DOI:10.13718/j.cnki.xsxb.2022.01.001

一种具有充分下降性的新混合型共轭梯度法[①]

郑宗剑[1]，　韩信[1,2]

1. 四川文理学院 数学学院，四川 达州 635000；2. 西南大学 电子信息工程学院，重庆 400715

摘要：对 PRP 法和 FR 法进行凸组合，提出了一种求解无约束优化问题的新共轭梯度法. 该方法总是能生成一个充分下降方向，且它的凸组合参数为 Babaie-Kafaki 和 Ghanbari 的推广形式. 在 Wolfe 线搜索条件下，新算法的全局收敛性得以建立，数值结果也说明提出的算法是有效的.

关　键　词：无约束优化；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**：1000-5471(2022)01-0001-07

A New Hybrid Conjugate Gradient Method with Sufficient Descent Property

ZHENG Zongjian[1]，　HAN Xin[1,2]

1. School of Mathematics, Sichuan University of Arts and Science, Dazhou Sichuan 635000, China;

2. College of Electronic and Information Engineering, Southwest University, Chongqing 400715, China

Abstract: In this paper, a new hybrid conjugate gradient algorithm has been developed for solving unconstrained optimization problems. The proposed method can be viewed as a convex combination of Polak-Ribière-Polyak method and Fletcher-Reeves method. In this method, a sufficient descent direction can always be generated, with its convex combination parameter as an extension of the parameter proposed by Babaie-Kafaki and Ghanbari. Under the Wolfe line search, the global convergence has been established. Some numerical results show that proposed method is effective.

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

本文考虑下面的无约束优化问题：

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

这里 $f: \mathbb{R}^n \longrightarrow \mathbb{R}$ 是一个光滑函数，$\mathbb{R}^n$ 表示 n 维欧氏空间，函数 f 的梯度记为 $\boldsymbol{g}(\boldsymbol{x}) := \nabla f(\boldsymbol{x})$. 在实际求

① 收稿日期：2020-03-08

基金项目：国家自然科学基金项目(11701470)；四川文理学院 2019 年度校级科研项目学科专业群发展研究专项重点资助项目(2019XKQ001)；四川文理学院 2017 年度科研基金自然科学研究一般项目(2017KZ010Y).

作者简介：郑宗剑，副教授，主要从事最优化理论与算法、动力系统分支理论与应用研究.

通信作者：韩信，助教.

解大规模无约束优化问题中，建立一些具有低存储的快速数值算法是十分重要的. 共轭梯度法是求解大型无约束优化问题(1) 的一类十分有效的方法[1]，其通常的迭代形式如下

$$\boldsymbol{x}_{k+1} := \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k \tag{2}$$

这里 $\boldsymbol{x}_k \in \mathbb{R}^n$ 是一个解的第 k 次逼近，$\alpha_k > 0$ 是由一种合适的线搜索所确定的步长，$\boldsymbol{d}_k \in \mathbb{R}^n$ 是搜索方向，其定义如下

$$\boldsymbol{d}_k := \begin{cases} -\boldsymbol{g}_k & k = 0 \\ -\boldsymbol{g}_k + \beta_k \boldsymbol{d}_{k-1} & k > 0 \end{cases} \tag{3}$$

其中：$\boldsymbol{g}_k$ 表示梯度 $\boldsymbol{g}(\boldsymbol{x}_k)$，$\beta_k \in \mathbb{R}$ 为共轭参数. 步长 α_k 由标准的 Wolfe 线搜索

$$\begin{cases} f(\boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k) \leqslant f(\boldsymbol{x}_k) + \delta \alpha_k \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{d}_k \\ \boldsymbol{g}(\boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k) T \boldsymbol{d}_k \geqslant \sigma \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{d}_k \end{cases} \tag{4}$$

或者强 Wolfe 线搜索

$$\begin{cases} f(\boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k) \leqslant f(\boldsymbol{x}_k) + \delta \alpha_k \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{d}_k \\ | \boldsymbol{g}(\boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k)^{\mathrm{T}} \boldsymbol{d}_k | \leqslant \sigma | \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{d}_k | \end{cases}$$

确定，其中 $0 < \delta < \sigma < 1$. 不同的共轭梯度法对应不同的共轭参数 β_k. 众所周知的共轭梯度法包括 Hestenes-Stiefel(HS)[2]，Polak-Ribière-Polyak(PRP)[3-4]，Dai-Yuan(DY)[5]，Liu-Storey(LS)[6]，Fletcher-Reeves(FR)[7] 和 Conjugate-Descent(CD)[8]，它们所对应的共轭参数如下

$$\beta_k^{\mathrm{HS}} := \frac{\boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{y}_{k-1}}{\boldsymbol{d}_{k-1}^{\mathrm{T}} \boldsymbol{y}_{k-1}}, \quad \beta_k^{\mathrm{PRP}} := \frac{\boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{y}_{k-1}}{\| \boldsymbol{g}_{k-1} \|^2}, \quad \beta_k^{\mathrm{DY}} := \frac{\| \boldsymbol{g}_k \|^2}{\boldsymbol{d}_{k-1}^{\mathrm{T}} \boldsymbol{y}_{k-1}}$$

$$\beta_k^{\mathrm{LS}} := -\frac{\boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{y}_{k-1}}{\boldsymbol{d}_{k-1}^{\mathrm{T}} \boldsymbol{g}_{k-1}}, \quad \beta_k^{\mathrm{FR}} := \frac{\| \boldsymbol{g}_k \|^2}{\| \boldsymbol{g}_{k-1} \|^2}, \quad \beta_k^{\mathrm{CD}} := -\frac{\| \boldsymbol{g}_k \|^2}{\boldsymbol{d}_{k-1}^{\mathrm{T}} \boldsymbol{g}_{k-1}}$$

其中 $\boldsymbol{y}_{k-1} := \boldsymbol{g}_k - \boldsymbol{g}_{k-1}$ 且 $\| \cdot \|$ 表示欧氏范数. 目标函数是一个严格凸二次函数并且步长由精确线搜索得到，且在一般情形下，它们的理论性质和数值表现不尽相同. 众所周知，DY 法、CD 法和 FR 法具有良好的收敛性质，但它们的数值计算效果一般；相反，PRP 法、HS 法和 LS 法具备出色的数值表现，但它们可能不收敛. 为了获得收敛性数值效果的算法，不少学者对共轭参数自身作了一些改进[9-10] 和一些混合[11-12]. 本文主要考虑共轭梯度法的混合形式. 文献[13] 通过对 PRP 法和 DY 法进行凸组合，得到一种新的共轭梯度法，其共轭参数如下

$$\beta_k^{\mathrm{N}} := (1 - \theta_k) \beta_k^{\mathrm{PRP}} + \theta_k \beta_k^{\mathrm{DY}}$$

其中参数 $\theta_k \in [0, 1]$ 由共轭条件 $\boldsymbol{y}_{k-1}^{\mathrm{T}} \boldsymbol{d}_k = 0$ 确定. 最近，文献[14] 对 PRP 法和 FR 法进行凸组合，引进了一种混合共轭梯度法：

$$\beta_k^{\mathrm{HCG}} := (1 - \theta_k) \beta_k^{\mathrm{PRP}} + \theta_k \beta_k^{\mathrm{FR}}$$

这里参数 $\theta_k \in [0, 1]$ 根据求解优化问题 $\min\limits_{\theta_k} \{ \| \boldsymbol{d}_k^{HCG} - \boldsymbol{d}_k^{ZZL} \|^2 : k \geqslant 1 \}$ 得到. 在标准的 Wolfe 线搜索条件下，该算法具有全局收敛性和良好的计算表现.

受文献[14] 的启发，本文构造了一种新的混合共轭梯度法

$$\beta_k^{\mathrm{NH+}} := (1 - \theta_k) \beta_k^{\mathrm{PRP}} + \theta_k \beta_k^{\mathrm{FR}} \tag{5}$$

这里

$$\theta_k = \begin{cases} \theta_k^* & \text{若 } \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{g}_{k-1} \cdot \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{p}_k \neq 0 \text{ 且 } \theta_k^* \in [0, 1] \\ 0 & \text{若 } \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{g}_{k-1} \cdot \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{p}_k = 0 \text{ 或 } \theta_k^* < 0 \\ 1 & \text{若 } \theta_k^* > 1 \end{cases}$$

其中

$$\theta_k^* = -\frac{\boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{d}_{k-1}}{\boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{g}_{k-1}} \frac{p_k^{\mathrm{T}} \boldsymbol{d}_{k-1}}{\| \boldsymbol{d}_{k-1} \|^2} \frac{\boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{y}_{k-1}}{\boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{p}_k}$$

为优化问题$\min\limits_{\theta_k}\{\|\boldsymbol{d}_k^{\mathrm{NH+}}-\boldsymbol{d}_k^{\mathrm{NYF}}\|^2: k\geqslant 1\}$的最优解. 基于共轭参数$\beta_k^{\mathrm{NH+}}$所得相应的共轭梯度法称作NH＋算法或NH＋法. 因为文献[15]提出的ZZL法是文献[16]提出的NYF法特殊情况. 因此，文献[14]提出的方法是NH＋法的特殊情况(当$\boldsymbol{p}_k=\boldsymbol{g}_k$即为文献[14]中的方法).

1　NH＋型共轭梯度算法

基于共轭参数$\beta_k^{\mathrm{NH+}}$的计算公式(5)，所得的NH＋算法如下：

NH＋算法

步骤1：给定初始点$\boldsymbol{x}_1$和精度ε. 置$k=1$，令$\boldsymbol{d}_1=-\boldsymbol{g}_1$.

步骤2：若$\|\boldsymbol{g}_k\|\leqslant\varepsilon$，终止. 否则，转步骤3.

步骤3：通过(3)式计算搜索方向$\boldsymbol{d}_k$.

步骤4：由标准的Wolfe线搜索(4)确定步长因子α_k.

步骤5：由(2)式计算下一个迭代点$\boldsymbol{x}_{k+1}$.

步骤6：令$k:=k+1$，转步骤2.

2　算法NH＋的全局收敛性分析

为了论证算法NH＋的收敛性，给出了如下的基本假设.

假设1　水平集$S=\{\boldsymbol{x}: f(\boldsymbol{x})\leqslant f(\boldsymbol{x}_0)\}$有界，即存在$\tilde{a}>0$，对于任意$\boldsymbol{x}\in S$，有$\|\boldsymbol{x}\|\leqslant\tilde{a}$.

目标函数的梯度函数在S的某邻域U内Lipschitz连续，即存在常数$L>0$，对任意$\boldsymbol{x},\boldsymbol{y}\in U$，有$\|\boldsymbol{g}(\boldsymbol{x})-\boldsymbol{g}(\boldsymbol{y})\|\leqslant L\|\boldsymbol{x}-\boldsymbol{y}\|$.

由假设可知存在常数$\bar{\gamma}>0$，对任意$\boldsymbol{x}\in S$，有

$$\|\boldsymbol{g}(\boldsymbol{x})\|\leqslant\bar{\gamma} \tag{6}$$

以Zoutendijk条件为基础给出如下引理1. 这个引理在论证共轭梯度算法的全局收敛性过程中具有重要的作用.

引理1　假设1成立，算法NH＋的搜索方向是下降的，步长满足标准Wolfe线搜索条件，则Zoutendijk条件即$\sum_{k=0}^{\infty}\dfrac{(\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k)^2}{\|\boldsymbol{d}_k\|^2}<\infty$成立.

证　由标准Wolfe线搜索条件(4)，知

$$(\boldsymbol{g}_k-\boldsymbol{g}_{k-1})^{\mathrm{T}}\boldsymbol{d}_{k-1}\geqslant(\sigma-1)\boldsymbol{g}_{k-1}^{\mathrm{T}}\boldsymbol{d}_{k-1}$$

另一方面，由假设1和(2)式，有

$$(\boldsymbol{g}_k-\boldsymbol{g}_{k-1})^{\mathrm{T}}\boldsymbol{d}_{k-1}\leqslant\alpha_{k-1}L\|\boldsymbol{d}_{k-1}\|^2$$

再结合算法NH＋的下降性，得

$$\alpha_{k-1}\geqslant\frac{\sigma-1}{L}\cdot\frac{\boldsymbol{g}_{k-1}^{\mathrm{T}}\boldsymbol{d}_{k-1}}{\|\boldsymbol{d}_{k-1}\|^2}>0$$

进一步由(4)式知

$$f(\boldsymbol{x}_{k-1})-f(\boldsymbol{x}_k)\geqslant\frac{\delta(1-\sigma)}{L}\cdot\frac{(\boldsymbol{g}_{k-1}^{\mathrm{T}}\boldsymbol{d}_{k-1})^2}{\|\boldsymbol{d}_{k-1}\|^2}>0 \tag{7}$$

对(7)式从$k=1,2,\cdots,\infty$求和，并注意目标函数$f(\boldsymbol{x})$有下界，即知Zoutendijk条件成立. 因此，结论得证.

下面的引理是论证算法NH＋全局收敛性的关键. 这个引理的具体论证过程可参见文献[17].

引理2[17]　假设1、充分下降性、标准Wolfe线搜索条件均成立. 若$\sum_{k=0}^{\infty}\dfrac{1}{\|\boldsymbol{d}_k\|^2}=\infty$，则

$\liminf\limits_{k\to\infty}\|\boldsymbol{g}_k\|=0$.

注 通常利用的是引理2的逆否命题，即若$\liminf\limits_{k\to\infty}\|\boldsymbol{g}_k\|\neq 0$，则$\sum_{k=0}^{\infty}\frac{1}{\|\boldsymbol{d}_k\|^2}<\infty$. 对于共轭梯度算法，文献[18]给出了性质(*)，当$s_k:=\boldsymbol{x}_{k+1}-\boldsymbol{x}_k$变小时，$\beta_k$也将变小.

性质(*)[18] 对于共轭梯度算法，若任意$k\geqslant 0$，都有$0<\gamma\leqslant\|\boldsymbol{g}_k\|$. 而且，若常数$b>1$，$\lambda>0$，使得$|\beta_k|\leqslant b$和$\|\boldsymbol{s}_k\|\leqslant\lambda\Rightarrow|\beta_k|\leqslant\frac{1}{2b}$都成立，就称算法具有性质(*).

接下来，假设全局收敛性不在有限步发生.

引理3 若假设1、充分下降性、标准Wolfe线搜索条件均成立. 如果性质(*)成立，则$\boldsymbol{d}_k\neq 0$且$\sum_{k=1}^{\infty}\|\boldsymbol{u}_k-\boldsymbol{u}_{k-1}\|^2<\infty$，其中$\boldsymbol{u}_k:=\frac{\boldsymbol{d}_k}{\|\boldsymbol{d}_k\|}$.

证 由$\boldsymbol{d}_k\neq 0$，$\boldsymbol{g}_k\neq 0$及充分下降性知，$\boldsymbol{u}_k$的定义有意义.

现定义$\boldsymbol{\rho}_k:=\frac{-\boldsymbol{g}_k}{\|\boldsymbol{d}_k\|}$和以下式子

$$\omega_k:=\frac{\beta_k^{\mathrm{NH+}}\|\boldsymbol{d}_{k-1}\|}{\|\boldsymbol{d}_k\|}\tag{8}$$

由(3)式和(5)式，对任意$k\geqslant 1$，有

$$\boldsymbol{u}_k=\boldsymbol{\rho}_k+\omega_k\boldsymbol{u}_{k-1}\tag{9}$$

基于$\|\boldsymbol{u}_k\|=\|\boldsymbol{u}_{k-1}\|=1$和(9)式，得

$$\|\boldsymbol{\rho}_k\|=\|\boldsymbol{u}_k-\omega_k\boldsymbol{u}_{k-1}\|=\|\omega_k\boldsymbol{u}_k-\boldsymbol{u}_{k-1}\|\tag{10}$$

由于$\beta_k^{\mathrm{NH+}}\geqslant 0$，则(8)式中的$\omega_k\geqslant 0$. 此外，利用三角不等性和(10)式，可得

$$\begin{aligned}\|\boldsymbol{u}_k-\boldsymbol{u}_{k-1}\|\leqslant(1+\omega_k)\|\boldsymbol{u}_k-\boldsymbol{u}_{k-1}\|=\\ \|\boldsymbol{u}_k-\omega_k\boldsymbol{u}_{k-1}+\omega_k\boldsymbol{u}_k-\boldsymbol{u}_{k-1}\|\leqslant\\ \|\boldsymbol{u}_k-\omega_k\boldsymbol{u}_{k-1}\|+\|\omega_k\boldsymbol{u}_k-\boldsymbol{u}_{k-1}\|=\\ 2\|\boldsymbol{\rho}_k\|\end{aligned}\tag{11}$$

由Zoutendijk条件、充分下降条件和(8)式知

$$\sum_{k=1}^{\infty}\frac{\|\boldsymbol{g}_k\|^4}{\|\boldsymbol{d}_k\|^2}=\sum_{k=1}^{\infty}\|\boldsymbol{g}_k\|^2\|\boldsymbol{\rho}_k\|^2<\infty\tag{12}$$

利用性质(*)、(11)式和(12)式，有$\sum_{k=1}^{\infty}\|\boldsymbol{u}_k-\boldsymbol{u}_{k-1}\|^2\leqslant 4\sum_{k=1}^{\infty}\|\boldsymbol{\rho}_k\|^2<\infty$，结论得证.

现令$\mathbb{N}_+$为非零自然数集. 给定正常数λ和正整数Δ，定义

$$K_{k,\Delta}^{\lambda}:=\{i\in\mathbb{N}_+:k\leqslant i\leqslant k+\Delta-1,\ \|\boldsymbol{s}_{i-1}\|>\lambda\}$$

其中$|K_{k,\Delta}^{\lambda}|$表示$K_{k,\Delta}^{\lambda}$的元素个数.

引理4 若引理3的所有假设条件都成立. 算法NH+满足性质(*)，则存在常数λ和k_0，对任意$\Delta\in\mathbb{N}_+$和$\tilde{k}\geqslant k_0$，使得

$$|K_{\tilde{k},\Delta}^{\lambda}|>\frac{\Delta}{2}$$

证 利用反证法论证. 假设存在$\Delta\in\mathbb{N}_+$和常数k_0，对任意正常数λ和任意$k\geqslant k_0$，有

$$|K_{k,\Delta}^{\lambda}|\leqslant\frac{\Delta}{2}\tag{13}$$

$b>1$和$\eta>0$是性质(*)中的常数. 令$\lambda=\eta$，由性质(*)和(13)式，有

$$\prod_{k=k_0+i\Delta+1}^{k_0+(i+1)\Delta}|\beta_k|=\prod_{k\in K_{k_0+i\Delta+1,\Delta}^{\lambda}}|\beta_k|\prod_{k\notin K_{k_0+i\Delta+1,\Delta}^{\lambda}}|\beta_k|\leqslant b^{\frac{\Delta}{2}}\left(\frac{1}{2b}\right)^{\frac{\Delta}{2}},\ \forall i\geqslant 0\tag{14}$$

因此

$$\prod_{j=k_0+1}^{k_0+i\Delta} 2\beta_j^2 = \prod_{j=k_0+1}^{k_0+\Delta} 2\beta_j^2 \cdots \prod_{j=k_0+(i-1)\Delta+1}^{k_0+i\Delta} 2\beta_j^2 \leqslant$$

$$2^{i\Delta}\left[\left(\frac{1}{2b}\right)^2\right]^{i\Delta} = \left(\frac{1}{2b^2}\right)^{i\Delta} \leqslant 1 \tag{15}$$

对任意 $i \geqslant 1$，$k_0 \leqslant l \leqslant k_0 + i\Delta$，存在 i'，使得

$$k_0 + i'\Delta \leqslant l \leqslant k_0 + (i'+1)\Delta \leqslant k_0 + i\Delta$$

$$\prod_{j=l}^{k_0+i\Delta} 2\beta_j^2 = \prod_{j=l}^{k_0+(i'+1)\Delta} 2\beta_j^2 \prod_{j=k_0+(i'+1)\Delta+1}^{k_0+(i'+2)\Delta} 2\beta_j^2 \cdots \prod_{j=k_0+(i-1)\Delta+1}^{k_0+i\Delta} 2\beta_j^2$$

再由 $b > 1$ 和(14) 式得

$$\prod_{j=l}^{k_0+i\Delta} 2\beta_j^2 \leqslant c_1 \tag{16}$$

其中 $c_1 = (2b^2)^{\Delta+1}$.

通过性质(*)、(6) 式、(15) 式和(16) 式，有

$$\begin{aligned}\|\boldsymbol{d}_{k_0+i\Delta}\|^2 &\leqslant (\|\boldsymbol{g}_{k_0+i\Delta}\| + |\beta_{k_0+i\Delta}|\ \|\boldsymbol{d}_{k_0+i\Delta-1}\|)^2 \leqslant \\ &2(\|\boldsymbol{g}_{k_0+i\Delta}\|^2 + |\beta_{k_0+i\Delta}|^2\|\boldsymbol{d}_{k_0+i\Delta-1}\|^2) \leqslant \\ &2\bar{\gamma}^2 + 2b^2\sum_{l=k_0+2}^{k_0+i\Delta}\Big(\prod_{j=l}^{k_0+i\Delta} 2\beta_j^2\Big) + \|\boldsymbol{d}_{k_0}\| \prod_{j=k_0+1}^{k_0+i\Delta} 2\beta_j^2 \leqslant \\ &2\bar{\gamma}^2 + 2b^2\sum_{l=k_0+2}^{k_0+i\Delta}\Big(\prod_{j=l}^{k_0+i\Delta} 2\beta_j^2\Big) + \|\boldsymbol{d}_{k_0}\| \leqslant \\ &2\bar{\gamma}^2 + 2b^2 c_1(i\Delta - 1) + c_2\end{aligned}$$

其中 $c_2 = \|\boldsymbol{d}_{k_0}\|$. 所以，$\sum_{k=0}^{\infty}\frac{1}{\|\boldsymbol{d}_k\|^2} \geqslant \sum_{i=1}^{\infty}\frac{1}{\|\boldsymbol{d}_{k_0+i\Delta}\|^2} \geqslant \sum_{i=1}^{\infty}\frac{1}{2\bar{\gamma}^2 + 2b^2c_1(i\Delta-1)+c_2} = \infty$.

由引理 2 知$\lim\limits_{k\to\infty}\inf\|\boldsymbol{g}_k\| = 0$，这与假设矛盾. 因此，原结论成立.

结合引理 3 和 4，利用文献[14] 的论证方法易得下面的定理 1. 限于篇幅，这里略去证明过程.

定理 1　若假设 1 和标准 Wolfe 线搜索都成立. 如果算法 NH＋满足下面 3 个性质：

(p1) 对任意 $k \geqslant 1$，有 $\beta_k^{\mathrm{NH}+} \geqslant 0$；

(p2) 充分下降性和 Zoutendijk 条件均满足；

(p3) 性质(*)成立，而且存在正常数 M，使得 $\theta_k \leqslant M\|s_{k-1}\|$.

则算法 NH＋全局收敛.

3　数值实验

为了比较算法 NH＋与算法 NYF[16]、算法 HCG＋[14]、算法 PRP＋[18]，对文献[19] 中的 63 个测试函数进行实验. 4 种算法均利用 Matlab 程序实现，且在 Windows 7 操作系统、AMD Athlon(tm) II Dual-Core M320 CPU 和 2 GB 内存环境测试运行. 常数 $\sigma = 0.95$，$\delta = 0.000\ 1$. 令 $\boldsymbol{p}_k = \boldsymbol{g}_k$. 终止准则为 $\|\boldsymbol{g}_k\| \leqslant 10^{-6}$ 或迭代时间超过 3 600 s. 部分数值结果见表 1. 所有数值实验具体结果请参见链接 https://weibo.com/2145331053/IsvzxrnEi?from=page_1005052145331053_profile&wvr=6&mod=weibotime&type=comment.

此外，利用文献[20]提出的性能理论刻画算法的计算效率和稳定性. 为此，以迭代时间、迭代次数为度量，纵轴为性能指标 $P_s(t)$，描绘出下面的性能图(图 1,图 2). 实验结果表明：NYF 和 PRP＋在 1.8 s 前收敛速度都比 NH＋稍慢，之后都与 NH＋接近，并都达到稳定；HCG＋比其他 3 种算法的收敛速度都慢；NH＋一直快于其他 3 种算法，最终与 PRP＋同时稳定. 原因可能是 NH＋法充分利用了 PRP 和 NYF 的加速特性以及 FR 的良好收敛性. 因此，NH＋是一个有效的算法.

表 1　部分数值实验结果

函数名	维度/维	算法	迭代次数/次	迭代时间/s	梯度值
LIARWHD	20	NH+	138	2 137.390	9.37E−07
	20	PRP+	94	1 303.897	9.93E−07
	20	NYF	345	11 117.170	9.02E−06
	20	HCG+	340	3 605.104	1.10E−05
DIAGONAL	50	NH+	179	1 957.943	8.47E−07
	50	PRP+	132	1 292.290	7.51E−07
	50	NYF	708	2 381.197	8.95E−07
	50	HCG+	740	3 615.365	1.04E−05
QUADRATICQF1	100	NH+	181	2 089.250	0.00E+00
	100	PRP+	136	1 315.878	0.00E+00
	100	NYF	231	4 130.011	1.15E−02
	100	HCG+	282	3 613.380	9.89E−03
QUARTC	400	NH+	17	1 859.335	9.91E−07
	400	PRP+	48	3 633.389	4.55E−03
	400	NYF	50	3 643.121	3.70E−03
	400	HCG+	17	1 956.870	9.91E−07

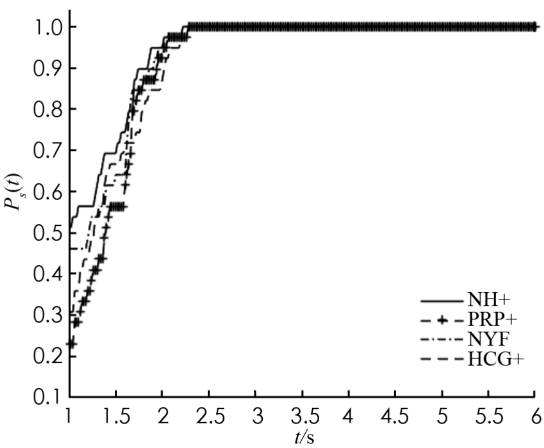

图 1　迭代时间算法性能图

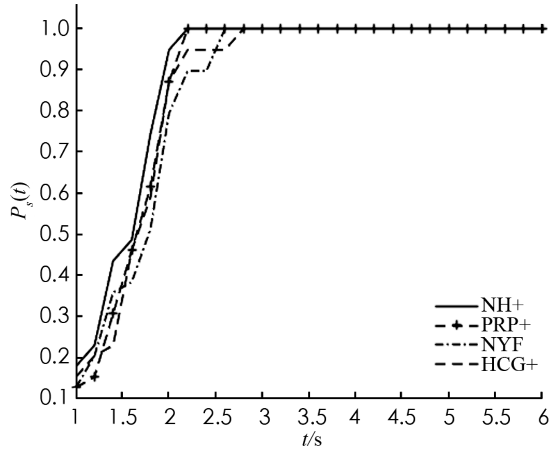

图 2　迭代次数算法性能图

参考文献:

[1]　唐天国. 一种求解无约束优化问题的新混合共轭梯度法 [J]. 西南师范大学学报(自然科学版), 2019, 44(9): 34-39.

[2]　HESTENES M R, STIEFEL E. Methods of Conjugate Gradients for Solving Linear Systems [J]. Journal of Research of the National Bureau of Standards, 1952, 49(6): 409-436.

[3]　POLAK E, RIBIERE G. Note Sur La Convergence de Méthodes de Directions Conjuguées [J]. Revue Française d'Informatique et De Recherche Opérationnelle Série Rouge, 1969, 3(16): 35-43.

[4]　POLYAK B T. The Conjugate Gradient Method in Extremal Problems [J]. USSR Computational Mathematics and Mathematical Physics, 1969, 9(4): 94-112.

[5]　DAI Y H, YUAN Y. A Nonlinear Conjugate Gradient Method with a Strong Global Convergence Property [J]. SIAM Journal on Optimization, 1999, 10(1): 177-182.

[6]　LIU Y, STOREY C. Efficient Generalized Conjugate Gradient Algorithms, Part 1: Theory [J]. Journal of Optimization Theory and Applications, 1991, 69(1): 129-137.

[7]　FLETCHER R, REEVES C M. Function Minimization by Conjugate Gradients [J]. The Computer Journal, 1964, 7(2):

149-154.
[8] FLETCHER R. Practical Method of Optimization，Unconstrained Optimization [M]. New York：John Wiley and Sons，1987.
[9] SHENGWEI Y，WEI Z X，HUANG H. A Note about WYL's Conjugate Gradient Method and Its Applications [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] 韩　信，张俊容，王森森. 一种新的混合共轭梯度算法 [J]. 西南大学学报(自然科学版)，2017，39(5)：132-138.
[12] LIU J K，DU X L. Global Convergence of an Efficient Hybrid Conjugate Gradient Method for Unconstrained Optimization [J]. Bulletin of the Korean Mathematical Society，2013，50(1)：73-81.
[13] ANDREI N. Hybrid Conjugate Gradient Algorithm for Unconstrained Optimization [J]. Journal of Optimization Theory and Applications，2009，141(2)：249-264.
[14] BABAIE-KAFAKI S，GHANBARI R. A Hybridization of the Polak-Ribière-Polyak and Fletcher-Reeves Conjugate Gradient Methods [J]. Numerical Algorithms，2015，68(3)：481-495.
[15] ZHANG L，ZHOU W J，LI D H. A Descent Modified Polak-Ribière-Polyak Conjugate Gradient Method and Its Global Convergence [J]. IMA Journal of Numerical Analysis，2006，26(4)：629-640.
[16] NARUSHIMA Y，YABE H，FORD J A. A Three-Term Conjugate Gradient Method with Sufficient Descent Property for Unconstrained Optimization [J]. SIAM Journal on Optimization，2011，21(1)：212-230.
[17] ZOUTENDIJK G. Nonlinear Programming，Computational Methods [J]. Integer and Nonlinear Programming，1970，143：37-86.
[18] GILBERT J C，NOCEDAL J. Global Convergence Properties of Conjugate Gradient Methods for Optimization [J]. SIAM Journal on Optimization，1992，2(1)：21-42.
[19] ANDREI N. An Unconstrained Optimization Test Functions Collection[J]. Environmental Science and Technology，2008，10(1)：6552-6558.
[20] DOLAN E D，MORÉ J J. Benchmarking Optimization Software with Performance Profiles [J]. Mathematical Programming，2002，91(2)：201-213.

责任编辑　张枸