Filomat 39:26 (2025), 9213–9234
https://doi.org/10.2298/FIL2526213M

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

Some improved nonlinear conjugate gradient methods and application to non-parametric estimation

Abd Elhamid Mehamdia[a,*], Yacine Chaib[b]

[a]*Laboratory Informatics and Mathematics (LIM), University of M'hamed Bougara, Boumerdes 35000, Algeria*
[b]*Laboratory Informatics and Mathematics (LIM), Mohamed Cherif Messaadia University, Souk Ahras 41000, Algeria*

Abstract.
The conjugate gradient method is one of the most important ideas in scientific computing, it is applied to solving linear systems of equations and nonlinear optimization problems. In this paper, based on a variant of Liu-Storey (LS) method, Hestenes-Stiefel (HS) method and Polak-Ribière-Polyak (PRP) method, three modified CG methods (named MCLS, MCHS and MCPRP) are presented and analyzed. The three presented methods generate a descent direction and possess good convergence properties under the strong Wolfe line search conditions. Preliminary elementary numerical experiment results are reported, which show that the proposed methods are promising and effective in minimizing some unconstrained optimization problems and each of these modifications outperforms the four famous conjugate gradient methods. Finally, the proposed methods were further extended to solve the problem of the conditional model regression function.

1. Introduction

Optimization methods are widely used to obtain the numerical solution of the optimal control problems arising in scientific and engineering computation, especially for solving large-scale problems, the nonlinear conjugate gradient (CG) method is welcomed due to its simple iteration and low storage. In this work, we focus our attention on the CG method for the following nonconvex unconstrained optimization:

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

where f is a smooth and nonlinear function. As we all know, the CG method is a line search method that we use the iterative method to solve (1.1) and the iterative formula is given by

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

2020 *Mathematics Subject Classification*. Primary 90C30; Secondary 65K05, 62G05.
Keywords. Unconstrained optimization, Conjugate gradient method, Line search, Sufficient descent, Global convergence, Numerical comparisons, regression function, Kernel estimator.
Received: 21 January 2024; Revised: 27 November 2024; Accepted: 18 July 2025
Communicated by Marko Petković
* Corresponding author: Abd Elhamid Mehamdia
Email addresses: a.mehamdia@univ-boumerdes.dz (Abd Elhamid Mehamdia), y.chaib@univ-soukahras.dz (Yacine Chaib)
ORCID iDs: https://orcid.org/0000-0001-9293-7922 (Abd Elhamid Mehamdia),
https://orcid.org/0000-0003-4151-8749 (Yacine Chaib)

where x_k is the current iteration point and $d_k \in \mathbb{R}^n$ is the search direction defined by the following formula:

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

where g_{k+1} the gradient of f at x_{k+1} and the parameter β_k is known as the CG coefficient. The step length α_k is very important for the global convergence of CG methods, one often requires the line search to satisfy the Wolfe conditions (WLS)

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

and

$$g_{k+1}^T d_k \geq \sigma g_k^T d_k. \tag{1.5}$$

Also, the strong Wolfe (SWLS) conditions consist of (1.4) and

$$\left| g_{k+1}^T d_k \right| \leq -\sigma g_k^T d_k, \tag{1.6}$$

where $0 < \delta < \sigma < 1$. For the scalar β_k many formulas have been proposed. Some of the classical algorithms for β_k are Fletcher-Reeves (FR) method [11], Dai-Yuan (DY) method [8], Conjugate Descent (CD) method proposed by Fletcher [12], Polak- Ribière and Polyak (PRP) method [25, 26], Hestenes-Stiefel (HS) method [15] and Liu and Storey (LS) method [19]. In which formulas for β_k are given, respectively, by

$$\begin{aligned}
\beta_k^{FR} &= \frac{\left\|g_{k+1}\right\|^2}{\left\|g_k\right\|^2}, \beta_k^{DY} = \frac{\left\|g_{k+1}\right\|^2}{y_k^T d_k}, \beta_k^{CD} = \frac{\left\|g_{k+1}\right\|^2}{-g_k^T d_k}, \\
\beta_k^{PRP} &= \frac{g_{k+1}^T y_k}{\left\|g_k\right\|^2},\ \ \beta_k^{HS} = \frac{g_{k+1}^T y_k}{y_k^T d_k},\ \ \beta_k^{LS} = \frac{g_{k+1}^T y_k}{-g_k^T d_k},
\end{aligned}$$

where $\|.\|$ denotes the Euclidean norm and $y_k = g_{k+1} - g_k$.

In the past few years, the PRP method is generally regarded to be one of the most efficient CG methods in practical computation. A wonderful property of the PRP method is that it automatically performs a restart if a bad direction occurs [14]. The numerical performances of the HS and LS methods are very similar to the PRP method since the coefficient β_k in these methods has the same numerator. However, the convergence properties of the PRP and HS methods are not so good [27]. In the latest years, based on the above six formulas and their hybridization, many works putting effort into seeking new CG methods with not only good convergence properties but also excellent numerical effects were published.

Based on the PRP method Zhang [30] gave a new CG formula, called NPRP method, where β_k is given by

$$\beta_k^{NPRP} = \frac{\left\|g_{k+1}\right\|^2 - \frac{\|g_{k+1}\|}{\|g_k\|}\left|g_{k+1}^T g_k\right|}{\left\|g_k\right\|^2}.$$

Also, this author [30] has given a modified HS method, in which β_k is defined by

$$\beta_k^{NHS} = \frac{\left\|g_{k+1}\right\|^2 - \frac{\|g_{k+1}\|}{\|g_k\|}\left|g_{k+1}^T g_k\right|}{d_k^T\left(g_{k+1} - g_k\right)}.$$

The NPRP and NHS methods possess sufficient descent conditions and converge globally if the SWLS is used and the parameter σ is restricted in $\left(0, \frac{1}{2}\right)$. Likewise, Du et al. [10] in 2016 give two modified CG methods, proposing the following formula

$$\beta_k^{VLS^*} = \frac{\left\|g_{k+1}\right\|^2 - \frac{\|g_{k+1}\|}{\|g_k\|}\left|g_{k+1}^T g_k\right|}{-g_k^T d_k},$$

and

$$\beta_k^{NVLS^*} = \frac{\left\|g_{k+1}\right\|^2 - \frac{\left|g_{k+1}^T g_k\right|}{\left\|g_k\right\|^2} g_{k+1}^T g_k}{-g_k^T d_k}.$$

The VLS* and NVLS* methods have sufficient descent conditions and are globally convergent if the SWLS is utilized. Jiang and Jian [16] gave a variant of the PRP method (called MN method) and its parameter β_k is yielded by

$$\beta_k^{MN} = \frac{\left\|g_{k+1}\right\|^2 - \frac{\left\|g_{k+1}\right\|}{\left\|d_k\right\|} g_{k+1}^T d_k}{\left(1-\eta_k\right)\left\|g_k\right\|^2},$$

where $\eta_k = \mu \min\left(0, \frac{g_{k+1}^T d_k}{\left\|g_{k+1}\right\|\left\|d_k\right\|}\right)$ and $\mu \geq 1$, Jiang and Jian [16] proved that the MN method produce the sufficient descent direction and globally convergent if SWLS is used.

The same authors in [17] presented two modified CG formula as follows, in which the conjugate parameters are computed by

$$\beta_k^{IFR} = \omega_k \beta_k^{FR} \text{ and } \beta_k^{IDY} = \omega_k \beta_k^{DY},$$

where $\omega_k = \frac{\left|g_{k+1}^T d_k\right|}{-g_k^T d_k}$. Under the SWLS, the two methods are proven to be sufficient descent and globally convergent [17].

The goal of this paper is to propose three modified conjugate gradient methods and prove that these methods with SWLS possess descent properties and are globally convergent. Numerical comparisons are reported by utilizing some test problems in the Cute library [1, 4]. Furthermore, the proposed methods were extended to solve the practical application problem of the regression model function.

The structure of the paper is as follows. In the next section, the MCLS, MCHS and MCPRP CG methods are presented and the descent property is analyzed. In section 3, the global convergence of the three proposed methods with an SWLS is proved. Numerical results for solving medium-large-scale unconstrained optimization are reported in section 4. The application of the modified methods in nonparametric estimation of conditional mode is discussed in section 5. Finally, we make a summary of our paper.

2. Motivation and some new formulas

Motivated by the construction of CG parameters β_k in [10], [16], [17] and [18], the paper proposes the three modified CG methods.

Firstly, the MCLS method, where β_k is defined by

$$\beta_k^{MCLS} = \frac{\left\|g_{k+1}\right\|^2 - \frac{\left|g_{k+1}^T d_k\right|}{\left\|g_k\right\|\left\|d_k\right\|}\left|g_{k+1}^T g_k\right|}{-g_k^T d_k + \varsigma\left\|g_{k+1}\right\|\left\|d_k\right\|}, \text{ where } \varsigma > 0. \tag{2.1}$$

Secondly, the MCHS method, where β_k is given by

$$\beta_k^{MCHS} = \tau_k \frac{\left\|g_{k+1}\right\|^2 - \frac{\left|g_{k+1}^T d_k\right|}{\left\|g_k\right\|\left\|d_k\right\|}\left|g_{k+1}^T g_k\right|}{d_k^T\left(g_{k+1} - g_k\right)}, \tag{2.2}$$

where

$$\tau_k = \begin{cases} \frac{g_{k+1}^T d_k}{-g_k^T d_k} & \text{if } g_{k+1}^T d_k > 0, \\ 1 & \text{otherwise.} \end{cases}$$

Finally, the MCPRP method, where β_k is defined by

$$\beta_k^{MCPRP} = \frac{1-\omega_k}{\rho_k} \frac{\left\|g_{k+1}\right\|^2 - \frac{\left|g_{k+1}^T d_k\right|}{\left\|g_k\right\|\left\|d_k\right\|}\left|g_{k+1}^T g_k\right|}{\left\|g_k\right\|^2}, \tag{2.3}$$

where $\rho_k = 1 - \min\left\{0, \frac{\mu\left(g_{k+1}^T d_k\right) g_{k+1}^T g_k}{\left\|g_{k+1}\right\|^2 \left\|g_k\right\| \left\|d_k\right\|}\right\}$ and $\mu \geq 1$.

2.1. Algorithms

In this part, we present the MCLS, MCHS and MCPRP Algorithms, with the SWLS.

Algorithm MCLS

Step 1: Initialization.

Choose an initial point $x_0 \in \mathbb{R}^n$ and the parameters $0 < \delta < \sigma < 1$. Compute $f(x_0)$ and g_0. Set $d_0 = -g_0$.

Step 2: Test for a continuation of iterations.

If $\left\|g_k\right\|_\infty \leq 10^{-6}$, then stop. Otherwise, go to the next step.

Step 3: Line search.

Calculate α_k satisfies the linear search conditions of stong Wolfe (1.4) and (1.6) and update the variables $x_{k+1} = x_k + \alpha_k d_k$.

Step 4: Compute β_k by the formula (2.1).

Step 5: Generate the search direction by the formula (1.3) .

Step 6: Set $k = k + 1$ and go to Step 2.

Algorithm MCHS

In Algorithm MCLS (2.1) in step 4 is remplaced by (2.2) .

Algorithm MCPRP

In Algorithm MCLS $0 < \delta < \sigma < 1$ in step 1 is remplaced by $0 < \delta < \sigma < \frac{1}{2}$, (2.1) in step 4 is remplaced by (2.3) .

2.2. The descent direction

The following Theorem shows that the search direction generated by MCLS Algorithm satisfies the sufficient descent conditions.

Theorem 2.1. Let the direction d_k be yielded by the MCLS Algorithm. Then, we get

$$g_k^T d_k \leq -c_1 \left\|g_k\right\|^2, \forall\, k \geqslant 0, \tag{2.4}$$

where $c_1 = 1 - \sigma$.

Proof. The following proof is by induction. For $k = 0$, $g_0^T d_0 = -\left\|g_0\right\|^2$, we conclude that the sufficient descent condition holds for $k = 0$. Now, we assume (2.4) holds for k and prove that for $k + 1$. Suppose that the ξ_k is the angle between the g_k and g_{k+1} vectors and the θ_k is the angle between the g_{k+1} and d_k vectors, then

$$\cos \xi_k = \frac{g_{k+1}^T g_k}{\left\|g_{k+1}\right\| \left\|g_k\right\|}, \ \cos \theta_k = \frac{g_{k+1}^T d_k}{\left\|g_{k+1}\right\| \left\|d_k\right\|}.$$

We have since the definition of β_k^{MCLS} and (2.4), that

$$0 \leq \beta_k^{MCLS} = \frac{\left\|g_{k+1}\right\|^2 \left[1 - \left|\cos \theta_k\right| \left|\cos \xi_k\right|\right]}{-g_k^T d_k} \leq \frac{\left\|g_{k+1}\right\|^2}{-g_k^T d_k}. \tag{2.5}$$

From (1.3), (1.6), (2.1) and (2.5), it is clear that

$$g_{k+1}^T d_{k+1} \quad \leq \quad -\left\|g_{k+1}\right\|^2 + \frac{\left\|g_{k+1}\right\|^2}{-g_k^T d_k} \left|g_{k+1}^T d_k\right| \leq -(1-\sigma) \left\|g_{k+1}\right\|^2.$$

Therefore, the proof is completed. □

The following Theorem lists the descent conditions and an interesting property of the MCHS method.

Theorem 2.2. Let the sequences $\{g_k\}_{k\geq 0}$ and $\{d_k\}_{k\geq 0}$ be generated by MCHS Algorithm. If $g_k \neq 0$, then

$$g_k^T d_k < 0, \forall\, k \geqslant 0, \tag{2.6}$$

and the relation

$$0 \leq \beta_k^{MCHS} \leq \frac{g_{k+1}^T d_{k+1}}{g_k^T d_k}, \tag{2.7}$$

holds.

Proof. We prove assertion (2.6) by induction. For $k = 0$, $g_0^T d_0 = -\left\|g_0\right\|^2$, we conclude that the descent condition holds for $k = 0$. Now, we assume (2.6) holds for k which implies that $g_k^T d_k < 0$. This together with (1.6) indicates that

$$d_k^T (g_{k+1} - g_k) \geq (1-\sigma)\left(-g_k^T d_k\right) > 0. \tag{2.8}$$

Now, we prove $g_{k+1}^T d_{k+1} < 0$ holds. We discuss in two cases

(i) If $g_{k+1}^T d_k > 0$. Using (1.6) and (2.6), we have

$$0 \leq \tau_k = \frac{g_{k+1}^T d_k}{-g_k^T d_k} \leq \sigma. \tag{2.9}$$

It follows from the definition of β_k^{MCHS}, (2.8) and (2.9), that

$$\beta_k^{MCHS} = \frac{g_{k+1}^T d_k}{-g_k^T d_k} \frac{\left\|g_{k+1}\right\|^2 [1 - \cos\theta_k \left|\cos\xi_k\right|]}{d_k^T (g_{k+1} - g_k)} \geq 0. \tag{2.10}$$

Using (1.3) and (2.2), we get

$$\begin{aligned} g_{k+1}^T d_{k+1} &= -\left\|g_{k+1}\right\|^2 + \tau_k \frac{\left\|g_{k+1}\right\|^2 [1 - \cos\theta_k \left|\cos\xi_k\right|]\, g_{k+1}^T d_k}{d_k^T (g_{k+1} - g_k)} \\ &= -\left\|g_{k+1}\right\|^2 + \tau_k \left\|g_{k+1}\right\|^2 [1 - \cos\theta_k \left|\cos\xi_k\right|] + \beta_k^{MCHS} g_k^T d_k. \end{aligned}$$

Since $d_k^T g_k < 0$, (2.9) and (2.10) we obtain

$$g_{k+1}^T d_{k+1} \leq -(1-\sigma)\left\|g_{k+1}\right\|^2 + \beta_k^{MCHS} g_k^T d_k < 0. \tag{2.11}$$

Then the descent condition is satisfied.

Now, according to (2.6) and (2.11), can be obtained

$$g_{k+1}^T d_{k+1} \leq \beta_k^{MCHS} g_k^T d_k,$$

then

$$\beta_k^{MCHS} \leq \frac{g_{k+1}^T d_{k+1}}{g_k^T d_k}. \tag{2.12}$$

(ii) If $g_{k+1}^T d_k \leq 0$. Using the (2.2) and (2.8), we get

$$\beta_k^{MCHS} = \frac{\left\|g_{k+1}\right\|^2 + \frac{g_{k+1}^T d_k \left|g_{k+1}^T g_k\right|}{\left\|g_k\right\| \left\|d_k\right\|}}{d_k^T (g_{k+1} - g_k)} = \frac{\left\|g_{k+1}\right\|^2 [1 + \cos\theta_k \left|\cos\xi_k\right|]}{d_k^T (g_{k+1} - g_k)} \geq 0. \tag{2.13}$$

From (1.3) and (2.2),

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

This together with (1.6), (2.6) and (2.8), implies that

$$g_{k+1}^T d_{k+1} \leq \frac{\left\|g_{k+1}\right\|^2 + \sigma \left\|g_{k+1}\right\|^2 \cos\theta_k \left|\cos\xi_k\right|}{d_k^T \left(g_{k+1} - g_k\right)} g_k^T d_k < 0. \tag{2.14}$$

Moreover, it follows from (2.2) and (2.14), that

$$\beta_k^{MCHS} = \frac{\left\|g_{k+1}\right\|^2 \left[1 + \cos\theta_k \left|\cos\xi_k\right|\right]}{d_k^T \left(g_{k+1} - g_k\right)} \leq \frac{\left\|g_{k+1}\right\|^2 + \sigma \left\|g_{k+1}\right\|^2 \cos\theta_k \left|\cos\xi_k\right|}{d_k^T \left(g_{k+1} - g_k\right)} \leq \frac{g_{k+1}^T d_{k+1}}{g_k^T d_k}. \tag{2.15}$$

From (2.11) and (2.14), we have proven (2.6).

Also, from the relations (2.10), (2.12), (2.13) and (2.15), we conclude that (2.7) is always satisfied. Therefore, the proof is completed. □

Now, we give a Theorem which shows that the MCPRP method possesses the sufficient descent property if the step size α_k is determined by the SWLS with $0 < \sigma < \frac{1}{2}$.

Theorem 2.3. Let the direction d_k be yielded by the MCPRP method. If parameter $\sigma < \frac{1}{2}$, then the relation

$$-\frac{1}{1-\sigma} \leq \frac{g_k^T d_k}{\left\|g_k\right\|^2} \leq -\frac{1-2\sigma}{1-\sigma}, \tag{2.16}$$

hold. So, the search direction d_k generated by the MCPRP method is sufficient descent.

Proof. From (2.3), it is clear that

$$\beta_k^{MCPRP} = (1-\omega_k) \frac{\left(1 - \left|\cos\xi_k\right| \left|\cos\theta_k\right|\right)}{\left[1 - \min\left(0, \mu\cos\xi_k\cos\theta_k\right)\right]} \beta_k^{FR}.$$

By the second inequality of the SWLS (1.6), we know that

$$0 < 1 - \sigma \leq 1 - \omega_k = 1 - \frac{\left|g_{k+1}^T d_k\right|}{-g_k^T d_k} < 1. \tag{2.17}$$

Again, in view of $\mu \geq 1$ and (2.17), we now prove $0 \leq \frac{1-\omega_k}{1-\min\left(0, \mu\cos\xi_k\cos\theta_k\right)} < 1$ by considering the following two cases.

Case (i), if $\cos\xi_k\cos\theta_k < 0$, we get

$$0 \leq \frac{1-\omega_k}{1 - \min\left(0, \mu\cos\xi_k\cos\theta_k\right)} = \frac{1-\omega_k}{1 - \mu\cos\xi_k\cos\theta_k} < 1. \tag{2.18}$$

Case (ii), if $\cos\xi_k\cos\theta_k \geq 0$, then

$$0 \leq \frac{1-\omega_k}{1 - \min\left(0, \mu\cos\xi_k\cos\theta_k\right)} = 1 - \omega_k < 1. \tag{2.19}$$

It follows from the definition of β_k^{MCPRP}, (2.18) and (2.19), that

$$0 \leq \beta_k^{MCPRP} \leq \left(1 - \left|\cos\xi_k\right| \left|\cos\theta_k\right|\right) \beta_k^{FR} \leq \beta_k^{FR}. \tag{2.20}$$

By the relations (2.20) and Lemma 3.1 in Glibert and Nocedal [13], we immediately obtain the MCPRP method satisfies the sufficiently descent condition (2.16).

3. Convergence analysis

In order to obtain the convergence properties of the proposed methods, the following basic assumptions for the objective function are always needed.

Assumption 3.1. The level set

$$S = \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\},$$

is bounded.

Assumption 3.2. In some open convex neighborhood $\mathcal{N}$ of S, the function f is continuously differentiable and its gradient is Lipschitz continuous, namely, there exists a constant $L > 0$, such that:

$$\left\|\nabla f(x) - \nabla f(y)\right\| \leq L \left\|x - y\right\|, \quad \forall x, y \in \mathcal{N}. \tag{3.1}$$

From Assumption 3.2, we can deduce that for all $x \in \mathcal{N}$, there exists a positive constant $\Gamma \geq 0$, such that

$$\| \nabla f(x) \| \leq \Gamma, \text{ for all } x \in \mathcal{N}. \tag{3.2}$$

It follows from Dai et al. [7] proved the sufficient condition for the convergence of CG methods with strong Wolfe line search.

Lemma 3.1. Let Assumptions 3.1 and 3.2 hold. Consider the method (1.2) and (1.3), where d_k is a descent direction and α_k is obtained by the SWLS. If

$$\sum_{k \geq 0} \frac{1}{\|d_k\|^2} = \infty,$$

then

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

The following Theorem establish the global convergence of the MCLS method with the SWLS.

Theorem 3.1. Consider that Assumptions 3.1 and 3.2 hold. Let the sequences $\{g_k\}_{k \geq 0}$ and $\{d_k\}_{k \geq 0}$ be generated by MCLS Algorithm. Then

$$\lim_{k \to \infty} \inf \left\|g_k\right\| = 0. \tag{3.3}$$

Proof. To prove Theorem 3.1, we use contradiction. That is, if equation (3.3) is not true. Then we can find a positive constant γ_1, such that

$$\left\|g_k\right\| \geq \gamma_1, \quad \text{for all } k. \tag{3.4}$$

In fact, from (2.4) and (2.5), we have

$$0 \leq \beta_k^{MCLS} \leq \frac{\| g_{k+1} \|^2}{\varsigma \left\|g_{k+1}\right\| \|d_k\|}. \tag{3.5}$$

Thus, it follows from (1.3), (3.2) and (3.5), that

$$\|d_{k+1}\| \leq \| g_{k+1} \| + \beta_k^{MCLS} \|d_k\| \leq \| g_{k+1} \| + \frac{\| g_{k+1} \|^2}{\varsigma \left\|g_{k+1}\right\| \|d_k\|} \|d_k\| \leq M.$$

where $M = \left(1 + \frac{1}{\varsigma}\right)\Gamma$.

Which implies that

$$\sum_{k\geq 0}\frac{1}{\|d_k\|^2}=\infty. \tag{3.6}$$

By applying Lemma 3.1, (3.3) is true. This is a contradiction with (3.4), so we have proved (3.3). □

The conclusion of the following Lemma, often called the Zoutendijk condition, plays a necessary role in the analysis of the global convergence properties for CG method. It was originally given by Zoutendijk [31].

Lemma 3.2. It is assumed that x_0 is a starting point for which Assumptions 3.1 and 3.2 hold. Consider any method in the form (1.2) and (1.3), where d_k is a descent direction and the step size α_k satisfies (1.4) and (1.5), then we have

$$\sum_{k=0}^{\infty}\frac{(g_k^T d_k)^2}{\|d_k\|^2}<\infty. \tag{3.7}$$

It is easy to get from (2.16) that the Zoutendijk condition (3.7) is equivalent to the following inequality

$$\sum_{k=0}^{\infty}\frac{\left\|g_k\right\|^4}{\|d_k\|^2}<\infty. \tag{3.8}$$

The following Theorem establish the global convergence of the MCHS method with the SWLS.

Theorem 3.2. Suppose that Assumptions 3.1 and 3.2 hold. Consider any CG method in the form (1.2) and (1.3), with the parameter $\beta_k=\beta_k^{MCHS}$, in which the step length α_k is determined to satisfy the SWLS condition (1.4) and (1.6), where d_k is a descent search direction. Then, this method converges in the sense that

$$\lim_{k\to\infty}\inf\left\|g_k\right\|=0. \tag{3.9}$$

Proof. Assume by contradiction that the formula (3.9) does not hold. For all k, there exists a constant $\gamma_2>0$, such that

$$\left\|g_k\right\|\geq\gamma_2,\ \forall k\geq 0. \tag{3.10}$$

Squaring both sides of $d_{k+1}+g_{k+1}=\beta_k^{MCHS}d_k$, we get

$$\|d_{k+1}\|^2=\left(\beta_k^{MCHS}\right)^2\|d_k\|^2-\left\|g_{k+1}\right\|^2-2g_{k+1}^T d_{k+1}. \tag{3.11}$$

Substituting (2.7) into (3.11), we obtain

$$\|d_{k+1}\|^2\leq\left(\frac{g_{k+1}^T d_{k+1}}{g_k^T d_k}\right)^2\|d_k\|^2-\left\|g_{k+1}\right\|^2-2g_{k+1}^T d_{k+1}. \tag{3.12}$$

Divided the both sides of (3.12) by $\left(g_{k+1}^T d_{k+1}\right)^2$, we obtain

$$\frac{\|d_{k+1}\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2}\leq\frac{\|d_k\|^2}{\left(g_k^T d_k\right)^2}-\frac{\left\|g_{k+1}\right\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2}-\frac{2}{g_{k+1}^T d_{k+1}}=\frac{\|d_k\|^2}{\left(g_k^T d_k\right)^2}-\left(\frac{1}{\left\|g_{k+1}\right\|}+\frac{\left\|g_{k+1}\right\|}{g_{k+1}^T d_{k+1}}\right)^2+\frac{1}{\left\|g_{k+1}\right\|^2}. \tag{3.13}$$

Combining with $\frac{\|d_0\|^2}{\left(g_0^T d_0\right)^2}=\frac{1}{\left\|g_0\right\|^2}$, by using (3.10) and a recurrence of relation (3.13), we have

$$\frac{\|d_{k+1}\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2}\leq\frac{\|d_k\|^2}{\left(g_k^T d_k\right)^2}+\frac{1}{\left\|g_{k+1}\right\|^2}\leq\sum_{i=0}^{k+1}\frac{1}{\left\|g_i\right\|^2}\leq\frac{k+2}{\gamma_2^2}.$$

This gives

$$\sum_{k\geq 0}\frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2}=\infty.$$

This contradicts the Zoutendjik condition (3.7), concluding the proof. □

Now, we can give the global convergence result of the MCPRP method.

Theorem 3.3. Consider that Assumptions 3.1 and 3.2 hold. Let the sequences $\{g_k\}_{k\geq 0}$ and $\{d_k\}_{k\geq 0}$ be generated by MCPRP Algorithm. Then

$$\liminf_{k\to\infty}\left\|g_k\right\|=0. \tag{3.14}$$

Proof. Suppose that (3.14) does not hold. Then there exists a constant $\gamma_3>0$, such that

$$\left\|g_k\right\|\geq\gamma_3,\ \forall k\geq 0. \tag{3.15}$$

Using (1.3) and (2.3), we get

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

Square both sides of the previous equation

$$\|d_{k+1}\|^2=\left(\beta_k^{MCPRP}\right)^2\|d_k\|^2-2\beta_k^{MCPRP}g_{k+1}^T d_k+\left\|g_{k+1}\right\|^2. \tag{3.16}$$

Also, by (1.6), (2.16) and (2.20),

$$-2\beta_k^{MCPRP}g_{k+1}^T d_k\leq 2\beta_k^{MCPRP}\left|g_{k+1}^T d_k\right|\leq\frac{-2\left\|g_{k+1}\right\|^2\sigma g_k^T d_k}{\left\|g_k\right\|^2}\leq\frac{2\sigma\left\|g_{k+1}\right\|^2}{1-\sigma}. \tag{3.17}$$

Substituting (2.20) and (3.17) into (3.16), we get

$$\|d_{k+1}\|^2\leq\frac{\left\|g_{k+1}\right\|^4}{\left\|g_k\right\|^4}\|d_k\|^2+\left(\frac{\sigma+1}{1-\sigma}\right)\left\|g_{k+1}\right\|^2. \tag{3.18}$$

Divided (3.18) by $\left\|g_{k+1}\right\|^4$, we obtain

$$\frac{\|d_{k+1}\|^2}{\left\|g_{k+1}\right\|^4}\leq\frac{\|d_k\|^2}{\left\|g_k\right\|^4}+\left(\frac{\sigma+1}{1-\sigma}\right)\frac{1}{\left\|g_{k+1}\right\|^2}. \tag{3.19}$$

Noting that $\frac{\|d_0\|^2}{\left(g_0^T d_0\right)^2}=\frac{1}{\|g_0\|^2}$ and using (3.19) recursively yields

$$\frac{\|d_{k+1}\|^2}{\left\|g_{k+1}\right\|^4}\leq\left(\frac{\sigma+1}{1-\sigma}\right)\sum_{i=0}^{k+1}\frac{1}{\left\|g_i\right\|^2}\leq\left(\frac{\sigma+1}{1-\sigma}\right)\frac{k+1}{\gamma_3^2}.$$

This implies that

$$\sum_{k\geq 0}\frac{\left\|g_k\right\|^4}{\|d_k\|^2}=\infty.$$

This contradicts the Zoutendjik condition (3.8), concluding the proof. □

4. Numerical results

To demonstrate more clearly the efficiency of the MCLS, MCHS and MCPRP Algorithms to some other conjugate gradient algorithms famous, we have run three groups of preliminary numerical experiments for the MCLS, MCHS and MCPRP CG methods, respectively.

• In Group A, we compare the MCLS with the NVLS* [10], VLS* [10], LS [19] and MLS [29] CG methods. There are 42 problems are tested.

• In Group B, the MCHS is compared with the IFR [17], IDY [17], MHS [29] and NHS [30] CG methods. Where 38 problems are tested.

• In Group C, the MCPRP is compared with the MN [16], IFR [17], IDY [17] and NPRP [30] CG methods. Using 41 test problems

In this numerical result, all Algorithms implement the SWLS condition with $\delta = 10^{-3}$ and $\sigma = 10^{-1}$. The iteration is terminated if one of the conditions below are met *(i)* $\left\|g_k\right\|_\infty < 10^{-6}$, where $\|.\|_\infty$ is the maximum absolute component of a vector, *(ii)* The number of iterations exceeded 2000, *(iii)* The computing time is more than 500 s. We show the performance difference clearly between our methods MCLS, MCHS, MCPRP and four conjugate gradient algorithms for each of them. We choose the performance profile introduced by Dolan and Morè [9] to compare the performance according to the number of iterations and CPU time to rule as follows. Let S is the set of methods and P is the set of the test problems with n_p, n_s is the number of the test problems and the number of the methods, respectively. For each problem $p \in P$ and solver $s \in S$, denote $\tau_{p,s}$ be the number of iterations or CPU time required to solve problems $p \in P$ by solver $s \in S$. Then a comparison between different solvers based on the performance ratio is given by

$$r_{p,s} = \frac{\tau_{p,s}}{\min\left\{\tau_{p,i},\, 1 \le i \le n_s\right\}}.$$

Suppose that a parameter $r_M \ge r_{p,s}$ for all problems and solvers chosen, and $r_M = r_{p,s}$ if and only if solver s does not solve problem p. The overall evaluation of the performance of the solvers is then given by the performance profile function given by

$$F_s(t) = \frac{size\left\{p:\ 1 \le p \le n_p,\, r_{p,s} \le t\right\}}{n_p},$$

where $t \ge 1$ and $size\left\{p:\ 1 \le p \le n_p,\, r_{p,s} \le t\right\}$ is the number of elements in the set $\left\{p:\ 1 \le p \le n_p,\, r_{p,s} \le t\right\}$. This function F_s: $[1,\infty[\ \to [0,1]$ is the distribution function for the performance ratio. The value of $F_s(1)$ is the probability that the solver will win the rest of the solvers.

In this numerical study, 'Dim' denotes the dimension of the problem, 'ITR' denotes the number of iterations, 'TIME' denotes the 'CPU' time and 'Inf' denotes the algorithm failed to yield a solution for the problem.

From Figure 1, we can see that the average CPU time of the MCLS method is much less than that of the LS and VLS* methods, while the average CPU time of the LS and VLS* methods are much less than that of the NVLS* and MLS methods.

Figure 2 illustrates the performance profiles for all methods in group A. Observing from Figure 2, the MCLS method is competitive and outperforms the tested methods in group A, with respect to the number of iterations.

From Table 1, it is clear that the average performance of the MCLS, with the NVLS*, VLS*, LS and MLS CG methods are very similar to the results obtained from Figure 1 and Figure 2.

Figure 3 shows the performance profile for the CPU time. Relative to this metric, MCHS achieves the top performance, followed by IDY, then IFR, NHS, then MHS. The IFR method behaves almost like the IDY method.

On the other side, Figure 4 is the performance profile of all methods in Group B. From this Figure, it is concluded that the MCHS method performs better than the IDY, IFR, NHS and MHS CG methods, from the viewpoint of the number of iterations.

From Table 2, it is clear that the average performance of the MCHS, IFR, IDY, MHS and NHS methods are very similar to the results obtained from Figure 3 and Figure 4.

Figure 5 gives a performance comparison of the MCPRP method versus IFR, IDY, NPRP and MN methods. As this figure indicates, the new algorithm prevailed over all other Methods, with respect to CPU time, this clearly confirms the effectiveness of the MCPRP method.

It can be seen from Figure 6 that the MCPRP curve is mostly at the top of the IFR, IDY, NPRP and MN CG curves, so it is indicating that the MCPRP algorithm outperforms the IFR, IDY, NPRP and MN methods based on the number of iterations.

From Table 3, it is clear that the average performance of the MCPRP, IFR, IDY, NPRP and MN methods are very similar to the results obtained from Figure 5 and Figure 6.

5. Application in Conditional mode regression

The conjugate gradient method has played an important role in solving large scale unconstrained optimization problems that may arise in statistics nonparametric [22-24], portfolio selection [2] and image restoration problems [21].

The regression function estimation is the most important tool for addressing nonparametric prediction problems. The study of the relationship between a variable of interest Y and a covariate X is one the most important problem in statistics. Recent years have witnessed a renewal of interest in regression modal estimation, we refer the reader to Boente and Fraiman [3] .

For any x denote by $f(. \mid x) = \frac{f(x,.)}{l(x)}$ the conditional probability density function (p.d.f) of Y given $X = x$, where $f(.,.)$ is the joint p.d.f. of (X, Y) and $l(.)$ is the marginal density of X. Assuming that $f(. \mid x)$ has a unique mode $\theta(x)$ the latter is given by

$$f(\theta(x) \mid x) = \max_{y \in \mathbb{R}^n} f(y \mid x) . \tag{5.1}$$

The estimation of the conditional mode has a long history and has been studied by many authors in the statistics literature. The nonparametric estimator of conditional mode has first been considered in the case of complete data. For independent and identically distributed (i.i.d) random variables, Samanta and Thavaneswaran [30], while Collomb et al. [5] in dependent case.

For the complete data is present, it is well known that the kernel estimator of the conditional mode $\theta(x)$ is defined as the random variable $\hat{\theta}_n(x)$ which maximizer the kernel estimator $\hat{f}_n(y \mid x)$ of $f(y \mid x)$, that is

$$\hat{f}_n\left(\hat{\theta}_n(x) \mid x\right) = \max_{y \in \mathbb{R}^n} \hat{f}_n(y \mid x) , \tag{5.2}$$

where

$$\hat{f}_n(y \mid x) = \frac{\hat{f}_n(x, y)}{l_n(x)},$$

with

$$\hat{f}_n(x, y) = \frac{1}{nh_n^{2n}} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right) H\left(\frac{y - Y_i}{h_n}\right),$$

and

$$l_n(x) = \frac{1}{nh_n^n} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right).$$

Here, the convention $\frac{0}{0} = 0$. The function K and H are a p.d.f. (so-called kernel) defined on $\mathbb{R}^n$ and (h_n) is a sequence of positive real numbers (so-called bandwidth) which goes to zero as n goes to infinity.

Simulation study

Let (X_1, Y_1), ..., (X_n, Y_n) be n independent pairs, identically distributed as (X, Y) which is a random pair valued in $\mathbb{R}^n \times \mathbb{R}^n$.

We first consider the classical linear model with normal errors

$$Y_i = X_i + \upsilon\epsilon_i.$$

Second, we consider nonlinear model (parabolic case) such that

$$Y_i = X_i^2 + \upsilon\epsilon_i.$$

where $(X_i)_{1\leq i\leq n}$ and $(\epsilon_i)_{1\leq i\leq n}$ are two i.i.d. sequences distributed as $N(0,1)$ and υ is an appropriately chosen constant (here we take $\upsilon = 0.2$).

In practice, some tuning parameters have to fixed: the kernel K is chosen by

$$K(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2}\sum_{j=1}^{n} x_j^2\right),$$

and the kernel H is defined by

$$H(y) = \left(\frac{3}{4}\right)^n \prod_{j=1}^{n}\left(1 - y_j^2\right).$$

The selection of the bandwidth h is an important and basic problem in kernel smoothing techniques. In this simulation, we choose the optimal bandwidth by the cross-validation method.

In this context, we employ the MCLS, MCHS and MCPRP algorithms to solve the problem (5.2) under strong Wolfe line search technique. According to Table 4, it is clear that the MCLS, MCHS and MCPRP are efficient for solving the problem (5.2) based on number of iterations and CPU time.

The estimation of the conditional mode has a long history and has been studied by many authors in the statistics literature. The nonparametric estimator of conditional mode has first been considered in the case of complete data. For independent and identically distributed (i.i.d) random variables, Samanta and Thavaneswaran [30], while Collomb et al. [5] in dependent case.

For the complete data is present, it is well known that the kernel estimator of the conditional mode $\theta(x)$ is defined as the random variable $\hat{\theta}_n(x)$ which maximizer the kernel estimator $\hat{f}_n(y \mid x)$ of $f(y \mid x)$, that is

$$\hat{f}_n\left(\hat{\theta}_n(x) \mid x\right) = \max_{y\in\mathbb{R}^n} \hat{f}_n(y \mid x), \tag{5.2}$$

where

$$\hat{f}_n(y \mid x) = \frac{\hat{f}_n(x, y)}{l_n(x)},$$

with

$$\hat{f}_n(x, y) = \frac{1}{nh_n^{2n}} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right) H\left(\frac{y - Y_i}{h_n}\right),$$

and

$$l_n(x) = \frac{1}{nh_n^n} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right).$$

Here, the convention $\frac{0}{0} = 0$. The function K and H are a p.d.f. (so-called kernel) defined on $\mathbb{R}^n$ and (h_n) is a sequence of positive real numbers (so-called bandwidth) which goes to zero as n goes to infinity.

Simulation study

Let (X_1, Y_1), ..., (X_n, Y_n) be n independent pairs, identically distributed as (X, Y) which is a random pair valued in $\mathbb{R}^n \times \mathbb{R}^n$.

We first consider the classical linear model with normal errors

$$Y_i = X_i + v\epsilon_i.$$

Second, we consider nonlinear model (parabolic case) such that

$$Y_i = X_i^2 + v\epsilon_i.$$

where $(X_i)_{1\le i\le n}$ and $(\epsilon_i)_{1\le i\le n}$ are two i.i.d. sequences distributed as $N(0,1)$ and v is an appropriately chosen constant (here we take $v = 0.2$).

In practice, some tuning parameters have to fixed: the kernel K is chosen by

$$K(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2}\sum_{j=1}^{n} x_j^2\right),$$

and the kernel H is defined by

$$H(y) = \left(\frac{3}{4}\right)^n \prod_{j=1}^{n}\left(1 - y_j^2\right).$$

The selection of the bandwidth h is an important and basic problem in kernel smoothing techniques. In this simulation, we choose the optimal bandwidth by the cross-validation method.

In this context, we employ the MCLS, MCHS and MCPRP algorithms to solve the problem (5.2) under strong Wolfe line search technique. According to Table 4, it is clear that the MCLS, MCHS and MCPRP are efficient for solving the problem (5.2) based on number of iterations and CPU time.

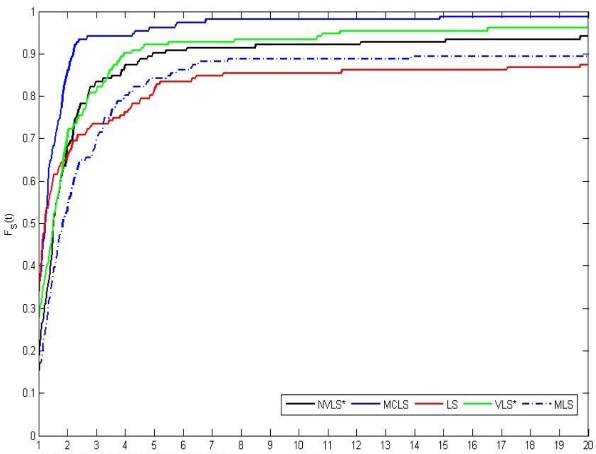

Figure 1. Performance profile on the CPU time.

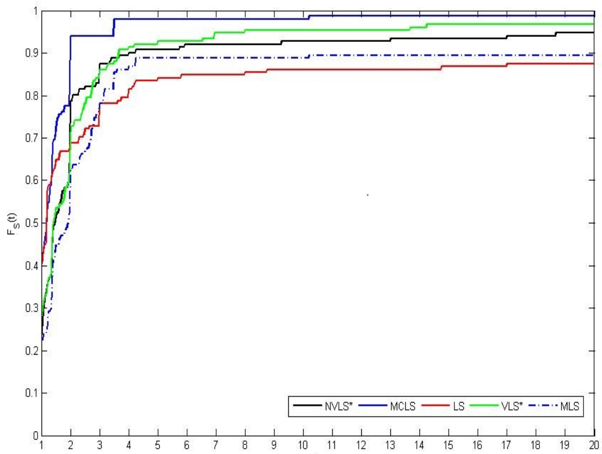

Figure 2. Performance profile on the number of iterations.

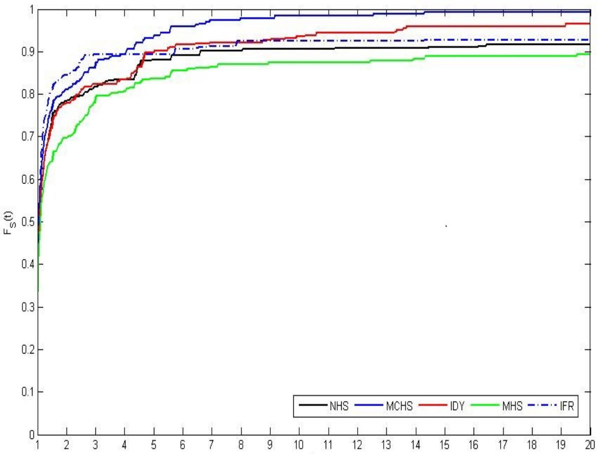

Figure 3. Performance profile on the CPU time.

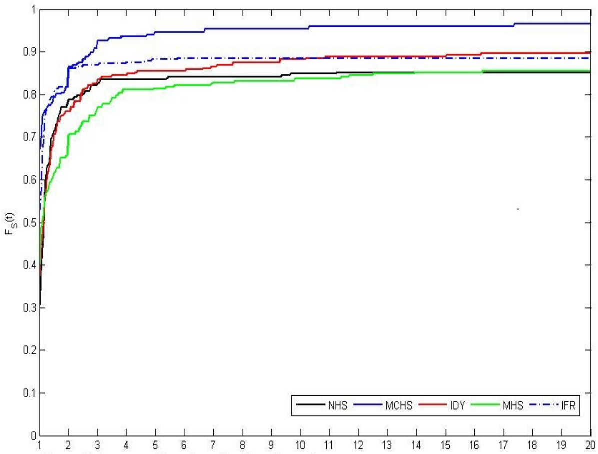

Figure 4. Performance profile on the number of iterations.

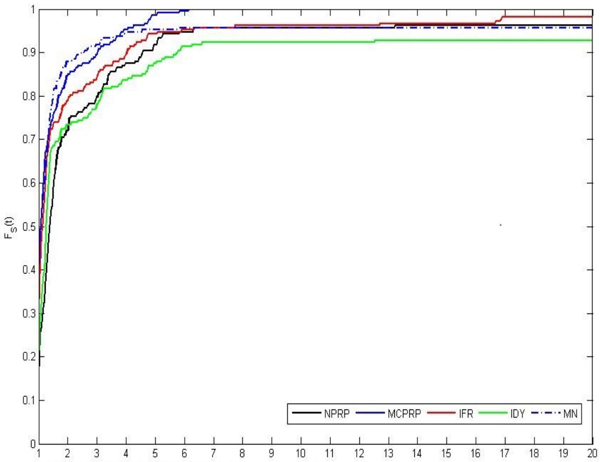

Figure 5. Performance profile on the CPU time.

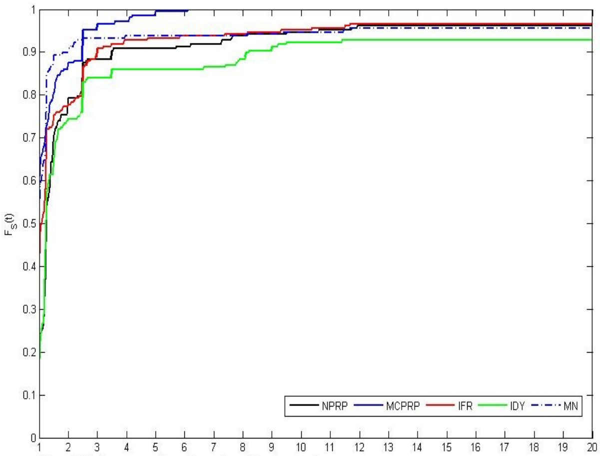

Figure 6. Performance profile on the number of iterations.

Table 1: The simulation results of MCLS, NVLS*, MLS, VLS*, and LS methods.

Function	Dim	MCLS		NVLS*		MLS		VLS*		LS	
		TIME	ITR	TIME	ITR	TIME	ITR	TIME	ITR	TIME	ITR
ZAKHAROV	1000	0.1090	3	0.1250	4	0.1090	3	0.1560	4	0.1250	3
	2800	0.2500	3	0.3750	4	0.2970	3	0.4220	4	0.2660	3
	6000	0.6130	3	0.7340	4	0.8060	3	0.7810	4	0.7340	3
	10000	1.0000	3	1.3910	4	0.9220	3	1.4220	4	1.0310	3
Sumsquares	1700	4.3900	96	3.5150	75	4.9060	106	4.5780	96	4.8900	106
	5000	27.7160	164	25.3500	151	46.1680	267	30.7090	173	46.0890	267
	6000	36.4660	177	36.2000	174	55.1630	240	36.4920	179	52.4850	240
	7000	41.7150	174	52.4000	177	75.2420	314	71.2250	281	78.6790	314
Extended Rosenbrock	3000	2.2650	47	2.0630	41	4.1250	75	4.9520	94	5.2800	95
	3500	1.7180	29	2.1720	40	3.4840	49	2.7520	38	3.1560	60
	4000	5.7810	69	3.2340	40	4.3280	96	8.0280	89	8.7960	102
	6000	2.6560	29	4.5770	50	7.1400	75	7.6240	60	10.6500	72
	14000	7.7970	33	11.5150	50	23.5300	99	8.8430	39	18.1550	78
Schwefel223	2000	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
	3500	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf

Function	Dim	MCLS		NVLS*		MLS		VLS*		LS	
		TIME	ITR	TIME	ITR	TIME	ITR	TIME	ITR	TIME	ITR
Raydan 2	800	0.1250	6	0.1090	6	0.1250	6	0.0780	4	0.1250	6
	3000	11.3590	51	1.7660	9	12.0150	51	1.3950	5	3.1560	20
	4000	4.9840	25	154.2860	474	6.6380	25	52.1500	193	17.1470	85
	5000	4.8400	33	123.7030	534	2.8270	13	106.8030	507	3.4220	20
Raydan 1	80	0.0780	56	0.0780	57	0.1400	104	0.0940	65	0.1350	102
	120	0.1410	59	0.1560	60	0.1720	71	0.5780	156	0.1800	80
	140	0.2930	131	Inf	Inf	Inf	Inf	Inf	Inf	0.2750	123
Styblinski	700	1.1520	25	1.3130	33	Inf	Inf	2.4570	59	1.1560	29
	900	3.3630	63	3.4760	65	Inf	Inf	5.4880	88	4.2080	77
Sphere	5000	0.7030	11	0.5620	11	0.5620	11	0.4220	8	0.6250	11
	6000	0.5560	10	0.6720	11	0.7180	11	0.5630	8	0.6560	11
	12000	1.1870	10	1.3360	11	1.2970	11	1.0620	9	1.2970	11
Rastrigin	2800	9.1360	92	Inf	Inf	7.9480	92	12.2910	132	14.1390	92
	3000	3.8590	44	Inf	Inf	4.7180	44	Inf	Inf	0.9180	12
	4000	16.0330	230	188.7660	1062	38.7660	232	19.9890	135	91.8200	590
	6000	16.0150	78	19.1700	88	17.9610	81	55.1920	218	18.3620	94
Quartic	900	1.9680	188	1.9690	188	1.9690	189	1.9710	189	1.9680	189
	1600	4.5150	26	4.6400	27	4.5200	27	4.5150	26	4.5310	27
	3000	7.8590	210	7.9370	213	9.5270	215	5.4990	123	10.6110	211
Quadratic	1800	4.8280	159	5.6280	166	18.3190	272	10.7050	242	9.5150	287
	3000	2.2880	36	2.4060	44	3.2830	47	2.5720	41	3.0370	48
	4000	3.3400	39	3.4580	42	4.2710	48	3.5660	43	4.0010	48
Qing	1000	0.0690	3	0.1470	3	0.1540	3	0.1310	3	0.0850	3
	3000	0.5330	4	0.4630	3	0.4830	3	0.5010	4	0.2920	3
Power	340	0.7200	72	1.1600	116	0.8150	83	0.7360	74	0.7280	74
	600	1.7830	116	2.0620	131	3.0470	158	2.0940	131	2.0700	133
Perquadratic	4400	2.9640	26	2.9660	28	3.2560	33	3.0970	28	3.9520	36
	4600	4.5810	39	4.2360	37	3.9210	38	3.3340	33	4.3760	39
	5000	4.8460	41	6.5550	53	6.1080	50	6.6430	54	5.8850	45
Penalty	200	0.0620	10	0.1090	16	0.1100	15	0.1870	24	0.0780	14
	800	0.3750	20	1.9220	84	0.8600	36	0.6090	29	0.5620	26
	2600	0.6090	18	1.6250	36	3.2660	65	2.4690	54	0.7660	20
Extended Himmelblau	2800	0.8090	17	1.0620	22	Inf	Inf	0.8750	19	0.8280	17
	3000	0.9840	17	1.2180	20	Inf	Inf	1.1250	20	0.9530	18
	4000	1.1090	17	1.3900	20	Inf	Inf	1.5060	21	1.2200	18
Hager	200	0.5160	68	0.2350	36	0.5000	66	0.2190	39	0.1880	33
	700	0.9690	62	1.2970	69	23.6610	711	3.0000	130	23.1660	738
	1600	13.8900	259	28.9070	483	Inf	Inf	10.7030	212	Inf	Inf
Griewank	3200	30.8750	836	32.1010	836	31.5540	836	31.2470	836	31.1010	836
	3600	33.2660	829	34.1880	829	33.4060	829	33.2820	829	34.1270	829
	5000	49.9620	879	50.8990	879	51.1890	879	52.2810	879	50.8900	879
Dixon	3000	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
	5000	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
Diagonal 4	20000	1.0830	8	1.3130	10	1.2500	10	1.3280	11	1.3590	11
	30000	1.3600	8	1.9590	9	2.0310	10	2.0270	10	2.0160	10
Ridge	600	0.9690	243	Inf	Inf	Inf	Inf	5.4840	780	Inf	Inf
	1000	2.5870	338	Inf	Inf	Inf	Inf	5.7970	610	Inf	Inf
	2000	2.9710	244	Inf	Inf	Inf	Inf	16.0440	726	Inf	Inf
Alpine1	100	0.1400	2	0.2190	5	0.1480	4	0.2340	6	0.2340	6
	400	0.6090	2	1.0330	3	0.8570	3	0.8550	3	0.8550	3
	500	0.1720	2	1.4600	3	1.2980	3	0.8910	3	0.7820	3
Quarticm	800	9.6580	988	9.7430	1092	9.7080	987	10.8510	984	Inf	Inf
	1800	26.2800	1279	26.7600	1280	26.5600	1290	27.5290	1290	26.9390	1289
	2100	32.3810	1333	33.9400	1335	32.8310	1355	17.0250	683	33.0210	1351
TR-Sum of quadratics	100	0.1250	22	0.1250	23	0.1400	24	0.1250	23	0.1250	24
	400	0.3740	24	0.3590	23	0.3600	24	0.3590	23	0.3750	24
	600	0.4530	22	0.4680	23	0.5000	24	0.5320	23	0.4840	24
	5000	3.4530	23	3.5380	23	3.8960	24	3.6530	23	3.8700	24
Linear perturbed	400	2.1090	255	6.6150	702	1.8280	232	2.2910	258	2.2860	262
	800	5.7340	374	5.7650	374	20.9690	984	5.7660	374	17.2810	918
NONDIA	7000	0.0780	2	3.1100	15	4.1750	15	3.3110	11	3.4690	15
	14000	2.6680	5	2.0790	5	2.4070	5	2.7930	5	2.2370	5
	18000	8.7340	15	8.9700	15	8.5480	15	8.3220	11	8.7340	15
Tridiagonal	100	1.9060	573	Inf	Inf	Inf	Inf	1.8050	418	9.3560	1960
	150	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf

Function	Dim	MCLS		NVLS*		MLS		VLS*		LS	
		TIME	ITR	TIME	ITR	TIME	ITR	TIME	ITR	TIME	ITR
Extended Hiebert	1400	1.2290	36	1.9690	106	1.3750	56	3.0630	56	1.7470	106
	2000	3.7030	92	3.2660	92	2.4220	53	2.1250	42	2.9530	65
	3000	4.4840	49	4.2340	102	4.8750	60	18.8410	188	5.7060	69
Extended White and Holst	5000	1.1870	5	1.7820	6	1.2030	5	2.0780	6	1.2180	5
	5400	1.2660	5	1.5470	6	1.3440	5	1.4220	6	1.2780	5
Double Border Arrow Up	1500	0.3910	9	Inf	Inf	0.4530	10	Inf	Inf	0.3940	9
	2000	0.6090	9	Inf	Inf	0.6100	11	Inf	Inf	0.4220	8
Almost Perturbed Quartic	1000	1.6850	122	2.8610	215	2.8540	212	3.6400	246	2.6360	215
	6000	19.9580	261	30.4460	300	17.1660	243	13.4140	148	16.5940	248
ENGVAL1	3000	0.6940	7	0.6570	7	0.5620	6	0.6250	7	0.5470	7
	10000	1.4690	5	1.4840	6	1.5000	6	1.9530	7	2.3440	7
Almost Perturbed Quadratic	600	2.1870	94	2.8780	127	2.7030	131	2.0460	97	2.8410	136
	1500	17.5620	391	23.2610	552	37.7110	692	11.0980	247	40.6110	783
	1800	11.2500	247	21.0050	426	11.4520	225	9.5470	176	10.4950	161
Prod 2	1400	0.0780	5	Inf	Inf	0.1720	6	0.2030	6	0.1710	6
	1800	Inf	Inf	Inf	Inf	Inf	Inf	0.2030	6	Inf	Inf
Prod 1	3000	2.1260	9	2.4390	10	2.3470	11	2.4130	10	2.3430	11
Fletcher	1000	0.1720	6	0.2190	7	0.2030	7	0.2030	7	0.2190	7
	1200	0.2500	7	0.2650	7	0.0780	7	Inf	Inf	0.2350	7
	1600	0.3600	6	0.3440	7	0.2970	7	0.2970	7	0.3120	7
Staircase S3	80	0.3750	314	0.3990	322	2.3440	1446	0.4530	371	1.0620	833
	110	0.8280	472	0.6560	407	Inf	Inf	1.7340	814	Inf	Inf
	1700	1.8910	667	2.0150	667	Inf	Inf	Inf	Inf	Inf	Inf
Staircase S2	30	0.0620	112	0.0780	118	0.2340	395	0.0780	148	0.1250	227
	100	0.7630	409	0.6410	392	Inf	Inf	1.5460	893	1.8340	1032
	300	10.7980	34	8.6190	36	Inf	Inf	Inf	Inf	Inf	Inf
Diagonal 2	400	0.9840	265	0.4850	177	Inf	Inf	1.8530	444	1.1650	339
	900	1.7340	262	2.7810	340	Inf	Inf	11.3130	972	6.3320	665
Diagonal 1	140	0.8750	157	1.4400	227	Inf	Inf	4.3640	630	4.3750	630
	160	0.6560	141	1.4710	227	Inf	Inf	7.2660	955	7.5030	955
Staircase S1	70	0.5680	450	0.6570	467	1.8440	1394	0.5290	447	2.2500	1474
	120	1.0420	580	1.1100	618	5.2810	1750	4.9360	1445	4.4370	1489
Liarwhd	5000	1.0870	8	0.8910	6	1.0790	7	0.8550	6	0.8550	2
	7000	1.4380	7	1.2340	6	1.3910	7	1.1280	6	0.2500	2

Table 2: The simulation results of MCHS, IDY, MHS, IFR and NHS methods.

Function	Dim	MCHS		IDY		MHS		IFR		NHS	
		TIME	ITR	TIME	ITR	TIME	ITR	TIME	ITR	TIME	ITR
Schwefel 223	3000	0.2720	5	0.1880	4	0.1870	4	0.2500	5	0.1720	3
	4000	0.2040	3	0.2190	3	0.3360	4	0.3440	3	0.2340	3
Extended Rosenbrock	1200	2.0820	62	2.0940	62	5.8600	97	7.1410	80	2.0000	61
	1250	2.0000	59	2.1720	63	Inf	Inf	14.7930	120	2.7350	63
	1300	2.8220	64	3.1850	65	Inf	Inf	2.9680	66	2.4540	65
	1600	3.2760	71	3.3130	74	Inf	Inf	3.5780	79	3.7810	80
	1800	3.8280	79	3.9070	78	Inf	Inf	3.9220	78	4.6270	88
Extended Rosenbrock	2000	1.0310	35	1.3440	45	1.1720	43	1.7650	61	1.8130	69
	2200	1.0780	32	1.5940	49	2.0160	57	2.6250	68	2.6100	76
	2700	1.9840	39	2.1410	43	1.5940	32	3.3910	69	3.0000	61
	3500	2.3910	38	2.8120	42	3.0620	45	3.5940	52	3.3280	60
Ridge	1000	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
	2000	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
Raydan 2	1000	0.0930	5	0.2030	9	0.1090	6	0.0940	5	0.1250	6
	3400	3.5630	28	2.3280	19	4.0470	33	2.2810	19	3.4530	28
	5000	2.5000	15	16.1040	89	217.7530	957	204.4080	914	3.6570	20

Function	Dim	MCHS		IDY		MHS		IFR		NHS	
		TIME	ITR	TIME	ITR	TIME	ITR	TIME	ITR	TIME	ITR
Raydan 1	110	0.2160	90	0.1870	88	1.0780	507	0.2500	106	0.1560	83
	120	0.1880	89	Inf	Inf	Inf	Inf	0.2500	124	0.2880	90
	140	0.2340	91	0.3130	129	Inf	Inf	0.3120	92	0.2500	97
Styblinski	800	0.7190	17	1.3750	31	0.9280	18	1.3440	30	1.3880	30
	1960	1.0150	21	3.3210	66	1.4200	25	1.3130	30	2.7960	56
Sphere	8000	0.9220	13	1.0470	12	1.0150	12	0.9370	12	0.9370	13
	10000	1.2240	13	1.3270	12	1.2500	12	1.3280	12	1.6500	13
Rastrigin	2000	0.6570	15	Inf	Inf	2.7970	50	6.5940	112	2.7970	50
	2300	0.5790	12	2.7340	43	5.4060	84	4.9220	75	1.9840	34
	2500	16.1100	224	7.6400	101	46.2500	548	1.4690	25	16.5290	226
	2600	1.0820	15	2.4580	30	11.6270	119	7.6210	86	11.4530	119
Quartic	1400	2.9860	187	3.3120	205	3.5470	215	3.2660	205	3.1720	95
	2000	3.9840	182	4.3590	201	4.6410	211	4.2650	196	4.4220	202
Quadratic	1800	5.8930	154	6.5780	157	5.9890	155	6.6420	157	12.5850	287
	1900	6.2970	158	6.3280	156	7.1470	169	16.4540	358	15.8020	351
	2700	13.1950	190	13.5800	195	14.2060	192	13.5800	197	29.0620	376
Qing	600	5.6480	243	5.6750	259	16.9050	608	15.0620	619	14.7470	608
	1000	12.1720	335	12.2500	338	37.7970	940	25.2410	655	37.2960	940
	2000	38.8590	494	35.1350	498	118.009	1395	117.014	1378	118.938	1395
Power	1000	9.2640	1062	17.3500	211	14.5990	937	15.2510	937	Inf	Inf
Perquadratic	1400	12.3280	512	14.2190	583	Inf	Inf	13.7750	537	18.4460	703
	1800	18.2630	557	21.4920	686	Inf	Inf	25.8070	746	31.7860	915
Penalty	1200	0.6870	21	8.3080	189	1.8960	49	2.7230	66	1.4610	38
	2000	2.5220	43	7.0570	109	5.5060	78	2.0740	38	2.3210	41
	3400	0.7340	7	63.4470	497	0.7190	7	5.5470	53	0.8590	8
Extended Himmelblau	2800	0.8090	18	1.0620	22	Inf	Inf	0.8750	19	0.8280	17
	3000	1.3840	20	1.2180	20	Inf	Inf	1.1250	20	0.9530	18
	4000	1.1090	17	1.3900	20	Inf	Inf	1.5060	21	1.2200	18
Hager	2600	22.7710	197	75.5450	610	Inf	Inf	232.257	1508	52.1760	466
	2700	60.0100	455	178.269	1163	Inf	Inf	75.8830	603	249.939	1653
	3000	19.6130	168	50.3130	327	167.216	1026	263.094	1560	273.112	1645
	4000	85.5000	456	262.803	1256	374.112	1599	247.534	1215	388.803	1758
Griewank	2000	0.1410	7	5.1410	140	72.0500	1213	5.5890	140	6.1620	140
	2500	0.1840	7	8.0000	225	25.5610	338	8.0790	225	8.0310	225
	3000	0.1100	7	13.5780	292	14.3590	175	16.2790	291	15.9070	291
Dixon	1800	0.4430	9	0.3590	7	0.3490	7	0.3440	7	0.3440	7
	7000	1.1250	6	1.3520	7	1.4220	7	1.3440	7	1.4060	7
	8000	1.4220	7	1.8590	7	1.8590	7	1.7970	7	1.7810	7
Diagonal 4	20000	1.0830	8	1.3130	10	1.2500	10	1.3280	11	1.3590	11
	30000	1.8600	9	1.9590	9	2.0310	10	2.0270	10	2.0160	10
Diagonal 2	3100	40.7740	819	24.9060	632	Inf	Inf	105.957	1577	Inf	Inf
	4000	46.3750	779	38.4800	708	Inf	Inf	108.287	1396	Inf	Inf
Alpine1	100	0.1400	2	0.2190	5	0.1480	4	0.2430	6	0.2430	6
	400	0.8909	2	1.0320	3	0.8570	4	0.8550	3	0.8550	3
	500	0.1720	2	1.4600	3	1.2980	3	0.8910	3	0.7820	3
Quarticm	1300	0.4840	6	0.7190	11	17.8130	1162	17.0920	1165	17.6410	1154
	1500	0.5290	7	0.7030	10	21.4530	1122	0.7340	10	20.6090	1214
	1600	0.7090	10	0.7630	9	0.7280	10	0.7030	9	22.8430	1239
T8-Sum of quadratics	2000	1.2980	23	1.3280	24	1.4060	25	1.3440	24	1.3290	24
	1200	2.6380	25	2.4680	24	2.6250	25	2.5630	24	2.4690	24
Linear perturbed	1500	3.0090	23	2.5950	24	3.5120	25	3.2810	24	3.2630	24
	400	1.8440	221	2.0920	229	2.1210	232	2.0990	265	4.1400	482
	900	5.7900	352	6.1910	367	5.4940	328	6.1870	365	17.8400	874
NONDIA	1000	0.1620	10	26.0160	719	27.0950	719	0.1630	10	26.4370	719
	1400	0.1730	12	35.0710	719	35.1940	719	0.1780	10	35.9530	719
	1800	0.1780	11	45.8100	719	44.3400	719	0.1780	10	45.0720	719
Tridiagonal	1000	1.9060	462	Inf	Inf	0.8750	304	1.9950	473	10.7490	1960
	1500	1.8900	470	Inf	Inf	2.6250	555	6.7420	945	Inf	Inf
Extended Hiebert	1400	1.4970	5	3.0310	6	9.3380	6	1.5000	6	1.4840	5
	2000	1.9480	43	3.0940	89	12.9840	245	1.8120	39	1.9570	43
	4000	4.1720	41	5.6320	67	4.7350	41	4.8750	45	7.3280	69
Extended White and Holst	5000	1.1870	5	1.7820	6	1.2030	5	2.0780	6	1.2180	5
	5400	1.5190	6	1.8310	6	1.3590	5	1.7530	6	1.3910	5
Double Border	1500	0.5910	9	Inf	Inf	0.4530	10	Inf	Inf	0.3940	9
Arrow Up	2000	0.6090	9	Inf	Inf	0.6100	11	Inf	Inf	0.4220	8
Almost Perturbed Quartic	700	1.2080	62	0.6800	46	2.5210	121	3.0070	149	2.6880	136
	1000	1.1100	64	1.6790	68	4.2530	142	3.5040	99	3.7500	128
ENGVAL1	3000	0.6250	5	0.7180	7	0.6800	6	0.8130	7	0.6720	7

Function	Dim	MCHS		IDY		MHS		IFR		NHS	
		TIME	ITR	TIME	ITR	TIME	ITR	TIME	ITR	TIME	ITR
Almost Perturbed Quadratic	2000	0.7030	6	0.8440	6	0.7970	6	0.9250	6	0.7560	6
	4000	0.7660	7	0.9660	6	0.9840	6	1.3030	6	1.4070	6
	6000	0.7030	7	1.4840	6	1.5780	6	1.4550	6	1.3910	6
Prod 2	1400	0.0790	5	Inf	Inf	0.1720	6	0.2030	6	0.1710	6
	1800	Inf	Inf	Inf	Inf	Inf	Inf	0.2030	6	Inf	Inf
Prod 1	3000	2.1260	9	2.4390	10	2.3470	11	2.4130	10	2.3430	11
	1000	0.1720	6	0.2050	7	0.2030	7	0.2030	7	0.2130	7
Fletcher	1200	0.2500	6	0.2550	7	0.2970	7	Inf	Inf	0.2250	7
	1600	0.3600	6	0.3440	7	0.2970	7	0.2970	7	0.3120	7
Staircase S3	60	0.1870	194	0.2350	221	0.2030	214	0.2970	332	0.6020	612
	70	0.2520	245	0.3820	265	0.2560	248	0.5940	577	1.3910	1138
	120	0.8120	448	0.8900	494	0.7130	398	3.8140	601	6.6930	1058
Staircase S2	10	0.0160	42	0.0160	36	0.0160	36	0.0160	42	0.0320	67

Table 3: The simulation results of MCPRP, NPRP, IDY, IFR and MN methods.

Function	Dim	MCPRP		NPRP		IDY		IFR		MN	
		TIME	ITR	TIME	ITR	TIME	ITR	TIME	ITR	TIME	ITR
Schwefel 223	1000	0.1220	4	0.1410	5	0.1250	5	0.1410	5	0.1560	5
	3000	0.2650	4	0.5950	5	0.2740	5	0.2350	5	0.3440	5
	4000	0.3320	4	0.6210	5	0.3750	5	0.3600	5	0.4730	5
ZAKHAROV	2000	0.2810	4	0.3910	5	0.3440	5	0.2960	5	0.3900	5
	2400	0.2970	5	0.4060	5	0.3750	5	0.3130	5	0.4220	5
	2800	0.4370	5	0.5630	5	0.5160	5	0.5250	5	0.5000	5
Extended Rosenbrock	2200	2.4310	84	2.6850	83	0.5000	17	2.9060	84	2.4590	74
	3500	2.8400	47	4.5720	60	4.5810	63	6.5470	90	6.3520	94
	4500	6.2350	86	6.4550	84	5.3590	73	4.9090	64	6.5510	86
Quartic	800	2.4110	192	2.2340	214	2.7810	248	2.7640	210	3.0080	205
	1500	3.9700	203	5.1250	246	1.1020	15	5.7540	234	4.3550	231
	1600	3.6720	183	4.4040	222	3.7630	194	4.9970	219	4.3580	194
Raydan 2	2000	0.3280	6	0.4070	7	0.5830	10	61.8000	955	0.3750	7
	6000	0.6400	7	0.6560	8	1.0100	10	0.6470	7	0.7340	8
	7000	3.5160	12	217.906	789	3.5410	15	3.1720	14	40.3910	153
	8000	30.5440	113	30.5570	113	2.0940	8	0.4690	2	18.2570	67
Raydan 1	100	0.0190	4	0.0780	38	0.0470	7	0.0310	5	0.0780	36
	1200	0.3030	9	Inf	Inf	0.2400	7	0.3130	9	Inf	Inf
	2000	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
Styblinski	600	1.5160	28	2.2450	49	Inf	Inf	2.0940	48	1.2030	28
	800	1.5460	31	1.7370	31	Inf	Inf	1.6670	38	2.7120	60
	900	3.1560	63	4.3040	84	Inf	Inf	3.3460	66	3.3530	65
Sphere	4000	0.4530	12	0.4370	12	0.3910	10	0.3440	10	0.4690	12
	16000	1.7810	12	1.4880	12	1.5940	10	1.8330	10	1.7030	12
	20000	2.2650	12	2.3900	12	2.0460	10	1.9620	10	2.3270	12
Rastrigin	2300	3.4530	47	5.1720	52	7.4060	108	2.5940	43	2.5790	43
	2800	0.7030	11	8.3920	92	9.2210	58	3.2520	34	8.0840	92
	4000	4.2870	37	42.1380	230	10.7400	54	4.9640	37	36.8400	255
	5000	13.7990	47	91.9610	342	206.1620	1022	13.5060	56	13.8440	48
Quadratic	1000	0.1720	6	18.7370	823	0.6720	23	0.2250	6	40.1560	1389
	2000	6.1720	188	5.9840	163	1.0620	18	6.7190	188	20.1090	524
	3500	1.2660	6	84.3590	906	2.6880	25	0.6090	6	99.4250	1045
	4000	0.6880	6	140.3140	958	2.9060	25	0.6250	6	156.6710	1113
Qing	600	0.0930	4	0.3120	8	0.1720	5	0.1720	5	20.4990	829
	900	0.1250	4	39.1400	1085	0.1720	5	0.1570	5	30.4370	875
	3000	0.4060	4	0.8280	5	0.5940	5	0.6090	5	0.7030	6
Power	260	0.3620	22	20.1180	1530	0.3360	22	0.0700	6	Inf	Inf
	320	0.0630	6	32.1420	1856	0.3850	26	0.0630	6	Inf	Inf
	340	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf

Function	Dim	MCPRP		NPRP		IDY		IFR		MN	
		TIME	ITR	TIME	ITR	TIME	ITR	TIME	ITR	TIME	ITR
Perquadratic	3000	0.7460	11	0.7960	15	1.9530	23	0.9220	19	0.8280	15
	4300	0.7500	11	0.9060	12	3.4530	30	0.9680	13	0.9530	12
	5000	1.4060	15	1.2650	12	3.1890	24	1.3750	14	1.0630	12
Penalty	1800	0.0940	5	0.1090	5	0.9680	24	0.4060	12	0.0930	5
	3000	2.1630	23	1.3600	20	2.1680	23	3.3530	30	1.7260	20
	4600	4.5310	35	24.9900	127	5.1760	33	15.3280	97	105.6910	533
Extended Himmelblau	320	0.1100	17	0.4710	19	0.2040	14	0.1720	22	0.0940	17
	1600	0.3810	17	0.8080	19	0.5870	17	Inf	Inf	0.3590	16
	2200	0.5000	15	0.6760	16	0.6650	16	0.9060	28	0.5470	19
	2800	0.8750	16	0.9820	19	1.0080	11	1.5630	26	0.9660	18
Hager	200	0.1880	33	0.2810	45	0.0630	20	0.2030	40	0.5840	70
Griewank	1500	1.1090	18	1.2340	20	3.1560	37	1.1410	19	1.2500	20
	2100	1.6940	18	1.5780	19	4.1880	36	1.5310	19	1.5780	19
	5000	8.4690	35	8.5990	36	12.9370	36	8.7810	36	8.7030	36
Dixon	4000	0.3210	4	0.5620	5	0.5150	5	0.5470	5	0.5620	5
	5000	0.5160	4	0.7500	5	0.6870	5	0.7810	5	0.7030	5
Diagonal 4	20000	0.8290	6	1.2180	7	1.1320	8	0.8590	7	0.9460	7
	30000	1.9860	11	1.9970	10	1.9240	10	1.8750	10	1.8440	10
Diagonal 2	500	0.2500	4	1.8130	104	1.3160	69	1.0940	69	1.1880	69
Alpine1	100	0.1410	2	0.2240	5	0.2190	5	0.2180	5	0.1880	5
	400	0.6360	2	0.7990	3	0.6250	3	1.2660	4	0.5310	3
	500	0.1720	1	0.9290	3	1.2780	3	1.2970	4	1.2970	3
TR-Sum of quadratics	1000	0.3580	10	0.5000	13	0.3710	10	2.5000	56	0.5470	14
	1800	0.4810	10	0.6410	13	0.4840	10	3.0310	61	0.6870	14
	3000	5.1570	58	1.3130	13	1.2180	13	5.6720	56	1.4640	14
Linear perturbed	500	2.4530	245	7.8290	593	Inf	Inf	2.5440	245	3.1440	295
	1200	11.5100	422	16.4590	619	Inf	Inf	12.4030	414	29.2970	1052
Nondia	1000	26.0880	719	29.7170	719	31.6890	731	27.2220	731	26.9000	720
	3000	81.7560	719	80.8500	719	89.1530	731	92.5330	731	89.4920	730
	4000	108.5620	720	107.5160	719	134.3340	731	113.4800	731	132.5500	730
Extended Hiebert	1400	1.7810	68	2.3820	127	Inf	Inf	2.7760	97	1.3790	56
	2000	2.8910	61	9.0640	199	Inf	Inf	5.4600	101	3.7310	79
	3000	5.3460	66	17.2590	211	Inf	Inf	10.3300	112	5.8470	72
Extended White and Holst	4000	1.0000	5	1.5580	7	1.1300	5	1.2400	5	1.5600	6
	6000	1.2600	5	1.8590	7	1.2110	5	1.3970	5	1.5280	6
	7000	1.7090	5	2.5480	7	1.7010	5	1.9670	5	2.3950	6
Double Border Arrow Up	2000	2.0470	41	1.4370	31	3.1270	51	2.9370	51	1.5940	34
	4500	9.5470	67	4.4690	35	7.6310	38	5.6190	38	3.9220	34
Almost Perturbed Quadratic	1200	1.6540	47	18.0730	456	1.7950	52	3.4110	95	1.9070	98
	1400	2.8210	88	22.4390	518	1.8130	51	1.8600	50	10.3250	256
ENGVAL1	7000	1.1720	6	1.4400	7	1.1880	6	1.2340	6	1.4370	7
	20000	3.7960	6	4.0220	7	3.5580	6	3.6410	6	4.0840	7
Almost Perturbed Quartic	2000	7.9540	95	8.2180	233	8.2180	232	8.7490	233	7.2660	224
	5000	3.7720	21	17.2360	246	17.0860	240	16.3540	241	14.5350	239
	6000	10.3970	27	51.3940	286	32.5190	250	36.8020	251	21.9130	218
Prod 2	2000	0.3280	4	0.5780	5	0.3740	6	0.3600	5	0.3440	5
	4000	0.7030	4	0.7590	5	1.0160	6	0.7500	5	0.7580	5
Prod 1	3000	2.0600	13	1.4380	9	2.0620	13	2.0620	13	2.1340	13
	5000	2.8440	11	3.0150	11	3.4990	11	3.1110	11	3.7120	11
Fletcher	1000	0.1410	6	0.1730	7	0.1780	7	0.2060	7	0.2030	7
	2000	0.2610	6	0.3430	6	0.3220	7	0.4980	7	0.3750	7
	4000	0.7710	7	0.7770	6	0.7380	7	0.8790	7	0.8590	7
Staircase S3	10	0.0150	44	0.0780	76	0.1400	372	0.0160	48	0.0310	89
	100	0.8440	577	5.0320	1917	1.1400	691	0.8900	605	5.4060	1988
	150	2.0620	839	4.5780	1524	3.3910	1193	1.8750	856	Inf	Inf
Staircase S2	100	0.9230	598	1.7840	977	1.2970	673	0.9700	628	3.6360	1417
	120	1.5560	749	Inf	Inf	0.9220	447	1.5940	778	Inf	Inf
Himmelbh	600	0.5000	21	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
	640	0.6710	23	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
Harkerp	1300	0.4190	12	5.4330	133	6.2820	133	5.4880	133	5.4880	133
	2000	0.5990	10	6.7650	159	Inf	Inf	6.8970	159	6.1130	159
Engval 8	100	2.3930	477	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
	200	15.6510	1474	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
	300	13.8760	987	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
Arwhead	1000	0.2650	9	0.1880	9	0.2660	9	0.2660	9	0.2660	9
	2000	0.4840	9	0.3750	9	0.4840	9	0.5150	9	0.4840	9
	3000	0.8140	9	0.6410	9	0.8600	9	1.0650	9	0.9840	9
TRIDIA	60	0.0940	30	0.3130	90	Inf	Inf	Inf	Inf	Inf	Inf
	80	0.1250	30	0.1560	38	Inf	Inf	Inf	Inf	Inf	Inf
	100	0.1570	30	0.1410	27	Inf	Inf	Inf	Inf	Inf	Inf
QUARTICM	1100	14.3070	1082	14.4290	1100	15.4290	1082	14.7940	1079	14.3470	1096
	1200	15.8380	1106	16.0340	1117	15.8170	1122	15.8960	1122	15.6750	1101
	1600	25.9460	1245	25.7220	1246	26.2220	1260	25.8520	1260	26.1480	1263

Table 4: The simulation result of MCLS, MCHS and MCPRP methods for solving problem (5.2)

Kernel	Initial Points	Dim	MCLS	MCHS	MCPRP
			TIME/ITR	TIME/ITR	TIME/ITR
linear	(0.2,...,0.2)	8	18/0.1100	130/0.7180	60/0.3590
		60	206 / 96,1570	41 / 14,2030	320 / 117,3580
		62	23 / 12,1630	114 / 43,5080	62 / 24,4960
		64	64 / 38,3930	67 / 41,2120	441 / 188,002
		68	46 / 22,7970	52 / 23,9690	353 / 170,063
		70	145 / 75,4740	571 / 278,652	173 / 88,1800
		74	32 / 27,0006	36 / 20,0500	32 / 27,3030
		78	226 / 240,868	31 / 5,7300	26 / 25,5250
		80	267 / 177,824	110 / 73,7760	13 / 8,13700
		86	434 / 338,729	83 / 73,4520	244 / 190,7440
		90	11 / 9,3920	85 / 69,6730	43 / 37,1100
		94	25 / 23,2000	94 / 80,3860	256 / 251,2810
		98	111 / 104,113	31 / 26,8440	144 / 16,1100
		100	445 / 89,5300	63 / 61,5480	40 / 42,5340
linear	(-0.5,...,-0.5)	30	327 / 93,5230	836 / 61,7458	847 / 16,1300
		40	138 / 34,7170	Inf / Inf	69 / 7,03100
		45	71 / 14,9240	402 / 94,9850	539 / 118,252
		50	274 / 80,7030	776 / 204,766	Inf / Inf
		55	Inf / Inf	224 / 69,5200	17 / 6,23400
		60	Inf / Inf	52 / 19,7810	39 / 4,67900
		70	139 / 91,7080	127 / 80,5020	142 / 18,2140
		75	208 / 119,677	187 / 109,875	194 / 112,440
		80	255 / 128,372	Inf / Inf	85 / 45,26400
		90	323 / 70,244	526 / 440,123	471 / 95,750
		95	613 / 453,333	399 / 175,912	291 / 192,656
Non linear	(1,.....,1)	10	10 / 0.6830	16 /0.1880	12 /0.1250
		20	21 / 0.8990	17 /0.6880	9 /0.3120
		30	9 / 0.7030	13 /1.0780	8 /0.6180
		40	26 / 2.4210	24 /2.1870	29 /2.4610
		50	32 / 3.5500	46 /3.5880	38 /4.1820
		70	20 / 6.2830	18 /3.3630	16 /6.7350
		80	24 / 12.6610	19 /10.1600	13 /6.4630
		90	38 / 16.3190	41 /19.2740	14 /7.6870
		100	13 / 12.6170	10/8.0470	11 /10.0650
		120	16 /25.2470	11/18.3240	12 /20.1680
		140	26 /47.5550	37/64.6570	13 / 23.6520

6. Conclusion

This paper presented three modified conjugate gradient methods for unconstrained optimization models, that is, MCLS, MCHS and MCPRP methods. Under basic assumptions, we prove that the three improved CG methods satisfy descent condition with the strong Wolfe line search and produces good convergence properties for unconstrained optimization problems.

Preliminary numerical results show that these improved methods are very robust and effective for given test problems. The practical applicability of our methods is also explored in the nonparametric estimation of the regression function.

Acknowledgments

The authors are very grateful to the anonymous referees for their valuable comments and useful suggestions, which improved the quality of this paper.

References

[1] N. Andrei, An unconstrained optimization test functions, Advanced modeling and optimization, 10 (2008) 147-161.
[2] A. M. Awwal, I. M. Sulaiman, M. Malik, M. Mamat, P. Kumam, K. Sitthithakerngkiet, A spectral RMIL conjugate gradient method for unconstrained optimization with applications in portfolio selection and motion control, IEEE access, 9 (2021) 75398-75414.
[3] G. Boente, R. Fraiman, Nonparametric regression estimation, Journal of multivariate analysis, 29 (1989) 180–198.
[4] I. Bongartz, A. R. Conn, N. Gould, P. L. Toint, Constrained and unconstrained testing environment, ACM transactions on mathematical software, 21 (1995) 123-160.
[5] G. Collomb, W. Hardle, S. Hassani, A note on prediction via estimation of the conditional mode function, Journal of statistical planning and inference, 15 (1987) 227-236.
[6] Y. H. Dai, Y. Yuan, Convergence properties of the Fletcher-Reeves method, IMA Journal of numerical analysis, 16 (1996) 155-164.
[7] Y. H. Dai, J.Y. Han, G.H. Liu, D.F. Sun, X. Yin, Y. Yuan, Convergence properties of nonlinear conjugate gradient methods, SIAM journal on optimization, 10 (1999) 348–358.
[8] Y. H. Dai, Y. Yuan, A nonlinear conjugate gradient method with a strong global convergence property, SIAM journal on optimization, 10 (1999) 177-182.
[9] E. D. Dolan, J. J. Morè, Benchmarking optimization software with performance profiles, Mathematical programming, 91 (2002) 201-213.
[10] X. W. Du, P. Zhang, W. Ma, Some modified conjugate gradient methods for unconstrained optimization, Journal of computational and applied mathematics, 305 (2016) 92-114.
[11] R. Fletcher, C. Reeves, Function minimization by conjugate gradients, The computer Journal, 7 (1964) 149-154.
[12] R. Fletcher, Practical Methods of Optimization, Wiley, New york, 1987.
[13] J. C. Glibert, J.Nocedal, Global covergence properties of conjugate gradient method for optimization. SIAM Journal on optimization, 2 (1992) 21-42.
[14] W. W. Hager, 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, Journal of Research of the national bureau of standards, 49 (1952) 409-436.
[16] X. Z. Jiang, J. B. Jian, Two modified nonlinear conjugate gradient methods with disturbance factors for unconstrained optimization, Nonlinear dynamics, 77 (2014) 387-397.
[17] X. Z. Jiang, J. B. Jian, Improved Fletcher-Reeves and Dai-Yuan conjugate gradient methods with the strong Wolfe line search, Nonlinear dynamics, 328 (2019) 525-534.
[18] J. Jian, P. Liu, X. Jiang , B. He, Two Improved Nonlinear Conjugate Gradient Methods with the Strong Wolfe Line Search, Bulletin of the iranian mathematical society, 48 (2020) 2297-2319.
[19] Y. Liu, C. Storey, Efficient generalized conjugate gradient algorithms, Journal of optimization theory and applications, 69 (1991) 129-137.
[20] J. K. Liu, S.J. Li, New hybrid conjugate gradient method for unconstrained optimization, Applied mathematics and computation, 245 (2014) 36-43.
[21] A. E. Mehamdia, H.M. Khudhur, A globally convergent of two conjugate gradient methods with application to image restoration problems, Numerical Algebra control and Optimization, (2024).
[22] A. E. Mehamdia, Y Chaib, Improved conjugate gradient methods and application to nonparametric estimation, Applicationes Mathematicae, (2024) 1-15.
[23] A. E. Mehamdia, Y. Chaib, T. Bechouat, Two improved nonlinear conjugate gradient methods with application in conditional model regression function, Journal of Industrial and Management Optimization 21 (2025) 658-675.
[24] A. E. Mehamdia, Y. Chaib, Two modified conjugate gradient methods for unconstrained optimization, Optimization Methods and Software, (2025) 1-14.
[25] E. Polak, G. Ribière, Note sur la convergence de directions conjuguée, Revue française d'informatique et de recherche opérationnelle, 16 (1969) 35-43.
[26] B. T. Polyak, The conjugate gradient method in extreme problems, USSR computational mathematics and mathematical physics, 9 (1969) 94-112.

[27] M. J. D. Powell, Nonconvex minimization calculations and the conjugate gradient method, Lecture notes in mathematics, 1066 (1984) 122-141.
[28] M. Samanta, A. Thavaneswaran, Nonparametric estimation of the conditional mode, Communications in statistics theory and methods, 16 (1990) 4515-4524.
[29] S. Yao, Z. Wei, H. Huang, A notes about WYL's conjugate gradient method and its applications, Applied mathematics and computation, 191 (2007) 381-388.
[30] L. Zhang, An improved Wei–Yao–Liu nonlinear conjugate gradient method for optimization computation, Applied mathematics and computation, 6 (2009) 2269-2274.
[31] G. Zoutendijk, Nonlinear programming computational methods, Journal integer and nonlinear programming, (1970) 37-86.