A Class of Nonmontone Line Search Method with Perturbations

Meixia Li
School of Mathematics and Information Science, Weifang University, Weifang, Shandong, 261061, China
Email: limeixia001@163.com

***Abstract*— In this paper, a new kind of nonmontone line search method which is called new hybrid projection method with perturbations is proposed. At the same time, global convergence of this kind of method is proved only in the case where the gradient function is uniformly continuous on an open convex set containing the iteration sequence. In doing so, we remove various boundedness conditions. Furthermore, we obtain that the convergence property of gradient-type method with new nonmontone linear search method will not be changed when search directions are perturbed slightly. Numerical examples are given in the third section of this paper.**

***Index Terms*— Nonmonotone line search method, Hybrid projection method, Perturbation, Global convergence**

I. INTRODUCTION

We consider the unconstrained optimization problem

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

where R^n denotes the $n-$dimensional *Euclidean* space and $f : R^n \to R$ is a continuously differentiable function. There are many iterative schemes for solving (1.1). Among them the line search method has the form

$$x_{k+1} = x_k + \alpha_k d_k, k = 0, 1, 2, \ldots, \tag{1.2}$$

where d_k is a descent direction of $f(x)$ at x_k and α_k is a step size. Denote x_0 the initial point and x_k the current iterate at the kth iteration. Generally, we denote $f(x_k)$ by f_k, $\nabla f(x)$ by $g(x)$, $\nabla f(x_k)$ by g_k and $f(x^*)$ by f^*, respectively. The search direction d_k is generally required to satisfy

$$g_k^T d_k < 0. \tag{1.3}$$

The set that consists of all the stationary points of problem (1.1) is denoted by Ω^*, that is, $\Omega^* = \{x \in R^n | g(x) = 0\}$.

There are many methods for solving (1.1), for example, gradient method, conjugate gradient method, Newton method, quasi-Newton method, trust region method, *et al* (see [1-7]). In line search methods, if the search direction d_k is given at the kth iteration then the next task is to find a step size α_k along the search direction. The ideal line search rule is the exact one that satisfies

$$f(x_k + \alpha_k d_k) = \min_{\alpha>0} f(x_k + \alpha d_k). \tag{1.4}$$

In fact, the exact step size is difficult or even impossible to seek in practical computation, and thus many researchers constructed some inexact line search rules, such as Armijo rule, Goldstein rule, Wolfe rule and nonmonotone line search rules(see [1,2,8]).

As to nonmonotone line search methods, the descent property is not guaranteed for every step. However, the nonmonotone line search rules are effective or even powerful at some iterations, especially when the iterates are trapped in a narrow curved valley of objective functions. Since Grippo, Lampariello, and Lucidi proposed the nonmonotone line search rule for Newton methods, the new line search approach has been studied by many authors (e.g. [8-12]). Although it has many advantages, especially in the case of iterates trapped in a narrow curved valley of objective functions, the nonmonotone line search rule has some drawbacks(see [13]). Therefore, Shi and Shen[13] proposed a new nonmonotone line search for general line search method, which is described as follows.

New nonmonotone line search (NNLS): Let M be a nonnegative integer. For each k, let $m(k)$ satisfy

$$m(0) = 0, 0 \le m(k) \le \min[m(k-1), M], \forall k \ge 1. \tag{1.5}$$

Given $\beta \in (0, 1)$, $\sigma \in (0, \frac{1}{2})$ and $\delta \in [0.5, 2)$, B_k is a symmetric positive definite matrix that approximates the Hessian of $f(x)$ at the iterate x_k and

$$s_k = -\frac{\delta g_k^T d_k}{d_k^T B_k d_k}.$$

Choose α_k to be the largest α in $\{s_k, s_k\beta, s_k\beta^2, ...\}$ such that

$$f(x_k + \alpha d_k) - \max_{0\le j\le m(k)} f_{k-j} \le \sigma\alpha[g_k^T d_k + \frac{1}{2}\alpha d_k^T B_k d_k]. \tag{1.6}$$

The new line search is a novel scheme of the nonmonotone Armijo line search and allows one to find a larger accepted step size and possibly reduces the function evaluations at each iteration.

In this paper, we propose a kind of nonmontone line search method with perturbations (see Algorithm 2.1). We prove the iteration sequence $\{x_k\}$ generated by the algorithm satisfies either $f_k \to -\infty$ or f_k converges to

This paper is based on “A Class of Hybrid Projection Methods with Perturbations,” by M.X. Li, which appeared in the Proceedings of the 2011 International Conference on Computer Education, Simulation and Modeling(CESM2011), Beijing, China, June 2011.

This research is supported by the National Natural Science Foundation of China (10901096) and the Natural Science Foundation of Shandong Province (ZR2009AL019) and the Project of Shandong Province Higher Educational Science and Technology Program (J09LA53).

doi:10.4304/jcp.7.4.941-946

finite value and $g_k \to 0$ only in the case where $g(x)$ is uniformly continuous on an open convex set containing the iteration sequence $\{x_k\}$. In doing so, we remove various boundedness conditions such as boundedness from below of $f(\cdot)$, boundedness of x_k, etc. By the analysis, we know that the convergence property of gradient-type method with new nonmontone linear search method will not be changed when search directions are perturbed slightly.

The rest of this paper is organized as follows. In the second section, we propose the kind of nonmontone line search method with perturbations and prove its convergence property. In the third section, we give the numerical examples. The conclusions of this paper are showed in the forth section.

II. New Hybrid Projection Methods with Perturbations

In this paper, the algorithms have the following iterative scheme

$$x_{k+1} = x_k + \alpha_k d_k \tag{2.1}$$

and

$$d_k = s_k + \omega_k. \tag{2.2}$$

In the above-mentioned two formulae, the main direction s_k satisfies the following conditions.

$(H_1) : \|s_k\| \le c_1\|g_k\|$.

$(H_2) : \langle g_k, s_k\rangle \le -c_2\|g_k\|^2$.

Perturbation term w_k satisfies

$(H_3) : \|w_k\| \le \gamma_k(q + p\|g_k\|)$

and $\gamma_k > 0$ satisfies

$(H_4) : \sum\limits_{k=1}^{\infty} \gamma_k^2 < +\infty$,

where c_1, c_2, q and p are positive constants.

Let $N = \{1, 2, \ldots\},\ I = \{k \in N|\langle g_k, s_k + \omega_k\rangle \ge 0\}$ and $J = N \setminus I$.

The new hybrid projection methods with perturbations are described as follows.

Algorithm 2.1.

Given a nonnegative integer $M \ge 1$, $x_0 \in R^n$, $\beta \in (0, 1)$, $\sigma \in (0, \frac{1}{2})$ and $\delta \in [0.5, 2)$, $\mu_0, \mu_1, \gamma \in (0, 1)$ and a symmetric positive definite matrix B_0. Set $k := 0$.

Step 1 If $g_k = 0$, then stop. x_k is a stationary point. Else, goto step 2.

Step 2 If $k \in I$, then let $x_{k+1} = x_k + \gamma_k d_k, k = k+1$, return to step 1. Else, goto step 3.

Step 3 Let $\alpha_k = \gamma^{m_k}$, where m_k is the smallest nonnegative integer satisfying

$$\langle g(x_k + \alpha_k d_k), d_k\rangle \le \mu_0\langle g_k, d_k\rangle \tag{2.1}$$

and

$$\langle g(x_k + \alpha_k d_k), g_k\rangle \ge \mu_1\|g(x_k + \alpha_k d_k)\|^2. \tag{2.2}$$

Step 4 Set $\zeta_k = x_k - x_{k-1}$, $\eta_k = g_k - g_{k-1}$ and modify B_{k-1} as B_k by using BFGS or DFP formula or other quasi-Newton formulae.

Step 5 Let $y_k = x_k + \alpha_k d_k$, $v_k = g(y_k)$, $P_k = -\frac{\langle v_k, x_k - y_k\rangle}{\|v_k\|^2} v_k$. $x_{k+1} = x_k + \lambda_k P_k$, where λ_k is defined by the NNLS.

Step 6 Set $k := k + 1$, return to step 1.

In the following, we prove the convergence property of Algorithm 2.1. We first assume that $\{x_k\}$ is an infinite sequence generated by Algorithm 2.1. The following assumptions are satisfied.

(H_5) The objective function $f(x)$ has a lower bound on R^n.

(H_6) The gradient $g(x)$ is uniformly continuous on an open convex set D that contains $\{x_k\}$.

(H_7) There exists $m > 0$, for any k,

$$P_k^T B_k P_k \ge m\|P_k\|^2.$$

The assumption $(H_5)$is very mild. Because the objective function $f(x)$ can be replaced by $e^{f(x)}$, if the assumption is not satisfied.

According to the related reference, we have the following Lemma.

Lemma 2.1[14]**.** Suppose that $(H_2), (H_3)$ and (H_4) hold for $s_k,\ w_k$ and γ_k. Then when $k \in I$ is sufficiently large, we have

$$\|g_k\| \le c_3\gamma_k,$$

where $c_3 > 0$ is a constant.

Lemma 2.2. Assume that (H_5) and (H_7) hold. Then

$$\max_{0\le j\le m(k)} f_{k-j} - f_{k+1} \ge 0, \forall k \in J.$$

proof. When $k \in J$, by Algorithm 2.1, there exists $l \in \{0, 1, \ldots\}$, such that $\lambda_k = s_k\beta^l$ and according to (1.6),we have

$$\begin{aligned} & \max_{0\le j\le m(k)} f_{k-j} - f_{k+1} \\ \ge\ & -\alpha_k\sigma[g_k^T d_k + \tfrac{1}{2}\alpha_k d_k^T B_k d_k] \\ =\ & \rho\beta^l \tfrac{\delta g_k^T d_k}{d_k^T B_k d_k}[g_k^T d_k - \tfrac{1}{2}\beta^l \tfrac{\delta g_k^T d_k}{d_k^T B_k d_k} d_k^T B_k d_k] \\ =\ & \tfrac{\rho\beta^l\delta(2-\delta\beta^l)}{2}\tfrac{(g_k^T d_k)^2}{d_k^T B_k d_k} \\ \ge\ & 0. \end{aligned}$$

Therefore, the conclusion is true.

Lemma 2.3. Let

$$k - m(k) \le l(k) \le k,$$

$$f_{l(k)} = \max_{0\le j\le m(k)} f_{k-j}$$

and (H_5) and (H_7) hold. Then when $k \in J$, $\{f_{l(k)}\}$ is monotone and non-increasing.

proof. When $k \in J$, by Lemma 2.2, we have

$$f_{k+1} \le \max_{0\le j\le m(k)} f_{k-j} = f_{l(k)}.$$

Therefore, by the definition of $f_{l(k+1)}$, we have

$$f_{l(k+1)} \le f_{l(k)}.$$

Lemma 2.4. Assume that $(H_1), (H_3) - (H_7)$ hold. Then there exists a constant $c_4 > 0$ such that $\{f_{l(k)} + c_4T_k\}$ is monotone and non-increasing, where $T_k = \sum\limits_{j=k}^{\infty} \gamma_j^2$. Further, the sequence $\{f_{l(k)}\}$ is convergent (which converges to a finite value or $-\infty$).

The conclusion may be obtained analogous to the proof of Lemma 2.3 in [14]. For the convenience of readers, we give the proof as follows.

proof. For $k \in I$ is sufficiently large, by $(H_1), (H_3)$ and Lemma 2.1, we have

$$\begin{aligned} \|s_k + w_k\| &\le c_1\|g_k\| + \gamma_k(q + p\|g_k\|) \\ &\le \gamma_k(c_1c_3 + q + pc_3\gamma_k). \end{aligned} \tag{2.3}$$

Hence

$$\lim_{k\in I, k\to\infty} \|s_k + \omega_k\| = 0. \tag{2.4}$$

It follows from the median value theorem, Cauchy-Schwartz inequality and (2.3) that

$$\begin{aligned} & f_{k+1} - f_k \\ = & \gamma_k(\langle g(x_k + \theta_k\gamma_k(s_k + w_k)) - g_k), s_k + w_k\rangle \\ & +\langle g_k, s_k + w_k\rangle) \\ \le & \gamma_k\|s_k + w_k\|(\|g(x_k + \theta_k\gamma_k(s_k + w_k)) - g_k\| \\ & +\|g_k\|) \\ \le & \gamma_k^2(c_1c_3 + q + pc_3\gamma_k)(\|g(x_k + \theta_k\gamma_k(s_k + w_k)) \\ & -g_k\| + c_3\gamma_k), \end{aligned}$$

where $\theta_k \in (0,1)$.

By (2.4), the above inequality and the uniform continuity of $g(x)$ on D, there exists a constant $c_4 > 0$ such that

$$f_{k+1} - f_k \le c_4\gamma_k^2$$

when $k \in I$ is sufficiently large.

It follows from the above inequality that when $k \in I$ is sufficiently large,

$$f_{k+1} + c_4T_{k+1} \le f_k + c_4T_k. \tag{2.5}$$

According to

$$f_{l(k)} = \max_{0\le j\le m(k)} f_{k-j},$$

we have

$$f_{k+1} + c_4T_{k+1} \le f_{l(k)} + c_4T_k.$$

Hence

$$f_{l(k+1)} + c_4T_{k+1} \le f_{l(k)} + c_4T_k. \tag{2.6}$$

When $k \in J$, (2.6) is obviously true by Lemma 2.3. Therefore the sequence $\{f_{l(k)} + c_4T_k\}$ is monotone and non-increasing. It follows from (H_4) that $T_k \to 0 (k \to \infty)$. Hence the sequence $\{f_{l(k)}\}$ is convergent (which converges to a finite value or $-\infty$). This completes the proof.

Denote

$$f^{'} = \lim_{k\to\infty} f_{l(k)}, N = \{1, 2, ...\}$$

and

$$K_i = \{l(k') - i - 1 | k' \in N, l(k') - i - 1 \ge 1\},$$

where $i = 0, 1, 2, ..., M-1$.

In addition, denote

$$K_{-1} = \{l(k^{'}) | k^{'} \in N\}.$$

According to the definition of $f_{l(k')}$, we have

$$\begin{aligned} l(k'+1) - l(k') &\le l(k'+1) - l(l(k'+1) - 1) \\ &\le l(k'+1) - \{l(k'+1) - 1 - \\ &\quad \min\{l(k'+1) - 1, M-1\}\} \\ &= 1 + \min\{l(k'+1) - 1, M-1\} \\ &\le M. \end{aligned}$$

Thus, $\bigcup_{i=0}^{M-1} K_i = N$.

Without loss of generality, we assume that

$$\lim_{k\to\infty} f_{l(k)} = f^*.$$

Lemma 2.5. Let $\{x_k\}$ be an infinite iteration sequence generated by Algorithm 2.1. If $g(x)$ is uniformly continuous on an open convex set D containing $\{x_k\}$ and

$$\lim_{k\in K_{i-1}, k\to\infty} f_k = f^*(0 \le i \le M-1),$$

then we have

(1) $\lim_{k\in K_i, k\to\infty} g_k = 0$

and

(2) $\lim_{k\in K_i, k\to\infty} f_k = f^*$.

proof. Now we show that (1) holds in two cases.

Case 1: When $I \cap K_i$ is an infinite index set, it follows from Lemma 2.1 that

$$\lim_{k\in I\cap K_i, k\to\infty} g_k = 0.$$

Case 2: When $J \cap K_i$ is an infinite index set, suppose, on the contrary, that there exist an infinite subset $\bar{K}_i \subseteq J \cap K_i$ and $\epsilon_0 > 0$ such that

$$\|g_k\| \ge \epsilon_0, \qquad \forall k \in \bar{K}_i. \tag{2.7}$$

Utilizing $(2.1), (2.2), (H_2), (H_3)$ and (2.7), we have

$$\begin{aligned} & -\langle g_k, P_k\rangle \\ = & \frac{\langle v_k, x_k - y_k\rangle}{\|v_k\|^2}\langle g_k, v_k\rangle \\ \ge & \mu_1\langle v_k, x_k - y_k\rangle \\ = & -\mu_1\alpha_k\langle g(y_k), d_k\rangle \\ \ge & -\mu_0\mu_1\alpha_k\langle g_k, d_k\rangle \\ = & -\mu_0\mu_1\alpha_k[\langle g_k, s_k\rangle + \langle g_k, \omega_k\rangle] \\ \ge & \mu_0\mu_1\alpha_k[c_2\|g_k\|^2 - \|g_k\|\|\omega_k\|] \\ \ge & \mu_0\mu_1\alpha_k[c_2\|g_k\|^2 - \gamma_k\|g_k\|(q + p\|g_k\|)] \\ \ge & \mu_0\mu_1(c_2 - \gamma_kp - \gamma_kq/\epsilon_0)\alpha_k\|g_k\|^2 \end{aligned} \tag{2.8}$$

for $k \in \bar{K}_i$.

On the other hand, by Algorithm 2.1, $(H_1), (H_3)$ and (2.7), we have

$$\begin{aligned} & \|P_k\| \\ = & \frac{|\langle v_k, x_k - y_k\rangle|}{\|v_k\|} \\ \le & \|x_k - y_k\| \\ = & \|\alpha_kd_k\| \\ \le & \alpha_k(c_1\|g_k\| + \gamma_k(q + p\|g_k\|)) \\ \le & \alpha_k(c_1 + \gamma_kp + \gamma_kq/\epsilon_0)\|g_k\| \end{aligned} \tag{2.9}$$

for $k \in \bar{K}_i$.

Note that for every $k \in \bar{K}_i$, there exists an integer k' such that $k = l(k') - i - 1$, that is, $k+1 = l(k') - (i-1) - 1 \in K_{i-1}$.

Therefore by (1.6), for $\forall k \in \bar{K}_i$,we have

$$\begin{aligned} & f_{l(k)} - f_{k+1} \\ \geq\ & -\sigma\lambda_k g_k^T P_k - \frac{1}{2}\sigma\lambda_k^2 P_k^T B_k P_k \\ \geq\ & -\sigma\lambda_k g_k^T P_k - \frac{1}{2}\sigma\lambda_k s_k P_k^T B_k P_k \\ =\ & -\sigma\lambda_k g_k^T P_k - \frac{1}{2}\sigma\lambda_k \frac{-\delta g_k^T P_k}{P_k^T B_k P_k} P_k^T B_k P_k \\ =\ & -\frac{\sigma(2-\delta)}{2}\lambda_k g_k^T P_k, \end{aligned} \tag{2.10}$$

which, together with (2.8) and $\sigma > 0, \delta < 2$, implies that

$$\begin{aligned} & f_{l(k)} - f_{k+1} \\ \geq\ & -\frac{\sigma(2-\delta)}{2}\lambda_k \langle g_k, P_k\rangle \\ \geq\ & \frac{\sigma(2-\delta)}{2}\mu_0\mu_1(c_2 - \gamma_k p - \gamma_k q/\epsilon_0)\alpha_k\lambda_k\|g_k\|^2. \end{aligned} \tag{2.11}$$

Taking limits on the both sides of the above inequality as $k \in \bar{K}_i, k \to \infty$, according to

$$\lim_{k\in K_{i-1}, k\to\infty} f_k = f^* = \lim_{k\to\infty} f_{l(k)},$$

we obtain

$$\lim_{k\in \bar{K}_i, k\to\infty} \alpha_k\lambda_k\|g_k\| = 0. \tag{2.12}$$

By (2.7), we have

$$\lim_{k\in \bar{K}_i, k\to\infty} \alpha_k\lambda_k = 0. \tag{2.13}$$

It follows from (2.13), Armijo rule and NNLS that if $\psi_k = \frac{\alpha_k}{\gamma}$ and $\psi'_k = \frac{\lambda_k}{\gamma}$, at least one of the following three inequalities holds for $k \in \bar{K}_i$ sufficiently large.

$$\langle g(x_k + \psi_k d_k), d_k\rangle > \mu_0\langle g_k, d_k\rangle.$$

$$\langle g(x_k + \psi_k d_k), g_k\rangle < \mu_1\|g(x_k + \psi_k d_k)\|^2.$$

$$\begin{aligned} & f(x_k + \psi'_k P_k) \\ >\ & \max_{0\le j\le m(k)} f_{k-j} + \sigma\psi'_k[g_k^T P_k + \frac{1}{2}\psi'_k P_k^T B_k P_k]. \end{aligned}$$

But whichever holds, similar to the proof of Lemma 2.4 in [14], by using $(2.7), (2.12)$ and (2.13), we can obtain the corresponding $\mu_i \geq 1 (i = 0, 1), \sigma > \frac{1}{2}$, which is a contradiction. Therefore we have

$$\lim_{k\in K_i, k\to\infty} g_k = 0.$$

(2) Similar to the proof of (2) in Lemma 2.4[14], we can obtain the conclusion easily.

Theorem 2.1. Suppose that $\{x_k\}$ is an infinite iteration sequence generated by Algorithm 2.1. If there exists an open convex set D containing $\{x_k\}$ such that $g(x)$ is uniformly continuous on D, then either

$$\lim_{k\to\infty} g_k = -\infty$$

or $\{f_k\}$ converges to a finite value and

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

proof. According to (H_5), we note that

$$\lim_{k\to\infty} f_k \neq -\infty.$$

Hence $\{f_{l(k)}\}$ converges to a finite value by Lemma 2.4. As a result, we have

$$\lim_{k\in K_{-1}, k\to\infty} f(x_k) = f' > -\infty.$$

By Lemma 2.5, for $i = 0$, we have

$$\lim_{k\in K_0, k\to\infty} g_k = 0,$$

$$\lim_{k\in K_0, k\to\infty} f_k = f'.$$

Similarly, we can use Lemma 2.5 repeatedly and get

$$\lim_{k\in K_i, k\to\infty} g_k = 0,$$

$$\lim_{k\in K_i, k\to\infty} f_k = f'$$

for $1 \leq i \leq M-1$.

Because of $\bigcup_{i=0}^{M-1} K_i = N$, we obtain

$$\lim_{k\to\infty} g_k = 0,$$

$$\lim_{k\to\infty} f_k = f' > -\infty,$$

which completes the proof.

By Theorem 2.1, the following corollary is obvious.

Corollary 2.1. Suppose that the assumption conditions of Theorem 2.1 hold. If the infinite iteration sequence $\{x_k\}$ generated by Algorithm 2.1 has cluster point x^*, then $x^* \in \Omega^*$.

III. Numerical Results

In this section, we report the numerical results obtained for a set of standard test problems, by means of the Algorithm 3.1[14] (denote by HNGP) and Algorithm 2.1 in this paper(denote by NHNGP). We utilize the negative gradient direction instead of the main direction s_k, that is, $s_k = -\nabla f(x_k)$. The perturbation term ω_k is obtained randomly in the case where it satisfies (H_3) and (H_4). And only one of the results is given in table 1-4.

In particular we report, for each problem, the number IT of iteration, the amount T of time and the value $f(\hat{x})$ of the objective function at the solution found $\hat{x}$.

Typical values for the parameters are: $\mu_0 = \mu_1 = 0.1, \sigma = 0.38, M = 3, \beta = 0.618, \delta = 1, p = 1, q = 0.1$.

The algorithm have been tested on the following set of problems.

Problem 1.

$$f(x) = 10(x_1^2 - x_2)^2 + (1 - x_1)^2 + 9(x_4 - x_3^2)^2 + (1 - x_3)^2 + 10.1((x_2 - 1)^2 + (x_4 - 1)^2) + 19.8(x_2 - 1)(x_4 - 1).$$

Problem 2.

$$f(x) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4.$$

Problem 3.

$$f(x) = \sum_{i=1}^{N/2}[(x_{2i} - x_{2i-1}^2)^2 + (1 - x_{2i-1})^2].$$

Problem 4.

$$f(x) = \sum_{i=1}^{N/4}[(x_{4i-1} + 10x_{4i-2})^2 + 5(x_{4i-1} - x_{4i})^2 + (x_{4i-2} - 2x_{4i-1})^2 + 10(x_{4i-3} - x_{4i})^4].$$

The numerical results have been showed in Table I-IV. For every method, we obtain the corresponding results for $\|\nabla f(\hat{x})\| \leq 0.1, 0.01, 0.001$. And $4.2(-5)$ means 4.2×10^{-5}.

Table I.
Numerical Results of Problem 1

Me.	$HNGP$	$NHNGP$
IT	405,606,1180	59,80,101
T	0.235s,0.401s,0.980s	0.078s,0.081s,0.125s
$f(\hat{x})$	2.5(-3),2.4(-5),2.3(-7)	4.2(-5),3.9(-8),1.3(-9)

Table II.
Numerical Results of Problem 2

Me.	$HNGP$	$NHNGP$
IT	484,1887,8441	51,57,108
T	0.500s,1.938s,9.282s	0.078s,0.082s,0.125s
$f(\hat{x})$	9.7(-3),4.2(-4),1.9(-5)	5.1(-4),1.6(-6),1.3(-7)

Table III.
Numerical Results of Problem 3
$N = 120$

Me.	$HNGP$	$NHNGP$
IT	401,650,386	26,40,71
T	3.273s,9.034s,4.578s	0.157s,0.266s,0.422s
$f(\hat{x})$	5.6(-3),5.3(-5),6.9(-7)	5.1(-4),6.9(-6),3.2(-7)

Table IV.
Numerical Results of Problem 4
$N = 60$

Me.	$HNGP$	$NHNGP$
IT	704,1519,7210	59,75,101
T	5.531s,12.703s,46.235s	0.172s,0.235s,0.344s
$f(\hat{x})$	3.2(-3),1.6(-4),6.6(-6)	6.9(-5),9.5(-6),1.4(-8)

Because the perturbation term is produced randomly, Table V-VI give the different three results of HNGP and NHNGP to problem 1 and problem 3 when $\|\nabla f(\hat{x})\| \leq 0.01$.

Table V.
Different Results of Problem 1
$\epsilon = 0.01$

Me.	$HNGP$	$NHNGP$
IT	606,805,890	80,70,82
T	0.401s,0.866s,1.021s	0.081s,0.093s,0.109s
$f(\hat{x})$	2.3(-5),2.4(-5),2.9(-5)	3.9(-8),3.7(-7),3.8(-7)

Table VI.
Different Results of Problem 3
$N = 120, \epsilon = 0.01$

Me.	$HNGP$	$NHNGP$
IT	650,336,28	40,52,55
T	9.034s,3.64s,0.312s	0.266s,0.296s,0.312s
$f(\hat{x})$	5.3(-5),6.8(-5),1.4(-4)	6.9(-6),1.4(-6),9.3(-6)

IV. Conclusions

This paper discusses a new kind of nonmontone line search method which is called new hybrid projection method with perturbations. The global convergence of this kind of method is proved only in the case where the gradient function is uniformly continuous on an open convex set containing the iteration sequence. Furthermore, numerical examples are given in the last section. From Table 1-6, we can see that the new hybrid projection method can improve the convergence rate of the problems if we choose proper perturbation term and it is superior than Algorithm 3.1 proposed in [14]. Certainly, more adequate test would be probably required.

References

[1] D.P. Bertsekas, ”Nonlinear Programing”, *Athena Scientific, Belmont, MA*, 1995.

[2] Y. J. Wang, and N. H. Xiu, ”Nonlinear Programming Theory and Algorithms”, *Shanxi Science and Technology Press*, 2004.

[3] Y. X. Yuan, and W. Y. Sun , ”Optimization Theory and Algorithms”, *Science Press*, 1997.

[4] W. Y. Cheng, and Q. F. Liu, ”Sufficient descent nonlinear conjugate gradient methods with conjugacy condition”, *Numer. Algor.*, no. 53, 2010, pp. 113-131.

[5] N. Andrei, ”Accelerated scaled memoryless BFGS preconditioned conjugate gradient algorithm for unconstrained optimization”, *Eur. J. Oper. Res.*, vol. 204, no. 3, 2010, pp. 410-420.

[6] G.H. Yu, J.H. Huang, and Y. Zhou, ”A descent spectral conjugate gradient method for impulse noise removal”, *Appl. Math. Lett.*, vol. 23, 2010, pp. 555-560.

[7] K. Yin, Y.H. Xiao, and M.L. Zhang, ”Nonlinear conjugate gradient method for l_1-norm regularization problems in compressive sensing:, *J. Comput. Infor. Sys.*, vol. 7, 2011, pp. 880-885.

[8] L. Grippo, F. Lampariello, and Lucidi S. , ”A nonmonotone line search technique for Newton’s method”, *SIAM J. Numer. Anal.*, vol.4, 1986, pp. 707-716.

[9] D.P. Bertsekas, and J.N. Tsitsiklis , ”Neuro-Dynamic programming”, *Athena Scientific, Belmont, MA*, 1996.

[10] A.A. Gaivoronski, ”Convergence properties of backpropagation for neural nets via theory of stochastic gradient methods”. Part 1, *Optim. Methods Software*, vol.4, 1994, pp. 117-134.

[11] Z. S. Yu, and D. G. Pu, ”A new nonmonotone line search technique for unconstrained optimization”, *Journal of Computational and Applied Mathematics*, vol.219, 2008, pp. 134-144.

[12] G. L. Yuan, and Z. X. Wei , ”New line search methods for unconstrained optimization”, *Journal of the Korean Statistical Society*, vol.38, 2009, pp. 29-39.

[13] Z. J. Shi, and J. Shen, ”Convergence of nonmontone line search method”, *J. Comput. Appl. Math.*, vol.193, 2006, pp. 397-412.

[14] M. X. Li, and C. Y. Wang , ”Convergence property of gradient-type methods with non-monotone line search in the presence of perturbation”, *Appl. Math. Optim.*, vol. 174, 2006, pp. 854-868.

ACKNOWLEDGMENT

We wish to thank the anonymous referees and the editor of the journal for their careful reading of the paper and for their constructive comments that contributed greatly to improving the paper.

Meixia Li was born in Shandong, China. She received her PhD degree in operations research from Dalian University of Technology in 2006, her MS degree in operations research from Qufu Normal University in 1999, and her BS degree in mathematics education from Qufu Normal University in 1994.

She is currently an Associate Professor of mathematics at Shandong, Weifang University, China. Her current research interests include nonlinear programming, semi-infinite programming and matrix programming.