Vol. 34 (2014)
No. 6

数 学 杂 志
J. of Math. (PRC)

Wolfe 步长规则下约束优化问题的共轭梯度投影算法

景书杰 , 赵海燕
(河南理工大学数学与信息科学学院, 河南 焦作 454003)

摘 要: 本文研究了约束优化问题 $\min\limits_{x\in\Omega} f(x)$. 利用共轭梯度算法与 GLP 梯度投影思想相结合的方法, 构造了一个新的共轭梯度投影算法, 并在 Wolfe 线搜索下获得了该算法的全局收敛性结果.

关 键 词: 约束优化问题; 共轭梯度法; GLP 梯度投影; Wolfe 线搜索; 全局收敛性
MR(2010) **主题分类号:** 90C30 **中图分类号:** O224
文献标识码: A **文章编号:** 0255-7797(2014)06-1193-07

1 引言

考虑约束优化问题

$$\min_{x\in\Omega} f(x), \tag{1.1}$$

其中 $\Omega \subseteq R^n$ 为非空闭凸集, $f: R^n \to R$ 为 Ω 上连续可微函数. 1987 年 Calamai P H 和 More J J[1] 借助投影算子沿可行域边界进行曲线搜索的思想给出了 GLP 梯度投影算法, 并讨论了该投影算法的收敛性质.

在约束优化问题中寻找可行的下降方向是构造新算法的关键, GLP 梯度投影算法有效地解决了这一问题. 但该算法收敛速度较慢并且需要多次投影, 计算量较大. 而共轭梯度投影算法具有算法简单, 计算量少, 收敛速度快的特点, 是求解无约束优化问题 $\min\limits_{x\in R^n} f(x)$ 的一种有效方法, 其迭代格式为:

$$x_{k+1} = x_k + \alpha_k d_k, \tag{1.2}$$

$$d_k = \begin{cases} -g_k, k = 1, \\ -g_k + \beta_k d_{k-1}, k \geq 2, \end{cases} \tag{1.3}$$

其中 d_k 为搜索方向, β_k 为标量, 参数 β_k 的不同选取对应不同的共轭梯度法, 见文献 [2]. 步长 α_k 满足 Wolfe 线搜索条件

$$f(x_k + \alpha_k d_k) - f(x_k) \leq \sigma_1 \alpha_k g_k^T d_k, \tag{1.4}$$

$$g_{k+1}^T d_k \geq \sigma_2 g_k^T d_k \tag{1.5}$$

*收稿日期: 2013-01-11 接收日期: 2013-04-17
基金项目: 国家自然科学基金资助 (10671057); 河南理工大学运筹学与控制论重点学科资助项目; 河南省数学一级重点学科资助项目.
作者简介: 景书杰 (1965–), 男, 河南新乡, 教授, 主要研究方向: 最优化理论与应用.

在无约束优化问题中, 为了得到目标函数在 x_k 点的下降的共轭梯度方向, 假设对 $\forall k > 1$, 参数 β_k 满足

$$\left\{\begin{array}{l} \|g_k\|^2 > \left|\beta_k g_k^T d_{k-1}\right|, \\ \left|g_k^T(-g_k + \beta_k d_{k-1})\right| \geq \lambda \left|\beta_k\right| \|g_k\| \left\|d_{k-1}\right\|, \end{array}\right. \tag{1.6}$$

其中 $\lambda > 1$. 记 $g_k = \nabla f(x_k)$ 是 $f(x)$ 在 x_k 点的梯度. 受文献 [3–6] 的启发, 限制 β_k 的范围, 使其满足 $\beta_k \in [-\beta_k^1, \beta_k^2]$, 其中 $\beta_k^1 = \frac{1}{\lambda - \cos\theta_k}\frac{\|g_k\|}{\|d_{k-1}\|}$, $\beta_k^2 = \frac{1}{\lambda + \cos\theta_k}\frac{\|g_k\|}{\|d_{k-1}\|}$, 其中 $\cos\theta_k$ 是 $-g_k$ 与 d_{k-1} 的夹角.

考虑到投影梯度法易于寻找可行下降方向, 共轭梯度法较快的收敛速度以及好的数值表现, 本文将这两种算法相结合用于求解约束优化问题, 并且对参数 β_k 进行适当的修正, 可以提高梯度投影算法的收敛速度.

2 定义与引理

定义 2.1 [2] 对任意 $x \in R^n$, 记 $P_\Omega(x) = \arg\min\{\|x - y\| : y \in \Omega\}$, $P_\Omega(x)$ 为 x 在 Ω 上的投影, $P_\Omega(\cdot)$ 称为从 R^n 到 Ω 的投影算子.

引理 2.2 [3] P 为 R^n 在 Ω 上的投影, 令 $x(\alpha, s) = P(x + \alpha s)$, 对 $x \in \Omega$ 和 $s \in R^n$ 有

(1) $\langle x(\alpha, s) - (x + \alpha s), y - x(\alpha, s)\rangle \geq 0$, 对任意 $y \in \Omega$, $\alpha > 0$.

(2) P 是一个非扩张算子, 则有 $\|P(y) - P(x)\| \leq \|y - x\|$, $x, y \in R^n$.

(3) $\langle -s, x - x(\alpha, s)\rangle \geq \frac{\|x(\alpha,s)-x\|^2}{\alpha}$, 对任意 $\alpha > 0$.

定义 2.3 [7] σ: $[0, +\infty) \to [0, +\infty)$ 称为强函数, 若对任意 $t_k \in [0, +\infty)$ 均有

$$\lim_{k\to\infty} \sigma(t_k) = 0, \Rightarrow \lim_{k\to\infty} t_k = 0.$$

定义 2.4 [7] 逆连续模函数 $f : R^n \to R$ 为一阶连续可微函数, 设

$$a = \sup\{\|\nabla f(x) - \nabla f(y)\| \,| x, y \in L\} > 0.$$

令

$$\rho(t) = \left\{\begin{array}{l} \inf\{\|x - y\| \,| x, y \in L, \|\nabla f(x) - \nabla f(y)\| \geq t\}, t \in [0, a), \\ \lim\limits_{r\to a} \rho(r), t \in [a, +\infty), \end{array}\right.$$

则称 $\rho(t)$ 为 $\nabla f(x)$ 在 L 上的逆连续模函数.

引理 2.5 [7] 若 $\nabla f(x)$ 在 L 上一致连续, 则 $\nabla f(x)$ 在 L 上的逆连续模函数为强函数.

引理 2.6 [8] 设 $x \in \Omega$, $x \in \Omega$ 处的切锥记为

$$T_\Omega(x) = \left\{ v \in R^n \,\middle|\, \exists x_t \in \Omega, \tau_t > 0, t \to \infty, x_t \to x, \tau_t \to \infty, v = \lim_{t\to\infty} \frac{x_t - x}{\tau_t} \right\}.$$

引理 2.7 [8] 令 $\nabla_\Omega f(x)$ 为 f 在 $x \in \Omega$ 上的梯度投影, 则有

(1) $\min\{\langle \nabla f(x), v\rangle \,| v \in T(x), \|v\| \leq 1\} = -\|\nabla_\Omega f(x)\|$.

(2) $\|\nabla_\Omega f(\cdot)\|$ 在 Ω 上是下半连续的, 若 $\lim\limits_{k\to\infty} x_k = x$, 则 $\|\nabla_\Omega f(x)\| \leq \lim\limits_{k\to\infty} \inf \|\nabla_\Omega f(x_k)\|$.

(3) $x \in \Omega$ 是约束优化问题的稳定点, 当且仅当 $\nabla_\Omega f(x) = 0$.

3 共轭梯度投影算法

设 $x_k \in \Omega$, $\alpha_k > 0$, 记 $\overline{x_k}(\alpha_k, s_k) = P(x_k - \alpha_k \nabla f(x_k))$, 令

$$x_{k+1} = x_k(\alpha_k, s_k) = P(x_k + \alpha_k s_k), \tag{3.1}$$

其中 $s_k = \begin{cases} -\nabla f(x_k), k = 1, \\ -\nabla f(x_k) + \beta_k d_{k-1}, k > 1. \end{cases}$ 取

$$|\beta_k| = \frac{1}{\lambda} \frac{\|\overline{x_k}(\alpha_k, s_k) - x_k\|^2}{\alpha_k^2 \|\nabla f(x_k)\| \|d_{k-1}\|}, \tag{3.2}$$

其中 $\lambda > 1 + \frac{1}{\varepsilon}$, $\varepsilon > 0$, 且 ε 充分小. 由 β_k 确定 s_k, 从而确定共轭梯度投影算法的搜索方向

$$d_k = \frac{x_k(\alpha_k, s_k) - x_k}{\alpha_k}. \tag{3.3}$$

结合 Wolfe 线搜索给出共轭梯度投影算法的步骤:

算法

步 1 对 $\forall x_1 \in \Omega$, 令 $d_0 = 0$, $\beta_1 = 0$, $k = 1$.

步 2 计算 $\alpha_k > 0$, 满足条件式 (1.4) 和式 (1.5).

步 3 令 $x_{k+1} = x_k + \alpha_k d_k$, 若 $\|x_k - x_k(\alpha_k, s_k)\| \le \varepsilon$, 停止, 则 $x_k(\alpha_k, s_k)$ 为问题 (1.1) 的近似最优解.

步 4 令 $k = k + 1$, 转步 2.

为了证明算法的全局收敛性, 需作如下假设 H [2]:

(a) 目标函数 $f(x)$ 在水平集 $L_0 = \{x_1 \in R / f(x) \le f(x_1)\}$ 上有下界, 其中 x_1 为初始点.

(b) 目标函数 $f(x)$ 在水平集 L_0 上连续可微, 且 $\nabla f(x)$ 是 Lipschitz 连续的, 即存在常数 $L > 0$, 使得 $\|g(x) - g(y)\| \le L\|x - y\|$, $\forall x, y \in L_0$.

引理 3.1 若 x_k 不是问题 (1.1) 的稳定点, 对 $\forall \alpha_k > 0$, $x_k \ne x_k(\alpha_k, s_k)$, 则由式 (3.3) 生成的搜索方向 d_k 为该算法的下降方向, 即 $\left\langle \nabla f(x_k), \frac{x_k(\alpha_k, s_k) - x_k}{\alpha_k} \right\rangle \le \frac{1-\lambda}{\lambda} \|\nabla f(x_k)\|^2 < 0$.

证 对 $\forall \alpha_k > 0$, 由式 (3.1), 引理 2.2(2), 引理 2.2(3) 和式 (3.2) 可得

$$\begin{aligned}
& \left\langle \nabla f(x_k), \frac{x_k(\alpha_k, s_k) - x_k}{\alpha_k} \right\rangle \\
= \ & \frac{1}{\alpha_k} \left[\langle \nabla f(x_k), x_k(\alpha_k, s_k) - \overline{x_k}(\alpha_k, s_k) \rangle + \langle \nabla f(x_k), \overline{x_k}(\alpha_k, s_k) - x_k \rangle \right] \\
\le \ & \frac{1}{\alpha_k} \left[\|\nabla f(x_k)\| \|x_k(\alpha_k, s_k) - \overline{x_k}(\alpha_k, s_k)\| - \langle \nabla f(x_k), x_k - \overline{x_k}(\alpha_k, s_k) \rangle \right] \\
\le \ & |\beta_k| \|\nabla f(x_k)\| \|d_{k-1}\| - \frac{\|\overline{x_k}(\alpha_k, s_k) - x_k\|^2}{\alpha_k^2} \le \left(\frac{1}{\lambda} - 1 \right) \frac{\|\overline{x_k}(\alpha_k, s_k) - x_k\|^2}{\alpha_k^2} \\
\le \ & \frac{1-\lambda}{\lambda} \|\nabla f(x_k)\|^2 < 0.
\end{aligned}$$

注 3.1 由引理 3.1 可得

$$\frac{1}{\alpha_k} \langle \nabla f(x_k), x_k - x_k(\alpha_k, s_k) \rangle \ge \frac{\lambda - 1}{\lambda} \frac{\|\overline{x_k}(\alpha_k, s_k) - x_k\|^2}{\alpha_k^2}. \tag{3.4}$$

引理 3.2 若 $x_k \in \Omega$ 不是稳定点, 则有

(1) $\|d_k\| \le \|s_k\| \le \frac{\lambda+1}{\lambda}\|\nabla f(x_k)\|$.

(2) $\langle -s_k, x_k - x_k(\alpha_k, s_k)\rangle - \langle \nabla f(x_k), x_k - x_k(\alpha_k, s_k)\rangle \le \frac{\lambda+1}{\lambda(\lambda-1)}\langle \nabla f(x_k), x_k - x_k(\alpha_k, s_k)\rangle$.

证 (1) 由式 (3.3)、式 (1.3) 和引理 2.2(2) 可得

$$\begin{aligned}\|d_k\| &= \frac{\|x_k(\alpha_k, s_k) - x_k\|}{\alpha_k} \le \|s_k\| = \|-g_k + \beta_k d_{k-1}\| \le \|g_k\| + |\beta_k|\,\|d_{k-1}\| \\ &\le \|\nabla f(x_k)\| + \frac{\|\overline{x_k}(\alpha_k, s_k) - x_k\|^2}{\lambda\alpha_k^2\|\nabla f(x_k)\|} \le \left(\frac{\lambda+1}{\lambda}\right)\|\nabla f(x_k)\|.\end{aligned}$$

(2) 由 s_k 的定义、式 (3.2)、引理 3.2(1) 和式 (3.4) 得

$$\begin{aligned}&\langle -s_k, x_k - x_k(\alpha_k, s_k)\rangle - \langle \nabla f(x_k), x_k - x_k(\alpha_k, s_k)\rangle \le |\beta_k|\,\|d_{k-1}\|\,\|x_k - x_k(\alpha_k, s_k)\| \\ \le\ & \frac{1}{\lambda}\frac{\|\overline{x_k}(\alpha_k, s_k) - x_k\|^2}{\alpha_k^2\|\nabla f(x_k)\|}\cdot\alpha_k\|s_k\| \le \frac{(\lambda+1)\alpha_k}{\lambda(\lambda-1)}\cdot\frac{\lambda-1}{\lambda}\frac{\|\overline{x_k}(\alpha_k, s_k) - x_k\|^2}{\alpha_k^2} \\ \le\ & \frac{\lambda+1}{\lambda(\lambda-1)}\langle \nabla f(x_k), x_k - x_k(\alpha_k, s_k)\rangle.\end{aligned}$$

4 收敛性

引理 4.1 若假设 H 成立, f 是连续可微函数, $\nabla f(x)$ 在 Ω 上一致连续, $\{x_k\}$ 是由共轭梯度投影算法产生的序列, 满足 Wolfe 线搜索条件

$$f(x_k) - f(x_k(\alpha_k, s_k)) \ge -\sigma_1\langle \nabla f(x_k), x_k(\alpha_k, s_k), s_k - x_k\rangle = \sigma(t_k),$$

其中 $t_k = \frac{-\langle \nabla f(x_k), x_k(\alpha_k, s_k) - x_k\rangle}{\|x_k(\alpha_k, s_k) - x_k\|}$, 则有

(1) $\sigma(t_k)$ 为强函数; (2) $\lim\limits_{k\to\infty} t_k = 0$.

证 (1) 由式 (1.5) 知

$$\begin{aligned}&(1-\sigma_2)\langle \nabla f(x_k), x_k - x_k(\alpha_k, s_k)\rangle \\ =\ & \langle \nabla f(x_k), x_k - x_k(\alpha_k, s_k)\rangle - \sigma_2\langle \nabla f(x_k), x_k - x_k(\alpha_k, s_k)\rangle \\ \le\ & \langle \nabla f(x_k), x_k - x_k(\alpha_k, s_k)\rangle + \langle \nabla f(x_k(\alpha_k, s_k)), x_k(\alpha_k, s_k) - x_k\rangle \\ =\ & \langle \nabla f(x_k) - \nabla f(x_k(\alpha_k, s_k)), x_k - x_k(\alpha_k, s_k)\rangle \\ \le\ & \|\nabla f(x_k) - \nabla f(x_k(\alpha_k, s_k))\|\,\|x_k - x_k(\alpha_k, s_k)\|,\end{aligned}$$

所以

$$\|\nabla f(x_k) - \nabla f(x_k(\alpha_k, s_k))\| \ge (1-\sigma_2)\frac{\langle \nabla f(x_k), x_k - x_k(\alpha_k, s_k)\rangle}{\|x_k - x_k(\alpha_k, s_k)\|} = (1-\sigma_2)t_k > 0. \quad (4.1)$$

由假设 H(2) 知 $\|\nabla f(x_k) - \nabla f(x_k(\alpha_k, s_k))\| \le L\|x_k - x_k(\alpha_k, s_k)\|$, 结合 (4.1) 得

$$\|x_k - x_k(\alpha_k, s_k)\| \ge \frac{1}{L}(1-\sigma_2)\frac{\langle \nabla f(x_k), x_k - x_k(\alpha_k, s_k)\rangle}{\|x_k - x_k(\alpha_k, s_k)\|}. \quad (4.2)$$

由定义 2.4 和引理 2.5 知 $\sigma(t_k)$ 为强函数.

(2) 由式 (1.4) 和式 (4.2) 可得

$$\begin{aligned} f(x_k) - f(x_k(\alpha_k, s_k)) \geq \sigma(t_k) &= \sigma_1 \langle \nabla f(x_k), x_k - x_k(\alpha_k, s_k)\rangle \\ &= \sigma_1 \|x_k - x_k(\alpha_k, s_k)\| \frac{\langle \nabla f(x_k), x_k - x_k(\alpha_k, s_k)\rangle}{\|x_k - x_k(\alpha_k, s_k)\|} \\ &\geq \frac{1}{L}\sigma_1((1-\sigma_2)t_k)t_k. \end{aligned}$$

两边对 k 取级数可得 $\sum\limits_{k\geq 1}(f(x_k) - f(x_k(\alpha_k, s_k))) \geq \sum\limits_{k\geq 1}\sigma(t_k)$.

由假设 H 知 $f(x)$ 在水平集 L 上单调不增有下界 $\sum\limits_{k\geq 1}(f(x_k) - f(x_k(\alpha_k, s_k))) < +\infty$, 所以 $\sum\limits_{k\geq 1}\sigma(t_k) < +\infty$, 则有 $\lim\limits_{k\to\infty}\sigma(t_k) = 0$. 即

$$\lim_{k\to\infty}\langle \nabla f(x_k), x_k - x_k(\alpha_k, s_k)\rangle = 0, \tag{4.3}$$

利用定义 2.3 强函数的定义知 $\lim\limits_{k\to\infty} t_k = 0$. 即

$$\lim_{k\to\infty}\frac{\langle \nabla f(x_k), x_k - x_k(\alpha_k, s_k)\rangle}{\|x_k - x_k(\alpha_k, s_k)\|} = 0. \tag{4.4}$$

引理 4.2 设 $f: R^n \to R$ 为 Ω 上连续可微函数且在 Ω 上有下界, $\nabla f(x)$ 在 Ω 上一致连续, 且 $\{\nabla f(x_k)\}$ 有界, 则由算法产生的迭代点列 $\{x_k\}$ 满足

$$\lim_{k\to\infty}\frac{\|x_k - x_k(\alpha_k, s_k)\|}{\alpha_k} = 0, \lim_{k\to\infty}\frac{\|x_k - \overline{x_k}(\alpha_k, s_k)\|}{\alpha_k} = 0.$$

证 由引理 3.2(2) 可得

$$\langle -s_k, x_k - x_k(\alpha_k, s_k)\rangle \leq \frac{\lambda^2+1}{\lambda(\lambda-1)}\langle \nabla f(x_k), x_k - x_k(\alpha_k, s_k)\rangle,$$

两边同时除以 $\|x_k - x_k(\alpha_k, s_k)\|$, 且取极限, 则上式可变为

$$0 \leq \lim_{k\to\infty}\sup\frac{\langle -s_k, x_k - x_k(\alpha_k, s_k)\rangle}{\|x_k - x_k(\alpha_k, s_k)\|} \leq \lim_{k\to\infty}\frac{\lambda^2+1}{\lambda(\lambda-1)}\frac{\langle \nabla f(x_k), x_k - x_k(\alpha_k, s_k)\rangle}{\|x_k - x_k(\alpha_k, s_k)\|},$$

由式 (4.4) 可得

$$\lim_{k\to\infty}\frac{\langle -s_k, x_k - x_k(\alpha_k, s_k)\rangle}{\|x_k - x_k(\alpha_k, s_k)\|} = 0. \tag{4.5}$$

由引理 2.2(3) 知 $\langle -s_k, x_k - x_k(\alpha_k, s_k)\rangle \geq \frac{\|x_k(\alpha_k, s_k) - x_k\|^2}{\alpha_k}$, 则有

$$0 \leq \frac{\|x_k(\alpha_k, s_k) - x_k\|}{\alpha_k} \leq \frac{\langle -s_k, x_k - x_k(\alpha_k, s_k)\rangle}{\|x_k - x_k(\alpha_k, s_k)\|},$$

则由式 (4.5) 可得 $\lim\limits_{k\to\infty}\frac{\|x_k - x_k(\alpha_k, s_k)\|}{\alpha_k} = 0$, 由式 (3.4) 知

$$0 \leq \frac{\lambda-1}{\lambda}\frac{\|\overline{x_k}(\alpha_k, s_k) - x_k\|^2}{\alpha_k^2} \leq \frac{1}{\alpha_k}\langle \nabla f(x_k), x_k - x_k(\alpha_k, s_k)\rangle \leq \|\nabla f(x_k)\|\frac{\|x_k - x_k(\alpha_k, s_k)\|}{\alpha_k},$$

两边取极限, 再利用 $\{\nabla f(x_k)\}$ 有界和式 (4.3) 可得 $\lim\limits_{k\to\infty}\frac{\|x_k-\overline{x_k}(\alpha_k,s_k)\|}{\alpha_k}=0$.

定理 4.3 设 $f: R^n\to R$ 为 Ω 上连续可微函数且在 Ω 上有下界, $\nabla f(x)$ 在 Ω 上一致连续, $\{x_k\}$ 是由算法产生的迭代点列, $0<\alpha_k\leq\gamma$, $k=1,2,\cdots$, 则 $\lim\limits_{k\to\infty}\|\nabla_\Omega f(x_k)\|=0$, 且 $\{x_k\}$ 的任意聚点都是优化问题的稳定点.

证 由引理 2.7(1), (3) 知, $\forall\varepsilon>0$, $\exists v_k\in T_\Omega(x_k)$, $\|v_k\|\leq 1$ 满足

$$\|\nabla_\Omega f(x_k)\|\leq\langle -\nabla f(x_k),v_k\rangle+\varepsilon, \tag{4.6}$$

对 $\forall z\in\Omega$ 由引理 2.2(1) 知 $\langle x_k(\alpha_k,s_k)-(x_k+\alpha_k s_k),z-x_k(\alpha_k,s_k)\rangle\geq 0$, 则有

$$\langle\alpha_k s_k,z-x_k(\alpha_k,s_k)\rangle\leq\langle x_k(\alpha_k,s_k)-x_k,z-x_k(\alpha_k,s_k)\rangle\leq\|x_k(\alpha_k,s_k)-x_k\|\,\|z-x_k(\alpha_k,s_k)\|,$$

所以

$$\langle s_k,z-x_k(\alpha_k,s_k)\rangle\leq\frac{\|x_k(\alpha_k,s_k)-x_k\|}{\alpha_k}\|z-x_k(\alpha_k,s_k)\|. \tag{4.7}$$

令 $v_{k+1}=z-x_k(\alpha_k,s_k)\in T_\Omega(x_{k+1})$, $\|v_{k+1}\|\leq 1$, 由式 (4.7) 可得

$$\langle s_k,v_{k+1}\rangle=\langle -\nabla f(x_k)+\beta_k d_{k-1},v_{k+1}\rangle\leq\frac{\|x_k(\alpha_k,s_k)-x_k\|}{\alpha_k},$$

所以将上式与式 (3.2)、引理 2.2(2) 结合可得

$$\begin{aligned}\langle -\nabla f(x_k),v_{k+1}\rangle &\leq \frac{\|x_k(\alpha_k,s_k)-x_k\|}{\alpha_k}+|\beta_k|\,\|d_{k-1}\|\\ &\leq \frac{\|x_k(\alpha_k,s_k)-x_k\|}{\alpha_k}+\frac{1}{\lambda\|\nabla f(x_k)\|}\left(\frac{\|\overline{x_k}(\alpha_k,s_k)-x_k\|}{\alpha_k}\right)^2\\ &\leq \frac{\|x_k(\alpha_k,s_k)-x_k\|}{\alpha_k}+\frac{1}{\lambda}\frac{\|\overline{x_k}(\alpha_k,s_k)-x_k\|}{\alpha_k},\end{aligned}$$

两边取极限, 再利用引理 4.2 得

$$\lim_{k\to\infty}\sup\langle -\nabla f(x_k),v_{k+1}\rangle=0, \tag{4.8}$$

又因为

$$\begin{aligned}\langle -\nabla f(x_k(\alpha_k,s_k)),v_{k+1}\rangle &= \langle\nabla f(x_k)-\nabla f(x_k(\alpha_k,s_k)),v_{k+1}\rangle+\langle -\nabla f(x_k),v_{k+1}\rangle\\ &\leq \|\nabla f(x_k)-\nabla f(x_k(\alpha_k,s_k))\|+\langle -\nabla f(x_k),v_{k+1}\rangle,\end{aligned} \tag{4.9}$$

由引理 4.2 和 $0<\alpha_k\leq\gamma$ 知

$$\lim_{k\to\infty}\|x_k-x_k(\alpha_k,s_k)\|=0, \tag{4.10}$$

利用式 (4.8)、式 (4.9)、式 (4.10) 和 $\nabla f(x)$ 在 Ω 上一致连续得

$$\lim_{k\to\infty}\sup\langle -\nabla f(x_k(\alpha_k,s_k)),v_{k+1}\rangle=0,$$

由式 (4.6) 和 ε 的任意性可得

$$\lim_{k\to\infty}\|\nabla_\Omega f(x_k)\|=0. \tag{4.11}$$

设 $\lim\limits_{k\in N_0,k\to\infty}x_k=x$, 其中 $N_0\subseteq N$, 由引理 2.7(2) 和式 (4.11) 知

$$\|\nabla_\Omega f(x)\|\le\lim_{k\in N_0,k\to\infty}\inf\|\nabla_\Omega f(x_k)\|=0.$$

由引理 2.7(3) 知, $\{x_k\}$ 的任一聚点 x 都是优化问题 (1.1) 的稳定点.

参 考 文 献

[1] Calamai P H, More J J. Projected gradient methods for linearly constrained problems[J]. Mathematical Programming, 1987, 39: 93–116.

[2] 王宜举, 修乃华. 非线性规划理论与算法 [M]. 西安: 山西科学技术出版社, 2008.

[3] Sun Qingying, Wang Changyu, Shi Zhenjun. Global convergence of a modified gradient projection method for convex constrained problems[J]. Acta Mathematicae Applicatae Sinica, 2006, 22(2): 227–242.

[4] 孙清滢, 高宝, 渐令, 王长钰. 约束优化问题的修正共轭梯度投影算法 [J]. 应用数学学报, 2010, 33(4): 640–651.

[5] 时贞军. 改进 HS 共轭梯度算法及其全局收敛性 [J]. 计算数学, 2001, 23(4): 393–406.

[6] 时贞军. 无约束优化的超记忆梯度算法 [J]. 工程数学学报, 2000, 17(2): 99–104.

[7] 王宜举, 江学军. 凸规划问题的一个梯度投影算法 [J]. 曲阜师范大学学报, 1996, 22(3): 7–12.

[8] Wang Changyu, Qu Biao. Convergence of the gradient projection method with a new stepsize rule [J]. Operations Research Transactions, 2002, 6(1): 36–44.

CONJUGATE GRADIENT PROJECTION METHOD OF CONSTRAINED OPTIMIZATION PROBLEMS WITH WOLFE STEPSIZE RULE

JING Shu-jie , ZHAO Hai-yan

(*School of Mathematics and Information Science, Henan Polytechnic University,*
Jiaozuo 454003, China)

Abstract: In this paper, a new conjugate gradient projection algorithm for the constrained optimization problem is presented. By Combining conjugate gradient algorithm with GLP gradient projection theory, global convergence properties of the new algorithm under the Wolfe stepsize rule are proved.

Keywords: constrained optimization problems; conjugate gradient method; GLP gradient projection; Wolfe line search; global convergence.

2010 MR Subject Classification: 90C30