Journal of Applied Mathematics and Physics, 2018, 6, 130-137
http://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

Global Convergence of an Extended Descent Algorithm without Line Search for Unconstrained Optimization

Cuiling Chen*, **Liling Luo**, **Caihong Han**, **Yu Chen**

College of Mathematics and Statistics, Guangxi Normal University, Guilin, China
Email: *mathchen@163.com

How to cite this paper: Chen, C.L., Luo, L.L., Han, C.H. and Chen, Y. (2018) Global Convergence of an Extended Descent Algorithm without Line Search for Unconstrained Optimization. *Journal of Applied Mathematics and Physics*, **6**, 130-137.
https://doi.org/10.4236/jamp.2018.61013

Received: December 16, 2017
Accepted: January 13, 2018
Published: January 16, 2018

Abstract

In this paper, we extend a descent algorithm without line search for solving unconstrained optimization problems. Under mild conditions, its global convergence is established. Further, we generalize the search direction to more general form, and also obtain the global convergence of corresponding algorithm. The numerical results illustrate that the new algorithm is effective.

Keywords

Unconstrained Optimization, Descent Method, Line Search, Global Convergence

1. Introduction

Consider an unconstrained optimization problem (UP)

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

where $f : \mathfrak{R}^n \to \mathfrak{R}$ is a continuously differentiable function. In general, the iterative algorithms for solving (UP) usually take the form:

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

where x_k, α_k and d_k are current iterative point, a positive step length and a search direction, respectively. For simplicity, we denote $\nabla f(x_k)$ by g_k and $f(x_k)$ by f_k.

The main task in the iterative formula (2) is to choose search direction d_k and determine step length α_k along the direction. There are many classic methods to choose search direction d_k, such as the steepest descent methods, Newton-type methods, Variable metric methods (see [1]), and conjugate gradient methods

$$d_k = \begin{cases} -g_k & \text{if } k = 1, \\ -g_k + \beta_k d_{k-1} & \text{if } k \geq 2, \end{cases} \tag{3}$$

where β_k is a parameter (see [2] [3] [4]). For step length α_k, it is usually determined by line search procedure, such as exact line search, Wolfe line search, Armijo line search, and so on. However, these line search procedures may involve extensive computation of objective functions and its gradients, which often becomes a significant burden for large-scale problems. Evidently, it is a good idea that line search procedure is avoided in algorithm design in order to reduce the evaluations of objective functions and gradients.

Based on the above consideration, some authors have started to study the algorithms without line search. Recently, some conjugate gradient algorithms without line search were investigated. In [5], Sun and Zhang studied some well-known conjugate gradient methods without line search, for instance, Fletcher-Reeves method, Hestenes-Stiefel method, Dai-Yuan method, Polak-Ribière method and Conjugate Descent method. In [6], Chen and Sun researched a two-parameter family of conjugate gradient methods without line search. In [7] [8], Wang and Zhu put forward to conjugate gradient path methods without line search. Shi, Shen and Zhou proposed descent methods without line search in [9] and [10], respectively. Further, Zhou presented the steepest descent algorithm without line search in [11].

Inspired by the above literatures, in this paper we will extend the descent algorithm without line search of [10] to more general case, and discuss its global convergence. The rest of this paper is organized as follows. In Section 2, we describe the extended descent algorithm without line search. In Section 3, we analyze its global convergence. Further, we generalize the search direction to more general form, and obtain global convergence of corresponding algorithm. Finally, numerical results are reported in Section 4.

2. Extended Descent Algorithm

To proceed, we first assume that [2]

(H_1) The function f has lower bound on $\pounds = \{x \in \Re^n \mid f(x) \leq f(x_1)\}$, where x_1 is available.

(H_2) The gradient g is Lipschitz continuous in an open convex set $\mathcal{B}$ that contains $\pounds$, *i.e.*, there exists $L > 0$ such that

$$\|g(x) - g(y)\| \leq L\|x - y\|, \quad \forall x, y \in \mathcal{B}. \tag{4}$$

Now we give the extended algorithm.

Algorithm 2.1. Given a starting point x_1, a positive constant ϵ, three parameters μ_1, μ_2 and ρ such that $0 < \mu_1 < \frac{1}{2} < \mu_2 < 1$, $\frac{1}{2} \leq \rho < 1$. Let $k := 1$.

Step 1. If $\|g_k\| < \epsilon$, then stop; otherwise go to Step 2.

Step 2. Compute

$$s_k = \begin{cases} \rho, & k = 1, \\ \dfrac{\rho\|g_k\|^2}{\rho\|g_k\|^2 + (1-\rho)\left|g_k^{\mathrm{T}} d_{k-1}\right|}, & k \geq 2. \end{cases} \tag{5}$$

Step 3. Set search direction

$$d_k=\begin{cases}-s_k g_k, & k=1,\\ -\left[\rho\left(1-\frac{\alpha_{k-1}s_k}{1+\alpha_{k-1}}\right)g_k+(1-\rho)\frac{\alpha_{k-1}s_k}{1+\alpha_{k-1}}d_{k-1}\right], & k\ge 2.\end{cases} \tag{6}$$

Step 4. Compute step length by the following rule. When $k=1$, α_k is determined by Wolfe line search, *i.e.*, it satisfies that

$$f(x_k+\alpha_k d_k)-f_k\le \mu_1\alpha_k g_k^{\mathrm{T}}d_k, \tag{7}$$

$$g(x_k+\alpha_k d_k)^{\mathrm{T}}d_k\ge \mu_2 g_k^{\mathrm{T}}d_k. \tag{8}$$

When $k\ge 2$,

$$\alpha_k=-\frac{g_k^{\mathrm{T}}d_k}{L_k\|d_k\|^2}, \tag{9}$$

where L_k satisfies that $\rho L\le L_k\le m_k L$ and $\{m_k,k=1,2,\cdots\}$ is a positive sequence which has a sufficient large upper bound.

Step 5. Set next iterative point

$$x_{k+1}=x_k+\alpha_k d_k. \tag{10}$$

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

Remark 2.1. Note that the formula of s_k and d_k in Algorithm 2.1 are the generalized forms of those in [10].

3. Global Convergence

Lemma 3.1. If Algorithm 2.1 generates an infinite sequence $\{x_k,k=1,2,\cdots\}$, then all search directions d_k are descent, and $\forall k\ge 2$, it holds that

$$-g_k^{\mathrm{T}}d_k\ge\frac{\rho\|g_k\|^2}{1+\alpha_{k-1}}. \tag{11}$$

Proof. If $k=1$, it is obvious that $-g_1^{\mathrm{T}}d_1=\rho\|g_1\|^2>0$. If $k\ge 2$, by (5) and (6), we have

$$\begin{aligned}-g_k^{\mathrm{T}}d_k&=\rho\left(1-\frac{\alpha_{k-1}s_k}{1+\alpha_{k-1}}\right)\|g_k\|^2+(1-\rho)\frac{\alpha_{k-1}s_k}{1+\alpha_{k-1}}g_k^{\mathrm{T}}d_{k-1}\\ &=\rho\|g_k\|^2-\frac{\alpha_{k-1}s_k}{1+\alpha_{k-1}}\left[\rho\|g_k\|^2-(1-\rho)g_k^{\mathrm{T}}d_{k-1}\right]\\ &\ge\rho\|g_k\|^2-\frac{\alpha_{k-1}s_k}{1+\alpha_{k-1}}\left[\rho\|g_k\|^2+(1-\rho)\left|g_k^{\mathrm{T}}d_{k-1}\right|\right]\\ &=\frac{\rho\|g_k\|^2}{1+\alpha_{k-1}}.\end{aligned} \tag{12}$$

This completes the proof. □

Lemma 3.2 (Mean value theorem, see [1]). Suppose that the objective function $f(x)$ is continuously differentiable on an open convex set $\mathcal{B}$, then

$$f(x_k+\alpha d_k)-f_k=\alpha\int_0^1 g(x_k+t\alpha d_k)^{\mathrm{T}}d_k\mathrm{d}t, \tag{13}$$

where $x_k, x_k + \alpha d_k \in \mathcal{B}$, $d_k \in \Re^n$. If $f(x)$ is twice continuously differentiable on $\mathcal{B}$, then

$$g(x_k + \alpha d_k) - g_k = \alpha \int_0^1 \nabla^2 f(x_k + t\alpha d_k) d_k \mathrm{d}t, \tag{14}$$

and

$$f(x_k + \alpha d_k) - f_k = \alpha g_k^{\mathrm{T}} d_k + \alpha^2 \int_0^1 (1-t) d_k^{\mathrm{T}} \nabla^2 f(x_k + t\alpha d_k) d_k \mathrm{d}t. \tag{15}$$

Lemma 3.3. $\forall k \ge 2$,

$$\|d_k\|^2 \le 3\rho^2 \cdot \sum_{1 \le i \le k} \|g_i\|^2. \tag{16}$$

Proof. Where $k \ge 2$, it holds that $(1-\rho) s_k \left| g_k^{\mathrm{T}} d_{k-1} \right| = \rho (1 - s_k) \|g_k\|^2$ by (5). Then $\forall k \ge 2$, we have

$$\begin{aligned}
\|d_k\|^2 &= \left\| \rho \left(1 - \frac{\alpha_{k-1} s_k}{1+\alpha_{k-1}} \right) g_k + (1-\rho) \frac{\alpha_{k-1} s_k}{1+\alpha_{k-1}} d_{k-1} \right\|^2 \\
&= \rho^2 \left(1 - \frac{\alpha_{k-1} s_k}{1+\alpha_{k-1}} \right)^2 \|g_k\|^2 + 2\rho \left(1 - \frac{\alpha_{k-1} s_k}{1+\alpha_{k-1}} \right) \\
&\quad \cdot (1-\rho) \frac{\alpha_{k-1} s_k}{1+\alpha_{k-1}} \cdot g_k^{\mathrm{T}} d_{k-1} + (1-\rho)^2 \left(\frac{\alpha_{k-1} s_k}{1+\alpha_{k-1}} \right)^2 \|d_{k-1}\|^2 \\
&\le \rho^2 \|g_k\|^2 + 2\rho(1-\rho) s_k \left| g_k^{\mathrm{T}} d_{k-1} \right| + \|d_{k-1}\|^2 \\
&= \rho^2 \|g_k\|^2 + 2\rho^2 (1 - s_k) \|g_k\|^2 + \|d_{k-1}\|^2 \le 3\rho^2 \|g_k\|^2 + \|d_{k-1}\|^2.
\end{aligned}$$

Using induction principle and noting that $\|d_1\|^2 = \rho^2 \|g_1\|^2$, it yields that

$$\|d_k\|^2 \le 3\rho^2 \|g_k\|^2 + 3\rho^2 \|g_{k-1}\|^2 + 3\rho^2 \|g_{k-2}\|^2 + \cdots + \rho^2 \|g_1\|^2.$$

Therefore (16) holds. The proof is completed. □

Theorem 3.1. If (H_1), (H_2) hold, and Algorithm 2.1 generates an infinite sequence $\{x_k, k = 1, 2, \cdots\}$, then

$$\sum_{k=2}^{+\infty} \frac{\|g_k\|^4}{(1+\alpha_{k-1})^2 \sum_{1 \le i \le k} \|g_i\|^2} < +\infty; \tag{17}$$

and

$$\sum_{k=2}^{+\infty} \frac{\alpha_k}{1+\alpha_{k-1}} \|g_k\|^2 < +\infty. \tag{18}$$

Proof. When $k \ge 2$, from (13), (4), Lemma 3.1, Lemma 3.3 and $\rho L \le L_k \le m_k L$, it yields that

$$\begin{aligned}
f_k - f_{k+1} &= -\alpha_k \int_0^1 g(x_k + t\alpha_k d_k)^{\mathrm{T}} d_k \mathrm{d}t \\
&= -\alpha_k g_k^{\mathrm{T}} d_k - \alpha_k \int_0^1 \left[g(x_k + t\alpha_k d_k) - g_k \right]^{\mathrm{T}} d_k \mathrm{d}t \\
&\ge -\alpha_k g_k^{\mathrm{T}} d_k - \alpha_k \int_0^1 \left\| g(x_k + t\alpha_k d_k) - g_k \right\| \cdot \|d_k\| \mathrm{d}t \\
&\ge -\alpha_k g_k^{\mathrm{T}} d_k - \alpha_k^2 L \int_0^1 t \|d_k\|^2 \mathrm{d}t = -\alpha_k g_k^{\mathrm{T}} d_k - \frac{1}{2} \alpha_k^2 L \|d_k\|^2
\end{aligned}$$

$$
\begin{aligned}
&= \left(\frac{1}{L_k} - \frac{1}{2L_k^2}\right)\frac{\left(g_k^{\mathrm{T}} d_k\right)^2}{\|d_k\|^2} \geq \frac{(2\rho - 1)\left(g_k^{\mathrm{T}} d_k\right)^2}{2Lm_k^2\|d_k\|^2} \\
&\geq \frac{(2\rho - 1)\cdot \rho^2 \|g_k\|^4}{2Lm_k^2(1+\alpha_{k-1})^2 \cdot 3\rho^2 \cdot \sum_{1\le i\le k}\|g_i\|^2} \\
&= \frac{(2\rho-1)\|g_k\|^4}{6Lm_k^2(1+\alpha_{k-1})^2 \sum_{1\le i\le k}\|g_i\|^2},
\end{aligned}
\tag{19}
$$

which implies that $\{f_k, k=1,2,\cdots\}$ is a decreasing sequence. And it is clear that the sequence $\{x_k, k=1,2,\cdots\}$ generated by Algorithm 2.1 is contained in $\mathcal{B}$ by (H_1), and there exists a constant f^* such that $\lim_{k\to\infty} f_k = f^*$. Therefore

$$
\sum_{k=2}^{+\infty}(f_k - f_{k+1}) = \lim_{N\to+\infty}\sum_{k=2}^{N}(f_k - f_{k+1}) = \lim_{N\to+\infty}(f_2 - f_{N+1}) = f_2 - f^*.
$$

Thus

$$
\sum_{k=2}^{+\infty}(f_k - f_{k+1}) < +\infty,
$$

which combining with (19) yields

$$
\sum_{k=2}^{+\infty}\frac{\|g_k\|^4}{m_k^2(1+\alpha_{k-1})^2\sum_{1\le i\le k}\|g_i\|^2} < +\infty. \tag{20}
$$

Since $\{m_k, k=1,2,\cdots\}$ has an upper bound, (17) holds.

On the other hand, by (9) and Lemma 3.1, we have

$$
\begin{aligned}
f_k - f_{k+1} &\geq -\alpha_k g_k^{\mathrm{T}} d_k - \frac{1}{2}\alpha_k^2 L\|d_k\|^2 \\
&= -\alpha_k g_k^{\mathrm{T}} d_k + \frac{L\alpha_k g_k^{\mathrm{T}} d_k}{2L_k} = -\frac{(2L_k - L)\left(\alpha_k g_k^{\mathrm{T}} d_k\right)}{2L_k} \\
&\geq -\frac{(2\rho-1)\left(\alpha_k g_k^{\mathrm{T}} d_k\right)}{2\rho} \geq \frac{(2\rho-1)\alpha_k\|g_k\|^2}{2(1+\alpha_{k-1})}.
\end{aligned}
\tag{21}
$$

By the same analysis as the above proof, (18) holds. The proof is completed. □

Lemma 3.4 (see [12]). If the conditions in Theorem 3.1 hold and $\sup_{k\ge1}\{\alpha_k\} < +\infty$, then both the sequence $\{g_k, k=1,2,\cdots\}$ and $\{d_k, k=1,2,\cdots\}$ have a bound.

Theorem 3.2. If the conditions in Theorem 3.1 hold, then

$$
\liminf_{k\to+\infty}\|g_k\| = 0. \tag{22}
$$

Proof. Suppose $\liminf_{k\to+\infty}\|g_k\| \neq 0$, then there exists a positive γ such that

$$
\|g_k\| \geq \gamma, \forall k \geq 1. \tag{23}
$$

In the following, we carry out our proofs in two cases.

Case 1. We complete the proof by utilizing reduction to absurdity. Suppose that $\sup_{k\ge1}\{\alpha_k\} < +\infty$. By (17), we have

$$\sum_{k=2}^{+\infty}\frac{\|g_k\|^4}{\sum_{1\le i\le k}\|g_i\|^2}<+\infty \tag{24}$$

From Lemma 3.4, we know that there exists $M>0$ such that $\|g_k\|\le M, \forall k\ge 1$. Combining (23), we have

$$\frac{\|g_k\|^4}{\sum_{1\le i\le k}\|g_i\|^2}\ge\frac{\gamma^4}{k\cdot M^2}.$$

It is known that

$$\sum_{k=2}^{+\infty}\frac{\gamma^4}{k\cdot M^2}=\frac{\gamma^4}{M^2}\sum_{k=2}^{+\infty}\frac{1}{k}=+\infty,$$

So

$$\sum_{k=2}^{+\infty}\frac{\|g_k\|^4}{\sum_{1\le i\le k}\|g_i\|^2}=+\infty, \tag{25}$$

which contradicts with (24). Therefore (22) holds.

Case 2. When $\sup_{k\ge1}\{\alpha_k\}=+\infty$, the proof is the same as that in [10] and here is omitted.

It follows from the proofs of Case 1 and Case 2 that (22) holds. This completes the proof. □

Remark 3.1. Search direction of Algorithm 2.1 can be extended to more general form as follows:

$$d_k=\begin{cases}-s_k g_k, & k=1,\\ -\rho\left(1-\varphi\left(\alpha_{k-1}\right)s_k\right)g_k\pm(1-\rho)\phi\varphi\left(\alpha_{k-1}\right)s_k d_{k-1}, & k\ge 2,\end{cases} \tag{26}$$

where the function $\varphi(\alpha)$ satisfies the following conditions(see [10]):

a) It is continuous and strictly monotone increasing when $\alpha\in[0,+\infty)$;

b) $\lim_{\alpha\to0^+}\varphi(\alpha)=\varphi(0)=0$ and $\lim_{\alpha\to+\infty}\varphi(\alpha)=1$;

c) $\alpha(1-\varphi(\alpha))$ is continuous, strictly monotone increasing when $\alpha\in[0,+\infty)$, and

$$\lim_{\alpha\to+\infty}\alpha\left(1-\varphi(\alpha)\right)=1.$$

Evidently, there are many functions satisfying the conditions (a)-(c). For example, $\frac{\alpha}{1+\alpha}$, $\frac{\alpha^2}{1+\alpha+\alpha^2}$, $\frac{\alpha^3}{1+\alpha^2+\alpha^3}$, etc (see [10]). We denote Algorithm 2.1 in which d_k is determined by (26) as Algorithm 3.1. By using proof technique of above Theorem 3.2, it is easy to get its convergence theorem.

4. Numerical Results

In this section, we report some preliminary numerical experiments. The test problems and their initial values are drawn from [13].

In numerical experiment, we take the parameter $L_k=100$, and stop the iteration if the inequality $\|g_k\|\le10^{-5}$ is satisfied. The detailed numerical results

Table 1. Numerical results.

Problem	Dim	NI	NF	NG	Problem	Dim	NI	NF	NG
ROSE	2	8	10	11	TRID	3	66	67	68
FROTH	2	6	7	8		50	84	85	86
GAUSS	3	126	128	129		100	86	87	88
BOX	3	1	51	52		200	85	87	89
SING	4	20	21	22	BAND	3	22	23	24
WOOD	4	6	7	8		50	27	28	29
BD	4	5	7	9		100	23	25	27
ROSEX	8	9	11	12		200	24	26	28
	50	8	9	10	LIN	2	1	3	5
	100	8	9	10		50	1	3	5
SINGX	4	20	21	22		500	1	3	5
PEN2	50	3	4	5		1000	1	3	5
VARDIM	2	99	100	101	LIN1	2	18	19	20
	50	4	5	6		10	1	3	5

are reported in Table 1, in which NI, NF and NG denote the total number of iterations, the total number of function evaluations and the total number of gradient evaluations, respectively. From Table 1, we can see the extended algorithm has good numerical results.

5. Conclusion

In this paper, we extended the descent algorithm without line search of [10] to more general case, and got its global convergence. Compared with [10], the extended algorithm has more effective numerical perfermance, so it is effective. In the future, we will further research the descent algorithms without line search, and try to get some new descent algorithms without line search, which not only convergence globally, but also have good numerical results.

Acknowledgements

We gratefully acknowledge the scholarship fund of education department of Guangxi Zhuang autonomous region, Guangxi basic ability improvement project fund for the middle-aged and young teachers of colleges and universities (2017KY0068, KY2016YB069), Guangxi higher education undergraduate course teaching reform project fund (2017JGB147), NNSF of China (11761014), Guangxi natural science foundation (2017GXNSFAA198243).

References

[1] Nocedal, J. and Stephen, J.W. (1999) Numerical Optimization. Springer-Verlag, New York. https://doi.org/10.1007/b98874

[2] Dai, Y.H. and Yuan, Y. X. (2000) Nonlinear Conjugate Gradient Methods. Shanghai Science and Technology Press of China, Shanghai. (In Chinese)

[3] Gilbert, J.C. and Nocedal, J. (1992) Global Convergence Properties of Conjugate Gradient Methods for Optimization. *SIAM Journal on Optimization*, **2**, 21-42. https://doi.org/10.1137/0802003

[4] Grippo, L. and Lucidi, S. (1997) A Globally Convergent Version of the Polak-Ribière Conjugate Gradient Method. *Mathematical Programming*, **78**, 375-391. https://doi.org/10.1007/BF02614362

[5] Sun, J. and Zhang, J.P. (2001) Global Convergence of Conjugate Gradient Methods without Line Search. *Annals of Operations Research*, **103**, 161-173. https://doi.org/10.1023/A:1012903105391

[6] Chen, X.D. and Sun, J. (2002) Global Convergence of a Two-Parameter Family of Conjugate Gradient Methods without Line Search. *Journal of Computational and Applied Mathematics*, **146**, 37-45. https://doi.org/10.1016/S0377-0427(02)00416-8

[7] Wang, J.Y. and Zhu, D.T. (2016) Conjugate Gradient Path Method without Line Search Technique for Derivative-Free Unconstrained Optimization. *Numerical Algorithms*, **73**, 957-983. https://doi.org/10.1007/s11075-016-0124-9

[8] Wang, J.Y. and Zhu, D.T. (2017) Derivative-Free Restrictively Preconditioned Conjugate Gradient Path Method without Line Search Technique for Solving Linear Equality Constrained Optimization. *Computers and Mathematics with Applications*, **73**, 277-293. https://doi.org/10.1016/j.camwa.2016.11.025

[9] Shi, Z. J. and Shen, J. (2005) Convergence of Descent Method without Line Search. *Applied Mathematics and Computation*, **167**, 94-107. https://doi.org/10.1016/j.amc.2004.06.097

[10] Zhou, G.M. (2009) A Descent Algorithm without Line Search for Unconstrained Optimization. *Applied Mathematics and Computation*, **215**, 2528-2533. https://doi.org/10.1016/j.amc.2009.08.058

[11] Zhou, G.M. and Feng, C.S. (2013) The Steepest Descent Algorithm without Line Search for p-Laplacian. *Applied Mathematics and Computation*, **224**, 36-45. https://doi.org/10.1016/j.amc.2013.07.096

[12] Shi, Z.J. and Shen, J. (2005) A New Descent Algorithm with Curve Search Rule. *Applied Mathematics and Computation*, **161**, 753-768. https://doi.org/10.1016/j.amc.2003.12.058

[13] Morè, J.J., Garbbow, B.S. and Hillstrom, K.E. (1981) Testing Unconstrained Optimization Software. *ACM Transactions on Mathematical Software*, **7**, 17-41. https://doi.org/10.1145/355934.355936