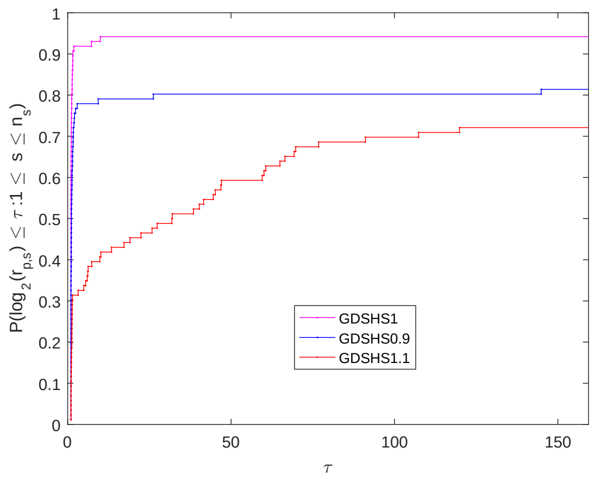

Figure 4: Performance profile by CPU time

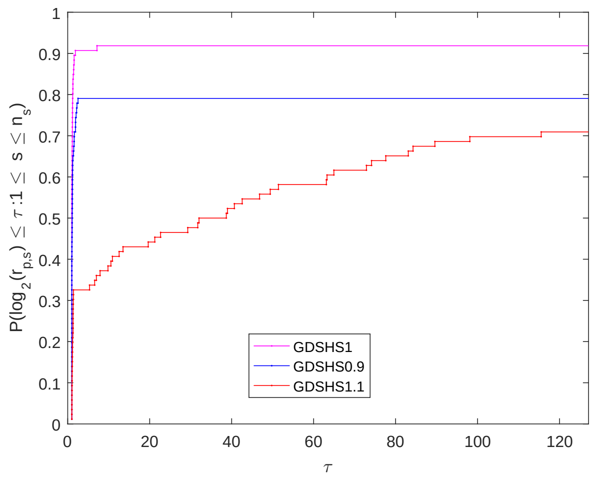

Figure 5: Performance profile by number of iterations

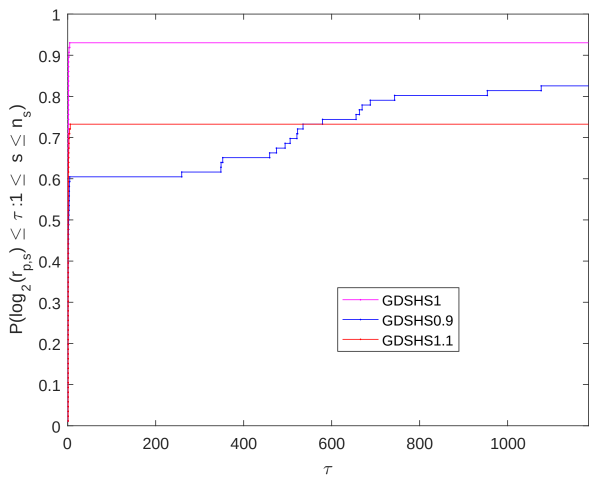

Figure 6: Performance profile by 2-norm of the gradient of the objective function

In Figures 4, 5 and 6, we see that GDSHS1 algorithm is better and more competitive than the GDSHS0.9 and GDSHS1.1 algorithms, especially in performance regarding the number of iterations.

6.3 The comparisons of the best numerical variant of GDSHS versus some other conjugate gradient algorithms

In [12] we have compared the performance of the new conjugate gradient methods GDSHS with the parameter $c = 1$ denoted GDSHS1, FR and PRP^+ respectively.

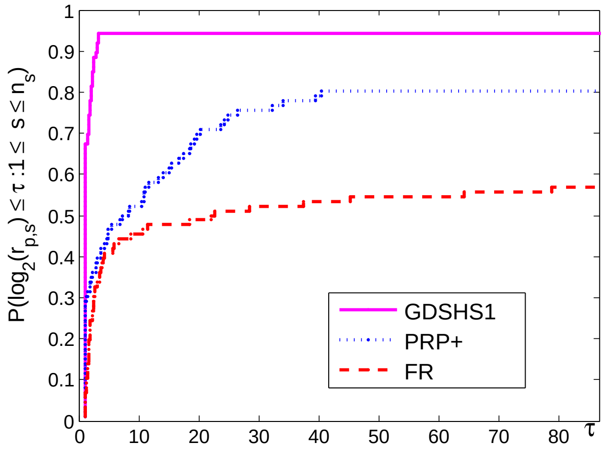

Figure 7: Performance profile by CPU time

In Figure 7, GDSHS1 is performed well from the viewpoint of CPU time.
But, the numerical performance should be compared by measures different from CPU time. For this reason we provide Figures 8 and 9.

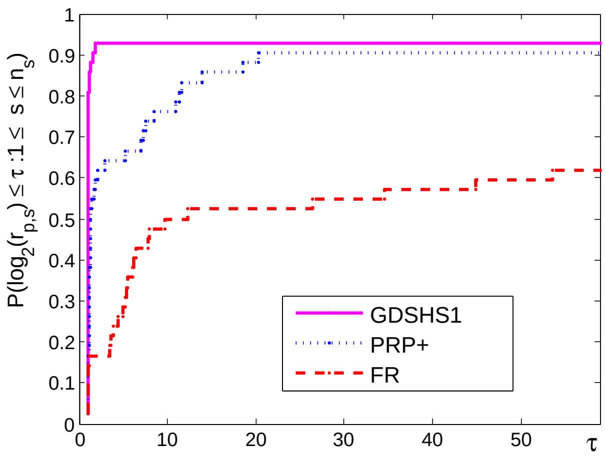

Figure 8: Performance profile by number of iterations

In Figures 8 and 9, we see that GDSHS1 algorithm is better and more competitive than the FR and PRP^+ algorithms especially in the number of iterations performance.

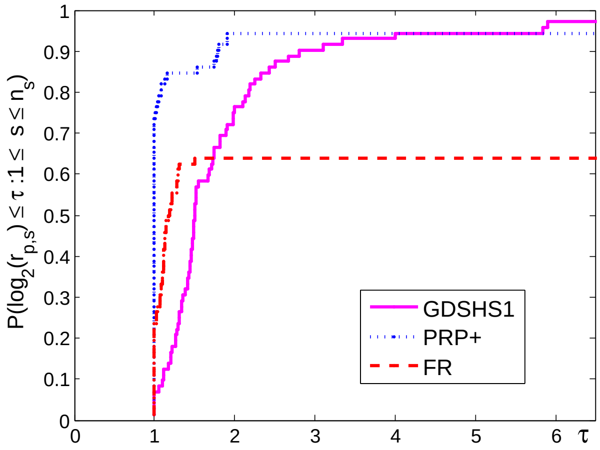

Figure 9: Performance profile by 2-norm of the gradient of the objective function

7 Concluding remarks

In this paper, we have proposed an optimal parameter for the Generalized Descent Symmetrical Hestenes-Stiefel algorithm . Some numerical results have been reported. These results showed the effectiveness of our method if we choose the parameter $c = 1$. The performance profile for our conjugate gradient GDSHS1 algorithm, implemented with our new line search, was higher than those of the GDSHS0.1, GDSHS2, GDSHS1.1 and GDSHS0.9 methods for a test set consisting of 86 problems from [18]. In our future research we would like further the theoretical properties of the parameter c.

8 Funding

This article was supported by the *Comité* National *ďEvaluation* et de Programmation de la Recherche Universitaire(CNEPRU: C00L03UN410120220001).

References

[1] Al-Baali, M., Descent property and global convergence of the Fletcher-Reeves method with inexact line-search , IMA J. Numer. Anal. 5, 121-124 (1985).

[2] Naima Hamel et . al , A new hybrid conjugate gradient algorithm based on the Newton direction to solve unconstrained optimization problems, JAMC (2023) 69,2531-2548

[3] Imane Hafaidia et. al , A new optimization method based on $PERRY\acute{S}$ idea , 10.1751 jamcm.2021.4.03 (2021)

[4] Jinwei Wang et. al, Forensics feature analysis in quaternion wavelet domain for distinguishing photographic images and computer graphics, Multimedia Tools and Applications, s11042-016-4153-0

[5] Xuezhi Wen et. al, A rapid learning algorithm for vehicle classification,Information Sciences, Information Sciences, vol. 295, no. 1, pp. 395-406, 2015.

[6] Yan Kong et. al, A belief propagation-based method for Task Allocation in Open and Dynamic Cloud Environments, Knowledge-based Systems, vol. 115, pp. 123-132, 2016.

[7] Yanhua Zhang et. al, Efficient Algorithm for K-Barrier Coverage Based on Integer Linear Programming, China Communications, vol. 13, no. 7, pp. 16-23, 2016.

[8] Gonglin Yuan, Zehong Meng, and Yong Li , A modified Hestenes and Stiefel conjugate gradient algorithm for large-scale nonsmooth minimizations and nonlinear equations ,Journal of Optimization Theory and Applications, 168 (2016), pp. 129-152.

[9] Gonglin Yuan, Maojun Zhang , A three-terms Polak-Ribiere-Polyak conjugate gradient algorithm for large-scale nonlinear equations , Journal of Computational and Applied Mathematics 286 (2015), pp. 186-195.

[10] Gonglin Yuan, Zengxin Wei, Xiwen Lu, Global convergence of BFGS and PRP methods under a modified weak Wolfe-Powell line search, Applied Mathematical Modelling, j.apm.2017.02.008.

[11] Liu, D., Xu, G., Applying Powells symmetrical technique to conjugate gradient methods, Computational Optimization and Applications, s10589-009-9302-1. (2009)

[12] Benrabia, N., Laskri,Y., Geubai, H., Al-Baali, M., Applying the*Powell's* Symmetrical Technique to Conjugate Gradient Methods with the Generalized Conjugacy Condition , NUMERICAL FUNCTIONAL ANALYSIS AND OPTIMIZATION 2016, VOL. 37, NO. 7, 839-849

[13] Hestenes, M. R. , Stiefel, E., Methods of conjugate gradients for solving linear systems, J. Res. Nat. Bur. Stand. 49(6), 409-439 (1952)

[14] Dai, Y. H., Liao, L. Z., An efficient hybrid conjugate gradient method for unconstrained optimization , Ann. Oper. Res., 103 (2001) 33-47

[15] Birgin, E. G. , Martinez, J. M., A spectral conjugate gradient method for unconstrained optimization. Appl, Math. Optim. 43, 117-128 (2001)

[16] J. C. Gilbert and J. Nocedal , Global convergence properties of conjugate gradient methods for optimization, SIAM J. Optim. 2(1):21-42.

[17] E. Polak and G. Ribiere , Note sur la convergence de directions conjuguées, Rev. Francaise Infomat Recherche Operatonelle, 3e Annee 16:35-43.

[18] Andrei, N., An Unconstrained Optimization Test Functions Collection, Advanced Modeling and Optimization, Volume 10, Number 1, 2008

[19] Dolan, E. D. , More, J. J., Benchmarking optimization software with performance profiles, Math. Program. Ser. A 91, 201-213 (2002).