doi:10.16055/j.issn.1672-058X.2015.0004.006

# 一种修正的 PRP 共轭梯度法的全局收敛性*

蔡 正 兰
(重庆师范大学 数学学院,重庆 401331)

**摘 要**:针对张丽提出的一种修正的 PRP 方法——NPRP 法,在广义 Wolfe 下证明了 NPRP 法的全局收敛性.

**关键词**:共轭梯度法;修正的 PRP 法;广义 Wolfe 线搜索;全局收敛性
**中图分类号**:O221 **文献标识码**:A **文章编号**:1672-058X(2015)04-0020-03

## 0 引 言

考虑无约束优化问题:

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

其中,$f:\mathbf{R}^n \to \mathbf{R}$ 是一个连续可微的函数.

众所周知,共轭梯度法是求解式(1)的一种重要方法,它具有如下的迭代格式

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

并且

$$d_k = \begin{cases} -g_k, & k = 0 \\ -g_k + \beta_k d_{k-1}, & k \geqslant 1 \end{cases} \tag{3}$$

其中,$\alpha_k$ 是利用某种线搜索获得的步长;$d_k$ 为搜索方向;纯量 $\beta_k$ 的选取应满足所谓的共轭性,即当 $f(x)$ 是严格凸二次函数,且采用精确线搜索时,由式(1)(2)产生的搜索方向关于 $f(x)$ Hessian 共轭.

不同的 $\beta_k$ 取法会产生不同的共轭梯度法,比较常见的 $\beta_k$ 有 FR(Fletcher-Revees[1]),PRP(Polak-Ribiere-Polyak[2]),HS(Hestenes-Stiefel[3]),LS(Liu-Storey[4]),CD(Conjugate Descent[5])和 DY(Dai-Yuan[6])共轭梯度法.它们分别为

$$\beta_k^{\mathrm{FR}} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2},\beta_k^{\mathrm{PRP}} = \frac{g_k^{\mathrm{T}} y_{k-1}}{\|g_{k-1}\|^2},\beta_k^{\mathrm{HS}} = \frac{g_k^{\mathrm{T}} y_{k-1}}{g_{k-1}^{\mathrm{T}} y_{k-1}}$$

$$\beta_k^{\mathrm{LS}} = \frac{g_k^{\mathrm{T}} y_{k-1}}{-d_{k-1}^{\mathrm{T}} g_{k-1}},\beta_k^{\mathrm{CD}} = \frac{\|g_k\|^2}{-d_{k-1}^{\mathrm{T}} g_{k-1}},\beta_k^{\mathrm{DY}} = \frac{\|g_k\|^2}{d_{k-1}^{\mathrm{T}} y_{k-1}}$$

步长 $\alpha_k$ 的选取应满足一定的下降条件,即 $g_k^{\mathrm{T}} d_k<0$,此处采用的线搜索为广义 Wolfe 线搜索

$$f(x_k + \alpha_k d_k) \leqslant f(x_k) + \delta \alpha_k g_k^{\mathrm{T}} d_k \tag{4}$$

$$\sigma_1 g_k^{\mathrm{T}} d_k \leqslant g(x_k + \alpha_k d_k)^{\mathrm{T}} d_k \leqslant -\sigma_2 g_k^{\mathrm{T}} d_k \tag{5}$$

其中 $0<\delta<\sigma_1<1$,$0<\sigma_2<+\infty$.可以看出只要 $\sigma_1=\sigma_2$,广义 Wolfe 线搜索可以简化为强 Wolfe 线搜索.

**收稿日期**:2014-08-24;**修回日期**:2014-09-15.
* **基金项目**:2012 年度重庆市软科学研究课题(cstc2012cx-rkxA0037).
**作者简介**:蔡正兰(1989-),女,重庆垫江人,硕士研究生,从事最优化方法及其应用研究.

2009 年，张丽[7]对 PRP 方法进行了一个修正，变为

$$\beta_k^{\mathrm{NPRP}}=\frac{\|g_k\|^2-\frac{\|g_k\|}{\|g_{k-1}\|}|g_k^{\mathrm{T}}g_{k-1}|}{\|g_{k-1}\|^2} \tag{6}$$

在文献[7]中，作者证明了 $\beta_k^{\mathrm{NPRP}}$ 方法在强 Wolfe 线搜索且满足 $\sigma<\frac{1}{2}$ 的条件下具有充分下降性和全局收敛性.

此处对证明过程进行一个推广，证明该方法在广义 Wolfe 下的全局收敛性，当 $\sigma_1=\sigma_2$ 时，就可以直接得出 $\beta_k^{\mathrm{NPRP}}$ 方法在强 Wolfe 线搜索下的全局收敛性.

## 1　算法

NPRP 算法：

初始步：给定 $x_0\in\mathbf{R}^n$，$\varepsilon\geqslant 0$，令 $d_0=-g_0$，$k:=0$；

第 1 步：若 $\|g_0\|_\infty\leqslant\varepsilon$，则停止；

第 2 步：求出满足广义 Wolfe 搜索式(4)(5)下的步长 $\alpha_k$；

第 3 步：计算 $x_{k+1}=x_k+\alpha_k d_k$，若 $\|g_{k+1}\|_\infty\leqslant\varepsilon$，则停止；

第 4 步：通过式(6)计算 $\beta_k$，并通过式(3)计算 $d_{k+1}$，令 $k:=k+1$，转第 2 步.

## 2　算法的收敛性

作如下假设[3]：

(A) 水平集 $\Omega=\{x\in\mathbf{R}^n \mid f(x)\leqslant f(x_0)\}$ 有下界，其中 $x_0\in\mathbf{R}^n$ 为初始点.

(B) $f$ 在水平集 $\Omega$ 的一个领域 $N$ 内连续可微，且其梯度 $g$ 满足 Lipschitz 连续，即存在常数 $L>0$，使得

$$\|g(x)-g(y)\|\leqslant L\|x-y\|,\ \forall x,y\in N \tag{7}$$

显然，在上述假设下，存在某个 $\gamma_1>0$，使得

$$\|g(x)\|\leqslant\gamma_1,\ \forall x\in\Omega \tag{8}$$

由假设(A)和(B)可知，$\sum_{k\geqslant 0}^{\infty}(1-\sigma)\alpha_k g_k^{\mathrm{T}}d_k<+\infty$，结合 $\|g(x)\|\leqslant\gamma_1$，可得 $\sum_{k\geqslant 0}^{\infty}\alpha_k\|d_k\|<+\infty$，从而有

$$\lim_{k\to\infty}\alpha_k\|d_k\|=0 \tag{9}$$

**引理 1**　在假设(A)和(B)成立的条件下，$\beta_k^{\mathrm{NPRP}}$ 满足 $0\leqslant\beta_k^{\mathrm{NPRP}}\leqslant\beta_k^{\mathrm{FR}}$.

$$\beta_k^{\mathrm{NPRP}}=\frac{\|g_k\|^2-\frac{\|g_k\|}{\|g_{k-1}\|}|g_k^{\mathrm{T}}g_{k-1}|}{\|g_{k-1}\|^2}\geqslant$$

$$\frac{\|g_k\|^2-\frac{\|g_k\|}{\|g_{k-1}\|}\|g_k\|\|g_{k-1}\|}{\|g_{k-1}\|^2}=0$$

并且

$$\beta_k^{\mathrm{NPRP}}\leqslant\frac{\|g_k\|^2}{\|g_{k-1}\|^2}=\beta_k^{\mathrm{FR}}$$

结论显然成立.

**引理 2**　如果存在常数 $\varepsilon>0$，使得 $\|g_k\|\geqslant\varepsilon$ 对任意 $k\geqslant 0$ 成立，则必存在常数 $M>0$，使得 $\|d_k\|\leqslant M$ 对任意 $k\geqslant 0$ 成立.

**证明** 当 $k=0$ 时，$\|d_0\|=\|-g_0\|=\|g_0\|\leqslant\gamma_1$，当 $k\geqslant 1$ 时，$\|d_k\|\leqslant\|g_k\|+\frac{2\|g_k\|(\|g_k-g_{k-1}\|)}{\|g_{k-1}\|^2}\|d_{k-1}\|$.

由式(7)(8)可知

$$\|d_k\|\leqslant \gamma_1 + \frac{2\gamma_1 L\alpha_{k-1|}\|d_{k-1}\|}{\varepsilon^2}\|d_{k-1}\|$$

由式(9)可知 $\exists q\in(0,1)$，并且存在整数 $k_0$，使得 $\forall k\geqslant k_0$，有

$$\frac{2\gamma_1 L}{\varepsilon^2}\alpha_{k-1}\|d_{k-1}\|\leqslant q$$

因此，对所有的 $k\geqslant k_0$，有

$$\begin{aligned}\|d_k\|&\leqslant \gamma_1 + q\|d_{k-1}\|\leqslant\\ &\gamma_1(1+q+q^2+\cdots+q^{k-k_0-1})+q^{k=k_0}\|d_{k_0}\|\leqslant\\ &\frac{\gamma_1}{1-q}+\|d_{k_0}\|\end{aligned}$$

令 $M=\max\left\{\|d_1\|,\|d_2\|,\cdots,\|d_{k_0}\|,\frac{\gamma_1}{1-q}+\|d_{k_0}\|\right\}$，就可以推断出对任意的 $k$，显然结论 $\|d_k\|\leqslant M$ 成立.

**定理 1**[8] 在假设(A)和(B)成立的条件下，如果 $g_k^{\mathrm{T}}d_k<0$，$0\leqslant\beta_k\leqslant\beta_k^{\mathrm{FR}}$，那么 $\liminf\limits_{k\to\infty}\|g_k\|=0$.

显然 $\beta_k^{\mathrm{NPRP}}$ 满足定理 1，即 $\beta_k^{\mathrm{NPRP}}$ 是全局收敛的.

**参考文献：**


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

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

[3] HESTENES M R,STIEFEL E.Methods of Conjugate Gradients for Solving Linear Systems[J].Research Nat.Bur.Standards,1952,49 (1952) :409-436

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

[5] FLETCHER R.Practical Methods of Optimization:Vol.2:Constrained Optimization[M].JOHN WILEY & SONS,INC,1981

[6] 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

[7] ZHANG L. An Improved Wei-Yao-Liu Nonlinear Conjugate Gradient Method for Optimization Computation [J]. Applied Mathematics and Computation,2009(6):2269-2274

[8] 杜学武，徐成贤.由 FR 共轭梯度法控制的下降算法的全局收敛性[J].西安交通大学学报,1998,32(6):100-102


# A Modified PRP Conjugate Gradient Method and Its Global Convergence


## CAI Zheng-lan

(School of Mathematics, Chongqing Normal University, Chongqing 401331)



**Abstract**: On the basis of NPRP method, a modified PRP method, was proposed by Zhang Li ,this paper proves the global convergence of NPRP with the generalized Wolfe line search.



**Key words**: Conjugate gradient method; modified PRP method; generalized Wolfe line search; global convergence