Advances in Applied Mathematics 应用数学进展, 2019, 8(4), 676-683
Published Online April 2019 in Hans. http://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2019.84076

Hans汉斯

A Modified Nonlinear Conjugate Gradient Algorithm Using Secant Conditions

Jing Liang*, Guodong Ma#, Yuting Shi, Zhimei Lin, Wenting Zou, Liuna Li

School of Mathematics and Statistics, Yulin Normal University, Yulin Guangxi
Email: #mgd2006@163.com

Received: Apr. 2nd, 2019; accepted: Apr. 17th, 2019; published: Apr. 24th, 2019

Abstract

In this paper, based on the idea of Ref. [1], we propose a modified conjugate gradient method using secant conditions for unconstrained optimization problems. The proposed algorithm improves the method in Ref. [1], which possesses the following properties: the search direction has the sufficient descent property; the global convergence of the given algorithm will be established under suitable assumptions; numerical results are reported to test its efficiency.

Keywords

Conjugate Gradient Method, Sufficient Descent Property, Global Convergence

基于拟牛顿方程一个改进的非线性共轭梯度算法

梁　静*，马国栋#，时瑜婷，林志梅，邹文婷，李柳娜

玉林师范学院，数学与统计学院，广西　玉林
Email: #mgd2006@163.com

收稿日期：2019年4月2日；录用日期：2019年4月17日；发布日期：2019年4月24日

摘　要

基于拟牛顿方程，借鉴文[1]的思想，本文给出一个改进的具有全局收敛非线性共轭梯度算法。该文的算法对文[1]中算法进行了改进，所产生的搜索方向具有充分下降性。在温和的假设下，新算法具有全局收

*第一作者。
#通讯作者。

文章引用: 梁静, 马国栋, 时瑜婷, 林志梅, 邹文婷, 李柳娜. 基于拟牛顿方程一个改进的非线性共轭梯度算法[J]. 应用数学进展, 2019, 8(4): 676-683. DOI: 10.12677/aam.2019.84076

敛性。最后，数值结果检验其有效性。

关键词

共轭梯度法，充分下降性，全局收敛性

1. 引言

本文考虑求解如下无约束优化问题

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

其中 $f(x):R^n\to R$ 连续可微，此问题在经济管理、生产运作、工程设计等领域具有广泛的应用背景。共轭梯度法是求解问题(1.1)的有效方法之一，该方法具有算法结构简单、计算存储量少、数值效果明显的优点，其迭代公式的一般形式为：

$$x_{k+1}=x_k+\alpha_k d_k, k=0,\ 1,\ 2,\cdots$$

其中 x_k 为第 k 次迭代点，$\alpha_k>0$ 为搜索步长，d_k 为搜索方向且其更新方式为

$$d_k=\begin{cases}-g_k+\beta_k d_{k-1}, k\ge 1,\\ -g_k, k=0,\end{cases} \tag{1.2}$$

其中 $\beta_k\in\Re$ 是方向调控参数。数值优化方法是根据搜索方向的选择来定义其方法类型，从上式可以看出，搜索方向 d_k 是由参数 β_k 决定的，所以根据 β_k 的选取方式可以定义不同的共轭梯度公式，下面给出几个著名的公式：

$$\beta_k^{PRP}=\frac{g_{k+1}^T(g_{k+1}-g_k)}{\|g_k\|^2}\ [2]\ [3],\quad \beta_k^{HS}=\frac{g_{k+1}^T(g_{k+1}-g_k)}{(g_{k+1}-g_k)^T d_k}\ [4],\quad \beta_k^{FR}=\frac{g_{k+1}^T g_{k+1}}{\|g_k\|^2}\ [5],$$

$$\beta_k^{DY}=\frac{g_{k+1}^T g_{k+1}}{(g_{k+1}-g_k)^T d_k}\ [6],\quad \beta_k^{CD}=\frac{g_{k+1}^T g_{k+1}}{-d_k^T g_k}\ [7],\quad \beta_k^{LS}=\frac{g_{k+1}^T(g_{k+1}-g_k)}{-g_k^{\ T} d_k}\ [8],$$

其中 $g_k=\nabla f(x_k)$ 和 $g_{k+1}=\nabla f(x_{k+1})$ 表示函数 $f(x)$ 在 x_k 和 x_{k+1} 的梯度值，$\|.\|$ 是欧氏向量范数。众多优化专家都希望能找到理论性质良好且数值表现又好的算法，已取得众多成果(见[9]-[25])，其中 Yuan [9]给出如下 PRP 修改公式：

$$\beta_k^{MPRP}=\beta_k^{PRP}-\min\left\{\beta_k^{PRP},\frac{\mu\|y_k\|^2}{\|g_k\|^4}g_{k+1}^T d_k\right\},$$

其中 $\mu>\frac{1}{4}$ 是常数，$y_k=g_{k+1}-g_k$，同时又将此公式推广到了其它 5 个公式中，得到了修改的 FR 公式、修改的 HS 公式、修改的 DY 公式、修改的 CD 公式和修改的 LS 公式，其中文[1]修改的 LS 为：

$$\beta_k^{MMLS+}=\beta_k^{MMLS}-\min\left\{\beta_k^{MMLS},\frac{\mu\|y_k^m\|^2}{(-d_k^T g_k)^2}g_{k+1}^T d_k\right\}, \tag{1.3}$$

其中 $y_k^m = g_{k+1} - \frac{\|g_{k+1}\|}{\|g_k\|} g_k$， $\beta_k^{MMLS} = \frac{g_{k+1}^T y_k^m}{-d_k^T g_k}$ 。

本文将在此公式的基础上进一步研究，从上式可以看出公式中不含函数值信息，这势必影响此方法的有效性。众所周知，拟牛顿方法不仅含有梯度值信息也含有函数值信息，且具有更好的理论和实际表现[10] [11]。也有许多学者将拟牛顿法的技术思想应用到共轭梯度公式中，从而获得更好的理论性质和实际数值效果[12] [13]。Zhang 等[11]给出了一个如下形式的拟牛顿方程

$$B_{k+1} S_k = y_k^m, \tag{1.4}$$

其中 $y_k^m = y_k + \gamma_k^1 s_k$， $s_k = x_{k+1} - x_k$ 及

$$\gamma_k^1 = \frac{3\left(g_{k+1} + g_k\right)^T s_k + 6\left(f\left(x_k\right) - f\left(x_{k+1}\right)\right)}{\|s_k\|^2},$$

此公式包含有梯度值信息和函数值信息，理论上能更高阶地近似目标函数的 Hesse 矩阵。基于公式(1.3)，(1.4)和文[12] [13]的思想，本文给出如下修改的 LS 共轭梯度公式

$$\beta_k^{MMLS*} = \beta_k^{MMLS} - \min\left\{\beta_k^{MMLS}, \frac{\mu\|y_k^m\|^2}{\left(-d_k^T g_k\right)^2} g_{k+1}^T d_k\right\}, \tag{1.5}$$

其中 $\beta_k^{MMLS} = \frac{g_{k+1}^T y_k^m}{-d_k^T g_k}$。相对于原 LS 方法，本文的创新点主要有：1) 新公式能保证方向调控参数 $\beta_k^{MMLS*} \ge 0$；2) 新方法具有充分下降性；3) 新方法不仅使用了梯度值信息还有函数值信息。

2. 算法的描述

基于搜索方向，给出一个基于拟牛顿方程改进的非线性共轭梯度算法。

算法 1：

步骤 0： 给定 $x_0 \in \Re^n, \delta \in (0, 1/2), \sigma \in (\delta, 1)$， $\varepsilon > 0$，令 $d_0 = -g_0 = -\nabla f(x_0)$，置 $k := 0$。

步骤 1： 若 $\|g_k\| \le \varepsilon$，停止。否则，进入步骤 2。

步骤 2： 利用弱 Wolfe-Powell(WWP)线搜索技术产生步长 α_k

$$f\left(x_k + \alpha_k d_k\right) \le f_k + \delta \alpha_k g_k^T d_k, \tag{2.1}$$

$$g\left(x_k + \alpha_k d_k\right)^T d_k \ge \sigma g_k^T d_k. \tag{2.2}$$

步骤 3： 令 $x_{k+1} = x_k + \alpha_k d_k$。如果 $\|g_{k+1}\| \le \varepsilon$，停止。

步骤 4： 利用公式(1.5)计算 β_k^{MMLS*}，并按如下公式计算搜索方向

$$d_{k+1} = -g_{k+1} + \beta_k^{MMLS*} d_k, \tag{2.3}$$

步骤 5： 置 $k := k + 1$，转步骤 2。

3. 算法的充分下降性和全局收敛性

引理 3.1： 对所有 $k \ge 0$，搜索方向 d_k 满足下面两式

$$d_k^T g_k \le -c\|g_k\|^2 \tag{3.1}$$

和

$$d_k^T y_k \ge c(1-\sigma)\|g_k\|^2 \tag{3.2}$$

$c>0$ 是常数。

证明：首先证明(3.1)成立。根据算法 1，如果 $k=0$，则 $g_0^T d_0 = -\|g_0\|^2$，式(3.1)显然成立。下面利用归纳法，当 $k \ge 1$，假如公式(2.3)满足(3.1)，对 $k+1$，有

$$
\begin{aligned}
g_{k+1}^T d_{k+1} &= -\|g_{k+1}\|^2 + \beta_k^{MMLS*} d_k^T g_{k+1} \\
&= -\|g_{k+1}\|^2 + \left[\frac{g_{k+1}^T y_k^m}{-d_k^T g_k} - \min\left\{\frac{g_{k+1}^T y_k^m}{-d_k^T g_k}, \frac{\mu\|y_k^m\|^2}{\left(-d_k^T g_k\right)^2} g_{k+1}^T d_k\right\}\right] d_k^T g_{k+1}
\end{aligned} \tag{3.3}
$$

由 $-d_k^T g_k > 0.$ 取 $u = \dfrac{\sqrt{-d_k^T g_k}}{\sqrt{2\mu}} g_{k+1}, v = \dfrac{\sqrt{2\mu} g_{k+1}^T d_k}{\sqrt{-d_k^T g_k}} y_k^m$。下面我们分两种情形分别讨论(3.3)式：

情形 I：如果 $\dfrac{g_{k+1}^T y_k^m}{-d_k^T g_k} < \dfrac{\mu\|y_k^m\|^2}{\left(-d_k^T g_k\right)^2} g_{k+1}^T d_k$，显然 $g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2$ 成立。

情形 II：如果 $\dfrac{g_{k+1}^T y_k^m}{-d_k^T g_k} \ge \dfrac{\mu\|y_k^m\|^2}{\left(-d_k^T g_k\right)^2} g_{k+1}^T d_k$，(3.3)式可化为

$$
\begin{aligned}
g_{k+1}^T d_{k+1} &= -\|g_{k+1}\|^2 + \left(\frac{g_{k+1}^T y_k^m}{-d_k^T g_k} - \frac{\mu\|y_k^m\|^2}{\left(-d_k^T g_k\right)^2} g_{k+1}^T d_k\right) d_k^T g_{k+1} \\
&= \frac{d_k^T g_{k+1} g_{k+1}^T y_k^m - \|g_{k+1}\|^2\left(-d_k^T g_k\right) - \dfrac{\mu\|y_k^m\|^2}{-d_k^T g_k}\left(g_{k+1}^T d_k\right)^2}{-d_k^T g_k} \\
&= \frac{u^T v - \frac{1}{2}\left(\|u\|^2 + \|v\|^2\right)}{-d_k^T g_k} + \frac{-\left(1-\dfrac{1}{4\mu}\right)\|g_{k+1}\|^2 d_k^T y_k^m}{-d_k^T g_k} \\
&\le -\left(1-\frac{1}{4\mu}\right)\|g_{k+1}\|^2,
\end{aligned}
$$

上述不等式利用了 $u^T v \le \frac{1}{2}\left(\|u\|^2 + \|v\|^2\right)$ 的关系。取 $c = 1-\frac{1}{4\mu}$，则(3.1)成立。根据线搜索技术(2.2)可推出(3.2)也成立。证毕。

我们将采用反证法来证明算法 1 的全局收敛性，假设存在常数 $\varepsilon_0 > 0$ 满足

$$
\|g_k\| \ge \varepsilon_0, \forall k \ge 0. \tag{3.4}
$$

根据(3.4)导出矛盾，从而证明 $\|g_k\| \to 0,\ k \to \infty$ 成立。

为证明算法 1 的收敛性，还需要下面的常规假设条件。

假设 A：1) 定义水平集 $\Omega = \left\{x \in \Re^n : f(x) \le f(x_0)\right\}$ 且有界，存在 $L_0 > 0$ 满足

$$
\left\|\nabla^2 f(x)\right\| \le L_0, x \in \Omega.
$$

2) 目标函数 f 在水平集 Ω 上连续可微并有下界，目标函数梯度满足 Lipschitz 条件，所谓 Lipschitz 条件就是存在常数 $L>0$ 满足下式

$$
\|g(x) - g(y)\| \le L\|x - y\|, \forall x, y \in \Omega。 \tag{3.5}
$$

下面引理 3.2 和引理 3.3 在共轭梯度算法文献中有类似的结论，本文只给出引理 3.3 证明过程。

引理 3.2：设假设 A 成立，序列 $\{g_k\}$ 和 $\{d_k\}$ 由算法 1 产生，则 $d_k \neq 0$ 及 $\sum_{k=0}^{\infty}\|u_{k+1}-u_k\|^2<\infty$，其中 $u_k=\frac{d_k}{\|d_k\|}$。

Gilbert 和 Nocedal [14]给出下面性质(P)，具体内容是：

性质(P)对于两参数 $\gamma_1,\gamma_2>0$，且满足

$$0<\gamma_1\leq\|g_k\|\leq\gamma_2。\tag{3.6}$$

若对所有 k，存在常数 $b>1$ 和 $\lambda>0$ 满足 $|\beta_k|\leq b$ 和 $\|s_k\|\leq\lambda$，有 $|\beta_k|\leq\frac{1}{2b}$ 成立。

引理 3.3：如果假设 A 满足且序列 $\{g_k\}$ 和 $\{d_k\}$ 由算法 1 产生，且存在常数 $M>0$ 使得 $\|d_k\|\leq M$ 成立，则 β_k^{MMLS*} 满足性质(P)。

证明：对于公式(2.3)，如果 $\frac{g_{k+1}^T y_k^m}{-d_k^T g_k}\leq\frac{\mu\|y_k^m\|^2}{\left(-d_k^T g_k\right)^2}g_{k+1}^T d_k$ 成立，结论显然成立。否则，利用假设 A(1)，可推出存在常数 $M_1>0$ 使下式

$$\|s_k\|\leq M_1 \tag{3.7}$$

成立。利用 $\|d_k\|\leq M$，(3.1)，(3.2)，(3.5)~(3.7)，有

$$\begin{aligned}
\left|\beta_k^{MMLS*}\right| &\leq \left|\frac{g_{k+1}^T y_k^m}{d_k^T g_k}\right| + \frac{\mu\|y_k^m\|^2}{\left(-d_k^T g_k\right)^2}\left|g_{k+1}^T d_k\right| \leq \frac{\|g_{k+1}\|\left\|\|g_{k+1}-g_k\|+3\left[\|g_{k+1}-g_k\|+\left\|\nabla^2 f(\varsigma_k)s_k\right\|\right]\right\|}{c\|g_k\|^2} \\
&+\frac{\mu\|g_{k+1}\|\|d_k\|\left\|\|g_{k+1}-g_k\|+3\left[\|g_{k+1}-g_k\|+\left\|\nabla^2 f(\varsigma_k)s_k\right\|\right]\right\|^2}{c^2\|g_k\|^4} \\
&\leq\frac{\left[L\gamma_2+3\left(L+L_0\right)\right]\|s_k\|}{c\gamma_1^2}+\frac{\left[\mu\gamma_2 M\left(L+3\left(L+L_0\right)\right)^2 M_1\right]\|s_k\|}{c^2\gamma_1^4} \\
&=\left(\frac{\left[L\gamma_2+3\left(L+L_0\right)\right]c\gamma_1^2+\left[\mu\gamma_2 M\left(L+3\left(L+L_0\right)\right)^2 M_1\right]}{c^2\gamma_1^4}\right)\|s_k\|,
\end{aligned}$$

取 $b=\max\left\{2,\frac{\left[L\gamma_2+3\left(L+L_0\right)\right]c\gamma_1^2+\left[\mu\gamma_2 M\left(L+3\left(L+L_0\right)\right)^2 M_1\right]}{c^2\gamma_1^4}\right\}$ 和

$$\lambda=\frac{c^2\gamma_1^4}{b\left\{\left[L\gamma_2+3\left(L+L_0\right)\right]c\gamma_1^2+\left[\mu\gamma_2 M\left(L+3\left(L+L_0\right)\right)^2 M_1\right]\right\}}.$$

利用(3.8)，λ 和 $b>1$，则 $\left|\beta_k^{MMLS*}\right|\leq b$ 和

$$\begin{aligned}
\left|\beta_k^{MMLS*}\right| &\leq\frac{\left[L\gamma_2+3\left(L+L_0\right)\right]c\gamma_1^2+\left[\mu\gamma_2 M\left(L+3\left(L+L_0\right)\right)^2 M_1\right]}{c^2\gamma_1^4}\|s_k\| \\
&\leq\frac{\left[L\gamma_2+3\left(L+L_0\right)\right]c\gamma_1^2+\left[\mu\gamma_2 M\left(L+3\left(L+L_0\right)\right)^2 M_1\right]}{c^2\gamma_1^4}\lambda=\frac{1}{2b}.
\end{aligned}$$

因此公式(2.3)满足性质(P)，引理得证。

利用假设 A，引理 3.1~3.3，与文献[15]中的定理 3.2 的证明类似，可得到算法 1 的全局收敛性定理，定理的证明请参照文献[15]中的定理 3.2 完成。

定理 3.1：序列 $\{d_k, g_k\}$ 由算法产生，若假设 A 成立，则 $\lim_{k\to\infty}\inf\|g_k\|=0$ 成立。

4. 数值结果

这部分将给出数值检验，Benchmar 问题[16]在工程领域具有广泛应用背景。

1) Sphere 函数：$f_{Sph}(x)=\sum_{i=1}^{n}x_i^2, x_i\in[-5.12,5.12], f_{Sph}(x^*)=0.$

2) Schwefel's 函数：$f_{SchDS}(x)=\sum_{i=1}^{n}\left(\sum_{j=1}^{i}x_j\right)^2, x_i\in[-65.536,65.536], f_{SchDS}(x^*)=0.$

3) Rastrigin 函数：$f_{Ras}(x)=10n+\sum_{i=1}^{n}\left(x_i^2-10\cos(2\pi x_i)\right), x_i\in[-5.12,5.12], f_{Ras}(x^*)=0.$

4) Griewank 函数：$f_{Gri}(x)=1+\sum_{i=1}^{n}\frac{x_i^2}{4000}-\prod_{i=1}^{n}\cos\frac{x_i}{i}, x_i\in[-600,600], f_{Gri}(x^*)=0.$

参数选取如下：$\varepsilon=10^{-5},\delta=0.1,\sigma=0.9$，停止准则采用 Himmeblau 准则：如果 $|f(x_k)|>e_1$，令 $\text{stop1}=\frac{|f(x_k)-f(x_{k+1})|}{|f(x_k)|}$；否则，令 $\text{stop1}=|f(x_k)-f(x_{k+1})|$。如果 $\|g(x)\|<\varepsilon$ 或者 $\text{stop1}<e_2$ 满足，算法终止，$e_1=e_2=10^{-5}$。若迭代次数超过 1000 次，程序也将终止。为比较算法 1 的有效性，也给出通常 LS 方法的数值检验，并进行比较。以下是各代码含义：

x_0：初始点；Dim：问题的维数；NI：迭代次数；NFG：函数值与梯度值迭代次数和；Time：以秒为单位的计算机 CPU 时间；$f(x)$：算法终止时的函数值。

为比较算法 1 与原 LS 算法的效率，我们采用下述技术即两算法的所有 NFG 的乘积的比值再开二十四分之一(结果的总个数)次方，以通常 LS 算法作为底，定义为 1，比值见表 2，此技术见文献[17]。

Table 1. Numerical results
表 1. 数值结果

		算法 1		原 LS 算法	
		NI/NFG/ $f(\bar{x})$ /Time	NI/NFG/ $f(\bar{x})$ /Time	NI/NFG/ $f(\bar{x})$ /Time	NI/NFG/ $f(\bar{x})$ /Time
Sphere	x_0	$(-4,-4,\cdots,-4)$	$(3,3,\cdots,3)$	$(-4,-4,\cdots,-4)$	$(3,3,\cdots,3)$
	10	2/6/7.888609e−030/0	2/6/7.888609e−030/0	2/6/7.888609e−030/0.0468003	2/6/7.888609e−030/0
Dim	100	2/6/3.478877e−026/0	2/6/1.262177e−027/0.0312002	2/6/3.478877e−026/0	2/6/1.262177e−027/0
	300	2/6/1.855401e−025/0	2/6/2.863565e−026/0	2/6/1.855401e−025/0.0312002	2/6/2.863565e−026/0
Schwefel's	x_0	$(-0.001,\cdots,-0.001)$	$(0.001,\cdots,-0.001)$	$(-0.001,\cdots,-0.001)$	$(0.001,\cdots,-0.001)$
	10	2/5/4.378593e−008/0	2/5/4.378593e−010/0	2/5/4.378593e−008/0	2/5/4.378593e−010/0
Dim	100	3/8/3.711002e−006/0.0468003	2/5/3.979128e−007/0.0312002	3/8/3.711002e−006/0.0936006	2/5/3.979128e−007/0.0312002
	300	4/11/1.871432e−005/2.043613	3/8/9.937482e−007/1.326009	4/11/1.871432e−005/1.950012	3/8/9.937482e−007/1.310408
Rastrigin	x_0	$(0.01,0.01,\cdots,0.01)$	$(0.001,\cdots,0.001)$	$(0.01,0.01,\cdots,0.01)$	$(0.001,\cdots,0.001)$

Continued

	10	4/12/0.000000e+000/0	2/5/3.208869e−006/0	3/10/0.000000e+000/0.0312002	2/5/3.208869e−006/0
Dim	100	3/8/3.612627e−009/0.0468003	3/8/0.000000e+000/0	3/9/0.000000e+000/0	3/9/0.000000e+000/0
	300	3/8/2.728484e−012/0	3/8/0.000000e+000/0	3/9/0.000000e+000/0.0312002	3/9/0.000000e+000/0
Griewank	x_0	$(-100,-100,\cdots,-100)$	$(30,30,\cdots,30)$	$(-100,-100,\cdots,-100)$	$(30,30,\cdots,30)$
	10	2/19/7.250909e+000/0	10/47/3.316156e−006/0.0312002	2/19/7.250909e+000/0.0468003	9/43/3.715464e−006/0
Dim	100	4/32/3.509751e−002/0.0468003	2/6/0.000000e+000/0	4/32/2.276154e−002/0.0624004	2/6/0.000000e+000/0.0312002
	300	19/54/8.028683e−005/0.5772037	2/6/0.000000e+000/0.0624004	23/64/8.635447e−005/0.6708043	2/6/0.000000e+000/0.0468003
总的 CUP 时间		4。243227		4。383628	

Table 2. The Efficiency of NFG
表 2. NFG 的效率

算法 1	原 LS 算法
0.98	1

从表 1 的结果可看，算法 1 和 LS 算法对上述 4 个问题都能有效地求解，迭代次数可以接受，最终函数值很接近最优函数值，新算法的 CPU 时间相对少一些，更具竞争力。表 2 的效率表明，算法 1 相对于原 LS 算法，效率提高 2%。

基金项目

2017 年国家级大学生创新创业训练计划项目(201710606041)，广西自然科学基金(2018JJA110039)和玉林师范学院校级科研项目(2019YJKY16)。

参考文献

[1] 陈海. 一个修改的 LS 非线性共轭梯度算法[J]. 应用数学进展, 2013(2): 48-54.

[2] Polak, E. and Ribiere, G. (1969) Note sur la xonvergence de directions conjugees. *Rev. Francaise informat Recherche Operatinelle*, 3e *Annee*, **16**, 35-43.

[3] Polyak, B.T. (1969) The Conjugate Gradient Method in Extreme Problems. *USSR Computational Mathematics and Mathematical Physics*, **9**, 94-112

[4] Hestenes, M.R. and Stiefel, E. (1952) Method of Conjugate Gradient for Solving Linear Equations. *Journal of Research of the National Bureau of Standards*, **49**, 409-436. https://doi.org/10.6028/jres.049.044

[5] Fletcher, R. and Reeves, C. (1964) Function Minimization bu Conjugate Gradients. *Compute Journal*, **7**, 149-154.

[6] Dai, Y. and Yuan, Y. (1999) A Nonlinear Conjugate Gradient with a Strong Global Convergence Properties. *SIAM Journal on Optimization*, **10**, 177-182. https://doi.org/10.1137/S1052623497318992

[7] Fletcher, R. (1980) Practical Method of Optimization, Vol. I: Unconstrained Optimization. 2nd Edition, Wiley, Hoboken.

[8] Liu, Y. and Storey, C. (1992) Efficient Generalized Conjugate Gradient Algorithms, Part 1: Theory. *Journal of Optimization Theory and Application*, **69**, 17-41.

[9] Yuan, G. (2009) Modified Nonlinear Conjugate Gradient Methods with Sufficient Descent Property for Large-Scale Optimization Problems. *Optimization Letters*, **3**, 11-21. https://doi.org/10.1007/s11590-008-0086-5

[10] Yuan, G. and Wei, Z. (2010) Convergence Analysis of a Modified BFGS Method on Convex Minimizations. *Computational Optimization and Applications*, **47**, 237-255. https://doi.org/10.1007/s10589-008-9219-0

[11] Zhang, J., Deng, N. and Chen, L. (1999) New Quasi-Newton Equation and Related Methods for Unconstrained Opti-

mization. *Journal of Optimization Theory and Application*, **102**, 147-167. https://doi.org/10.1023/A:1021898630001

[12] Yuan, G., Wei, Z. and Zhao, Q. (2014) A Modified Polak-Ribière-Polyak Conjugate Gradientalgorithm for Large-Scale Optimization Problems. *IIE Transactions*, **46**, 397-413. https://doi.org/10.1080/0740817X.2012.726757

[13] Yuan, G. and Zhang, M. (2013) A Modified Hestenes-Stiefel Conjugate Gradient Algorithm for Large-Scale Optimization. *Numerical Functional Analysis and Optimization*, **34**, 914-937. https://doi.org/10.1080/01630563.2013.777350

[14] 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

[15] Hager, W.W. and Zhang, H. (2005) A New Conjugate Gradient Method with Guaranteed Descent and an Efficient Line Search. *SIAM Journal on Optimization*, **16**, 170-192. https://doi.org/10.1137/030601880

[16] Yao, S., He, D. and Shi, L. (2018) An Improved Perry Conjugate Gradient Method with Adaptive Parameter Choice. *Numerical Algorithms*, **78**, 1255-1269. https://doi.org/10.1007/s11075-017-0422-x

[17] Yuan, G. and Lu, X. (2009) A Modified PRP Conjugate Gradient Method. *Annals of Operations Research*, **66**, 73-90. https://doi.org/10.1007/s10479-008-0420-4

[18] Dai, Y. and Liao, L. (2001) New Conjugacy Conditions and Related Nonlinear Conjugate Methods. *Applied Mathematics and Optimization*, **43**, 87-101. https://doi.org/10.1007/s002450010019

[19] Hager, W.W. and Zhang, H. (2006) Algorithm 851: CGDESENT, A Conjugate Gradient Method with Guaranteed Descent. *ACM Transactions on Mathematical Software*, **32**, 113-137. https://doi.org/10.1145/1132973.1132979

[20] Li, G., Tang, C. and Wei, Z. (2007) New Conjugacy Condition and Related New Conjugate Gradient Methods for Unconstrained Optimization Problems. *Journal of Computational and Applied Mathemathics*, **202**, 532-539.

[21] Wei, Z., Li, G. and Qi, L. (2006) New Nonlinear Conjugate Gradient Formulas for Large-Scale Unconstrained Optimization Problems. *Applied Mathematics and Computation*, **179**, 407-430. https://doi.org/10.1016/j.amc.2005.11.150

[22] Wei, Z., Li, G. and Qi, L. (2008) Global Convergence of the PRP Conjugate Gradient Methods with Inexact Line Search for Nonconvex Unconstrained Optimization Problems. *Mathematics of Computation*, **77**, 2173-2193. https://doi.org/10.1090/S0025-5718-08-02031-0

[23] Wei, Z., Yao, S. and Lin, L. (2006) The Convergence Properties of Some New Conjugate Gradient Methods. *Applied Mathematics and Computation*, **183**, 1341-1350. https://doi.org/10.1016/j.amc.2006.05.150

[24] Yuan, G., Lu, X. and Wei, Z. (2009) A Conjugate Gradient Method with Descent Direction for Unconstrained Optimization. *Journal of Computational and Applied Mathematics*, **102**, 519-530. https://doi.org/10.1016/j.cam.2009.08.001

[25] 李向荣. 修改的 DY 和 HS 共轭梯度算法及其全局收敛性[J]. 理论数学, 2011(1): 1-7.