

RESEARCH

Open Access

# New cautious BFGS algorithm based on modified Armijo-type line search

Zhong Wan\*, Shuai Huang and Xiao Dong Zheng

\*Correspondence:  
wanmath@163.com  
School of Mathematics and  
Statistics, Central South University,  
Changsha, Hunan, China

## Abstract

In this paper, a new inexact line search rule is presented, which is a modified version of the classical Armijo line search rule. With lower cost of computation, a larger descent magnitude of objective function is obtained at every iteration. In addition, the initial step size in the modified line search is adjusted automatically for each iteration. On the basis of this line search, a new cautious Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm is developed. Under some mild assumptions, the global convergence of the algorithm is established for nonconvex optimization problems. Numerical results demonstrate that the proposed method is promising, especially in comparison with the existent methods.

**Keywords:** unconstrained optimization; inexact line search rule; global convergence; BFGS

## 1 Introduction

Consider the following unconstrained optimization problem:

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

where  $f : R^n \rightarrow R$  is a twice continuously differentiable function.

Amongst the variant methods to solve problem (1), it is well known that the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method has obtained great success either in the aspect of the theoretical research or in the field of engineering applications. In this connection, it is referred to, for example, the literature [1–11] and the references therein.

Summarily, in the framework of the BFGS method, a quasi-Newton direction  $d_k$  at the current iterate point  $x_k$  is first obtained by solving the following linear system of equation:

$$B_k d_k = -g_k, \tag{2}$$

where  $B_k$  is a given positive definite matrix,  $g : R^n \rightarrow R^n$  is the gradient function of  $f$ , and  $g_k$  is the value of  $g$  at  $x_k$ . At the new iterate point  $x_{k+1}$ ,  $B_k$  is updated by

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k}, \tag{3}$$

where  $s_k = x_{k+1} - x_k$ ,  $y_k = g_{k+1} - g_k$ .

Next, along the search direction  $d_k$ , we choose a suitable stepsize  $\alpha_k$  by employing some line search strategy. Thus, the iterate point  $x_k$  is updated by

$$x_{k+1} = x_k + \alpha_k d_k. \tag{4}$$

Actually, it has been reported that a choice of suitable line search rule is important to the efficiency and convergence of the BFGS method (see, for example, [7, 12–15] and [16]). It is well known that the Armijo line search is the cheapest and most popular algorithm to obtain a step length among all line search methods. However, when the Armijo line search algorithm is implemented to find a step length  $\alpha_k$ ,  $B_{k+1}$  in (3) may not be positive definite even if  $B_k$  is a positive definite matrix [17]. For this, in [5] a cautious BFGS method (CBFGS) associated with the Armijo line search is presented to solve the nonconvex unconstrained optimization problems. The update formula of  $B_k$  in [5] is

$$B_{k+1} = \begin{cases} B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k}, & \text{if } \frac{y_k^T s_k}{\|s_k\|^2} \geq \epsilon \|g_k\|^\gamma, \\ B_k, & \text{otherwise,} \end{cases} \tag{5}$$

where  $\epsilon$  and  $\gamma$  are positive constants.

In this paper, we shall first present a modified Armijo-type line search rule. Then, on the basis of this line search, a new cautious BFGS algorithm is developed. It will be shown that in our line search, a larger descent magnitude of an objective function is obtained with lower cost of computation at every iteration. In addition, the initial step size is adjusted automatically at each iteration.

The rest of this paper is organized as follows. In Section 2, a modified Armijo-type inexact line search rule is presented and a new cautious BFGS algorithm is developed. Section 3 is devoted to establishing the global convergence of the proposed algorithm under some suitable assumptions. In Section 4, numerical results are reported to demonstrate the efficiency of the algorithm. Some conclusions are given in the last section.

## 2 Modified Armijo-type line search rule and new cautious BFGS algorithm

The classical Armijo line search is to find  $\alpha_k$  such that the following inequality holds:

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \sigma_1 \alpha_k g_k^T d_k, \tag{6}$$

where  $\sigma_1 \in (0, 1)$  is a given constant scalar. In a computer procedure,  $\alpha_k$  in (6) is obtained by searching in the set  $\{\beta, \beta\rho, \beta\rho^2, \dots\}$  such that  $\alpha_k$  is the largest component satisfying (6), where  $\rho \in (0, 1)$  and  $\beta > 0$  are given constant scalars.

Compared with other line search methods, the computer procedure of the Armijo line search is simplest, and the computational cost to find a feasible stepsize is very low, especially for  $1 > \rho > 0$  being close to 0. Its drawback lies in that at each iteration, there may be only little reduction of an objective function to be obtained.

Inspired by this observation, we present a modified Armijo-type line search (MALS) rule as follows. Suppose that  $g$  is a Lipschitz continuous function. Let  $L$  be the Lipschitz constant. Let  $L_k$  be an approximation of  $L$ . Set

$$\beta_k = -\frac{g_k^T d_k}{(L_k \|d_k\|^2)}.$$

Different from the classical Armijo line search (6), we find a step size  $\alpha_k$  as the largest component in the set  $\{\beta_k, \beta_k\rho, \beta_k\rho^2 \dots\}$  such that the inequality

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \sigma \alpha_k \left( g_k^T d_k - \frac{1}{2} \alpha_k \mu L_k \|d_k\|^2 \right) \tag{7}$$

holds, where  $\sigma \in (0, 1)$ ,  $\mu \in [0, +\infty)$ ,  $\rho \in (0, 1)$  are given constant scalars.

In the following proposition, we show that the new line search (7) is well defined.

**Proposition 1** *Let  $f : R^n \rightarrow R$  be a continuously differentiable function. Suppose that the gradient function  $g$  of  $f$  is Lipschitz continuous. Let  $L_k > 0$  be an approximation value of the Lipschitz constant. If  $d_k$  is a descent direction of  $f$  at  $x_k$ , then there is an  $\alpha > 0$  in the set  $\{\beta_k, \beta_k\rho, \beta_k\rho^2 \dots\}$  such that the following inequality holds:*

$$f(x + \alpha d_k) \leq f(x_k) + \sigma \alpha \left( g_k^T d_k - \frac{1}{2} \alpha \mu L_k \|d_k\|^2 \right) \tag{8}$$

where  $\sigma \in (0, 1)$ ,  $\mu \in [0, +\infty)$ ,  $\rho \in (0, 1)$  are given constant scalars.

*Proof* In fact, we only need to prove that a step length  $\alpha$  is obtained in finitely many steps. If it is not true, then for all sufficiently large positive integer  $m$ , we have

$$f(x + \beta_k \rho^m d_k) - f(x_k) > \sigma \alpha \left( g_k^T d_k - \frac{1}{2} \beta_k \rho^m \mu L_k \|d_k\|^2 \right). \tag{9}$$

By the mean-theorem, there is a  $\theta_k \in (0, 1)$  such that

$$\beta_k \rho^m g(x_k + \theta_k \beta_k \rho^m d_k)^T d_k > \sigma \beta_k \rho^m \left( g_k^T d_k - \frac{1}{2} \beta_k \rho^m \mu L_k \|d_k\|^2 \right). \tag{10}$$

Thus,

$$(g(x_k + \theta_k \beta_k \rho^m d_k) - g_k)^T d_k > (\sigma - 1) g_k^T d_k - \frac{1}{2} \sigma \beta_k \rho^m \mu L_k \|d_k\|^2. \tag{11}$$

As  $m \rightarrow \infty$ , it is obtained that

$$(\sigma - 1) g_k^T d_k \leq 0.$$

From  $\sigma \in (0, 1)$ , it follows that  $g_k^T d_k \geq 0$ . This contradicts the fact that  $d_k$  is a descent direction. □

**Remark 1** Since the third term on the right-hand side of (7) is negative, it is easy to see that the obtained step size  $\alpha$  ensures a larger descent magnitude of the objective function than that in (6).

It is noted that (7) reduces to (6) when  $\mu = 0$ .

**Remark 2** In the MALS, the parameter  $L_k$  should be estimated at each iteration. In this paper, for  $k > 1$ , we choose

$$L_k = \frac{s_{k-1}^T y_{k-1}}{\|s_{k-1}\|^2}, \tag{12}$$

where  $s_{k-1} = x_k - x_{k-1}$ ,  $y_{k-1} = g_k - g_{k-1}$ . Actually,  $L_k$  in (12) is a solution of the minimization problem

$$\min \|L_k s_{k-1} - y_{k-1}\|. \tag{13}$$

Therefore, it is acceptable that  $L_k$  is an approximation of  $L$ .

Based on Proposition 1, Remarks 1 and 2, a new cautious BFGS algorithm is developed for solving problem (1).

**Algorithm 1** (New cautious BFGS algorithm)

Step 1. Choose an initial point  $x_0 \in R^n$  and a positive definite matrix  $B_0$ . Choose  $\sigma \in (0, 1)$ ,  $\mu \geq 0$ ,  $\epsilon > 0$  and  $L_0 > 0$ . Set  $k := 0$ .

Step 2. If  $\|g_k\| \leq \epsilon$ , the algorithm stops. Otherwise, go to Step 3.

Step 3. Find  $d_k$  that is a solution of the following system of linear equations:

$$B_k d_k = -g_k.$$

Step 4. Determine a step size  $\alpha_k$  satisfying (7).

Step 5. Set  $x_{k+1} := x_k + \alpha_k d_k$ . Compute  $s_k$  and  $y_k$ . Update  $B_k$  as  $B_{k+1}$  by (5). Set  $k := k + 1$ , return to Step 2.

**3 Global convergence**

In this section, we are going to prove the global convergence of Algorithm 1.

We need the following conditions.

**Assumption 1**

1. The level set  $\Omega = \{x \in R^n | f(x) \leq f(x_0)\}$  is bounded.
2. In some neighborhood  $N$  of  $\Omega$ ,  $f$  is continuously differentiable and its gradient is Lipschitz continuous, namely there exists a constant  $L > 0$  such that

$$\|g(x) - g(y)\| \leq L \|x - y\|, \quad \forall x, y \in N. \tag{14}$$

3. The sequence  $\{L_k\}$ ,  $0 < L_k \leq ML$ , where  $M$  is a positive constant.

Before the statement of the global convergence, we first prove the following useful lemmas.

**Lemma 1** Let  $\{x_k\}$  be a sequence generated by Algorithm 1. If  $L_k > 0$  for each  $k \geq 0$ , then for any given initial point  $x_0$ , the following results hold:

1.  $\{f_k\}$  is a decreasing sequence.
2.  $\{x_k\} \in \Omega$ .
3.  $\sum_{k=1}^{\infty} (f_k - f_{k+1}) < +\infty$ .

*Proof* The first and second results are directly from the condition  $L_k > 0$  for each  $k \geq 0$ , Proposition 1 and the definition of  $\Omega$ . We only need to prove the third result.

Since  $\{f_k\}$  is a decreasing sequence and is bounded below, it is clear that there exists a constant  $f^*$  such that

$$\lim_{k \rightarrow \infty} f_k = f^*. \tag{15}$$

From (15), we have

$$\sum_{k=1}^{\infty} (f_k - f_{k+1}) = \lim_{N \rightarrow \infty} \sum_{k=1}^N (f_k - f_{k+1}) = \lim_{N \rightarrow \infty} (f_1 - f_{N+1}) = f_1 - f^*. \tag{16}$$

Thus,

$$\sum_{k=1}^{\infty} (f_k - f_{k+1}) < +\infty. \tag{17}$$

□

**Lemma 2** *Let  $\{x_k\}$  be a sequence generated by Algorithm 1. Let  $\{d_k\}$  be the sequence of search direction. If Assumption 1 holds, then*

$$\sum_{k=1}^{\infty} \left( \frac{g_k^T d_k}{\|d_k\|} \right)^2 < +\infty. \tag{18}$$

*In particular,*

$$\lim_{k \rightarrow \infty} \frac{g_k^T d_k}{\|d_k\|} = 0. \tag{19}$$

*Proof* Denote

$$K_1 = \{k \mid \alpha_k = \beta_k\}, \quad K_2 = \{k \mid \alpha_k < \beta_k\}.$$

For  $k \in K_1$ , we have

$$\begin{aligned} f_k - f_{k+1} &\geq -\sigma \alpha_k \left( g_k^T d_k - \frac{1}{2} \alpha_k \mu L_k \|d_k\|^2 \right) \\ &= -\sigma \frac{-g_k^T d_k}{L_k \|d_k\|^2} \left( g_k^T d_k - \frac{1}{2} \frac{-g_k^T d_k}{L_k \|d_k\|^2} \mu L_k \|d_k\|^2 \right) \\ &= \sigma \left( 1 + \frac{1}{2} \mu \right) \frac{(g_k^T d_k)^2}{L_k \|d_k\|^2} \\ &\geq \frac{\sigma(2 + \mu)}{2ML} \left( \frac{g_k^T d_k}{\|d_k\|} \right)^2. \end{aligned} \tag{20}$$

For  $k \in K_2$ , it follows from (7) that

$$f(x_k + \rho^{-1} \alpha_k d_k) - f(x_k) > \sigma \rho^{-1} \alpha_k \left( g_k^T d_k - \frac{1}{2} \rho^{-1} \alpha_k \mu L_k \|d_k\|^2 \right). \tag{21}$$

By the mean-theorem, there is a  $\theta_k \in (0,1)$  such that

$$\rho^{-1}\alpha_k g(x_k + \theta_k \rho^{-1}\alpha_k d_k)^T d_k > \sigma \rho^{-1}\alpha_k \left( g_k^T d_k - \frac{1}{2}\rho^{-1}\alpha_k \mu L_k \|d_k\|^2 \right). \tag{22}$$

Hence,

$$(g(x_k + \theta_k \rho^{-1}\alpha_k d_k) - g_k)^T d_k > (\sigma - 1)g_k^T d_k - \frac{1}{2}\sigma \rho^{-1}\alpha_k \mu L_k \|d_k\|^2. \tag{23}$$

From the Lipschitz continuity of  $g$ , it is obtained that

$$\left( L + \frac{1}{2}\sigma \mu L_k \right) \rho^{-1}\alpha_k \|d_k\|^2 > (\sigma - 1)g_k^T d_k. \tag{24}$$

It reads

$$\alpha_k > -\frac{2\rho(1-\sigma)}{2L + \sigma \mu L_k} \frac{g_k^T d_k}{\|d_k\|^2}. \tag{25}$$

From (7) and (25), it is deduced that

$$\begin{aligned} f_k - f_{k+1} &\geq -\sigma \alpha_k \left( g_k^T d_k - \frac{1}{2}\alpha_k \mu L_k \|d_k\|^2 \right) \\ &\geq -\sigma \frac{-2\rho(1-\sigma)}{2L + \sigma \mu L_k} \frac{g_k^T d_k}{\|d_k\|^2} g_k^T d_k \\ &\geq \frac{2\rho\sigma(1-\sigma)}{2L + \sigma \mu ML} \left( \frac{g_k^T d_k}{\|d_k\|} \right)^2. \end{aligned} \tag{26}$$

Denote

$$\eta = \min \left\{ \frac{\sigma(2 + \mu)}{2ML}, \frac{2\rho\sigma(1-\sigma)}{2L + \sigma \mu ML} \right\}. \tag{27}$$

Then from (26), we obtain

$$f_k - f_{k+1} \geq \eta \left( \frac{g_k^T d_k}{\|d_k\|} \right)^2. \tag{28}$$

From Lemma 1, it is clear that

$$\sum_{k=1}^{\infty} \eta \left( \frac{g_k^T d_k}{\|d_k\|} \right)^2 < +\infty. \tag{29}$$

That is to say

$$\sum_{k=1}^{\infty} \left( \frac{g_k^T d_k}{\|d_k\|} \right)^2 < +\infty, \tag{30}$$

since  $\eta > 0$ . It is certain that

$$\lim_{k \rightarrow \infty} \frac{g_k^T d_k}{\|d_k\|} = 0. \tag{31}$$

□

**Lemma 3** Let  $\{x_k\}$  be a sequence generated by Algorithm 1. Suppose that there exist constants  $a_1, a_2 > 0$  such that the following relations hold for infinitely many  $k$ :

$$\|B_k s_k\| \leq a_1 \|s_k\|, \quad a_2 \|s_k\|^2 \leq s_k^T B_k s_k. \tag{32}$$

Then

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0. \tag{33}$$

*Proof* Let  $\Lambda$  be the indices set of  $k$  satisfying (32).

From (32) and  $g_k = -B_k d_k$ , it follows that for each  $k \in \Lambda$ ,

$$a_2 \|d_k\|^2 \leq d_k^T B_k d_k = -g_k^T d_k. \tag{34}$$

Thus,

$$a_2 \|d_k\| \leq -\frac{g_k^T d_k}{\|d_k\|}. \tag{35}$$

Combined with (31), it yields

$$\lim_{k \in \Lambda, k \rightarrow \infty} \|d_k\| = 0. \tag{36}$$

On the other hand, from (32) and  $g_k = -B_k d_k$ , it is deduced that for each  $k \in \Lambda$ ,

$$0 \leq \|g_k\| = \|B_k d_k\| \leq a_1 \|d_k\|. \tag{37}$$

From (36) and (37), it is easy to see that

$$\lim_{k \in \Lambda, k \rightarrow \infty} \|g_k\| = 0. \tag{38}$$

The desired result (33) is proved. □

Lemma 3 indicates that for the establishment of the global convergence, it suffices to show that (32) holds for infinitely many  $k$  in Algorithm 1. The following lemma gives sufficient conditions for (32) to hold (see Theorem 2.1 in [18]).

**Lemma 4** Let  $B_0$  be a symmetric and positive matrix and  $B_k$  be updated by (3). Suppose that there are positive constants  $m_1, m_2$  ( $m_1 < m_2$ ) such that for all  $k \geq 0$ ,

$$\frac{y_k^T s_k}{\|s_k\|^2} \geq m_1, \quad \frac{\|y_k\|^2}{y_k^T s_k} \leq m_2. \tag{39}$$

Then there exist constants  $a_1, a_2$  such that for any positive integer  $t$ , (32) holds for at least  $\lceil t/2 \rceil$  values of  $k \in \{1, 2, \dots, t\}$ .

Now, we come to establish the global convergence for Algorithm 1. For the sake of convenience, we define an index set

$$\tilde{K} = \left\{ i \mid \frac{y_i^T s_i}{\|s_i\|^2} \geq \epsilon \|g_i\|^\gamma \right\}. \tag{40}$$

**Theorem 1** *Let  $\{x_k\}$  be a sequence generated by Algorithm 1. Under Assumption 1, (33) holds.*

*Proof* From Lemma 3, we only need to show that (32) holds for infinitely many  $k$ .

If  $\tilde{K}$  is a finite set, then  $B_k$  remains a constant matrix after a finite number of iterations. Hence, there are constants  $a_1, a_2$  such that (32) holds for all  $k$  sufficiently large. The proof of the result is completed.

In the following, we prove (33) in the case that  $\tilde{K}$  is a infinite set.

Suppose that (33) is not true. Then there is a constant  $\delta > 0$  such that  $\|g_k\| \geq \delta$  for all  $k$ . From (40), the inequality

$$\frac{y_k^T s_k}{\|s_k\|^2} \geq \epsilon \delta^\gamma \tag{41}$$

holds for all  $k \in \tilde{K}$ .

Combined with (14), it is obtained that

$$\frac{\|y_k\|^2}{y_k^T s_k} \leq \frac{\|y_k\|^2}{\epsilon \delta^\gamma \|s_k\|^2} \leq \frac{L^2}{\epsilon \delta^\gamma}. \tag{42}$$

From Lemma 4, it follows that there exist constants  $a_1, a_2$  such that (32) holds for infinitely many  $k$ . It contradicts the result in Lemma 3.

The proof is completed. □

#### 4 Numerical experiments

In this section, we report the numerical performance of Algorithm 1. The numerical experiments are carried out on a set of 16 test problems from [19]. We make comparisons with the cautious BFGS method associated with the ordinary Armijo line search rule.

In order to study the numerical performance of Algorithm 1, we record the run time of CPU, the total number of function evaluations required in the process of line search and the total number of iterations for each algorithm.

All MATLAB procedures run in the following computer environment: 2GHz CPU, 1GB memory based operating system of WINDOWS XP. The parameters are chosen as follows:

$$\epsilon = 10^{-6}, \quad B_0 = I_{n \times n}, \quad \rho = 0.3, \quad \sigma = 0.2, \quad \mu = 1, \quad L = 1.$$

As to the parameters in the cautious update (5), we first let  $\gamma = 0.01$  if  $\|g_k\| \geq 1$ , and  $\gamma = 3$  if  $\|g_k\| < 1$ .

The performance of algorithms and the solution results are reported in Table 1. In this table, we use the following denotations:

*Dim*: the dimension of the objective function;

**Table 1 Comparison of efficiency with other method**

Functions	Algorithm	Dim	GV	NI	NF	CT
Rosenbrock	CBFGS	2	6.2782e-007	35	74	0.0310s
	NCBFGS	2	1.1028e-007	40	70	0.0310s
Freudenstein and Roth	CBFGS	2	7.9817e-007	28	82	0.0310s
	NCBFGS	2	2.7179e-007	11	25	0.0320s
Beale	CBFGS	2	7.2275e-007	40	55	0.0310s
	NCBFGS	2	3.1136e-007	18	23	0.0470s
Brown badly	CBFGS	2	7.7272e-007	36	223	0.0310s
	NCBFGS	2	0	29	50	0.0620s
Broyden tridiagonal	CBFGS	4	7.5723e-007	26	126	0.0320s
	NCBFGS	4	3.8712e-007	15	21	0.0310s
Powell singular	CBFGS	4	9.9993e-007	13,993	14,031	2.4530s
	NCBFGS	4	9.4607e-007	31	38	0.0320s
Kowalik and Osborne	CBFGS	4	9.9783e-007	3126	3128	2.1250s
	NCBFGS	4	4.4454e-007	30	45	0.0470s
Brown almost-linear	CBFGS	6	9.5864e-007	263	300	0.1100s
	NCBFGS	6	1.2290e-007	22	30	0.0160s
Discrete boundary	CBFGS	6	8.6773e-007	79	85	0.0470s
	NCBFGS	6	3.3650e-007	14	17	0.0320s
Variably dimensioned	CBFGS	8	3.4688e-008	7	51	0.0470s
	NCBFGS	8	3.1482e-007	10	21	0.0320s
Extended Rosenbrock	CBFGS	8	8.2943e-007	91	190	0.0470s
	NCBFGS	8	7.7959e-007	99	149	0.0320s
Extended Powell singular	CBFGS	8	9.9975e-007	6154	6199	1.4690s
	NCBFGS	8	6.5685e-007	42	55	0.0630s
Brown almost-linear	CBFGS	8	9.8392e-007	364	379	0.1880s
	NCBFGS	8	4.8080e-007	20	27	0.0780s
Broyden tridiagonal	CBFGS	9	4.4261e-007	38	86	0.0470s
	NCBFGS	9	6.2059e-007	41	56	0.0310s
Linear-rank1	CBFGS	10	-	-	-	-
	NCBFGS	10	2.6592e-007	4	15	0.0310s
Linear-full rank	CBFGS	12	9.5231e-007	18	36	0.0160s
	NCBFGS	12	9.4206e-016	2	3	0.0150s

GV: the gradient value of the objective function when the algorithm stops;

NI: the number of iterations;

NF: the number of function evaluations;

CT: the run time of CPU;

CBFGS: the CBFGS method associated with Armijo line search rule;

NCBFGS: the new BFGS method proposed in this paper.

In Table 1 it is shown that the developed algorithm in this paper is promising. In some cases, it requires less number of iterations, less number of function evaluation or less CPU time to find an optimal solution with the same tolerance than another algorithm.

## 5 Conclusions

A modified Armijo-type line search with an automatical adjustment of initial step size has been presented in this paper. Combined with the cautious BFGS method, a new BFGS algorithm has been developed. Under some assumptions, the global convergence was established for nonconvex optimization problems. Numerical results demonstrate that the proposed method is promising.

### Competing interests

The authors declare that they have no competing interests.

#### Authors' contributions

ZW carried out the studies of the new Armijo line search studies, participated in the sequence alignment, the proof of the main results and drafted the manuscript. SH participated in the proving some main results. XZ participated in the design of algorithm and performed the numerical experiments. All authors read and approved the final manuscript.

#### Acknowledgements

Supported by the National Natural Science Foundation of China (Grant No. 71210003, 71071162).

Received: 8 October 2011 Accepted: 20 September 2012 Published: 17 October 2012

#### References

1. Al-Baali, M: Quasi-Newton algorithms for large-scale nonlinear least-squares. In: Di Pillo, G, Murlı, A (eds.) High Performance Algorithms and Software for Nonlinear Optimization, pp. 1-21. Kluwer Academic, Dordrecht (2003)
2. Al-Baali, M, Grandinetti, L: On practical modifications of the quasi-Newton BFGS method. *Adv. Model. Optim.* **11**(1), 63-76 (2009)
3. Gill, PE, Leonard, MW: Limited-Memory reduced-Hessian methods for large-scale unconstrained optimization. *SIAM J. Optim.* **14**, 380-401 (2003)
4. Guo, Q, Liu, JG: Global convergence of a modified BFGS-type method for unconstrained nonconvex minimization. *J. Appl. Math. Comput.* **21**, 259-267 (2006)
5. Li, DH, Fukushima, M: On the global convergence of the BFGS method for nonconvex unconstrained optimization problems. *SIAM J. Optim.* **11**(4), 1054-1064 (2001)
6. Mascarenhas, WF: The BFGS method with exact line searches fails for non-convex objective functions. *Math. Program.* **99**, 49-61 (2004)
7. Nocedal, J: Theory of algorithms for unconstrained optimization. *Acta Numer.* **1**, 199-242 (1992)
8. Xiao, YH, Wei, ZX, Zhang, L: A modified BFGS method without line searches for nonconvex unconstrained optimization. *Adv. Theor. Appl. Math.* **1**(2), 149-162 (2006)
9. Zhou, W, Li, D: A globally convergent BFGS method for nonlinear monotone equations without any merit function. *Math. Comput.* **77**, 2231-2240 (2008)
10. Zhou, W, Zhang, L: Global convergence of the nonmonotone MBFGS method for nonconvex unconstrained minimization. *J. Comput. Appl. Math.* **223**, 40-47 (2009)
11. Zhou, W, Zhang, L: Global convergence of a regularized factorized quasi-Newton method for nonlinear least squares problems. *Comput. Appl. Math.* **29**(2), 195-214 (2010)
12. Cohen, AI: Stepsize analysis for descent methods. *J. Optim. Theory Appl.* **33**, 187-205 (1981)
13. Dennis, JE, Schnable, RB: *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice-Hall, Englewood Cliffs (1983)
14. Shi, ZJ, Shen, J: New inexact line search method for unconstrained optimization. *J. Optim. Theory Appl.* **127**(2), 425-446 (2005)
15. Sun, WY, Han, JY, Sun, J: Global convergence of nonmonotone descent methods for unconstrained optimization problems. *J. Comput. Appl. Math.* **146**, 89-98 (2002)
16. Wolfe, P: Convergence conditions for ascent methods. *SIAM Rev.* **11**, 226-235 (1969)
17. Nocedal, J, Wright, JS: *Numerical Optimization*. Springer, New York (1999)
18. Byrd, R, Nocedal, J: A tool for the analysis of quasi-Newton methods with application to unconstrained minimization. *SIAM J. Numer. Anal.* **26**, 727-739 (1989)
19. More, JJ, Garbow, BS, Hillstrome, KE: Testing unconstrained optimization software. *ACM Trans. Math. Softw.* **7**, 17-41 (1981)

doi:10.1186/1029-242X-2012-241

Cite this article as: Wan et al.: New cautious BFGS algorithm based on modified Armijo-type line search. *Journal of Inequalities and Applications* 2012 2012:241.

Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)