第 28 卷第 3 期　　四川理工学院学报(自然科学版)　　Vol. 28　No. 3
2015 年 6 月　　Journal of Sichuan University of Science & Engineering(Natural Science Edition)　　Jun. 2015
文章编号:1673-1549(2015)03-0086-04　　DOI:10.11863/j. suse. 2015. 03. 18

Hilbert 空间中解凸集约束优化问题的梯度投影算法

杨　丽

(西华师范大学数学与信息学院, 四川　南充　637002)

摘　要:梯度投影算法是求解非线性约束最优化问题的基本方法之一,多年来一直吸引着许多学者对其进行研究。在 Hilbert 空间 H 中,利用梯度投影算法解决有约束条件的凸集 C 上的凸函数 f 的最优问题,引入 CKQ 方法,与以往研究的差异是在定理中新增加了集合 K_n, 并证明了改进的梯度投影算法的强收敛性。所得结果将文献中的梯度投影算法推广为 Ishikawa 形式。

关键词:梯度投影算法;CKQ 方法;强收敛

中图分类号:O224　　　　**文献标志码**:A

引　言

近年来,利用梯度投影算法解决有限制凸集优化的问题受到了广泛的关注[1-6]。设 H 是 Hilbert 空间, C 是 H 的一个非空闭凸子集,考虑有约束条件的凸集 C 上 f 的最优问题:

$$\min_{x\in C} f(x) \tag{1}$$

其中, $f:H\rightarrow R$ 是一实值凸函数,如果 f 是可导的,那么由梯度投影算子生成的序列 $\{x_n\}_{n=0}^{\infty}$, 对于任意的初始迭代点 $x_0\in C$,

$$x_{n+1}=P_C(x_n-\gamma\nabla f(x_n)),n\geqslant 0 \tag{2}$$

或者更一般地

$$x_{n+1}=P_C(x_n-\gamma_n\nabla f(x_n)),n\geqslant 0 \tag{3}$$

其中, γ 和 γ_n 都是正实数。(2)式和(3)式是否收敛取决于梯度函数 ∇f。事实上,如果 ∇f 是利普希兹连续的且是强单调的,即存在 $L>0$ 和 $\alpha>0$, 使得

$$\|\nabla f(x)-\nabla f(y)\|\leqslant L\|x-y\|,\forall x,y\in C \tag{4}$$

且

$$\langle\nabla f(x)-\nabla f(y),x-y\rangle\geqslant \alpha\|x-y\|^2,\forall x,y\in C \tag{5}$$

若 $0<\gamma<\frac{2\alpha}{L^2}$, 那么由算法(2)式产生的迭代序列 $\{x_n\}_{n=0}^{\infty}$ 是强收敛于问题(1)的唯一解的。对于一般的情况,如果序列 $\{\gamma_n\}_{n=0}^{\infty}$ 满足条件:

$$0\leqslant\liminf_{n\to\infty}\gamma_n\leqslant\limsup_{n\to\infty}\gamma_n<\frac{2}{L} \tag{6}$$

则由算法(3)式产生的迭代序列 $\{x_n\}_{n=0}^{\infty}$ 是强收敛于问题(1)的唯一解。但是如果梯度函数 ∇f 不是强单调的,则由算法(2)式或(3)式产生的迭代序列 $\{x_n\}_{n=0}^{\infty}$ 的强收敛性就有可能不成立[7]。为了得到强收敛,Xu[1]结合 CQ 方法进行了变形:

$$\begin{cases}\text{任意 } x_0\in C\\ y_n=P_C(x_n-\gamma_n\nabla f(x_n))\\ C_n=\{z\in C:\|y_n-z\|\leqslant\|x_n-z\|\}\\ Q_n=\{z\in C:\langle x_n-z,x_0-x_n\rangle\geqslant 0\}\\ x_{n+1}=P_{C_n\cap Q_n}x_0\end{cases} \tag{7}$$

并证明了当 ∇f 利普希兹连续且 γ_n 满足条件(6)式时, $x_n\rightarrow P_S x_0$, 这里 S 是问题(1)的解集。

收稿日期:2015-04-21

基金项目:国家自然科学基金项目(11371015)

作者简介:杨　丽(1980-),女,四川大邑人,讲师,硕士,主要从事非线性分析及最优化方面的研究,(E-mail)yangli@cwnu.edu.cn

本文在实 Hibert 空间中改进(7)式,引入了 CKQ 方法,证明了改进的梯度投影算法的强收敛性。本文中,H 都为 Hilbert 空间。

1　预备知识

定义 1　令 C 是 H 的非空闭凸子集,映象 $T: C \to C$ 称为非扩张映象,如果对任意的 $x,y \in C$,都有 $\|Tx - Ty\| \leqslant \|x - y\|$, $x \in C$ 被称为 T 的一个不动点,如果 $Tx = x$,用 $F(T)$ 表示 T 的不动点的集合,即 $F(T) = \{x \in C: Tx = x\}$。

定义 2　映射 $T: H \to H$ 称为平均映射,当且仅当它能用恒等映射与非扩张映射表示,即 $T = (1-\alpha)I + \alpha U$,这里 $\alpha \in (0,1)$,且 $U: H \to H$ 是非扩张的。更精确地,当该式满足时,T 是 α—平均映射,那么固定非扩张映射(特别地,投影算子)是 $\frac{1}{2}$ 平均映射。

定义 3　假设 $\gamma > 0$, T 是 γ—逆强单调($\gamma - \mathrm{ism}$),当且仅当 $\langle x - y, Ax - Ay\rangle \geqslant \gamma\|Ax - Ay\|^2$, $x,y \in H$。

引理 1[8]　对任意的 $u,v \in H$,有:$\|u - v\|^2 = \|u\|^2 - \|v\|^2 - 2\langle u - v, v\rangle$。

引理 2[8]　设 C 是 H 的非空闭凸子集,T:是 $C \to C$ 的一非扩张映象。如果序列 $\{x_n\}_{n=0}^{\infty} \subset C$,且满足 $x_n \to z$(弱)和 $x_n - Tx_n \to 0$,那么 $z \in F(T)$。

引理 3[9]　令 C 是 H 的非空闭凸子集,点 $x,y,z,w \in H$, a 是实数,那么集合

$$D = \{v \in C: \|y - v\|^2 \leqslant \|x - v\|^2 + \langle w, v\rangle + a\}$$

是闭凸集。

引理 4　令 C 是 H 的非空闭凸子集,P_C 表示从 H 到 C 上的投影,即对 $\forall x \in H$,有 $\|x - P_C x\| = \inf\{\|x - z\|, z \in C\}$,给定 $x \in H$ 和 $z \in C$,那么 $z = P_C x$ 的充要条件是 $< x - z, y - z > \leqslant 0$ 或 $\|x - z\|^2 \leqslant \|x - y\|^2 - \|y - z\|^2$,对任意的 $y \in C$ 都成立。

引理 5[10]　令 C 是 H 的非空闭凸子集,序列 $\{x_n\}_{n=0}^{\infty} \subset H$, $u \in H$,记 $q = P_C u$。如果 $\{x_n\}_{n=0}^{\infty}$ 的弱 ω -极限集合 $\omega_w(x_n) = \{x: \exists x_{n_j} \to x\} \subset C$,且满足对所有的 n 有 $\|x_n - u\| \leqslant \|u - q\|$ 成立,那么有 $x_n \to q$。

2　主要结果

定理 1　假设最优问题(1)是可解的,令 S 是它的解集。假设 ∇f 满足条件(4)式,并且满足 $0 \leqslant \gamma < \frac{1}{L}$,给定任意 $x_0 \in C$,定义序列 $\{x_n\}_{n=0}^{\infty} \subset C$:

$$\begin{cases} y_n = P_C(x_n - \gamma\nabla f(x_n)) \\ z_n = P_C(x_n - \gamma\nabla f(y_n)) \\ C_n = \{z \in C: |y_n - z| \leqslant |x_n - z|\} \\ K_n = \{z \in C: |z_n - z| \leqslant |x_n - z|\} \\ Q_n = \{z \in C: \langle x_n - z, x_0 - x_n\rangle \geqslant 0\} \\ x_{n+1} = P_{C_n \cap K_n \cap Q_n} x_0, \quad \forall n \geqslant 1 \end{cases} \tag{8}$$

那么有 $x_n \to P_S x_0$。(当 $n \to \infty$)。

证明　由引理 3 知,C_n 和 K_n 是凸集。由于 ∇f 满足利普希兹条件,则 ∇f 是 $\frac{1}{L} - \mathrm{ism}$[1],则 $\gamma\nabla f$ 是 $\frac{1}{\gamma L} - \mathrm{ism}$,易证 $I - \gamma\nabla f$ 是 $\frac{\gamma L}{2}$ 平均算子,$P_C(I - \gamma\nabla f)$ 是 $\frac{2+\gamma L}{4}$—平均算子。令 $V = P_C(I - \gamma\nabla f)$,那么 V 是一个非扩张映射。

先证 $S \subset C_n \cap K_n \cap Q_n$,对任意的 $p \in S$,因 $Vp = p$,即得到

$$\|y_n - p\| = \|Vx_n - Vp\| \leqslant \|x_n - p\| \tag{9}$$

从而 $p \in C_n$,对所有的 $n \geqslant 0$,因此 $S \subset C_n$。又因 $z_n = P_C(x_n - \gamma\nabla f(y_n))$ 及 ∇f 的利普希兹连续性,由引理 4 并结合引理 1 有

$$\begin{aligned} &\|z_n - p\|^2 \leqslant \|x_n - \gamma\nabla f(y_n) - p\|^2 - \\ &\|x_n - \gamma\nabla f(y_n) - z_n\|^2 = \\ &\|x_n - p\|^2 - \|x_n - z_n\|^2 + 2\gamma\nabla f(y_n), p - z_n\rangle = \\ &\|x_n - p\|^2 - \|x_n - z_n\|^2 + 2\gamma(\langle\nabla f(y_n) - \\ &\nabla f(p), p - y_n\rangle + \langle\nabla f(p), p - y_n\rangle + \\ &\langle\nabla f(y_n), y_n - z_n\rangle \leqslant \\ &\|x_n - p\|^2 - \|x_n - z_n\|^2 + 2\gamma\langle\nabla f(y_n), y_n - z_n\rangle = \\ &\|x_n - p\|^2 - \|x_n - y_n\|^2 - 2\langle x_n - y_n, y_n - z_n\rangle - \\ &\|y_n - z_n\|^2 + 2\gamma\langle\nabla f(y_n), y_n - z_n\rangle = \\ &\|x_n - p\|^2 - \|x_n - y_n\|^2 - \|y_n - z_n\|^2 + \\ &2\langle x_n - \gamma\nabla f(y_n) - y_n, z_n - y_n\rangle \end{aligned} \tag{10}$$

又因为

$$\begin{aligned} &\langle x_n - \gamma\nabla f(y_n) - y_n, z_n - y_n\rangle = \\ &\langle x_n - \gamma\nabla f(x_n) - y_n, z_n - y_n\rangle + \end{aligned}$$

$$\langle \gamma \nabla f(x_n) - \gamma \nabla f(y_n), z_n - y_n \rangle \leqslant$$
$$\langle \gamma \nabla f(x_n) - \gamma \nabla f(y_n), z_n - y_n \rangle \leqslant$$
$$\gamma L \|x_n - y_n\| \|z_n - y_n\| \quad (11)$$

结合(10)式和(11)式,得到:

$$\|z_n - p\|^2 \leqslant \|x_n - p\|^2 - \|x_n - y_n\|^2 - \|y_n - z_n\|^2 + 2\gamma L \|x_n - y_n\| \|z_n - y_n\| \leqslant$$
$$\|x_n - p\|^2 - \|x_n - y_n\|^2 - \|y_n - z_n\|^2 + \gamma^2 L^2 \|x_n - y_n\|^2 + \|z_n - y_n\|^2 =$$
$$\|x_n - p\|^2 + (\gamma^2 L^2 - 1)\|x_n - y_n\|^2 \leqslant \|x_n - p\|^2$$

从而 $\|z_n - p\| \leqslant \|x_n - p\|$, 这表明 $p \in K_n$, 因此 $S \subset K_n$。

证明 $S \subset Q_n$, 对 n 用数学归纳法:(1)当 $n = 0$ 时, $S \subset Q_0 = C$, 显然成立。(2)假设 $S \subset Q_n$, 需要证明 $S \subset Q_{n+1}$。由序列 $\{x_n\}_{n=0}^{\infty}$ 的定义 $x_{n+1} = P_{C_n \cap K_n \cap Q_n} x_0$ 以及引理4,有: $< x_{n+1} - z, x_0 - x_{n+1} > \geqslant 0$ 对任意的 $z \in C_n \cap K_n \cap Q_n$ 成立。因为 $S \subset C_n \cap K_n \cap Q_n$, 从而对所有的 $z \in S$, $< x_{n+1} - z, x_0 - x_{n+1} > \geqslant 0$ 也成立。又由 Q_n 的定义可知 $z \in Q_{n+1}$, 从而有 $S \subset Q_{n+1}$, 这表明 $S \subset Q_n$, 得到 $S \subset C_n \cap K_n \cap Q_n$。

注意到

$$Q_n = \{z \in C: < x_n - z, x_n - x_0 > \leqslant 0\}$$

即说明 $x_n = P_{Q_n} x_0$, 结合 $S \subset Q_n$, 得 $\|x_n - x_0\| \leqslant \|t - x_0\|$, 对所有的 $t \in S$。

特别地,有

$$\|x_n - x_0\| \leqslant \|q - x_0\| \quad (12)$$

这里 $q = P_S x_0$, 也就是 $\{x_n\}_{n=0}^{\infty}$ 有界,从而由集合 C_n 的定义知 $\{y_n\}$ 也有界。

证明 $\|x_{n+1} - x_n\| \to 0$。注意到 $Q_n = \{z \in C: \langle x_n - z, x_n - x_0 \rangle \leqslant 0\}$, 即说明 $x_n = P_{Q_n}(x_0)$, 因为 $x_{n+1} \in C_n \cap K_n \cap Q_n \subset Q_n$, 隐含着 $\|x_0 - x_n\| \leqslant \|x_0 - x_{n+1}\|$, 这表明 $\{\|x_n - x_0\|\}_{n=0}^{\infty}$ 是递增的,再由集合 C 是有界的,从而其极限存在。由 $x_n = P_{Q_n}(x_0)$ 以及 $x_{n+1} \in Q_n$, 故 $\langle x_{n+1} - x_n, x_n - x_0 \rangle \geqslant 0$, 根据引理1有

$$0 \leqslant \|x_{n+1} - x_n\|^2 = \|(x_{n+1} - x_0) - (x_n - x_0)\|^2 =$$
$$\|x_{n+1} - x_0\|^2 - \|x_n - x_0\|^2 - 2 < x_{n+1} - x_n, x_n - x_0 > \leqslant$$
$$\|x_{n+1} - x_0\|^2 - \|x_n - x_0\|^2 \to 0, \text{ 当 } n \to \infty \text{ 。}$$

又注意到 $x_{n+1} \in C_n$, 从而有

$$\|x_{n+1} - y_n\| \leqslant \|x_{n+1} - x_n\| \to 0$$

得到

$$\|x_n - y_n\| \leqslant \|x_{n+1} - x_n\| + \|x_{n+1} - y_n\| \to 0 \quad (13)$$

证明 $\omega_\omega(x_n) \subset S$。让 $\tilde{x} \in \omega_\omega(x_n)$, 取 $\{x_n\}_{n=0}^{\infty}$ 的子集 $\{x_{n_j}\}_{j=1}^{\infty}$, 且 $x_{n_j} \xrightarrow{\text{弱}} \tilde{x}$。令 $V = P_C(I - \gamma \nabla f)$, 那么 V 是非扩张的。注意到 $y_n = P_C(x_n - \gamma \nabla f(x_n))$, 结合(13)式得到

$$\|x_{n_j} - Vx_{n_j}\| \leqslant \|x_{n_j} - y_{n_j}\| + \|y_{n_j} - Vx_{n_j}\| =$$
$$\|x_{n_j} - y_{n_j}\| + \|P_C(x_{n_j} - \gamma \nabla f(x_{n_j})) - P_C(I - \gamma \nabla f) x_{n_j}\| \leqslant$$
$$\|x_{n_j} - y_{n_j}\| \to 0, (n \to \infty)$$

由引理2, $\tilde{x} \in F(V)$, 又因为 $F(V) = S$, 所以, $\tilde{x} \in S$, 从而 $\omega_\omega(x_n) \subset S$。

最后结合(12)式和引理5 得到 $x_n \to q(n \to \infty)$, 即 $x_n \to P_S x_0$, (当 $n \to \infty$)。

3 结束语

本文主要研究了 Hilbert 空间中解凸集约束优化问题的梯度投影算法,并证明了算法的强收敛性,所得结果将文献[1]中的梯度投影算法推广为 Ishikawa 的形式。与以往研究不同的是,在定理 1 中新增加了一个集合 K_n, 这样的好处是使收敛速率有所提高,今后可以考虑在收敛速率这方面作进一步的研究。

参 考 文 献:

[1] Xu H K.Averaged Mappings and the Gradient-Projection Algorithm[J]. J. Optim. Theory. Appl, 2011, 150 (2): 360-378.

[2] Su M,Xu H K.Remarks on the gradient-Projection algorithm[J].J.Nonl.Anal.Optim,2011,1(1):35-43.

[3] Ceng L C,Guu S M,Yao J C.Hybrid methods with regularization for minimization problems and asymptotically pseudocontractive mappings in the intermediate sense[J]. J Glob Optim,2014,60(4):617-637.

[4] Ryu S,Chen A,Choi K.A modified gradient projection algorithm for solving the elastic demand traffic assignment problem[J]. Computer& Operations Research,2014,1 (47):61-71.

[5] Liu Z Y,Wei Z H,Sun W Y.An iteratively approximated gradient projection algorithm for sparse signal recon-

struction[J].Applied Mathematics and computation,2014, 2(228):454-462.

[6] Ceng L C,Guu S M,Yao J C.Hybrid methods with regularization for minimization problems and asymptotically pseudocontractive mappings in the intermediate sense[J]. J Glob Optim,2014,60(4):617-637.

[7] Levitin E.S,Polyak B T.Constrained minimization methods[J].Zh.Vychisl.Mat.Fiz,1966,6:787-823.

[8] Goebel K,Kirk W A.Topics in Metric Fixed Point Theory[M].Cambridge:Cambridge University Press,1990.

[9] Halpern B.Fixed points of nonexpanding maps[J].Bull. Am.Math.Soc,1967(73):957-961.

[10] Baillon J B,Haddad G.Quelques proprietes des operateurs angle-bornes et n-cycliquement monotones[J]. Israel J.Math,1977,26(2):137-150.

Gradient Projection Algorithms for Solution Convex Sets Constraints Optimization Problem In Hilbert Space

YANG Li

(School of Mathematics and Information, China West Normal University, Nanchong 637002, China)

Abstract: The gradient projection operator is one of the basic approaches for solving nonlinear constrained optimization problem, so it has been attracting many scholars to research. In Hilbert space, the gradient projection algorithm is used to solve optimization problems of convex function f on convex set with constraint condition, and the CKQ method is introduced, a set K_n is added in theorem which is different from previous study, and the strong convergence of improved gradient projection algorithm is proved. The obtained results make the gradient projection algorithm in literatures generalized to be Ishikawa form.

Key words: gradient projection operator; CKQ method; strong convergence