

DOI: 10.13718/j.cnki.xdzk.2017.05.020

一种新的混合共轭梯度算法^①

韩 信, 张俊容, 王森森

西南大学 数学与统计学院, 重庆 400715

摘要: 根据现有的共轭梯度算法, 提出了一种新的求解无约束优化问题的混合共轭梯度法. 在每一步迭代过程中, 新算法总是能生成一个充分下降方向. 在 Wolfe 线搜索下, 提出的算法具有全局收敛性. 数值实验表明该算法具有良好的计算性能.

关键词: 无约束优化; 混合共轭梯度法; 充分下降; 全局收敛性

中图分类号: O224

文献标志码: A

文章编号: 1673-9868(2017)05-0132-07

考虑下面的无约束优化问题

$$\min\{f(\mathbf{x}): \mathbf{x} \in \mathbb{R}^n\} \quad (1)$$

其中 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 是连续可微函数, $\mathbb{R}^n$ 表示 n 维欧氏空间, $\mathbf{g}$ 是 f 的梯度向量, 即 $\mathbf{g}(\mathbf{x}) = \nabla f(\mathbf{x})$. 共轭梯度法是求解无约束优化问题(1)的一种十分有效的算法. 该算法的一般迭代格式如下:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k \quad (2)$$

这里 α_k 是由某种线搜索获得的迭代步长, 并且

$$\mathbf{d}_k = \begin{cases} -\mathbf{g}_k & k=1 \\ -\mathbf{g}_k + \beta_k \mathbf{d}_{k-1} & k>1 \end{cases} \quad (3)$$

为搜索方向, β_k 为共轭参数. 迭代步长 α_k 通常由如下的 Wolfe 线搜索

$$\begin{aligned} f(\mathbf{x}_k + \alpha_k \mathbf{d}_k) - f(\mathbf{x}_k) &\leq \delta \alpha_k \mathbf{g}_k^T \mathbf{d}_k \\ \mathbf{g}(\mathbf{x}_k + \alpha_k \mathbf{d}_k)^T \mathbf{d}_k &\geq \sigma \mathbf{g}_k^T \mathbf{d}_k \end{aligned} \quad (4)$$

或者强 Wolfe 线搜索

$$\begin{aligned} f(\mathbf{x}_k + \alpha_k \mathbf{d}_k) - f(\mathbf{x}_k) &\leq \delta \alpha_k \mathbf{g}_k^T \mathbf{d}_k \\ |\mathbf{g}(\mathbf{x}_k + \alpha_k \mathbf{d}_k)^T \mathbf{d}_k| &\leq \sigma |\mathbf{g}_k^T \mathbf{d}_k| \end{aligned} \quad (5)$$

获得. 这里 δ 与 σ 均为参数, 且 $0 < \delta < \sigma < 1$. 对于共轭参数 β_k 的选取^[1-7], 有如下 6 种经典的形式:

$$\begin{aligned} \beta_k^{\text{FR}} &= \frac{\|\mathbf{g}_k\|^2}{\|\mathbf{g}_{k-1}\|^2} & \beta_k^{\text{DY}} &= \frac{\|\mathbf{g}_k\|^2}{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}} & \beta_k^{\text{CD}} &= -\frac{\|\mathbf{g}_k\|^2}{\mathbf{d}_{k-1}^T \mathbf{g}_{k-1}} \\ \beta_k^{\text{PRP}} &= \frac{\mathbf{g}_k^T \mathbf{y}_{k-1}}{\|\mathbf{g}_{k-1}\|^2} & \beta_k^{\text{HS}} &= \frac{\mathbf{g}_k^T \mathbf{y}_{k-1}}{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}} & \beta_k^{\text{LS}} &= -\frac{\mathbf{g}_k^T \mathbf{y}_{k-1}}{\mathbf{d}_{k-1}^T \mathbf{g}_{k-1}} \end{aligned}$$

其中: $\mathbf{y}_{k-1} = \mathbf{g}_k - \mathbf{g}_{k-1}$, $\|\cdot\|$ 代表欧氏范数.

上面的 6 种算法的全局收敛性和数值表现有所不同. 众所周知, 虽然 FR 法、DY 法和 CD 法实际计算效

① 收稿日期: 2016-03-27

基金项目: 国家自然科学基金项目(11401487).

作者简介: 韩 信(1989-), 男, 四川平昌人, 硕士研究生, 主要从事最优化理论、算法及应用的研究.

通信作者: 张俊容, 博士, 副教授.

果一般,但是它们具有很好的全局收敛性;另外,尽管 PRP 法、HS 法和 LS 法在一般情形下可能不收敛,但是它们具有良好的数值表现.在以上方法的基础上,很多学者试着构造一种既具有好的收敛性质又具有优秀的数值表现的新算法.一种思路是直接对上面的共轭参数 β_k 进行改良,例如文献[8—10].另外就是将收敛性好和数值计算性能优良的方法进行有效混合,例如文献[11].

2006 年,文献[8]提出了 FR 法的一种改进形式:

$$\beta_k^{\text{VFR}} = \frac{\mu_1 \| \mathbf{g}_k \|^2}{\mu_2 | \mathbf{g}_k^T \mathbf{d}_{k-1} | + \mu_3 \| \mathbf{g}_{k-1} \|^2}$$

其中: $\mu_1 \in (0, +\infty)$, $\mu_2 \in (\mu_1 + \epsilon_1, +\infty)$, $\mu_3 \in (0, +\infty)$, 且 ϵ_1 为给定的正参数.在任意线搜索条件下,该方法都具有充分下降性.

此后,文献[9]给出了 HS 法的一种改良公式:

$$\beta_k^{\text{YWH}} = \frac{\| \mathbf{g}_k \|^2 - \frac{\| \mathbf{g}_k \|^2}{\| \mathbf{g}_{k-1} \|^2} \mathbf{g}_k^T \mathbf{g}_{k-1}}{\mathbf{d}_{k-1}^T \mathbf{y}_{k-1}} \quad (6)$$

在强 Wolfe 线搜索条件(5)下,若 $0 < \sigma < \frac{1}{3}$,则该算法能生成充分下降方向并且具有全局收敛性.

最近,基于文献[8]的想法,文献[10]引进了 PRP 法的一种修正形式:

$$\beta_k^{\text{DPRP}} = \frac{\| \mathbf{g}_k \|^2 - \frac{\| \mathbf{g}_k \|^2}{\| \mathbf{g}_{k-1} \|^2} | \mathbf{g}_k^T \mathbf{g}_{k-1} |}{\mu | \mathbf{g}_k^T \mathbf{d}_{k-1} | + \| \mathbf{g}_{k-1} \|^2} \quad (7)$$

其中参数 $\mu > 1$.在任意线搜索条件下,该方法都具有充分下降性.在 Wolfe 线搜索条件(4)下,该算法具有全局收敛性.

受文献[8]和[10]的启示,本文提出了一种新的混合共轭梯度方法 HZW 法,其共轭参数如下:

$$\beta_k^{\text{HZW}} = \frac{\| \mathbf{g}_k \|^2 - \frac{\| \mathbf{g}_k \|^2}{\| \mathbf{g}_{k-1} \|^2} \max\{0, \mathbf{g}_k^T \mathbf{g}_{k-1}\}}{\max\{u | \mathbf{g}_k^T \mathbf{d}_{k-1} | + \| \mathbf{g}_{k-1} \|^2, \mathbf{d}_{k-1}^T \mathbf{y}_{k-1}\}} \quad (8)$$

其中参数 $u > 1$.由(8)式易得, β_k^{HZW} 为 β_k^{DPRP} , β_k^{YWH} , β_k^{VFR} 和 β_k^{DY} 的一种混合.新的混合算法 HZW 法继承了 DPRP 法、YWH 法、VFR 法和 DY 法的优点,既具有好的全局收敛性又有出色的数值表现.

1 新的混合型共轭梯度算法及其性质

基于(8)式,我们将给出新的混合共轭梯度法 HZW 的算法步骤.

HZW 算法

步骤 0 给定初始点 $\mathbf{x}_1$ 和精度 $\epsilon > 0$.令 $k := 1$, 并计算 $\mathbf{d}_1 = \mathbf{g}_1$.

步骤 1 若 $\| \mathbf{g}_k \| \leq \epsilon$, 停止.否则,转步骤 2.

步骤 2 按 Wolfe 线搜索(4)式,计算迭代步长 α_k .

步骤 3 由 $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$, $\mathbf{g}_{k+1} = \mathbf{g}(\mathbf{x}_{k+1})$ 生成新的迭代点并根据(8)式计算 β_{k+1}^{HZW} .

步骤 4 由 $\mathbf{d}_{k+1} = -\mathbf{g}_{k+1} + \beta_{k+1}^{\text{HZW}} \mathbf{d}_k$ 计算新的搜索方向.令 $k := k + 1$, 转步骤 1.

下面,我们将给出算法 HZW 中的共轭参数 β_k^{HZW} 所具有的性质.

引理 1 若共轭参数 β_k^{HZW} 由(8)式定义,则

$$0 \leq \beta_k^{\text{HZW}} \leq \beta_k^{\text{FR}} \quad \forall k \geq 2$$

证 根据 β_k^{HZW} 的定义,有

$$\beta_k^{\text{HZW}} = \frac{\| \mathbf{g}_k \|^2 - \frac{\| \mathbf{g}_k \|^2}{\| \mathbf{g}_{k-1} \|^2} \max\{0, \mathbf{g}_k^T \mathbf{g}_{k-1}\}}{\max\{u | \mathbf{g}_k^T \mathbf{d}_{k-1} | + \| \mathbf{g}_{k-1} \|^2, \mathbf{d}_{k-1}^T \mathbf{y}_{k-1}\}} \geq$$

$$\frac{\| \mathbf{g}_k \|^2 - \frac{\| \mathbf{g}_k \|^2}{\| \mathbf{g}_{k-1} \|^2} \| \mathbf{g}_k \|^2 \| \mathbf{g}_{k-1} \|^2}{\max\{u | \mathbf{g}_k^T \mathbf{d}_{k-1} | + \| \mathbf{g}_{k-1} \|^2, \mathbf{d}_{k-1}^T \mathbf{y}_{k-1}\}} = 0 \quad (9)$$

另一方面, 对 β_k^{HZW} 适当放大, 就有

$$\begin{aligned} \beta_k^{\text{HZW}} &= \frac{\| \mathbf{g}_k \|^2 - \frac{\| \mathbf{g}_k \|^2}{\| \mathbf{g}_{k-1} \|^2} \max\{0, \mathbf{g}_k^T \mathbf{g}_{k-1}\}}{\max\{u | \mathbf{g}_k^T \mathbf{d}_{k-1} | + \| \mathbf{g}_{k-1} \|^2, \mathbf{d}_{k-1}^T \mathbf{y}_{k-1}\}} \leqslant \\ &= \frac{\| \mathbf{g}_k \|^2}{u | \mathbf{g}_k^T \mathbf{d}_{k-1} | + \| \mathbf{g}_{k-1} \|^2} \leqslant \frac{\| \mathbf{g}_k \|^2}{\| \mathbf{g}_{k-1} \|^2} = \beta_k^{\text{FR}} \end{aligned} \quad (10)$$

综上, 结论得证.

下面的引理将说明算法 HZW 具有充分下降性.

引理 2 若目标函数 $f(\mathbf{x})$ 是光滑的. 设 $\{\mathbf{g}_k\}$ 和 $\{\mathbf{d}_k\}$ 是由算法 HZW 生成的序列. 则

$$\mathbf{g}_k^T \mathbf{d}_k \leqslant -c \| \mathbf{g}_k \|^2 \quad (11)$$

对任意的 $k \geqslant 1$ 都成立, 其中 c 和 u 均为参数, 且 $c = 1 - \frac{1}{u}$, $u > 1$.

证 由 β_k^{HZW} 的定义和(9)式, 可知

$$0 \leqslant \beta_k^{\text{HZW}} = \frac{\| \mathbf{g}_k \|^2 - \frac{\| \mathbf{g}_k \|^2}{\| \mathbf{g}_{k-1} \|^2} \max\{0, \mathbf{g}_k^T \mathbf{g}_{k-1}\}}{\max\{u | \mathbf{g}_k^T \mathbf{d}_{k-1} | + \| \mathbf{g}_{k-1} \|^2, \mathbf{d}_{k-1}^T \mathbf{y}_{k-1}\}} \leqslant \frac{\| \mathbf{g}_k \|^2}{u | \mathbf{g}_k^T \mathbf{d}_{k-1} |} \quad (12)$$

结合(3)式, 有

$$\mathbf{g}_k^T \mathbf{d}_k \leqslant -\| \mathbf{g}_k \|^2 + \beta_k^{\text{HZW}} | \mathbf{g}_k^T \mathbf{d}_{k-1} | \leqslant -\| \mathbf{g}_k \|^2 + \frac{\| \mathbf{g}_k \|^2}{u | \mathbf{g}_k^T \mathbf{d}_{k-1} |} | \mathbf{g}_k^T \mathbf{d}_{k-1} | = -c \| \mathbf{g}_k \|^2$$

从而, 结论得证.

2 全局收敛性

为了证明算法 HZW 的全局收敛性, 下面给出一些标准假设:

假设 A 目标函数 $f(\mathbf{x})$ 的水平集 $\mathcal{L} = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \leqslant f(\mathbf{x}_1)\}$ 有界, 其中 $\mathbf{x}_1$ 为所给的初始点.

假设 B 在 $\mathcal{L}$ 的某个邻域 $\mathcal{N}$ 内, 目标函数 $f(\mathbf{x})$ 在 $\mathcal{N}$ 上连续可微且导函数 $\mathbf{g}(\mathbf{x})$ 在 $\mathcal{N}$ 上 Lipschitz 连续, 即存在常数 $L > 0$, 使得

$$\| \mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{y}) \| \leqslant L \| \mathbf{x} - \mathbf{y} \|, \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{N}$$

接下来给出著名的 Zoutendijk 条件. 该引理经常被用来证明共轭梯度算法的全局收敛性.

引理 3^[12] 假设 A 和 B 都成立. 考虑形如(2)式和(3)式的任意迭代算法, 若搜索方向 $\mathbf{d}_k$ 是一个下降方向, 步长 α_k 满足 Wolfe 线搜索条件(4), 则

$$\sum_{k=1}^{+\infty} \frac{(\mathbf{g}_k^T \mathbf{d}_k)^2}{\| \mathbf{d}_k \|^2} < +\infty \quad (13)$$

根据(11)式, 不难得到与 Zoutendijk 条件(13)等价的不等式

$$\sum_{k=1}^{+\infty} \frac{\| \mathbf{g}_k \|^4}{\| \mathbf{d}_k \|^2} < +\infty \quad (14)$$

现在, 我们建立算法 HZW 的全局收敛定理.

定理 1 假设 A 和 B 都成立. $\{\mathbf{g}_k\}$ 和 $\{\mathbf{d}_k\}$ 是由算法 HZW 生成的序列, 则

$$\liminf_{k \rightarrow +\infty} \| \mathbf{g}_k \| = 0 \quad (15)$$

证 假设(15)式不成立, 即存在常数 $r > 0$ 以及无穷多个 k , 有

$$\| \mathbf{g}_k \| \geqslant r \quad (16)$$

在(3)式中, 令 $\beta_k = \beta_k^{\text{HZW}}$. 当 $k \geq 2$ 时, 对(3)式两边取模的平方, 可得

$$\| \mathbf{d}_k \|^2 = \| \mathbf{g}_k \|^2 + (\beta_k^{\text{HZW}})^2 \| \mathbf{d}_{k-1} \|^2 - 2\beta_k^{\text{HZW}} \mathbf{g}_k^T \mathbf{d}_{k-1} \quad (17)$$

由 β_k^{HZW} 的定义和(12)式, 可知

$$-2\beta_k^{\text{HZW}} \mathbf{g}_k^T \mathbf{d}_{k-1} \leq 2\beta_k^{\text{HZW}} | \mathbf{g}_k^T \mathbf{d}_{k-1} | \leq 2 \frac{\| \mathbf{g}_k \|^2}{u} \quad (18)$$

将(10)式和(18)式应用到(17)式中, 有

$$\begin{aligned} \| \mathbf{d}_k \|^2 &\leq \| \mathbf{g}_k \|^2 + \frac{\| \mathbf{g}_k \|^4}{\| \mathbf{g}_{k-1} \|^4} \| \mathbf{d}_{k-1} \|^2 + 2 \frac{\| \mathbf{g}_k \|^2}{u} = \\ &\frac{\| \mathbf{g}_k \|^4}{\| \mathbf{g}_{k-1} \|^4} \| \mathbf{d}_{k-1} \|^2 + \frac{u+2}{u} \| \mathbf{g}_k \|^2 \end{aligned} \quad (19)$$

(19)式两边同时除以 $\| \mathbf{g}_k \|^4$, 可得

$$\frac{\| \mathbf{d}_k \|^2}{\| \mathbf{g}_k \|^4} \leq \frac{\| \mathbf{d}_{k-1} \|^2}{\| \mathbf{g}_{k-1} \|^4} + \frac{u+2}{u} \frac{1}{\| \mathbf{g}_k \|^2} \quad (20)$$

对(20)式, 使用递归法. 并注意到

$$\| \mathbf{d}_1 \|^2 = -\mathbf{g}_1^T \mathbf{d}_1 = \| \mathbf{g}_1 \|^2$$

就有

$$\frac{\| \mathbf{d}_k \|^2}{\| \mathbf{g}_k \|^4} \leq \frac{u+2}{u} \sum_{j=1}^k \frac{1}{\| \mathbf{g}_j \|^2} \quad (21)$$

结合(16)式和(21)式, 有

$$\frac{\| \mathbf{g}_k \|^4}{\| \mathbf{d}_k \|^2} \geq \frac{r^2}{k} \frac{u}{u+2} \quad (22)$$

(22)式说明

$$\sum_{k=1}^{+\infty} \frac{\| \mathbf{g}_k \|^4}{\| \mathbf{d}_k \|^2} = +\infty$$

这与(14)式相矛盾, 也即与引理 3 矛盾. 因此, 原结论成立.

3 数值实验

为了清楚的检验算法 HZW 的实际计算效果利用文献[13]中的 11 个测试函数进行数值测验. 现将要比较的算法, 一一列举如下:

DY: 戴或虹和袁亚湘^[2]提出的经典算法, 共轭参数 $\beta_k = \beta_k^{\text{DY}}$;

YWH: 文献[9]引进的算法, 共轭参数 $\beta_k = \beta_k^{\text{YWH}}$ 由(6)式给出;

DPRP: 文献[10]定义的算法, 共轭参数 $\beta_k = \beta_k^{\text{DPRP}}$ 由(7)式给出;

HZW: 本文所提出的算法, 共轭参数 $\beta_k = \beta_k^{\text{HZW}}$ 由(8)式确定.

实验采用 Wolfe 线搜索条件(4), 并且均利用 MATLAB 编程实现. 算法测试的环境是 Windows 7 操作系统, AMD Athlon(tm) II Dual-Core M320 2.00 GB 内存. 算法中参数取值为: $u=1.1$, $\delta=0.0001$, $\sigma=0.1$. 终止条件为: $\| \mathbf{g}_k \| \leq 10^{-7}$, 或者算法的迭代时间超过 3 600 s. 将 HZW 算法与 DPRP, YWH 以及 DY 法进行比较, 20 个算例详细报告见表 1. 在表 1 中: F 为测试函数的名称; M 为共轭梯度算法的名称; n 为测试函数的维数; I 为算法解决某一测试问题的迭代次数; T 为算法解决某一测试问题的计算时间; N_g 为算法解决某一测试问题的最终梯度范数, 即 $\| \mathbf{g}_k \|$; S 为若算法的终止条件满足, 令 $S=1$, 否则, 令 $S=0$.

表 1 的测试结果表明, 算法 HZW 的计算性能比另外 3 种算法要好得多. 另外, DPRP 法在求解 120 维 dixmaane 问题与 dixmaang 问题时迭代失败, DY 法在求解 120 维 dixmaang 问题和 1 000 维 raydan2 问题时迭代失败, 但 HZW 法和 YWH 法均迭代成功.

表 1 数值结果

F	M	n	I	T/s	N_g	S
almost perturbed quadratic	HZW	200	133	1 287.3296	5.043 9e-07	1
	DPRP	200	218	2 137.986 1	8.406 9e-07	1
	YWH	200	140	1 380.108 6	9.525 6e-07	1
	DY	200	143	1 376.910 3	9.474 7e-07	1
almost perturbed quadratic	HZW	300	139	2 300.400 1	6.697 3e-07	1
	DPRP	300	185	3 099.127 5	6.394 6e-07	1
	YWH	300	187	3 207.550 9	9.708 4e-07	1
	DY	300	156	2 666.627 2	9.011 7e-07	1
bdexp	HZW	50	5	401.361 0	8.994 9e-07	1
	DPRP	50	5	409.253 5	8.994 9e-07	1
	YWH	50	5	414.313 9	9.068 9e-07	1
	DY	50	5	407.241 4	4.323 2e-07	1
bdexp	HZW	100	6	1 538.670 6	6.782 3e-08	1
	DPRP	100	6	1 543.582 1	6.782 3e-08	1
	YWH	100	6	1 549.907 9	6.696 4e-08	1
	DY	100	6	1 544.107 6	5.174 5e-08	1
diagonal 2	HZW	50	46	229.907 7	8.224 0e-07	1
	DPRP	50	65	330.462 3	6.206 8e-07	1
	YWH	50	51	256.398 9	8.203 8e-07	1
	DY	50	85	430.923 1	8.724 5e-07	1
diagonal 2	HZW	100	72	1 180.823 7	6.570 0e-07	1
	DPRP	100	107	1 808.820 1	6.228 3e-07	1
	YWH	100	77	1 265.915 5	7.555 6e-07	1
	DY	100	114	1 913.785 1	8.984 1e-07	1
diagonal 7	HZW	50	11	32.708 5	2.448 0e-07	1
	DPRP	50	36	109.504 3	4.067 3e-07	1
	YWH	50	13	37.897 5	9.442 5e-07	1
	DY	50	75	214.572 5	8.426 7e-08	1
diagonal 7	HZW	100	11	75.330 9	3.462 1e-07	1
	DPRP	100	35	246.199 9	8.479 5e-07	1
	YWH	100	14	94.927 7	9.351 7e-07	1
	DY	100	71	471.018 7	2.489 4e-07	1
dixmaane	HZW	45	49	193.611 2	9.374 0e-07	1
	DPRP	45	168	621.496 1	9.600 3e-07	1
	YWH	45	56	214.102 9	8.991 7e-07	1
	DY	45	65	238.271 9	9.623 8e-07	1
dixmaane	HZW	120	86	1 707.232 9	8.872 7e-07	1
	DPRP	120	186	3 608.628 5	1.366 6e-05	0
	YWH	120	87	1 713.985 7	7.833 6e-07	1
	DY	120	107	2 026.692 2	9.359 3e-07	1
dixmaang	HZW	51	55	688.351 1	5.922 9e-07	1
	DPRP	51	156	1 816.333 1	7.589 6e-07	1
	YWH	51	61	751.846 7	9.787 7e-07	1
	DY	51	63	741.821 7	9.742 9e-07	1

续表 1

<i>F</i>	<i>M</i>	<i>n</i>	<i>I</i>	<i>T</i> /s	<i>N_g</i>	<i>S</i>
dixmaang	HZW	120	80	3 478.780 1	7.269 6e-07	1
	DPRP	120	86	3 641.523 9	1.300 0e-03	0
	YWH	120	79	3 577.304 2	7.2217e-07	1
	DY	120	88	3 609.186 8	1.559 6e-05	0
dqdrtic	HZW	50	90	208.688 1	4.627 6e-07	1
	DPRP	50	365	872.746 3	7.603 0e-07	1
	YWH	50	123	293.283 4	7.288 9e-07	1
	DY	50	216	542.869 3	9.926 9e-07	1
dqdrtic	HZW	150	88	701.850 9	5.379 4e-07	1
	DPRP	150	309	2 495.1985	7.940 7e-07	1
	YWH	150	115	917.457 2	9.064 1e-07	1
	DY	150	111	894.757 2	4.201 2e-07	1
himmelbg	HZW	100	4	272.922 8	9.325 4e-07	1
	DPRP	100	6	449.434 1	1.446 6e-07	1
	YWH	100	6	444.254 6	3.667 6e-07	1
	DY	100	4	277.838 4	5.134 2e-07	1
himmelbg	HZW	500	5	2 737.837 1	1.196 4e-07	1
	DPRP	500	6	3 347.309 4	3.234 6e-07	1
	YWH	500	6	3 326.077 1	8.200 9e-08	1
	DY	500	5	2 755.178 9	3.718 9e-08	1
liarwhd	HZW	20	40	167.762 9	8.287 9e-07	1
	DPRP	20	79	342.065 5	5.603 9e-07	1
	YWH	20	52	210.689 8	8.837 7e-07	1
	DY	20	108	466.299 1	3.989 3e-07	1
raydan 1	HZW	60	72	327.976 1	9.502 9e-07	1
	DPRP	60	91	436.074 8	7.375 8e-07	1
	YWH	60	98	513.077 9	6.147 0e-07	1
	DYW	60	73	339.826 4	9.975 7e-07	1
raydan 2	HZW	200	5	61.820 7	5.674 8e-07	1
	DPRP	200	6	75.304 7	5.657 0e-12	1
	YWH	200	6	79.011 4	3.245 3e-09	1
	DY	200	8	99.840 1	9.307 3e-07	1
raydan 2	HZW	1 000	6	2 951.566 3	2.457 6e-14	1
	DPRP	1 000	6	2 991.777 1	1.264 6e-11	1
	YWH	1 000	6	3 004.285 7	7.256 8e-09	1
	DY	1 000	8	3 976.370 8	2.081 2e-06	0

4 结 论

本文构造的算法 HZW，在任何线搜索条件下，都具有充分下降性. 采用 Wolfe 线搜索时，算法 HZW 具有良好的全局收敛性，且在实际计算中，具有出色的数值表现.

参考文献：

[1] FLETCHER R, REEVES C. Function Minimization by Conjugate Gradients [J]. Computer Journal, 1964, 7(2): 149—154.

- [2] DAI Y H, YUAN Y X. A Nonlinear Conjugate Gradient Method with a Strong Global Convergence Property [J]. SIAM Journal of Optimization, 2000, 10(1): 177—182.
- [3] FLETCHER R. Practical Methods of Optimization, Unconstrained Optimization [M]. New York: John Wiley and Sons, 1987.
- [4] POLAK E, RIBIÉRE G. Note Sur La Convergence De Directions Conjuguée [J]. Rev Francaise Infomat Recherche Operationelle 3e Année, 1969, 16(16): 35—43.
- [5] POLIAK B T. The Conjugate Gradient Method in Extreme Problems [J]. USSR Computational Mathematics and Mathematical Physics, 1969, 9(4): 94—112.
- [6] HESTENES M R, STIEFEL E L. Methods of Conjugate Gradients for Solving Linear Systems [J]. Journal of Research of National Bureau of Standards, 1952, 49(6): 409—436.
- [7] LIU Y, STOREY C. Efficient Generalized Conjugate Gradient Algorithms [J]. Journal of Optimization Theory Application, 1991, 69(1): 129—137.
- [8] WEI Z, LI G, QI L. New Nonlinear Conjugate Gradient Formulas for Large-Scale Unconstrained Optimization Problems [J]. Applied Mathematics and Computation, 2006, 179(2): 407—430.
- [9] YAO S W, WEI Z X, HUANG H. A Note about WYL's Conjugate Gradient and Its Application [J]. Applied Mathematics and Computation, 2007, 191(2): 381—388.
- [10] DAI Z F, WEN F H. Another Improved Wei-Yao-Liu Nonlinear Conjugate Gradient Method with Sufficient Descent Property [J]. Applied Mathematics and Computation, 2012, 218(14): 7421—7430.
- [11] JIAN J B, HAN L, JIANG X Z. A Hybrid Conjugate Gradient Method with Descent Property for Unconstrained Optimization [J]. Applied Mathematical Modelling, 2014, 39(3—4): 1281—1290.
- [12] GILBERT J C, NOCEDAL J. Global Convergence Properties of Conjugate Gradient Methods for Optimization [J]. SIAM Journal of Optimization, 1992, 2(1): 21—42.
- [13] ANDREI N. An Unconstrained Optimization Test Functions Collection [J]. Advanced Modeling and Optimization, 2008, 10(1): 147—161.

A New Hybrid Conjugate Gradient Algorithm

HAN Xin, ZHANG Jun-rong, WANG Sen-sen

School of Mathematics and Statistics, Southwest University, Chongqing 400715, China

Abstract: Based on some previous conjugate gradient algorithms, a new hybrid conjugate gradient algorithm is put forward for solving unconstrained optimization problems. This algorithm always generates a sufficient descent direction at each iteration. Furthermore, we show that the proposed algorithm possesses global convergence under the Wolfe line search. Finally, some numerical results are reported, which also demonstrate that our algorithm possesses good computational performance.

Key words: unconstrained optimization; hybrid conjugate gradient method; sufficient descent; global convergence

责任编辑 张 枸

